

THÈSE

PRÉSENTÉE A

L'UNIVERSITÉ BORDEAUX 1

ÉCOLE DOCTORALE DES SCIENCES CHIMIQUES

Par

Luc NEGRONI

POUR OBTENIR LE GRADE DE DOCTEUR

SPÉCIALITÉ : Chimie analytique

**Analyse phosphoprotéomique pour la recherche de biomarqueurs:
développements et applications**

Soutenue le : 13 décembre 2013

Devant la commission d'examen :

Pr. Jean-Marie Schmitter	Directeur de Thèse
Dr. Daniel Lafitte	Rapporteur
Dr. Gérard Bolbac	Rapporteur
Pr. Marc Bonneu	Président
Dr. Jean Rosenbaum	Examineur

Remerciements

En premier lieu, je tiens à remercier Jean-Marie Schmitter qui m'a accueilli dans son équipe, il y a déjà plus de 4 ans, les années passent vites. A mon arrivée, je lui ai demandé de pouvoir faire une thèse, un vieux rêve, presque oublié, il y a bien longtemps. J'ai trouvé une écoute bienveillante et un soutien efficace. Sans son implication, je n'aurais pas présenté ce sujet devant un jury de thèse.

Je remercie les membres du jury pour leur investissement. Faire passer une thèse à un vieil ingénieur peut être pénible. Merci de courir le risque, merci de vous être plongé dans une prose pas très rose, d'avoir rédigé des rapports, d'être venus jusqu'à Bordeaux.

Merci aux chercheurs et directeurs qui m'ont laissé un bout de leur thématique pour construire le projet phosphoprotéomique et cette thèse ; Christian Baudoin et ses cellules transgéniques, Jean Rosenbaum et son foie non fibreux. Merci à Marijà Vlaski pour ses précieuses cellules pluripotentes CD34+. Jonathan Gerbore et Patrice Rey pour le bois, les racines et les feuilles de vigne (j'aurais préféré une bonne bouteille...).

Merci au groupe spectrométrie de masse du Centre de Génomique Fonctionnelle, Stéphane & Stéphane, Anne Marie, Jean-William, Delphine, Katel, Céline. Un merci tout particulier à Corinne pour sa bonne humeur et sa relecture consciencieuse du manuscrit.

Merci à Claude Manigand pour ses synthèses peptidiques et les bonnes vieilles hydrolyses HCL 6N, on arrivera bien à en tirer quelque chose.

Merci aussi à l'équipe «mon chef ne met que des bermudas» de l'U1053 : Eric Chevet, Saïd Taouji, Daniela Arma. Merci aux anatomopathologistes de l'U 1053 ; Paulette Bioulac-Sage, Charles et son foie.

Par la somme de chacun, vous avez rendu ce travail possible.

Merci.

INTRODUCTION.....	4
CHAPITRE 1 : ANALYSE BIBLIOGRAPHIQUE	8
1.1 L'analyse des peptides par spectrométrie de masse	9
1.1.1 Mesures de masse d'ions par spectrométrie de masse.....	9
1.1.2 L'ionisation électrospray (ESI).....	10
Sources d'ionisation pour l'analyse de peptides.....	10
Principe de l'ESI.....	11
Miniaturisation des sources ESI	12
Spectres de masse en mode ESI et calcul des masses	13
1.1.3 Analyseurs hybrides Q TOF et IT Orbitrap	14
Spectrométrie de masse hybride de type Q-TOF.....	14
Spectrométrie de masse hybride de type IT Orbitrap.....	17
1.1.4 Fragmentation par CID des peptides en mode positif	19
Nomenclature	19
Interprétation des spectres MS/MS.....	20
Vue d'ensemble des voies de fragmentation.....	20
Fragmentation ChR	22
Fragmentation ChD, modèle du proton mobile.....	22
Voie b_x-y_z	22
Voie dicétopipérazine Y_{N-n}	23
Voie a_1-y_x	24
Autres voies de fragmentation	25
Perte du phosphate en CID	27
Prévision des intensités des ions fragments	28
1.2 L'analyse protéomique.....	30
1.2.1 Généralité	30
1.2.2 Analyse protéomique par électrophorèse bidimensionnelle	32
Electrophorèse SDS-PAGE et LC-MS/MS.....	35
1.2.3 Chromatographie liquide et spectrométrie MS/MS	35
Paramètres d'optimisations des colonnes HPLC.....	35
Chromatographie monodimensionnelle	39
Chromatographie multidimensionnelle	39
1.2.4. Analyse protéomique quantitative	41
Quantification label-free.....	43

