

SPECIAL

n°50

Protéomique et

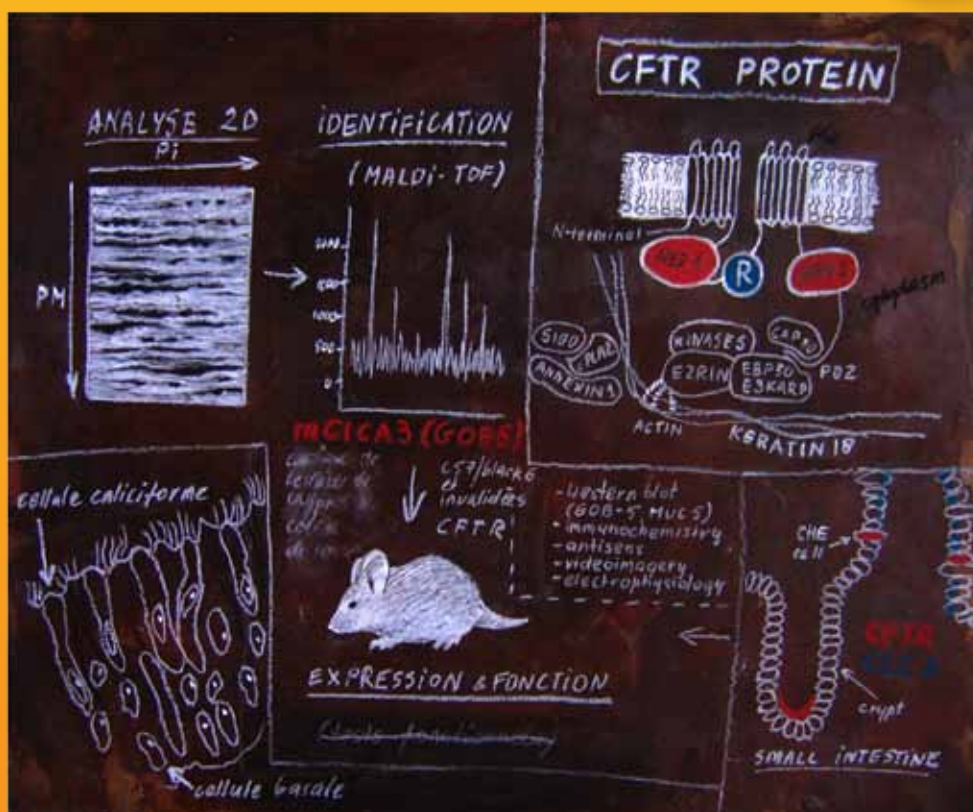
Maladies Rares

sfeap

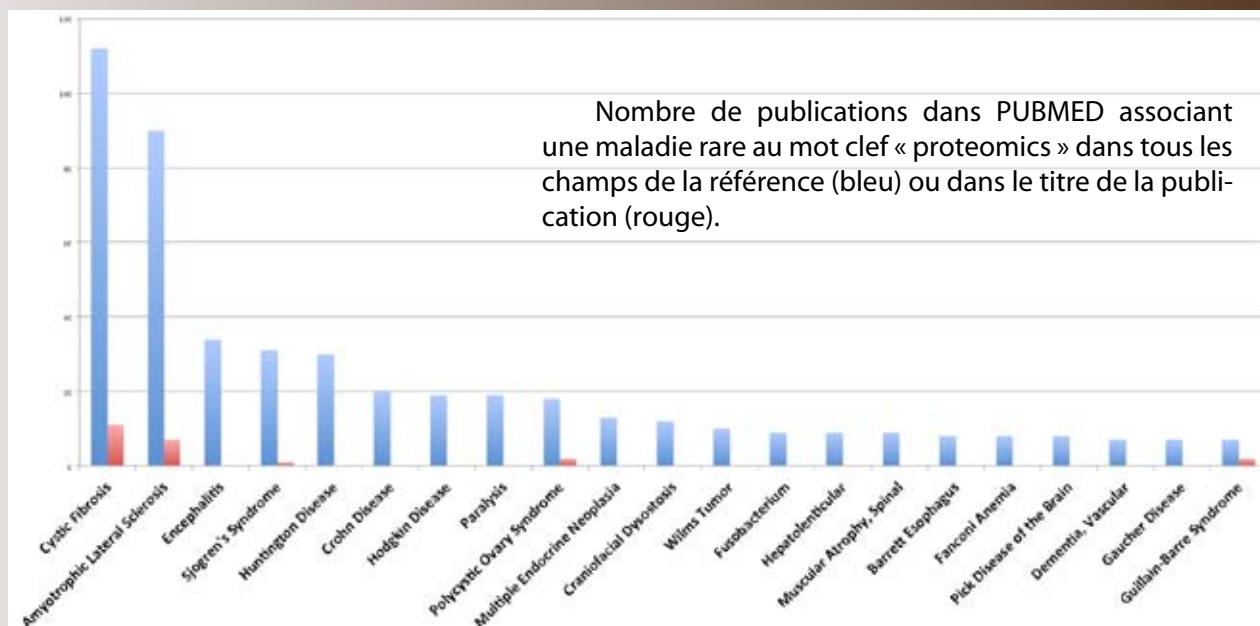
société française d'électrophorèse  
& d'analyse protéomique



# Le Mag



*Les textes publiés via la SFEAP engagent la  
responsabilité de leurs seuls auteurs, et  
restent leur propriété pleine et entière*



Ce numéro spécial est consacré aux maladies rares. Parmi les 184 maladies rares répertoriées ([www.hon.ch/HONselect/RareDiseases/](http://www.hon.ch/HONselect/RareDiseases/)), seulement quelques maladies ont été associées à des études protéomiques comme le montre la figure ci-dessus. Seulement 24 publications associent une maladie rare et le mot « proteomics » dans le titre de l'article.

Bonne lecture de ce numéro spécial.

Luc Camoin

<i>Le Mot de la présidente et editos</i>	page 2
<i>Editos</i>	page 3
<i>Aviesan</i>	page 4
<i>La protéomique : état de l'art et perspectives - Odile Schiltz</i>	page 5
<i>Quels échantillons biologiques pour l'analyse protéomique dans l'étude des maladies rares ? -Thierry Rabilloud</i>	page 7
<i>Bases de données et ressources pour la protéomique - Lydie Lane</i>	page 9
<i>La protéomique au service de la mucoviscidose : l'identification d'une nouvelle cible thérapeutique pour la mutation <math>\Delta F508CFTR</math>. - Aleksander Edelman</i>	page 10
<i>Analyse protéomique du muscle dystrophique canin et évaluation de nouvelles stratégies thérapeutiques - Laetitia Guével et Karl Rouger</i>	page 11
<i>Stratégies d'Analyse des Assemblages Supramoléculaires de Protéines - Marc Bonneu</i>	page 12
<i>Interactome in vivo - Romain Roncagalli</i>	page 13
<i>Méthodes de fouille et outils de visualisation pour l'analyse de données "omics" - Yves Vandembrouck</i>	page 14
<i>Imagerie par spectrométrie de masse MALDI : applications, limites et potentiel - Charles Pineau et Daniel Lafitte</i>	page 15
<i>Analyse protéomique de fluides biologiques - Anne Gonzalez de Peredo</i>	page 16
<i>Plate-formes Protéomiques labellisées IBISA</i>	page 18
<i>Nos Partenaires</i>	page 19

## Le mot de la Présidente



Une des missions de la SFEAP est de soutenir sous différentes formes les actions visant à promouvoir la protéomique. Notre magazine est à ce titre un moyen de communication efficace au sein de notre communauté mais il peut également servir de vecteur de diffusion de l'information vers d'autres communautés.

Ainsi, je suis très heureuse que la SFEAP édite ce numéro spécial du "Mag de la SFEAP", qui sera diffusé largement, et qui sera une mémoire du Workshop "Protéomique et Maladies rares", qui s'est tenu à Paris le 25 septembre 2012. Cette manifestation est la résultante d'une volonté de rapprochement de deux communautés scientifiques désireuses de travailler ensemble. Elle a été organisée par l'ITMO « Génétique, génomique et bio-informatique » et l'ITMO « Bases moléculaires et structurales du vivant » de l'alliance nationale AVIESAN (Alliance nationale pour les sciences de la vie et de la santé) en partenariat avec la Fondation maladies rares. Chaque communauté a présenté ses problématiques et ses objectifs. Une large place a ensuite été laissée à la discussion. Je pense que cette journée a été très enrichissante et j'espère qu'elle permettra d'initier de nouvelles collaborations.

Je vous souhaite une bonne lecture.

Odile Schiltz  
Présidente de la SFEAP

## Editos

Le Workshop « Protéomique et Maladies rares » qui s'est tenu à Paris au mois de septembre dernier a réuni des membres de deux communautés peu habituées à interagir : d'une part des spécialistes d'une discipline à forte composante technologique, la Protéomique, et d'autre part des équipes de recherche actives dans le domaine des Maladies rares. Cet événement, parrainé par l'alliance AVIESAN, avait pour origine une demande forte de ces deux communautés qui était, pour l'une, de mieux appréhender l'offre technologique offerte par la Protéomique, et pour l'autre de cerner les défis qui existent dans le domaine des Maladies rares afin d'évaluer dans quelle mesure les outils de la Protéomique pourraient permettre des avancées significatives dans ce domaine. Cette journée a été ponctuée de présentations didactiques des technologies, de quelques exemples de projets protéomiques en cours dans le domaine de maladies rares, et enfin d'une table ronde au cours de laquelle diverses problématiques ont pu être discutées. Ce premier Workshop a eu le succès escompté et nous espérons qu'il aura permis de catalyser de nouveaux projets de recherche interdisciplinaires.

Elisabeth Tournier-Lasserve  
Jérôme Garin  
Charles Pineau

### **Pourquoi un numéro spécial du Mag ?**

Les membres de la SFEAP se sont fortement investis dans la journée scientifique « Protéomique et maladies rares », tant pour son organisation que comme conférenciers ou participants. Cependant, les enjeux scientifiques et médicaux évoqués à cette occasion nous ont semblé mériter un début d'approfondissement et surtout une diffusion qui ne se limite pas aux scientifiques qui ont pu participer à cette journée. C'est là tout l'objectif de ce numéro spécial du Mag, et nous vous en souhaitons bonne lecture.

Thierry Rabilloud  
Charles Pineau

**Programme et téléchargement des présentations sur**

[www.aviesan.fr/aviesan/accueil/toute-l-actualite/workshop-proteomique-et-maladies-rares](http://www.aviesan.fr/aviesan/accueil/toute-l-actualite/workshop-proteomique-et-maladies-rares)



Créée en avril 2009, l'Alliance pour les Sciences de la vie et de la Santé (Aviesan) rassemble les grands opérateurs de la recherche en sciences de la vie et de la santé en France (<http://www.aviesan.fr/fr>). Aviesan est organisée en dix instituts thématiques multiorganismes (ITMO) couvrant l'ensemble des grands domaines des sciences de la vie et de la santé. Un ITMO correspond à une communauté d'animation et de coordination scientifique analogue à une société savante. Les objectifs principaux des ITMOs sont de mettre en place une politique scientifique globale, concertée et coordonnée d'appui à la recherche dans le domaine concerné.