Quantification par marquage isotopique.....	43
1.3 L'analyse des phosphopeptides.....	47
1.3.1 Sites de phosphorylation	47
1.3.2 Elimination du groupement phosphate.....	48
Elimination du groupement phosphate par β --élimination	49
Addition nucléophile.....	49
1.3.3 Techniques d'enrichissement des phosphopeptides.....	52
Précipitation des phosphopeptides	53
Chromatographie d'affinité par métal chélaté (IMAC)	53
Chromatographie d'affinité sur oxyde de métal (MOAC)	56
Autres chromatographies	58
Chromatographie multidimensionnelle appliquée aux phosphopeptides.	58
1.3.4. Dosage des acides aminés phosphorylés.....	60
Généralités.....	60
Dérivation et chromatographie des acides aminés	61
Dosage des acides aminés phosphorylés	64

CHAPITRE 2 : ANALYSE PROTEOMIQUE DE CELLULES COLIQUES CANCEREUSES TRAITEES PAR THERAPIE GENIQUE CD/5FC. 66

2.1 Présentation du sujet.....	67
2.1.1 Epidémiologie des cancers.....	67
2.1.2 Mécanismes moléculaires du cancer colorectal	68
2.1.3 Traitement du cancer colorectal par le 5-fluorouracile (5FU)	69
2.1.4 Thérapie génique par la 5-Fluorocytosine (5FC).....	70
2.1.5 Objectif de l'étude	70
2.2 Choix analytiques.....	71
2.3 Résultats	72
2.3.1 Publication	72
2.3.2 Résultats complémentaires	83
Electrophorèse bidimensionnelle	83
Retraitement des données de spectrométrie de masse.....	84
Identification de l'HSP90.....	85
Identification de nouvelles modifications post-traductionnelles.	86
2.4 Discussion	89

**CHAPITRE 3 : ETUDE COMPARATIVE DES CHROMATOGRAPHIES D’AFFINITE IMAC
ET MOAC..... 92**

3.1 Introduction93

3 Choix méthodologiques93

3.2.1 Chromatographie en colonne 93

3.2.2 Choix des échantillons 94

3.2.3 Préparation des hydrolysats tryptiques..... 95

3.2.4 Chromatographie d’affinité en colonne 96

3.2.5 Analyse LC-MS/MS..... 96

3.3 Résultats97

3.3.1 Publication 99

3.3.2 Résultats complémentaires 102

Stabilité des colonnes IMAC..... 102

Influence du tampon de charge pour la chromatographie TiO₂ 104

Purification des phosphopeptides, extrait de foie de souris 109

Purification des phosphopeptides, extrait de biopsie humain 109

3.4 Discussion110

**CHAPITRE 4 : ANALYSE PROTEOMIQUE ET PHOSPHOPROTEOMIQUE DE BIOPSIES
HEPATIQUES DANS LE CADRE DU CARCINOME HEPATOCELLULAIRE SUR FOIE NON
FIBREUX 116**

4.1 Présentation du sujet117

4.1.1 Le foie et ses fonctions essentielles..... 117

Organisation du lobule hépatique 117

Types cellulaires 118

Fonctions essentielles 118

4.1.2 Le carcinome hépatocellulaire..... 119

Introduction 119

Mécanismes Moléculaires du carcinome hépatocellulaire..... 120

Analyses protéomiques et phosphoprotéomiques 122

4.1.3 Analyse quantitative par marquage iTRAQ..... 125

4.2 Matériel et Méthodes127

4.2.1 Matériel 127

4.2.2 Echantillons.....	127
4.2.3. Méthodes.....	129
Broyage, réduction et alkylation, trypsinolyse.....	129
HPLC Sep Pak tC18	129
HPLC TiO ₂	130
Dosage colorimétrique des peptides	130
Marquage ITRAQ.....	131
Analyse LC-MS/MS (nanoLC-LTQ-Orbitrap)	131
Traitement des données	132
4.3 Résultats	133
4.3.1 Préparation des échantillons	133
Digestion trypsique et dessalage des peptides	133
Chromatographie TiO ₂	135
Dosage colorimétrique d'une solution peptidique	136
Marquage iTRAQ.....	137
4.3.2 Répétabilité du protocole	138
Estimation du taux de faux positifs	138
Gestion des valeurs manquantes et normalisation	139
4.3.3 Analyse quantitative de 24 biopsies de tumeurs hépatiques sur foie non fibreux	141
Analyse LC-MS/MS	141
Caractéristiques des peptides identifiés	143
Analyse quantitative des données relatives aux références C1 et C2.	145
Détermination des faux positifs à partir de la série 4	146
Identification des peptides communs aux différentes séries	147
Caractéristiques générales des protéines identifiées	148
Classification hiérarchique.....	149
Niveaux d'expression protéique significativement différents entre tumeurs et tissus normaux.....	151
Analyse des motifs d'acides aminés.....	153
4.4 Discussion	154
CONCLUSION GENERALE	162
Valider les peptides candidats par MRM	164
De l'échantillon modèle à l'échantillon clinique : une frontière visible.....	164
Analyser les peptides à phosphotyrosine	166
Développer un approche Top-Down.....	167

ANNEXES	170
1. Fraction peptides totaux.....	170
Ratios d'expression pour les tumeurs (série 1 à 3) et les non-tumeurs (série 4).....	170
Classification hiérarchique par biopsie et par peptide.	174
Test non paramétrique <i>Rank Product</i>	175
2. Fraction purifiée par TiO2.	177
Classification hiérarchique.	178
Test non paramétrique <i>Rank Product</i>	179
3. Communications par poster	180
RÉFÉRENCES BIBLIOGRAPHIQUES	183

Liste des principales abréviations

AA : acide acétique
ACN : Acétonitrile
AQUA : Absolute QUAntification
AF : acide formique
AU : Absorbance Unit
BN-PAGE : Blue Native PolyAcrylamide Gel Electrophoresis
BSA : Bovin Serum Albumin
CF : Chromato-Focalisation
CHCA : α -cyano-4-hydroxycinnamic acid
CID : Collision-Induced Dissociation
Da : Dalton (masse moléculaire)
2-DE : Two-dimensional gel electrophoresis
DIGE : Differential in gel electrophoresis
DTT : Dithiothréitol
ERLIC: Electrostatic Repulsion Hydrophilic Interaction Chromatography
ESI : Electrospray Ionization
FT-ICR : Fourier transform-ion cyclotron resonance
FWHM : Full-Width Half-Maximum, largeur du pic à mi hauteur
HILIC: Hydrophilic Interaction Liquid Chromatography
HPLC : High-Performance Liquid Chromatography
ICAT : Isotope-Coded Affinity Tag
ID : Inside Diameter
IEF : IsoElectroFocalisation
IMAC : Immobilized Metal Affinity Chromatography
IPG : Immobilized pH gradient
ITRAQ : Isobaric Tag for Relative and Absolute Quantitation
LC : Liquid Chromatography
MALDI : Matrix-Assisted Laser Desorption/Ionization
MOAC: Metal Oxide Affinity Chromatography
MRM : Multiple Reaction Monitoring
MS : Mass spectrometry
MS/MS : Tandem mass spectrometry
m/z: Ratio mass / charge
pI : Point isoléctrique
pS, pT, pY : phosphoSérine, phosphoThréonine, phosphoTyrosine
PMF : Peptide Mass Fingerprint
ppm : partie par million (part per million)
PSAQ : Protein Standard Absolute Quantification
PTM : Post Translational Modification
QconCAT : Quantitative Concatemer
RP : Reverse Phase
S/N : Signal to Noise ratio
SAX: Strong Anion eXchange
SCX : Strong Cation eXchange
SDS PAGE: Sodium Dodecyl Sulfate PolyAcrylamide Gel Electrophoresis
SILAC : Stable Isotope Labeling with Amino acid in Cell Culture
TCA : TriChloroacetic Acid
TEAB : Triethylammonium bicarbonate
TFA : TriFluoroacetic acid