L'Institut Thématique Multi-Organismes (ITMO) Bases Moléculaires et Structurales du Vivant (BMSV) s'intéresse à l'étude des systèmes biologiques, de leur dynamique, des interactions et interconversions, du niveau moléculaire au niveau cellulaire. Elle inclut aussi la compréhension de l'organisation de la cellule et l'identification de composés naturels ou synthétiques permettant d'interférer avec son fonctionnement.

La protéomique est par conséquent l'un des domaines couverts par l'ITMO BMSV. A la suite du séquençage du génome humain, il a été lancé un programme similaire pour le protéome humain (The Human Proteome Project, HPP). L'ITMO BMSV a ainsi soutenu le colloque inaugural de HPP France organisé à Paris en juin 2011. Lors de cette journée, la communauté "Maladies rares" a sollicité la communauté des protéomistes pour une collaboration renforcée. L'ITMO BMSV s'est donc fait porteur de l'organisation d'un workshop « Protéomique et maladies rares » en partenariat avec l'ITMO GGB.

L'Institut Thématique Multi-Organismes (ITMO) Génétique, génomique et bioinformatique (GGB) couvre la génétique, la génomique de tous les organismes vivants depuis les virus, microorganismes, plantes, animaux, jusqu'à l'Homme. Les domaines couverts vont de la recherche fondamentale abordant les problèmes de l'organisation, stabilité, variation du matériel génétique et de la régulation de l'expression et de l'évolution des génomes, à la génétique des populations, jusqu'à la recherche translationnelle pour la mise en application médicale de ces recherches.

L'ITMO GGB couvre notamment le domaine des maladies rares, dont 80% sont d'origine génétique et qui implique une approche pluridisciplinaire associant des équipes de recherche clinique, génétique, physiopathologique, thérapeutique et/ou en sciences humaines et sociales. Le nombre très important de ces affections (plusieurs milliers) et leur hétérogénéité nécessitent un effort spécifique, mené sur le long terme, et une coordination des actions menées dans ce domaine. Les Plans Nationaux Maladies Rares (2005-2008 et 2011-2014) ont permis de donner une impulsion très forte à ce domaine, notamment dans le cadre du volet recherche avec la création en février 2012 de la Fondation maladies Rares. C'est donc en interaction étroite avec cette Fondation et les autres ITMO concernés que l'ITMO GGB mène des actions dans un domaine où les défis restent importants : il existe en effet une très grande hétérogénéité de l'état des recherches selon les maladies concernées et très peu d'entre elles bénéficient d'approches thérapeutiques. **Une grande réactivité vis à vis des nouveaux outils et approches technologiques est indispensable pour faire face à ces défis.** L'objectif de ce colloque était donc de rapprocher les communautés « Maladies rares » et « Protéomique » en France afin de voir dans quelle mesure certaines problématiques de recherche sur les maladies rares pourraient bénéficier de l'apport de la Protéomique.

Dominique Daegelen  
Directrice adjointe ITMO GGB

Jean-Claude Michalski  
Directeur ITMO BMSV

# La protéomique : état de l'art et perspectives

Odile Schiltz

Institut de Pharmacologie et de Biologie Structurale (IPBS), UMR 5089, CNRS, Université Paul Sabatier, Toulouse  
(Odile.Schiltz@ipbs.fr)

Cet article introductif a pour objectif de donner une vue d'ensemble des approches protéomiques aujourd'hui. Il s'agira de poser les bases de l'analyse protéomique, de décrire ses potentialités et ses limitations, et d'évoquer ses différents domaines d'applications.

La définition de la protéomique n'a que peu évolué depuis une dizaine d'années. Elle vise à identifier et quantifier un ensemble donné de protéines mais également à déterminer leur localisation, leurs modifications et leurs interactions pour finalement accéder à leur fonction (1). Pour cela, diverses approches sont mises en oeuvre, qui ont quant à elles évolué de façon impressionnante et constante, et pour lesquelles des défis majeurs restent encore à surmonter malgré les progrès déjà remarquables. Un protéome étant hautement dynamique, les enjeux résident dans la mesure la plus précise et reproductible possible des variations qualitatives et quantitatives des protéines entre plusieurs conditions afin de comprendre les processus biologiques et de définir les états pathologiques de la cellule.

Le schéma expérimental des approches protéomiques basées sur la spectrométrie de masse (MS) peut être décomposé en trois grandes étapes : la préparation de l'échantillon, l'analyse par MS et l'analyse bioinformatique des données. Chacune de ces étapes est critique pour l'obtention de données exploitables de façon pertinente pour la compréhension des systèmes biologiques et sera définie en fonction de l'objectif poursuivi. Ainsi, deux grandes voies se dégagent aujourd'hui. On distingue, d'une part, une protéomique dite globale, sans a priori, qui vise à caractériser et quantifier un maximum de protéines dans un système donné, et d'autre part, une protéomique dite ciblée, qui est orientée vers le dosage de protéines connues d'intérêt. La première approche concerne principalement des études qualifiées de découverte, alors que la deuxième approche est tournée vers des études de validation. Les approches protéomiques peuvent être appliquées à des échantillons biologiques de diverses natures (cellules en culture, fluides biologiques, tissus), pour lesquels des protocoles spécifiques seront utilisés pour extraire les protéines à étudier. Ces protéines sont ensuite digérées à l'aide d'une protéase telle que la trypsine en peptides, qui sont analysés par couplage LC-MS/MS (chromatographie liquide-MS en tandem) permettant d'obtenir des informations sur l'abondance (spectre MS) et sur la séquence des peptides (spectre MS/MS).

Dans le cas des analyses globales, des spectromètres de masse très sensibles, très résolutifs et à haute vitesse d'acquisition sont généralement utilisés. Grâce aux progrès constants de l'instrumentation, des dizaines de milliers de spectres MS et MS/MS sont aujourd'hui enregistrés en une seule analyse. Cette énorme masse de données est ensuite traitée à l'aide de plusieurs outils bioinformatiques dédiés conduisant à l'identification et la quantification de plus d'un millier de protéines. Ce chiffre peut être significativement augmenté en ajoutant des étapes de fractionnement

préalable de l'échantillon ou bien en allongeant les gradients chromatographiques utilisés, ce qui permet de décomplexifier le mélange peptidique à analyser. Des études globales récentes ont ainsi dépassé le seuil des 5000 protéines identifiées et quantifiées (2,3).

Les études ciblées utilisent une instrumentation également très sensible permettant de plus une grande sélectivité des espèces d'intérêt à quantifier. Pour cela, un mode d'acquisition particulier est utilisé, appelé SRM (Selected Reaction Monitoring) ou MRM (Multiple Reaction Monitoring), qui consiste à sélectionner une (SRM) ou plusieurs (MRM) transitions de masse caractéristiques d'un peptide d'intérêt (4). Ces transitions sont définies par un couple de masse (masse de l'ion peptidique d'intérêt et masse d'un ion fragment caractéristique de sa séquence) et sont sélectionnées pour des peptides tryptiques protéotypiques, signature de la protéine d'intérêt. Des peptides analogues isotopiquement alourdis sont ensuite utilisés comme standards pour la quantification. Ces méthodes atteignent aujourd'hui la sensibilité des tests ELISA pour certaines protéines (5). Elles ont de plus l'avantage de pouvoir être multiplexées et mises en oeuvre pour le dosage de protéines modifiées (6,7).

Les applications des approches protéomiques sont nombreuses. On peut citer comme principaux objectifs la compréhension des mécanismes moléculaires mis en jeu par la cellule dans différents contextes, la détermination de réseaux d'interactions protéiques, la caractérisation de voies de signalisation, la découverte et la validation de candidats biomarqueurs. Dans chaque cas, des stratégies dédiées seront mises en place pour répondre le plus efficacement possible à la question biologique posée (8).

Par exemple, l'étude de certaines voies de signalisation nécessitera l'étude d'un sous-protéome particulier comme le phosphoprotéome (9,10). Des méthodes d'enrichissement spécifique des protéines et/ou des peptides modifiés seront alors utilisées afin d'accéder à la caractérisation des protéines d'intérêt (identification, position de la modification, quantification), qui peuvent être présentes en faible abondance dans l'échantillon. En effet, la phosphorylation est une modification dynamique, qui peut conduire à des taux de protéines phosphorylées très faibles. Les quantités de matériel disponibles constitueront ainsi dans ces études un facteur déterminant pour leur réussite.

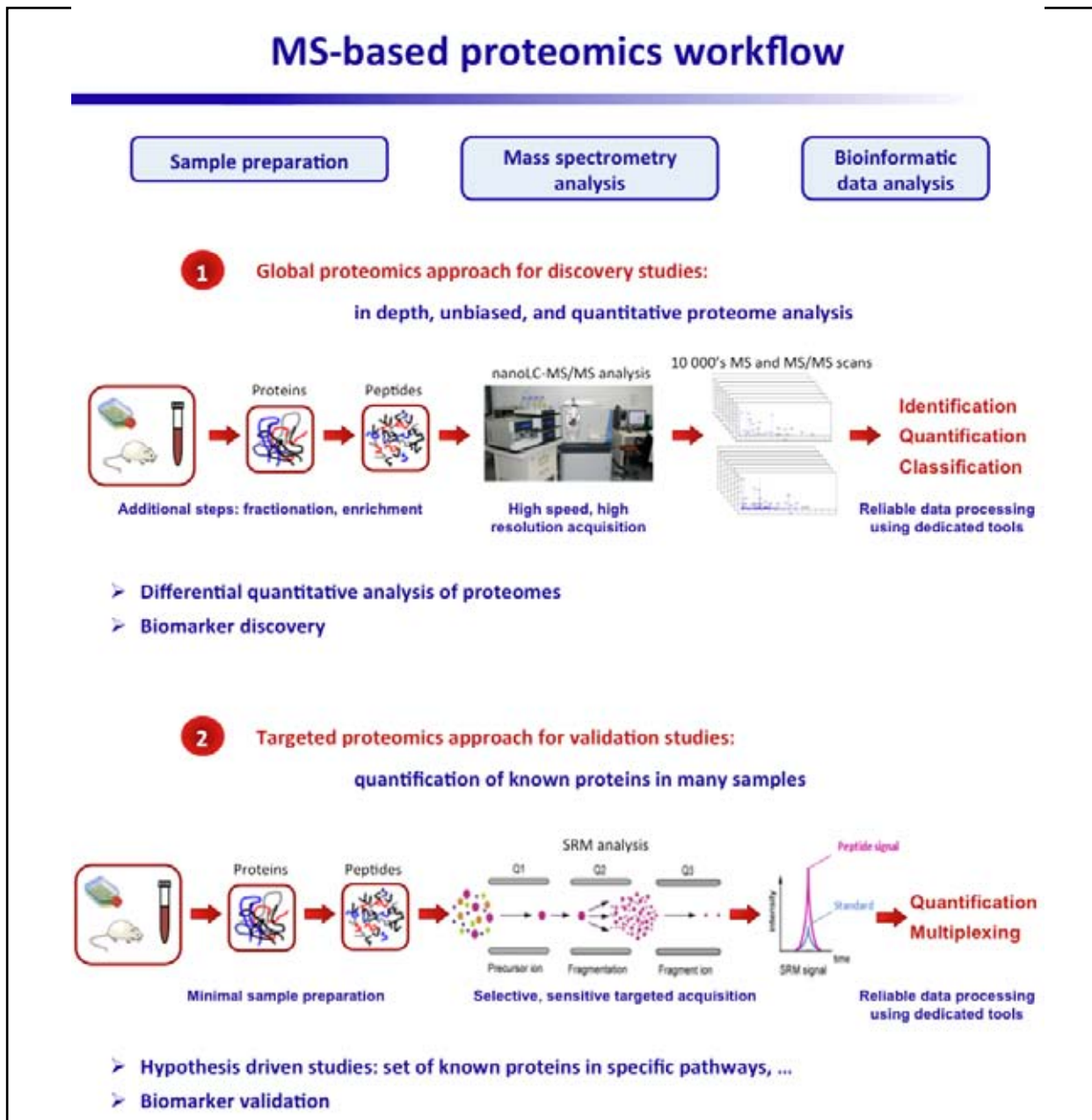
Un autre exemple d'application concerne l'analyse de l'interactome d'une protéine d'intérêt afin d'appréhender sa fonction (11,12). Dans ce cas, l'efficacité des étapes d'isolement du complexe, en amont de l'analyse par MS, sera capitale. En effet, la difficulté majeure de ces analyses réside dans la détermination des vrais interactants et l'élimination des protéines contaminantes. Ainsi, en complément de l'optimisation des étapes de purification des complexes, le choix d'un contrôle adéquat peut s'avérer critique. En effet, ce contrôle sera utilisé pour l'analyse quantitative des données protéomiques afin