TMT : Tandem Mass Tags

TOF : Time Of Flight

UHPLC : Ultra High-performance liquid chromatography

WAX: Weak Anion eXchange Chromatography

INTRODUCTION

La spectrométrie de masse est une technique d'analyse qui a été longtemps réservée aux physiciens et aux chimistes. Avec les techniques d'ionisation permettant la mesure d'ions moléculaires de composés peu volatils, les techniques de fragmentation sous vide et l'automatisation liée à l'informatique, la mise en œuvre des spectromètres de masse s'est considérablement simplifiée et les capacités d'analyse se sont étendues aux biomolécules. L'utilisation de la génomique et la création de base de données de séquences protéiques a ensuite permis d'identifier des peptides en grand nombre, permettant alors l'étude de protéomes. Plus récemment, les techniques de marquage isotopique ont permis l'émergence de l'analyse protéomique quantitative, avec la volonté de corrélérer les niveaux d'expression des protéines avec des mécanismes biologiques.

Ma thèse s'inscrit dans cette approche de protéomique quantitative, à la recherche de marqueurs biologiques. La partie expérimentale se divise en trois parties. Une première étude mettra en évidence l'intérêt de focaliser la recherche de biomarqueurs sur les protéines phosphorylées. Une deuxième étude montrera les problèmes associés à la purification des phosphopeptides et établira un protocole d'enrichissement, la troisième étude présentera une analyse phosphoprotéomique de biopsies hépatiques cancéreuses.

La sélection des phosphopeptides est l'idée directrice du projet et se base sur deux observations communément admises. D'une part, la phosphorylation est une modification post-traductionnelle qui régule l'activité des protéines. En ce sens, le dosage des protéines phosphorylées ne peut être corrélé à une activité biologique que s'il différencie la forme phosphorylée de la forme non phosphorylée. Sachant que près d'un tiers des protéines exprimées par un génome possèdent des sites de phosphorylation potentiels, il semble indispensable d'orienter l'analyse protéomique vers la détection des protéines phosphorylées. D'autre part, la dynamique d'expression d'une cellule est supérieure de plusieurs ordres de grandeur à la dynamique de quantification d'un spectromètre de masse. De ce fait, l'analyse protéomique par spectrométrie de masse surreprésente les protéines majoritaires, ces protéines abondantes étant généralement des protéines de ménage peu représentatives des mécanismes biologiques. La purification d'une classe particulière de peptides comme les phosphopeptides permet d'éliminer de l'étude une partie de ces protéines abondantes, rendant ainsi accessibles à l'analyse de nouvelles protéines, dans l'espoir qu'elles soient plus représentatives des mécanismes biologiques mis en œuvre.

Il faut cependant reconnaître que la phosphorylation est une modification post-traductionnelle parmi les nombreuses autres modifications qui permettent de réguler l'activité d'une protéine (glycosylation, acétylation, ubiquitination, etc...); il est donc peu probable que la phosphorylation explique à elle seule les mécanismes cellulaires complexes comme la différenciation cellulaire ou la cancérogénèse. Il est par contre raisonnable d'espérer identifier des différences caractéristiques, c'est-à-dire des biomarqueurs, entre les protéines phosphorylées extraites d'échantillons correspondant à des états physiologiques différents. Cette thèse illustrera à travers deux études de cas la recherche de biomarqueurs phosphorylés.

La première étude met en œuvre une électrophorèse bidimensionnelle pour séparer et caractériser les extraits protéiques de cellules de colon transformées pour une thérapie génique. Si la technique d'électrophorèse est ancienne et peu sensible, ici sa mise en œuvre permettra d'identifier des biomarqueurs parmi lesquels une protéine phosphorylée. La deuxième étude met en œuvre une technique plus récente de multiplexage de phosphopeptides extrait de biopsies hépatiques cancéreuses ou non cancéreuses. Dans ce cas, les protéines subissent une hydrolyse enzymatique avant toute séparation. Les phosphopeptides sont alors dilués dans un mélange complexe qui contient majoritairement (en nombre et en quantité) des peptides non phosphorylés. Il s'agit donc de purifier les phosphopeptides de manière quantitative avant leur analyse par spectrométrie de masse. Ce besoin d'enrichissement des phosphopeptides explique la partie expérimentale qui présente une évaluation des deux chromatographies les plus couramment utilisées : la chromatographie IMAC (Immobilized Metal Affinity Chromatography) et la chromatographie MOAC (Metal Oxide Affinity Chromatography). A cette occasion, un biais important de la chromatographie IMAC aura été identifié ; la chélation du fer ferrique est sensible aux tampons, l'ion métallique est élué à pH faible et se transforme au cours du temps. Pour la deuxième étude, le multiplexage ITRAQ (Isobaric Tag for Relative and Absolute Quantification) a été choisi afin de permettre l'analyse d'un grand nombre d'échantillons, soit 24 biopsies. Dans ce cas, les quantités de biopsies disponibles sont de l'ordre de 100 mg, soit 5-10 mg de protéines et quelques microgrammes de phosphopeptides. L'étude aura donc été confrontée à des problèmes liés à l'analyse des petites quantités.

Ce manuscrit détaillera ces différents points et débutera par une analyse bibliographique. L'objet de cette première partie est d'aborder l'ensemble des points nécessaires à la maîtrise du sujet, elle décompose l'analyse phosphoprotéomique à travers ses outils et ses principes : la spectrométrie de masse, les techniques de purification des phosphopeptides, les techniques de

marquage et de quantification. Dans ces différents domaines, la vue ne sera pas exhaustive mais pointera sur les applications mises en œuvre : l'électrospray, la fragmentation CID, les phosphopeptides, la quantification par marquage isotopique. Une deuxième partie présentera les résultats obtenus lors de l'analyse des cellules de colons sous traitement génique et illustrera l'apport de l'étude du phosphoprotéome par rapport au protéome global ; enfin, une troisième partie présentera l'étude du carcinome hépatocellulaire par analyse phosphoprotéomique. Les développements seront l'occasion de pointer les difficultés liées à ce type d'analyse et la mise en œuvre sur des biopsies hépatiques démontrera la faisabilité du protocole pour des échantillons rares et peu abondants. Ce travail identifiera des marqueurs dont la pertinence sera discutée. Une conclusion présentera une analyse critique de ce travail pour dégager ses atouts et ses faiblesses.

CHAPITRE 1 :
Analyse bibliographique

1.1 L'analyse des peptides par spectrométrie de masse

1.1.1 Mesures de masse d'ions par spectrométrie de masse

Des premiers spectrographes aux spectromètres de masse actuels, les différents éléments qui composent l'instrument sont restés identiques et répondent aux besoins essentiels de la technique : (1) le passage en phase gazeuse et l'ionisation des analytes dans une partie source (2) la séparation des ions dans une partie analyseur (3) la détection des ions dans une partie détecteur (**Fig. 1**). Dès les premières mesures de masses réalisées au début du 20^e siècle, il est apparu que les molécules ionisées pouvaient se fragmenter dans la source et ne permettaient pas forcément de mesurer la masse de la molécule entière. Dans ce cas, le spectre observé correspond au mélange des ions présents dans la source : ions parents (les ions moléculaires) et ions fils (les ions fragments). Deux évolutions techniques ont permis de résoudre ce problème majeur : d'une part, la mise au point de sources qui déposent une faible quantité d'énergie dans les ions formés et de ce fait n'entraînent pas ou peu de fragmentation lors de l'ionisation des analytes, d'autre part la mise au point de la spectrométrie de masse en tandem (MS/MS) capable de fragmenter de manière contrôlée des ions parents préalablement isolés par un premier analyseur.