# La protéomique : état de l'art et perspectives (suite)

d'éliminer les associations non spécifiques. Ces approches sont aujourd'hui très efficaces et peuvent être appliquées à l'étude de différents types de complexes. Des méthodes de réticulation des complexes in vivo peuvent également être ajoutées afin de stabiliser les complexes en amont de leur purification, ce qui facilite ensuite l'identification des partenaires les plus labiles et/ou transitoirement associés (13).

## Références

- (1) Fields S. *Proteomics in genomeland*. *Science*, 2001, 291, 1221-1224.
- (2) Boisvert F.M. et al. *Quantitative spatial proteomics analysis of proteome turnover in human cells*. *Mol Cell Proteomics* 2012, 11, M111.011429.
- (3) Thakur S.S. et al. *Deep and highly sensitive proteome coverage by LC-MS/MS without prefractionation*. *Mol Cell Proteomics* 2011, 10, M110.003699.
- (4) Lange V. et al. *Selected reaction monitoring for quantitative proteomics: a tutorial*. *Mol. Syst. Biol.* 2008, 4, 222.
- (5) Fortin T. et al. *Clinical quantitation of prostate-specific antigen biomarker in the low nanogram/milliliter range by conventional bore liquid chromatography-tandem mass spectrometry (multiple reaction monitoring) coupling and correlation with ELISA tests*. *Mol Cell Proteomics* 2009, 8, 1006-1015.
- (6) Hüttenhain R. et al. *Perspectives of targeted mass spectrometry for protein biomarker verification*. *Curr. Opin. Chem. Biol.* 2009, 13, 518-525.
- (7) Picotti P. and Aebersold R. *Selected reaction monitoring-based proteomics: workflows, potential, pitfalls and future directions*. *Nature Methods*, 2012, 9, 555-566.
- (8) Mallick P. and Kuster B. *Proteomics: a pragmatic perspective*. *Nature Biotechnol.*, 2010, 28, 695-709.
- (9) Stasyk T. and Huber A. *Mapping in vivo signal transduction defects by phosphoproteomics*. *Trends in Mol. Med.* 2012, 18, 43-51.
- (10) Rigbolt K.T.G. and Blagoev B. *Quantitative phosphoproteomics to characterize signaling networks*. *Semin. Cell Dev. Biol.*, 2012, 23, 863-871.
- (11) Kaake R.M. et al. *Profiling of protein interaction networks of protein complexes using affinity purification and quantitative mass spectrometry*. *Mol. Cell Proteomics* 9, 1650-65 (2010).
- (12) Trinkle-Mulcahy L. *Resolving protein interactions and complexes by affinity purification followed by label-based quantitative mass spectrometry*. *Proteomics* 2012, 12, 1623-1638.
- (13) Fabre B. et al. *Subcellular distribution and dynamics of active proteasome complexes unraveled by a workflow combining in vivo complex cross-linking and quantitative proteomics*. *Mol. Cell Proteomics*, 2013, 12, M112.023317.



## Quels échantillons biologiques pour l'analyse protéomique dans l'étude des maladies rares ?

Thierry Rabilloud

IRTSV/LCBM, UMR CNRS 5249 - CEA-Grenoble - 17, rue des martyrs 38054 Grenoble Cedex 9  
(thierry.rabilloud@cea.fr)

Le but de ce bref article est de rappeler quelques points fondamentaux qui doivent guider les chercheurs souhaitant utiliser l'analyse protéomique pour étudier des maladies rares. Certains points pourront apparaître comme des truismes, et l'indulgence du lecteur est sollicitée sur ce point.

Au rebours des analyses génomiques, l'analyse protéomique est très sensible au contexte biologique et à la physiologie cellulaire. Cette sensibilité en explique la valeur dans un but pathognomonique, mais introduit en retour des contraintes particulières sur le plan du choix des échantillons, en particulier dans le cas de l'étude de pathologies.

Même si les techniques modernes d'analyse protéomique permettent de creuser de plus en plus profondément dans les protéomes, explorant une part toujours plus grande de la gamme d'expression des protéines qui est elle-même de mieux en mieux connue [1], l'analyse protéomique reste très sensible aux variations dans les grandes fonctions cellulaires comme le métabolisme, l'architecture cellulaire (cytosquelette) et la gestion des protéines (production, repliement, dégradation). Cette sensibilité introduit des contraintes particulières quand il s'agit d'étudier une maladie qui touche un tissu complexe et/ou une maladie qui met en jeu une altération progressive du tissu, que ce soit une dégénérescence ou une inflammation invasive. Dans tous ces cas, le problème posé va être celui de l'homogénéité des prélèvements et donc de la dispersion des mesures de protéines qui vont être ensuite réalisées par l'analyse protéomique. Dans le cas des tissus complexes, c'est la variabilité des proportions des différents types cellulaires constituant le tissu dans les différentes biopsies qui va être la base de la dispersion des mesures de protéines. Dans le cas des maladies progressives, il s'agira de pouvoir disposer de prélèvements au même stade d'évolution du tissu lésé afin, là encore, de minimiser la dispersion des mesures de quantités de protéines ou de leurs produits de dégradation dans le cas de maladies où une forte dégradation du tissu est observée. Disposer de tels prélèvements "synchrones" peut s'avérer être un vrai défi dans le cas de maladies rares.

Du fait de cette variabilité, qui se surajoute à la variabilité interindividuelle et aux variabilités collectives liées à l'âge et au genre, par exemple, et qui posent problème pour obtenir des séries d'échantillons cohérentes et appariées autant que possible en termes de physiologie entre les cas et les témoins, il peut s'avérer intéressant d'étudier des systèmes modèles, qui sont aussi indispensables quand des raisons éthiques évidentes empêchent l'accès à des prélèvements chez les malades. Se pose alors toute la question de la pertinence de ces modèles, qui est un sujet en soi. Cependant, il est intéressant de souligner que l'analyse protéomique peut révéler des phénotypes moléculaires dans des types cellulaires qui ne sont pas ceux touchés par l'expression de la pathologie. Un exemple

classique est fourni par les pathologies mitochondriales, qui sont souvent étudiées sur des fibroblastes ou des lymphocytes immortalisés issus de patients, alors même que ces maladies mitochondriales s'expriment essentiellement dans des tissus grands consommateurs d'énergies (cerveau, muscles, cœur). Dans ce cas, c'est le côté ubiquitaire de la fonction mitochondriale qui est mis à profit pour l'étudier dans d'autres types cellulaires que ceux touchés par la pathologie. Un autre exemple, très récent, est fourni par l'étude protéomique de la maladie de Cornélia de Lange chez des lymphocytes immortalisés [2]. Là encore, c'est l'aspect transversal de l'anomalie moléculaire (des cohésines dans ce cas) qui est mis à profit. Il s'avèrera donc souvent judicieux de commencer par une analyse génomique, et la connaissance du ou des gènes touchés permettra de mieux définir, dans un second temps, les modèles optimaux pour une étude protéomique.

Ce point central de la dispersion des mesures de quantités de protéines, qui provient pour une part de la technologie protéomique elle-même, et pour l'autre part de la variabilité biologique, explique aussi pourquoi il est impossible de répondre de façon simple et a priori à la question du nombre de points expérimentaux nécessaires pour réaliser une analyse protéomique pertinente. Si la dispersion des mesures est faible, un petit nombre de points peut suffire, mais ce nombre nécessaire peut augmenter très rapidement si la dispersion des mesures est forte, ce qui est problème évident dans le cadre des maladies rares. De plus, il est impossible de prédire cette dispersion des mesures, qui va apparaître seulement quand une étude au moins pilote sera réalisée. A ce problème de la dispersion des mesures s'ajoute celui des tests multiples, qui fait que le taux de faux positifs augmente avec le nombre d'objets testés, peptides ou protéines en analyse protéomique. Parmi les méthodes statistiques permettant de contrôler cette explosion des faux positifs, les méthodes de Diz et coll. [3] et de Karp et coll. [4], qui ont été optimisées pour l'analyse protéomique, méritent d'être examinées attentivement. De même, il sera souvent intéressant de porter attention, dès le début, au choix d'une méthode d'analyse protéomique qui soit performante dans le cas d'analyses étalées dans le temps, et qui soit donc surtout constante dans ses performances (profondeur et reproductibilité d'analyse en particulier) sur une période pouvant se chiffrer en mois ou en années. Il sera ainsi possible de réaliser des analyses d'une façon incrémentale, en réalisant les tests de puissance statistique au fur et à mesure et en s'arrêtant quand une puissance suffisante est obtenue. Dans cette optique, les méthodes sans marquage semblent être plus robustes que les méthodes avec marquage et multiplexage [5], et sont très prometteuses [6].

Enfin, pour ce qui est des échantillons issus d'une préparation biochimique plus ou moins complexe (fractionnement subcellulaire, préparations de complexes



## Quels échantillons biologiques pour l'analyse protéomique dans l'étude des maladies rares ? (Suite)

protéiques), il conviendra de garder à l'esprit que de nombreuses méthodes d'analyse protéomique sont à la fois très sensibles et non-linéaires dans leurs réponses quantitatives, en particulier quand d'un peptide à l'autre. La conséquence en est que ces analyses protéomiques sont très sensibles à des phénomènes de contamination non-spécifiques (voir par exemple [7]) qui ne sont pas facilement mis en évidence par des techniques biochimiques plus classiques et qui doivent être pris en compte dès le départ afin de mettre en place les meilleurs contrôles expérimentaux possibles. A titre d'exemple, une bille de chromatographie n'est jamais totalement inerte en termes de liaison de protéines [8], et un anticorps ne se limite pas, en termes de liaison de protéines, à ses deux sites antigéniques.

La conclusion évidente de ce bref article est que l'échantillon biologique optimal va être celui qui va permettre une expression d'un phénotype moléculaire pertinent tout en minimisant la complexité moléculaire et les sources biologiques de dispersion des mesures. De même, la stratégie expérimentale optimale en analyse protéomique sera avant tout celle qui offrira la plus grande reproductibilité de mesure, plutôt qu'une stratégie qui sacrifiera par exemple la reproductibilité à la sensibilité.

[2] Gimigliano A et al. Proteomic Profile Identifies Dysregulated Pathways in Cornelia de Lange Syndrome Cells with Distinct Mutations in SMC1A and SMC3 Genes. *J Proteome Res.* (2012) Nov 5. [Epub ahead of print]

[3] Diz AP et al. Multiple hypothesis testing in proteomics: a strategy for experimental work. *Mol Cell Proteomics.* (2011) 10(3):M110.004374.

[4] Karp NA & Lilley KS. Design and analysis issues in quantitative proteomics studies. *Proteomics.* (2007) Sep;7 Suppl 1:42-50.

[5] Wang H et al. Comprehensive comparison of iTRAQ and label-free LC-based quantitative proteomics approaches using two *Chlamydomonas reinhardtii* strains of interest for biofuels engineering. *J Proteome Res.* (2012) 11: 487-501

[6] Gautier V et al. Label-free quantification and shotgun analysis of complex proteomes by one-dimensional SDS-PAGE/NanoLC-MS: evaluation for the large scale analysis of inflammatory human endothelial cells. *Mol Cell Proteomics.* (2012) 11:527-539

[7] Boulon S et al. Establishment of a protein frequency library and its application in the reliable identification of specific protein interaction partners. *Mol Cell Proteomics.* (2010) 9:861-879.

[8] Trinkle-Mulcahy L et al. Identifying specific protein interaction partners using quantitative mass spectrometry and bead proteomes. *J Cell Biol.* (2008) 183: 223-239



### Références :

[1] Schwanhäusser B et al. Global quantification of mammalian gene expression control. *Nature.* (2011) 473: 337-342

# Bases de données et ressources pour la protéomique

Lydie Lane

Institut Suisse de Bioinformatique (SIB), groupe CALIPHO, Genève, Suisse (lydie.lane@isb-sib.ch)

Le traitement des grandes quantités de données brutes issues d'expériences de protéomique nécessite différents types de bases de données et outils répondant à quatre catégories de besoins :

- le stockage de données brutes (spectres) et des protocoles correspondants (« metadata »)
- l'analyse des données brutes (identification et quantification des protéines)
- l'analyse des résultats (comparaison avec d'autres jeux de données, contextualisation)
- la visualisation et la diffusion des résultats.

Pour le stockage des données brutes, deux entrepôts de données sont majoritairement utilisés : Pride ([ebi.ac.uk/pride/](http://ebi.ac.uk/pride/)) et PeptideAtlas ([peptideatlas.org/](http://peptideatlas.org/)). A la différence de Pride, PeptideAtlas ré-analyse toutes les données avec un protocole unifié, «the Trans-Proteomic Pipeline» (TPP). La soumission des données aux deux ressources peut se faire maintenant à travers un portail commun appelé ProteomeXchange ([proteomexchange.org/](http://proteomexchange.org/)).

Les bases de données de séquences utilisées pour l'identification des peptides à partir des spectres doivent être «complètes». Or, contrairement au génome, il n'est pas simple de définir un protéome complet : pour un gène donné, plusieurs centaines de formes protéiques peuvent être générées par épissage alternatif et/ou modifications post-traductionnelles (PTM). Les bases de données comme UniProtKB ([uniprot.org/](http://uniprot.org/)) mettent à la disposition des chercheurs en protéomique un ensemble de séquences de référence pour chaque organisme modèle. Ces « protéomes complets » vérifiés manuellement permettent d'éviter la redondance tout en prenant en compte cette diversité.

Pour comprendre la signification biologique des listes de protéines identifiées, il est nécessaire de déterminer si celles-ci comportent des biais d'enrichissement en termes de fonction, expression etc. Cette analyse s'effectue à l'aide d'outils statistiques comme DAVID ([david.abcc.ncifcrf.gov](http://david.abcc.ncifcrf.gov)). Ceux-ci utilisent les annotations fournies par les bases de données comme UniProtKB/Swiss-Prot, et en particulier les annotations qui utilisent les termes contrôlés de type GO (Gene Ontology).

neXtProt ([nextprot.org/](http://nextprot.org/)) est une ressource libre d'accès spécifique pour les protéines humaines. Elle ne remplace pas UniProtKB/Swiss-Prot qui reste la base de référence pour les séquences humaines, mais intègre des informations additionnelles résultant d'expériences à haut débit en génomique (polymorphismes), transcriptomique (expression tissulaire), protéomique (PTMs), immunohistochimie (localisation) filtrées manuellement. Elle propose une visualisation thématique de ces données, et notamment, une vue dédiée aux informations de protéomique. neXtProt va proposer prochainement des outils qui feront appel à ces informations supplémentaires et permettront d'aller plus loin dans l'analyse des données de protéomique.

Dans le cadre du projet HUPO « Human Proteome Project » ([hupo.org/research/hpp/](http://hupo.org/research/hpp/)), neXtProt a été choisie comme base de référence pour l'annotation des protéines humaines. Pour accomplir cette tâche, elle s'appuiera sur les analyses des données brutes réalisées sous l'égide de ProteomeXchange.

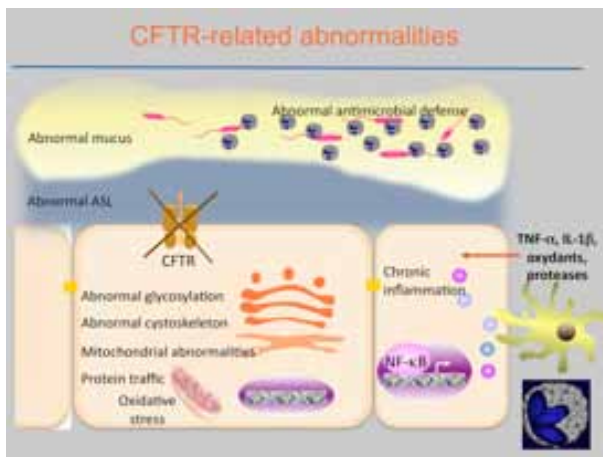


# La protéomique au service de la mucoviscidose: l'identification d'une nouvelle cible thérapeutique pour la mutation $\Delta F508CFTR$

Aleksander Edelman

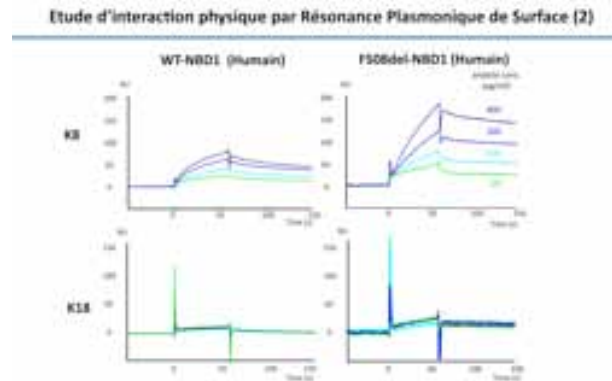
Inserm U845, Université Paris Descartes, Paris, France (aleksander.edelman@inserm.fr)

La mucoviscidose (MV) est une maladie génétique, récessive et autosomale. Elle est due aux mutations dans le gène CFTR (cystic fibrosis transmembrane protein). La maladie touche les organes formés par les épithéliums (poumons, pancréas, intestin, glandes sudoripares etc.). Plus de 1600 mutations de CFTR ont été identifiées, la mutation la plus fréquente étant une délétion de Phe en position 508,  $\Delta F508$ . Le produit de CFTR, la protéine CFTR, est une protéine multifonctionnelle, avec comme fonction principale, une fonction de canal  $Cl^-$ . Dans les épithéliums CFTR est un facteur limitant pour l'absorption des fluides. La plupart des mutations dans le gène CFTR conduisent à des anomalies du transport transépithélial de  $Cl^-$  et à des défauts de transport des fluides. Ces mutations sont également associées à des dysfonctionnements intracellulaires liés aux réponses des épithéliums à l'inflammation et au stress oxydatif (figure ci-dessous). Rapidement après



l'identification de CFTR, il est devenu évident que les mutations seules ne peuvent expliquer la diversité des phénotypes cliniques observés chez les patients, ouvrant la voie aux études par des approches omiques. Les premiers travaux de protéomique dans le contexte de la MV datent de 2003-2004[1]. Il s'agit de travaux de protéomique différentielle (gels 2D associés à la spectrométrie de masse (SM) de type MALDI-TOF)[2]. L'exemple que nous présentons montre que les approches protéomiques associées aux approches de biophysique, biologie cellulaire, ainsi que de modélisation peuvent conduire à l'identification de nouvelles cibles thérapeutiques et de nouveaux médicaments pour la MV. En 2004, les premiers résultats ont montré que 2 filaments intermédiaires, les kératines 8/18 (K8/18), sont impliqués dans le dysfonctionnement de  $\Delta F508CFTR$ [3]. Au cours des années suivantes (2005-2009), les recherches de modélisation par dynamique moléculaire[4] d'une part, de l'interactome différentiel d'autre part, ont permis de conclure que l'interaction directe entre K8 et un fragment de CFTR, le NBD1/ $\Delta F508NBD1$  est responsable de la rétention de  $\Delta$

$F508CFTR$  dans le cytoplasme (figure 2). Cette interaction



empêche l'adressage de  $\Delta F508CFTR$  à la membrane[5]. L'interruption de l'interaction entraîne chez la souris  $\Delta F508/\Delta F508$  le rétablissement du transport de  $Cl^-$  dépendant de CFTR[5]. Plus récemment, une autre approche biophysique de SM permettant l'analyse des structures des protéines (l'échange hydrogène-deutérium couplé à la SM (SM à transformée de Fourier)) a permis de préciser les sites d'interaction entre K8 et  $\Delta F508NBD1$ . Très récemment, l'association des techniques de dock-in avec les approches de biologie cellulaire, de physiologie et de biophysique a conduit à l'identification de molécules « interruptrices » des interactions et correctrices de l'activité de  $\Delta F508CFTR$  dans les cellules provenant des malades atteints de MV. En conclusion, les approches de protéomique et de SM, en association avec d'autres méthodologies s'avèrent être indispensables pour comprendre les mécanismes physiopathologiques, et contribuer utilement à la recherche de correcteurs des dysfonctionnements observés au cours des maladies. L'exemple des recherches sur la MV le montre très clairement.