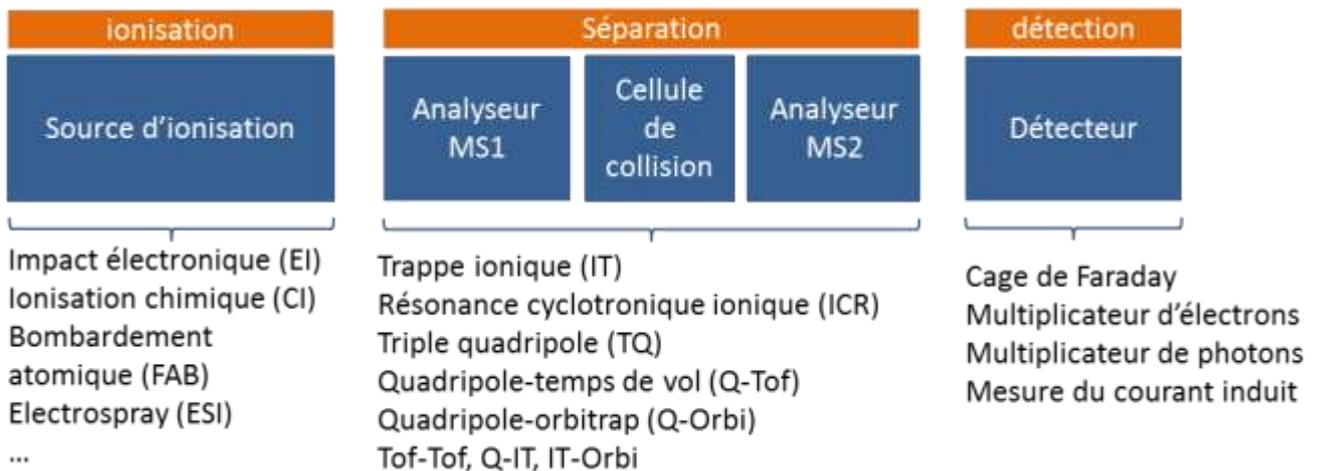


Figure 1. Fonctions essentielles d'un spectromètre de masse en tandem. Un spectromètre de masse réalise trois fonctions : l'ionisation des analytes, la séparation des ions et la détection des ions. Dans le cas de la spectrométrie de masse en tandem, la partie séparation peut se décomposer en 3 parties : l'analyseur MS1 permet la sélection d'un ion de m/z donné, la cellule de collision fragmente l'ion sélectionné, la troisième partie sépare les ions fragments produits. Ces trois fonctions peuvent être réalisées dans des régions distinctes de l'instrument (MS/MS dans l'espace pour les TQ, Q-Tof, Tof-Tof, etc) ou dans une même région de l'analyseur (MS/MS dans le temps pour les IT et ICR).

Le terme de spectromètre de masse en tandem désigne donc une évolution de la partie analyseur qui peut être alors divisée en trois sous-parties ayant chacune une fonction précise ; une

fonction de sélection d'ions (analyseur MS1), une fonction de fragmentation (cellule de collision), et une dernière fonction d'analyse des ions issus de la fragmentation (analyseur MS2) (**Fig. 1**).

Ces trois fonctions peuvent être regroupées au sein d'un même analyseur dans le cas de la résonance cyclotronique ionique ou des trappes ioniques, elles peuvent également être réalisées dans trois zones distinctes de l'instrument. Cette dernière configuration, réalisée avec deux analyseurs du même type, à savoir des quadripôles est à l'origine du terme de spectrométrie de masse en tandem (Yost and Enke, 1979). Dans le cas de l'association d'analyseurs différents, on parle alors de spectromètres hybrides.

Il existe une grande diversité de sources, d'analyseurs et de détecteurs ; la figure 1 présente une liste des principales sources et analyseurs afin d'illustrer la diversité des instruments. Seule la source électrospray (ESI) sera présentée pour la partie ionisation ; la partie séparation sera abordée à travers la présentation de deux instruments hybrides utilisés dans le cadre de ce travail : le Q TOF et l'IT Orbitrap.

1.1.2 L'ionisation électrospray (ESI)

Sources d'ionisation pour l'analyse de peptides

L'analyse des peptides utilise principalement deux modes d'ionisation : l'ionisation par désorption laser assistée par matrice (MALDI) (Karas and Hillenkamp, 1988; Tanaka et al., 1988) et l'ionisation par électrospray (ESI) (Fenn et al., 1989). Le MALDI produit majoritairement des ions monochargés qui sont produits par une irradiation laser de l'analyte cristallisé en présence d'une matrice alors que la source ESI produit des ions multichargés à partir d'échantillons dissous dans une phase liquide ; elle permet donc un couplage facile de la source à des systèmes de chromatographie liquide (LC-ESI). Le couplage LC-MALDI est par contre plus complexe à mettre en œuvre. Quelques articles ont comparé les deux techniques. Dans le cadre d'identification de protéines issues d'électrophorèse bidimensionnelle, le MALDI permet une identification rapide alors que la LC-ESI permet d'identifier plusieurs protéines dans un même spot (Lim et al., 2003). L'utilisation conjointe des deux techniques permet d'augmenter le nombre de peptides identifiés et donc la couverture de séquence, l'ESI permettant d'identifier des peptides de plus hautes masses et plus hydrophobes alors que le MALDI favorise l'ionisation des petits peptides basiques (Kuzyk et al., 2009; Molle et al., 2009) ou des peptides contenant de l'arginine (Belghazi et al., 2001).

Par contre, pour des échantillons complexes, la configuration LC-ESI présente des performances clairement supérieures. A titre d'exemple, une chromatographie multidimensionnelle associant échange d'ions et phase inverse (SCX+RP) couplée à un système ESI-IT-Orbitrap a permis d'identifier une moyenne de 29952 peptides ($n=3$), soit 3844 protéines de levure pour 8 h de chromatographie multidimensionnelle (Webb et al., 2013). Aucun équivalent n'a été présenté avec un couplage LC-MALDI. La technique LC-ESI est donc la technique de choix pour l'analyse de mélange complexes.

Principe de l'ESI

D'un point de vue pratique, l'ionisation électrospray consiste à appliquer une tension de 1 à 5kV entre une aiguille amenant l'échantillon et l'orifice d'entrée du spectromètre de masse. Un flux de gaz entoure l'aiguille pour favoriser l'évaporation du solvant. D'un point de vue théorique, l'ionisation électrospray consiste à désolvater un analyte ionisé en solution. Deux mécanismes expliquent le processus ESI ; le modèle de la charge résiduelle ; *Charged Residue Model ou CRM*, proposé par Dole et al. en 1968 (Dole et al., 1968) et l'évaporation ionique (*Ion Evaporation Model ou IEM*), de Iribarne et Thomson (Iribarne and Thomson, 1976).

Le mécanisme CRM tire ses origines d'un article de Lord Rayleigh publié en 1882 qui porte sur l'effet des charges portées par des cations ou des anions sur une gouttelette qui s'évapore. Lorsque la gouttelette diminue de diamètre, les ions de charge identique appliquent une force sur la gouttelette qui est fonction de leur répulsion électrostatique. La distance minimale entre ces ions en phase liquide correspond à la limite de Rayleigh. A la limite de Rayleigh, les forces de répulsion sont suffisantes pour faire exploser la gouttelette en une multitude de gouttelettes plus petites dont la surface totale est supérieure. L'idée de Dole et collaborateurs est d'appliquer cette théorie à l'ionisation électrospray et d'expliquer que le processus d'explosion coulombienne se reproduit jusqu'à la formation d'une gouttelette ultime ne contenant plus que l'analyte chargé : l'ion est alors en phase gazeuse.

D'après Iribarne et Thomson, lorsque la taille des gouttelettes atteint un diamètre suffisamment petit, de l'ordre de 10 à 20 nm, l'analyte est directement extrait de la gouttelette pour former un ion en phase gazeuse. Il semblerait que le mécanisme IEM corresponde au cas général, le mécanisme CRM étant plutôt observé pour les molécules de très haut poids moléculaire (Nguyen and Fenn, 2007).