1. Roxo-Rosa M, Davezac N, Bensalem N, Majumder M, Heda GD, Simas A, Penque D, Amaral MD, Lukacs GL, Edelman A: Proteomics techniques for cystic fibrosis research. *J Cyst Fibros* 2004, 3 Suppl 2:85-89.
2. Ollero M, Brouillard F, Edelman A: Cystic fibrosis enters the proteomics scene: new answers to old questions. *Proteomics* 2006, 6(14):4084-4099.
3. Davezac N, Tondelier D, Lipecka J, Fanen P, Demaugre F, Debski J, Dadlez M, Schratzenholz A, Cahill MA, Edelman A: Global proteomic approach unmasks involvement of keratins 8 and 18 in the delivery of cystic fibrosis transmembrane conductance regulator (CFTR)/ $\Delta F508$ -CFTR to the plasma membrane. *Proteomics* 2004, 4(12):3833-3844.
4. Wieczorek G, Zielenkiewicz P:  $\Delta F508$  mutation increases conformational flexibility of CFTR protein. *J Cyst Fibros* 2008, 7(4):295-300.
5. Colas J, Faure G, Sausseure E, Trudel S, Rabeh WM, Bitam S, Guerrero IC, Fritsch J, Sermet-Gaudelus I, Davezac N et al: Disruption of cytokeatin-8 interaction with F508del-CFTR corrects its functional defect. *Hum Mol Genet* 2012, 21(3):623-634.

# Analyse protéomique du muscle dystrophique

Laetitia Guével<sup>1, 2</sup> – Karl Rouger<sup>1</sup>

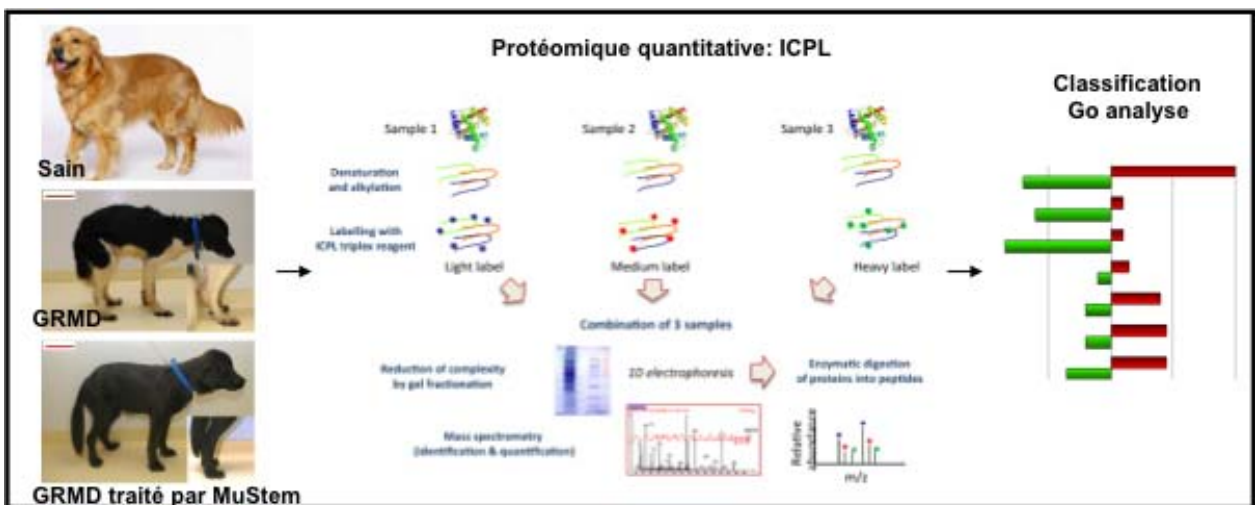
1. UMR 703 INRA/ONIRIS- 2. Université de Nantes (Laetitia.Guevel@univ-nantes.fr)

La Dystrophie Musculaire de Duchenne (DMD) correspond à la forme la plus fréquente des myopathies pour laquelle il n'existe pas de traitement aujourd'hui. Elle résulte de l'absence complète de la dystrophine à l'origine d'une nécrose des fibres et d'une faiblesse musculaire progressive. Les analyses protéomiques du muscle dystrophique présentent un grand intérêt afin d'affiner la physiopathologie de la maladie de Duchenne et d'identifier ainsi de nouveaux bio-marqueurs protéiques. Une première étude utilisant des puces à protéines nous a permis de comparer le profil de phosphorylation de kinases dans le modèle animal cliniquement relevant de la DMD, le chien GRMD, révélant ainsi un défaut de phosphorylation dans le muscle du chien GRMD (Féron et al., Am.J.Pathol, 2009). De façon à approfondir l'identification et la quantification des protéines dérégulées dans la DMD et ainsi de définir les voies ou familles protéiques altérées, une large étude protéomique, non biaisée, a été réalisée. Une approche de protéomique quantitative par ICAT (marquage des cystéines) puis par ICPL (marquage des lysines/couplée à une analyse par Orbitrap) ont été effectuées (Guével et al., J. Proteome. Res, 2011). Des différences significatives ont été identifiées dans le muscle canin dystrophique vs sain, et nous avons démontré que parmi les protéines sous-exprimées, appartenant au métabolisme glucidique, énergétique ou présentant une activité oxydo-réductase, 30% sont des cibles du co-activateur transcriptionnel PGC1- $\alpha$  qui apparaît comme un nouveau bio-marqueur de la DMD. Dans un second temps, les expérimentations ont été conduites en combinant des outils de transcriptomique

(puces à ADN, miRNA), de protéomique (ICPL) et de bioinformatique sur des prélèvements musculaires de chiens sains, GRMD et de chiens GRMD après transplantation de cellules souches somatiques dérivées du muscle, les cellules MuStem (Rouger et al., Am. J. Pathol, 2011). Ces dernières, se sont révélées capables de stabiliser l'activité locomotrice des chiens myopathes avec un recul de plus de trois ans. Notre objectif est d'appréhender les mécanismes cellulaires à l'origine de l'impact clinique de l'administration systémique des cellules MuStem, par l'analyse du phénotype tissulaire.

Des résultats préliminaires indiquent que l'analyse protéomique par ICPL a permis de quantifier 540 protéines (n=3). Parmi elles, 130 sont dérégulées dans le muscle dystrophique et 18 sont statistiquement différentielles dans le muscle canin traité par les cellules MuStem. Ce projet structurant réunit autour d'une même thématique, une équipe de recherche, un centre d'expérimentation gros animal et trois plateformes biotechnologiques du réseau Biogenouest.

**Mots-clés :** Dystrophie Musculaire de Duchenne, Cellules MuStem, Thérapie cellulaire, Protéomique quantitative, Génomique intégrative.



# Stratégies d'Analyse des Assemblages Supramoléculaires de Protéines

Marc Bonneau

Plate-forme Protéome - UMR 5248 CNRS, ENSTBB-IPB, 33076 Bordeaux Cedex (Marc.Bonneu@ipb.fr)



La plateforme Protéome du Centre de Génomique Fonctionnelle Bordeaux est une plateforme scientifique et technologique ouverte à l'ensemble des laboratoires de recherche académiques ou privés. Cette plateforme est adossée à un laboratoire de recherche spécialisé en spectrométrie de masse (équipe dirigée par Jean-Marie Schmitter au sein de l'UMR CNRS 5248 Chimie Biologie des

Membranes et Nanoobjets). Cet adossement a facilité le développement technologique de la plateforme qui propose aujourd'hui toutes les technologies majeures dans le domaine de l'analyse protéomique : séparation des protéines par électrophorèse, identification des protéines par spectrométrie de masse, quantification non ciblée des protéines avec ou sans marquage isotopique, quantification ciblée des protéines par MRM ...

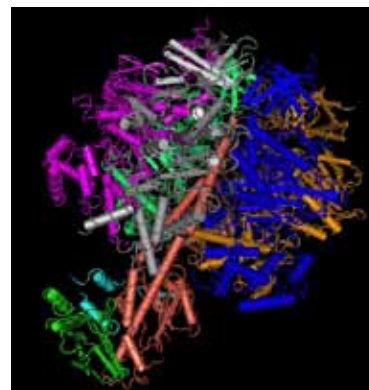
Cet adossement a également permis de développer une expertise spécifique pour l'étude des assemblages supramoléculaires de protéines. Plusieurs approches complémentaires sont mises en œuvre, avec notamment :

- La séparation de complexes selon deux dimensions par BN-PAGE/SDS-PAGE
- La modification chimique par des réactifs de réticulation (cross-link)
- L'analyse topologique au moyen de l'échange isotopique Hydrogène/Deutérium (HXMS)

La méthodologie **BN-PAGE/SDS-PAGE** permet de séparer par électrophorèse dans une première dimension les complexes de protéines à l'état natif, avant d'en résoudre les sous-unités constitutives par SDS-PAGE dans la seconde dimension, puis de les identifier par MS/MS. Cette méthodologie a notamment été appliquée à la caractérisation des complexes supramoléculaires chez *Escherichia coli* et *Helicobacter pylori* [1,2,3].



L'approche utilisant le **cross-link** consiste à modifier au moyen d'agents bifonctionnels des groupements réactifs de surface au sein d'un complexe. Les produits de réticulation inter-protéines permettent ensuite d'identifier les sous-unités membres d'un complexe après une étape de protéolyse et



une analyse par MS/MS. L'utilisation de réactifs de cross-link spécifiquement élaborés permet d'identifier les acides aminés réactifs situés à l'interface des sous-unités. Cette méthodologie a par exemple été appliquée à l'étude topologique du complexe de l'ATP synthase de levure [4].

**L'échange isotopique H/D** analysé par spectrométrie de masse (**HXMS**) exploite l'échange spontané de protons amides des protéines avec des deutérons fournis par le solvant. Une carte topologique d'accessibilité au solvant peut ainsi être établie après protéolyse et analyse par spectrométrie de masse. Cette méthodologie a notamment été appliquée à l'étude de prions et à leur transition amyloïde [5,6].

## REFERENCES

1. Lassère et al. *Electrophoresis* 2006, 27, 3306–3321
2. Pyndiah et al. *Mol. Cell Proteomics* 2007, 6, 193–206
3. Bernardé et al. *Mol. Cell Proteomics* 2010, 9, 2796–2826
4. Fronzes et al. *Biochemistry* 2003, 42, 12038–12049
5. Nazabal et al. *Biochemistry* 2003, 42, 8852–8861.
6. Nazabal et al. *J. Biol. Chem.*, 2005, 280, 13220–13228



# Interactome in vivo

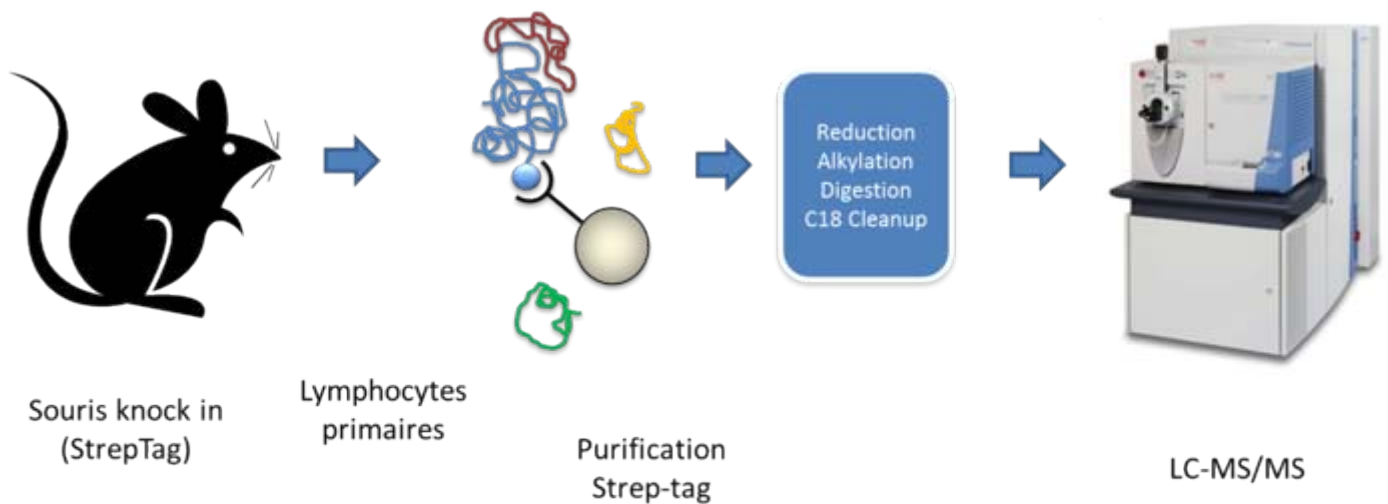
Romain Roncagalli

CNRS, Centre d'Immunologie de Marseille Luminy (roncagalli@ciml.univ-mrs.fr)

Les lymphocytes T constituent les effecteurs majeurs de la réponse immunitaire adaptative. Le récepteur pour l'antigène des lymphocytes T (TCR) contrôle la physiologie de ces cellules. Afin de propager le signal d'activation, le TCR coopère avec de nombreuses protéines intracellulaires incluant par exemple des adaptateurs et des protéines à activité enzymatique (kinases, lipases...). L'étude de la complexité des interactions impliquées dans cette signalisation amène à repenser les modes de signalisations selon un concept de 'system biology' où des interactomes de protéines se constituent et se déstabilisent en fonction du temps et du type d'activation. Afin d'appréhender ce système, notre laboratoire a généré des souris génétiquement modifiée (knock-in), dans lesquelles une étiquette (StrepTag) a été rajouté sur la partie C-terminale de certaines protéines impliquées dans la signalisation des

cellules T via le TCR. Ces modèles de souris permettent ainsi de 'pêcher' les molécules qui s'associent de manière spécifique à la protéine appât étiquetée dans les cellules T primaires. Les complexes ainsi formés peuvent être de ce fait purifiés et combinés à une analyse par spectrométrie de masse (LC-MS/MS).

Les résultats obtenus par cette approche ont permis de confirmer mais aussi d'identifier de nouvelles protéines effectrices impliquées dans la signalisation et l'activation des cellules T primaires. Ces modèles de souris sont ainsi un point de départ qui permettra de comprendre et d'identifier les perturbations qui caractérisent les interactomes des environnements sains de ceux qui sont pathologiques.



Etude de l'interactome de cellules T primaires: stratégie



# Méthodes de fouille et outils de visualisation pour l'analyse de données "omics"

Yves Vandembrouck

Laboratoire de Biologie à Grande Echelle, U1038, CEA-Grenoble - 17, rue des martyrs 38054 Grenoble Cedex 9  
(yves.vandembrouck@cea.fr)

La production à haut-débit de données en Biologie et en Médecine induit un besoin croissant en capacités de traitement analytique, d'intégration et de visualisation de l'information générée. L'analyse protéomique par spectrométrie de masse possède son propre niveau de complexité de par la nature des échantillons étudiés (gamme dynamique d'expression des protéines élevée, espèces moléculaires de l'ordre de 10<sup>6</sup>) et la diversité des chaînes de traitement des données selon les questions à résoudre (recherche de candidats bio-marqueurs, expression différentielle, caractérisation de complexes,...). Dans ce contexte, la « fouille de données » (ou data mining) est un champ disciplinaire situé au croisement de la statistique et des technologies de l'information (bases de données, intelligence artificielle, classification, apprentissage, etc.) dont le but est d'extraire des connaissances dans de vastes ensembles de données. L'analyse descriptive permet de distinguer les biais technologiques des variabilités biologiques et de qualifier l'information avec une fiabilité associée alors que l'exploration des données permet de dégager des corrélations entre variables d'intérêts. Ces techniques peuvent aussi déboucher sur l'élaboration de modèles qui résume des relations entre variables, permettant de comprendre des phénomènes et d'émettre des prédictions (e.g., identification de candidats bio-marqueurs). Chaque étape du flot de données protéomiques, de l'acquisition jusqu'à l'obtention d'un résultat étayé, conduit à des réductions successives des volumes de données imposant des traitements explicites et adaptés (filtrage, gestion des données manquantes, tests statistiques, gestion de l'incomplétude et des imprécisions) qui doivent être planifiées et calibrées dès le démarrage d'un projet. Il est crucial de considérer la mise en place d'une chaîne de compétences alliant médecins, biologistes, informaticiens et mathématiciens si l'on souhaite maximiser les chances de réussite d'un projet. Il est clair qu'une analyse statistique rigoureuse ne saura corriger les

faiblesses d'un mauvais design expérimental (e.g., effectifs déséquilibrés, manque de réplicats, choix des contrôles, effet indésirable d'une variable du dispositif,...). Dans cet esprit, l'étendue des méthodes de data mining impose un dialogue permanent entre les différentes parties prenantes permettant ainsi de s'accaparer les concepts sur lesquelles elles reposent ; il s'agit ici non seulement d'acquérir une terminologie commune mais aussi de garantir l'adéquation entre le choix d'une méthode précise (i.e. les concepts et leur condition d'application) et les hypothèses biologiques à tester. En complément, les techniques d'analyse visuelle soutiennent le data mining en proposant des représentations graphiques variées et des fonctionnalités qui viennent idéalement enrichir l'interprétation de l'information: elles offrent la possibilité d'explorer de façon globale ce qu'un cerveau humain ne saurait réaliser à partir d'une liste de protéines étiquetées par des mesures et des annotations dans une feuille Excel. L'enjeu des logiciels de visualisation de données pour la biologie moderne est de permettre à terme d'interroger, d'annoter, de croiser les données avec d'autres sources d'information de façon interactive et intuitive tout étant capable de « communiquer » avec d'autres systèmes de gestion de l'information. Dans ce but, la plate-forme logicielle communautaire Cytoscape (<http://www.cytoscape.org>) représente une avancée significative pour explorer et analyser les entités moléculaires sous forme de réseaux biologiques.

• Cappadona S, Baker PR, Cutillas PR, Heck AJ, van Breukelen B. Current challenges in software solutions for mass spectrometry-based quantitative proteomics. *Amino Acids*. 2012. 43:1087-108

• Käll L, Vitek O. Computational mass spectrometry-based proteomics. *PLoS Comput Biol*. 2011. 7:e1002277

• Special issue in *Nature Methods*, Visualizing biological data, 2010, Volume 7 No 3s ppS1-S68



# Imagerie par spectrométrie de masse MALDI

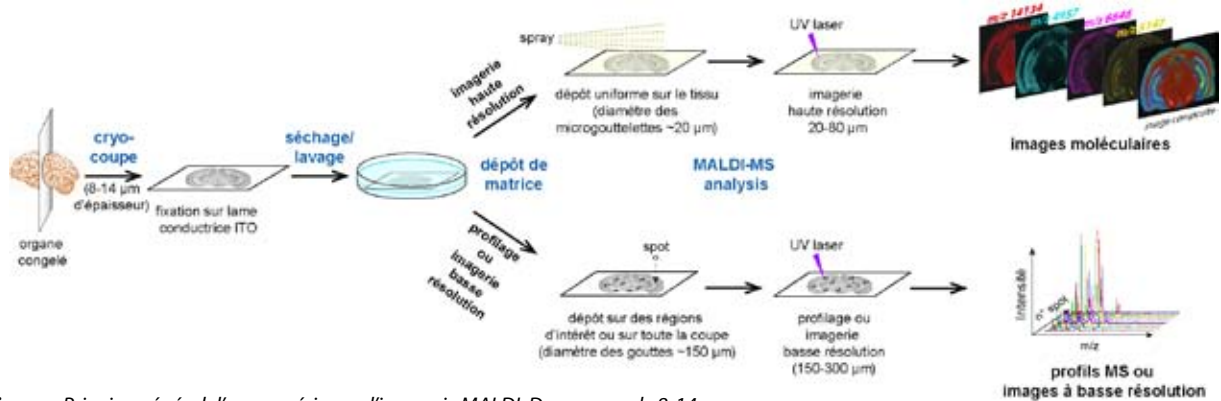
## Applications, limites et potentiels

Charles Pineau<sup>1</sup> & Daniel Lafitte<sup>2</sup>

1. Plate-forme Protéomique Biogenouest, IRSET – Inserm U1085, Campus de Beaulieu, 35042 Rennes cedex  
 2. Marseille Protéomique - Plate-forme protéomique & innovation technologique Timone, 13385 Marseille cedex 05  
 (charles.pineau@inserm.fr ou daniel.lafitte@univ-amu.fr)

Depuis son introduction en 1997 par l'équipe de Richard Caprioli [1], l'imagerie MALDI n'a cessé de se développer. Le principe de cette technologie repose sur l'acquisition de spectres de masse directement à partir de coupes de tissus (Figure) et permet de localiser simultanément des centaines de composés sans a priori et sans marquage qui pourrait altérer la fonction ou la localisation native des espèces d'intérêt. L'un des atouts majeurs de l'imagerie MALDI repose sur la vaste gamme de biomolécules qui peuvent être analysées telles que les protéines et les peptides, les sucres [2], les lipides [3], les petites molécules et leurs métabolites [4], etc. Le potentiel de l'imagerie MALDI s'est donc révélé dans des champs d'application très variés comme le domaine clinique et pré-clinique [5], la biologie de la reproduction [6], la lipidomique [3], la biologie végétale [7], etc. Des travaux très intéressants ont également démontré, en complément des techniques standard, l'énorme potentiel de l'imagerie MALDI pour l'aide au pronostic [8], au suivi thérapeutique [9] et à la classification de tumeurs [10]. Toutefois, deux applications parmi les plus prometteuses actuellement concernent la découverte de nouveaux marqueurs de pathologies [5] et la biodistribution de petites molécules [4]. Cependant, malgré les progrès importants réalisés récemment, des limitations restent à surmonter. L'imagerie MALDI a récemment bénéficié d'avancées technologiques impor-

tantes permettant d'atteindre en routine sur un instrument commercial une résolution latérale de 20 µm [6]. Il n'est pas encore possible à ce jour de caractériser directement sur les coupes toutes les molécules correspondant aux signaux d'intérêt mis en évidence. Toutefois des méthodes d'identification ont été développées à partir des extraits de tissus ayant permis la réalisation des coupes histologiques [6]. La digestion enzymatique in situ suivi d'une analyse protéomique permet de relier information spatiale et identification. L'utilisation des techniques de fragmentation en source (ISD) a contribué à la caractérisation et à la localisation spatiale d'isotypes de protéines [12]. La spectrométrie de masse en tandem (MS/MS, MSn) couplée à l'ISD permet aussi une identification et une quantification des biomolécules in situ. Les spectromètres de masse utilisés sont de plus en plus rapides, résolutifs et sensibles et des méthodes d'analyses statistiques dédiées ont vu le jour [11]. L'imagerie MALDI présente donc des avantages indéniables pour la recherche in situ de marqueurs potentiels de pathologies dans des échantillons qui ne se prêtent pas à des approches plus conventionnelles de Protéomique. Toutefois d'importants développements restent encore nécessaires pour améliorer les méthodes d'identification et de quantification, les outils d'analyse statistique ou l'imagerie tridimensionnelle.



**Figure :** Principe général d'une expérience d'imagerie MALDI. Des coupes de 8-14 µm d'épaisseur sont généralement préparées à partir d'organes congelés. Les coupes sont fixées sur des lames conductrices ITO puis séchées. Pour l'analyse de protéines, les coupes sont ensuite lavées dans des solutions appropriées. Cette étape de lavage n'est généralement pas requise pour l'analyse de lipides ou de petites molécules. Pour l'imagerie MALDI « haute résolution » (20 µm), la coupe est entièrement recouverte de matrice en utilisant un nébuliseur automatique. Pour le profilage ou l'imagerie MALDI « basse résolution » (150-300 µm) la matrice est déposée avec un micro-spotter. Les images moléculaires correspondant à chaque rapport m/z peuvent ensuite être reconstruites. Les approches de profilage ou d'imagerie « basse résolution » fournissent des profils MS généralement plus riches.

### References

- [1] Caprioli RM et al. *Anal Chem* 1997, 69,4751-4760.
- [2] Harvey DJ. *Mass Spectrom Rev* 2012, 31,183-311.
- [3] Goto-Inoue N et al. *Biochim Biophys Acta* 2011, 1811,961-969.
- [4] Castellino S et al. *Bioanalysis* 2011, 3,2427-2441.
- [5] Balluff B et al. *Histochem Cell Biol* 2011, 136,227-244.
- [6] Lagarrigue M et al. *Mol Cell Proteomics* 2011, 10,M110.005991.
- [7] Kaspar S et al. *Proteomics* 2011, 11,1840-1850.
- [8] Elsner M et al. *J Proteomics* 2012, 75,4693-4704.
- [9] Seeley EH & Caprioli RM. *Trends Biotechnol* 2011, 29,136-143.
- [10] Meding S et al. *J Proteome Res* 2012, 11,1996-2003.
- [11] Lagarrigue M et al. *J Proteome Res* 2012, 11(11): 5453-5463.
- [12] Ait-Belkacem et al. *J proteomics* 2012, 79C: 172-179.



# Analyse protéomique de fluides biologiques

Anne Gonzalez de Peredo

Institut de Pharmacologie et de Biologie Structurale (IPBS), UMR 5089, CNRS, Université Paul Sabatier, Toulouse  
(Anne.Gonzalez-de-Peredo@ipbs.fr)

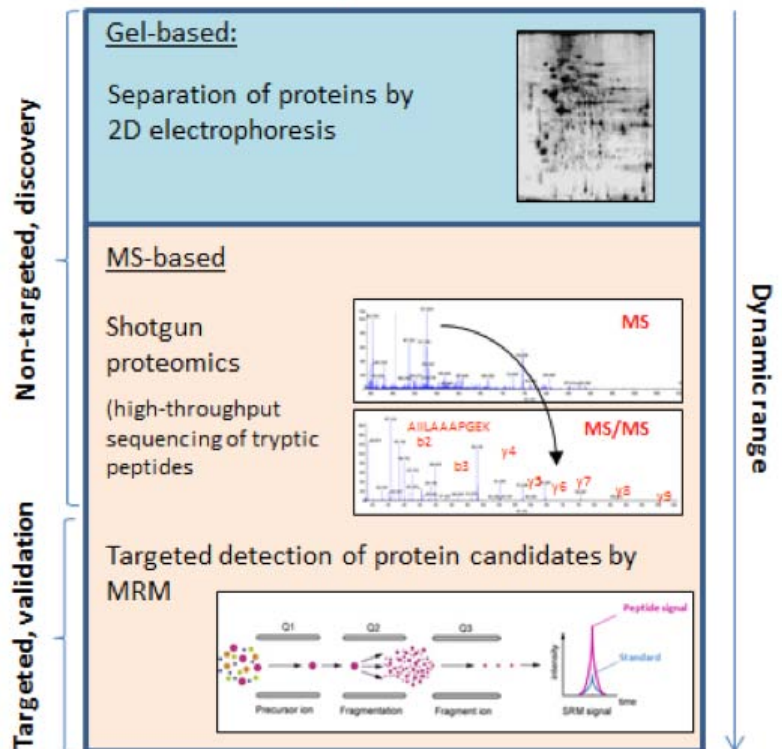
Les fluides biologiques, plus faciles à prélever chez les patients que des tissus, sont des échantillons de choix pour rechercher des marqueurs permettant de diagnostiquer une maladie ou d'aider au choix d'un traitement. Dans la mesure où ils sont en contact direct avec des tissus pathologiques susceptibles d'y relarguer leurs composants protéiques, ils constituent un réservoir de biomarqueurs potentiels. La recherche de biomarqueurs cliniques dans les fluides biologiques est cependant un domaine où l'analyse protéomique est confrontée à des enjeux techniques importants, liés à la grande gamme dynamique en concentration des protéines présentes dans ces liquides (où un faible nombre de protéines très majoritaires comme l'albumine rend difficile la détection des espèces minoritaires), et à la nécessité de comparer quantitativement, de façon la plus fiable possible, des séries d'échantillons.

Les méthodes pour réaliser ces analyses sont soit l'électrophorèse bidimensionnelle, où l'on cherche à séparer chaque isoforme protéique et la quantifier par des mesures de colorimétrie ou de fluorescence, soit des méthodes basées principalement sur l'analyse quantitative du signal par spectrométrie de masse. Parmi ces dernières, les plus utilisées ont longtemps été des approches dites « shotgun », basées sur un séquençage systématique à haut-débit des peptides tryptiques directement générés à partir des mélanges protéiques. Plus récemment se sont développées des approches de protéomique ciblée utilisant un mode d'acquisition particulier de type MRM (Multiple Reaction Monitoring), dans lesquelles plusieurs candidats protéiques préalablement définis peuvent être quantifiés de façon plus sensible et précise au sein de groupes importants d'échantillons. L'évolution de ces techniques répond à la nécessité de gagner en gamme dynamique lors de l'analyse protéomique d'échantillons complexes, particulièrement dans le cas des fluides biologiques où la gamme dynamique représente un défi majeur.

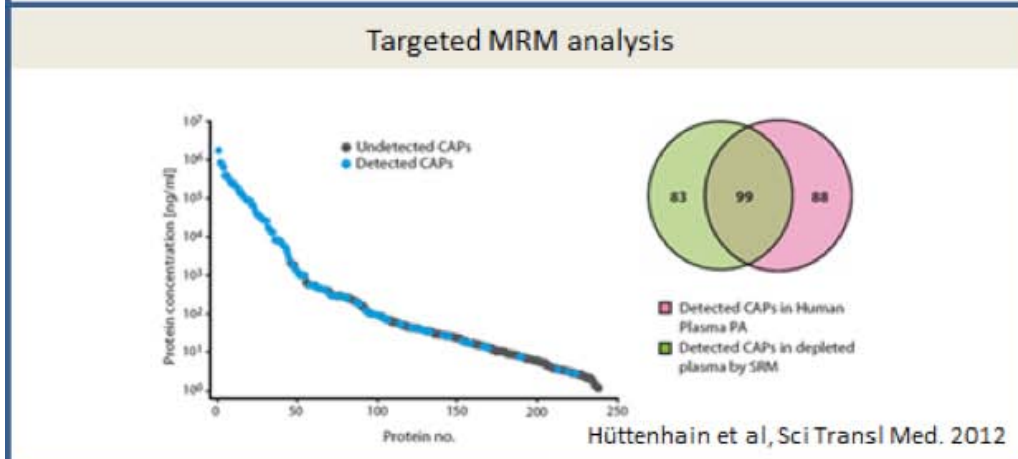
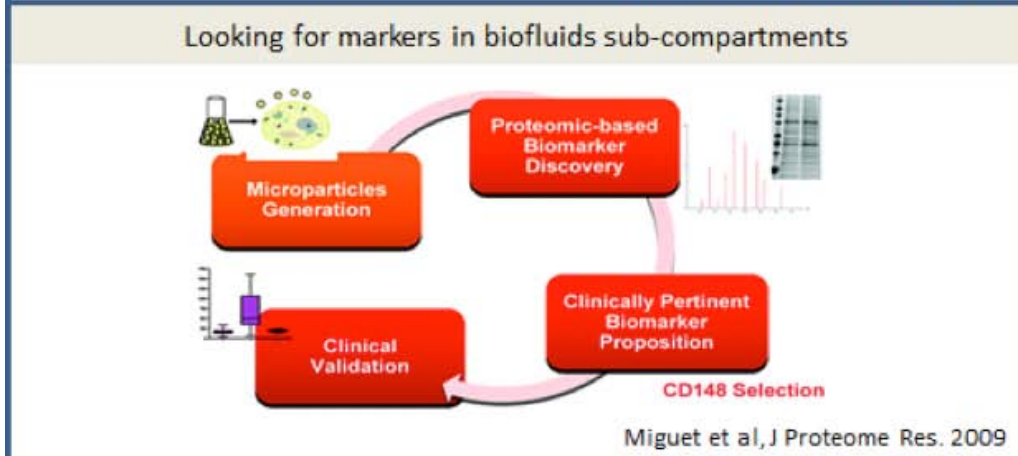
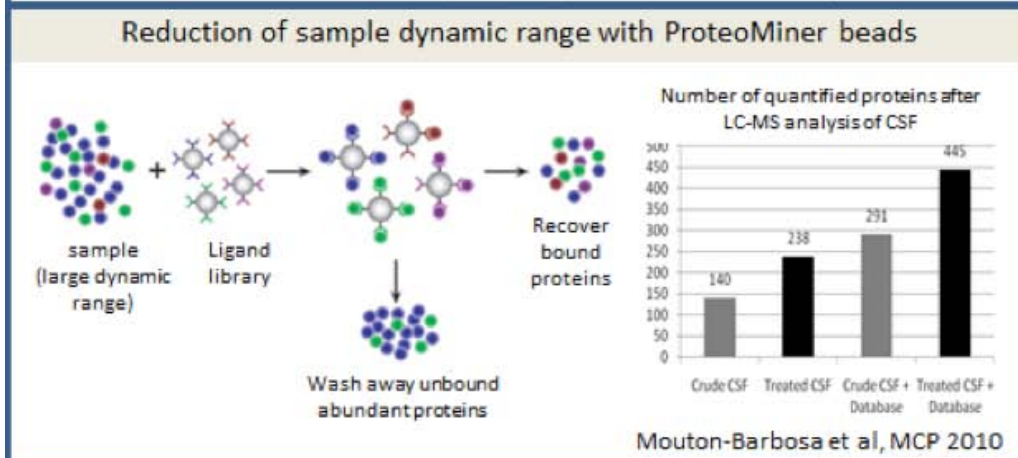
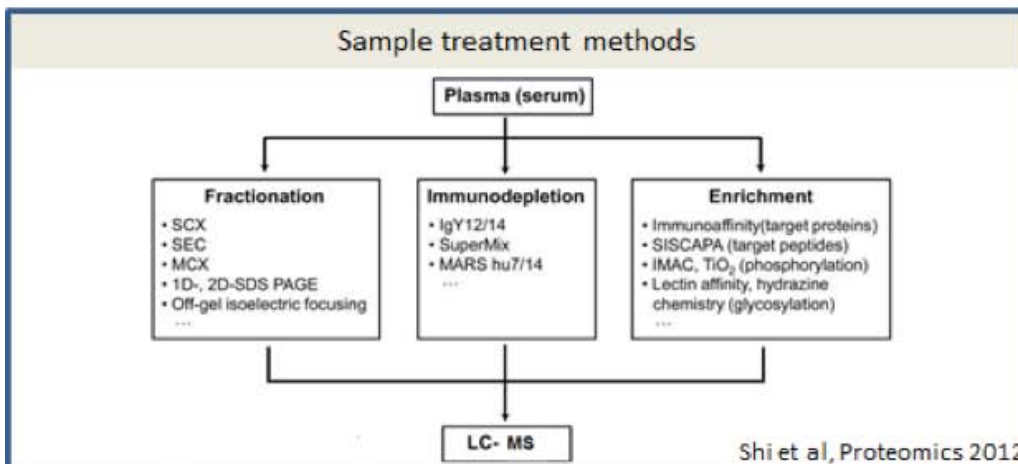
Parallèlement à l'amélioration des techniques analytiques et de l'instrumentation, des stratégies en amont ont également été développées pour répondre à cet enjeu et tenter de détecter des marqueurs de faible abondance dans les échantillons cliniques de fluides. On peut citer en premier lieu les méthodes multiples de préfractionnement et d'immunodéplétion utilisées pour simplifier par exemple le sérum ou le plasma, éliminer les protéines majoritaires présentes dans ces échantillons et améliorer ainsi la profondeur de l'analyse protéomique. Des méthodes destinées à enrichir certaines espèces d'intérêt peuvent également être employées (immunoaffinité,

enrichissement d'espèces glycosylées ou phosphorylées), ainsi que des approches détournées visant à analyser des compartiments particuliers de gamme dynamique réduite comme les microparticules du plasma, qui peuvent constituer des niches de biomarqueurs spécifiques. Au sein de l'équipe « Protéomique et Spectrométrie de masse des Biomolécules » (IPBS, Toulouse), nous avons plus particulièrement travaillé à l'optimisation de méthodes pour l'analyse du liquide céphalo-rachidien, en utilisant la technique « ProteoMiner » pour traiter les échantillons et réduire leur gamme dynamique, et en développant des approches bioinformatiques dédiées à l'analyse quantitative sans marquage, actuellement utilisées pour l'analyse de cohortes de patients présentant ou non des rechutes cérébrales de lymphomes. Ces techniques quantitatives ont également été mise en œuvre pour l'analyse protéomique d'échantillons urinaires, avec pour objectif de définir de marqueurs prédictifs de l'évolution clinique chez des enfants présentant des néphropathies obstructives, et de mieux comprendre la physiopathologie de ces maladies. Ce type d'étude permet d'établir des listes de protéines significativement modulées entre les différents groupes analysés, constituant des candidats pertinents pour une seconde phase validation par une stratégie de protéomique ciblée de type MRM.

## Analytical methods for proteomic analysis of biological fluids



# Analyse protéomique de fluides biologiques (suite)



## Plateformes Protéomiques labellisées IBiSA



IBiSA : Le GIS « IBiSA » (Infrastructures en Biologie Santé et Agronomie) a été créé en mai 2007, à la suite de la dissolution du GIP «CNRG » (Consortium de Recherche en Génomique, regroupant CNS, CNG et RNG). Les membres actuels du GIS sont l'INSERM, le CNRS, l'INRA, le CEA, l'INRIA, l'INCa (Institut National du Cancer), la CPU (Conférence des Présidents d'Université), et les deux directions du Ministère de l'Enseignement Supérieur de la Recherche la DGRI et la DGES. La Présidence du GIS est assurée par Gilles Bloch (CEA). <http://www.ibisa.net/presentation.php#presentation>

### **Plateforme Paris Descartes (3P5)**

Responsable : P. Mayeux Institut Cochin  
<http://3p5.medecine.univ-paris5.fr/>

### **Plateforme Paris Montagne Ste Geneviève**

Responsable J. Vinh  
<http://www.mass-spec.espci.fr/-Accueil,50-.html>

### **Plateforme Gif-sur -Yvette (PAPPSO)**

Responsable Michel Zivy  
<http://pappso.inra.fr/>

### **Plateforme Mont Saint Aignan (PISSARO)**

Responsable Thierry Jouenne  
<http://plateforme-proteomique.crihan.fr/ptp/home/>

### **Plateforme Lille**

Responsable Ch. Rollando  
<http://prot.univ-lille1.fr/>

### **Plateforme Marseille (MaP)**

Responsables : Borg/Bougis/Laffitte/Meunier-Gontero  
<http://map.univmed.fr/>

### **Plateforme Bordeaux (Proteome)**

Responsable M. Bonneu  
<http://www.cgfb.u-bordeaux2.fr/fr/proteome>

### **Plateforme Rennes (Biogenouest)**

Responsable Ch. Pineau  
<http://www.proteome.univ-rennes1.fr>

### **Pôle Protéomique de Montpellier LR**

Responsable Ph. Marin  
<http://www.ppm.cnrs.fr/>

### **Grenoble Proteomics Infrastructure**

Responsable Ch. Bruley  
<http://www.dsv.cea.fr/edyp>

### **Toulouse Proteomics Infrastructure**

Responsables O. Schiltz / B. Monsarrat  
<http://proteomique.ipbs.fr>

### **Plateforme Grand Est Strasbourg**

Responsable A. Van Dorsselaer  
<http://www.proteomique.u-strasbg.fr>

## Nos Partenaires

Les sociétés commerciales suivantes accordent leur soutien à la SFEAP en souscrivant comme partenaires.  
Nous les remercions pour leur concours



**Proteinsimple**  
3040 Oakmead Village Drive  
Santa Clara, CA 95051 USA  
Tél 06 31 51 23 55  
Courriel :  
thierry.salomon@proteinsimple.com  
www.proteinsimple.com



**ABSciex**  
Life Sciences Holdings France SAS  
Parc Technopolis - Bâtiment Sigma  
3 avenue du Canada - 91940 Les Ulis  
Tél 01 60 19 86 00; Fax 01 60 19 86 01  
Courriel: france.orders@absciex.com  
www.absciex.com



**BIO-RAD**  
3 bd Raymond Poincaré  
92430 Marnes-la-Coquette  
Tél 01 47 95 69 65 ; Fax 01 47 95 61 21  
www.discover.bio-rad.com



**PROTEOMIC SOLUTIONS**  
BP 2223  
7 rue Léo Lagrange  
27950 Saint-Marcel  
Tél 02 32 54 16 28 ; Fax 02 32 54 03 77  
Courriel: sarazin@proteomicsolutions.fr  
www.proteomicsolutions.fr



**BRUKER**  
34 rue de l'Industrie  
67166 Wissembourg Cedex  
Tél 03 88 73 68 30 ; Fax 03 88 73 68 79  
www.bruker.fr



**SEBIA**  
Parc Technologique Léonard de Vinci  
CP 8010 Lisses - 91008 EVRY cedex  
Tél 01 69 89 80 80 ; Fax 01 69 89 78 78  
www.sebia.com



**FRANCE: EURISO-TOP SAS**  
Parc des Algorithmes, Bât. Homère -  
Route de l'Orme - 91194 Saint-Aubin  
Cedex  
Tel : 01 69 41 95 96; Fax : 01 69 41 93 52  
Courriel : eurisotop@eurisotop.com  
www.eurisotop.com



**SERVA Electrophoresis GmbH**  
Carl-Benz-Str. 7 P.O.B. 10 52 60  
69115 Heidelberg Germany  
Tel.: +49-(0)6221-13840-0  
Fax: +49-(0)6221-13840-10  
info@serva.de



**GE Healthcare Europe GmbH**  
Parc Technologique, rue René Razel  
91898 Orsay Cedex  
Tél 01 69 35 67 00 ; Fax 01 69 41 96 77  
Courriel : orderfr@ge.com  
www.gelifesciences.com



**THERMO FISHER SCIENTIFIC**  
16 Avenue du Québec  
Silic 765  
91963 Courtaboeuf Cedex  
Tél 01 60 92 48 24 ; Fax 01 60 92 48 29  
Courriel : jocelyn.dupuy@thermofisher.com  
http://www.thermoscientific.com



**WATERS S. A. S.**  
BP 608  
78056 St-Quentin-en-Yvelines Cedex  
Tél 0820 885 885 ; Fax 01 30 48 72 01  
Courriel : france@waters.com  
www.waters.com