Miniaturisation des sources ESI

La miniaturisation des sources électrospray a permis d'augmenter considérablement la sensibilité de cette technique d'ionisation. Deux modèles mathématiques permettent de calculer la taille des gouttelettes à la sortie d'une source ESI et définissent le rayon des gouttelettes par une relation de proportionnalité avec le débit à la puissance $1/3$ ou $2/3$ (Schmidt et al., 2003). En d'autres termes, diminuer l'orifice de la source permet de diminuer le débit et la taille des gouttelettes. De petites gouttelettes ont un ratio surface/ volume important qui favorise la désorption des analytes d'après le mécanisme IEM. D'un point de vue pratique, le faible débit permet de mettre la source dans l'axe et à proximité de l'entrée du spectromètre de masse, limitant ainsi la dispersion des ions et favorisant ainsi leur introduction. Wilms et Mann présentent en 1996 un article qui détaille la mise en œuvre de la source nanoelectrospray (nanoES) (Wilm and Mann, 1996). L'échantillon est placé directement dans le capillaire nanoES dont l'orifice de $1-2 \mu\text{m}$ permet un débit d'environ 20 nl/min lorsque la tension est appliquée. Les auteurs déterminent une transmission des analytes de la gouttelette au détecteur de 1 pour 390, soit 510 fois meilleur que pour une source conventionnelle (**Tab. 1**).

Tableau 1. Caractéristiques des différentes sources ESI (d'après (Wilm and Mann, 1996))

	ESI conventionnelle	microESI	nanoESI
Débit	2-1000 $\mu\text{l/min}$	200-2000 nl/min	20-200 nl/min
Pointe (diamètre)	500 μm	20-50 μm	1-20 μm
Taille des gouttes	1 μm	-	200nm
Efficacité (*)	1/200 800	-	1/390

*l'efficacité est définie comme la fraction des analytes transmis jusqu'au détecteur.

L'emploi des sources nanoES présente deux autres avantages moins connus. D'une part, leur tolérance vis-à-vis des sels est supérieure aux sources conventionnelles (Wilm and Mann, 1996), d'autre part, les phénomènes de suppression ioniques sont considérablement réduits à des débits nanoES inférieurs à 50 nl/min (Schmidt et al., 2003). La **Figure 2** illustre cette propriété intéressante du nanoESI ; le signal du maltoheptaose (sous la forme d'un adduit sodium) est proportionnellement plus intense que le signal de la neurotensine avec une aiguille nanoESI de diamètre interne $1\mu\text{m}$ (figure de gauche) par rapport à cette même aiguille dont le diamètre a été élargi à $10 \mu\text{m}$ (figure de droite).

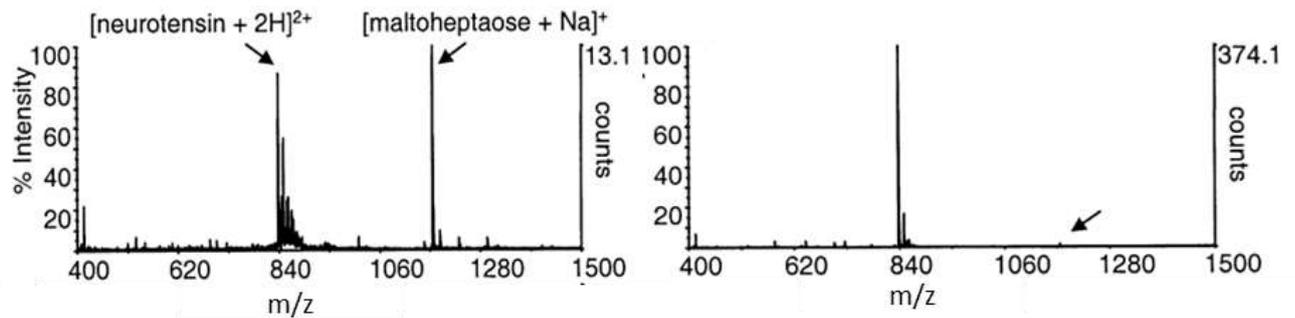


Figure 2. Spectre MS d'un mélange en nanoESI et ESI. Un mélange équimolaire de neurotensine et de maltoheptaose en acétate d'ammonium 5mM, 50 % méthanol est analysé avec une aiguille nanospray de diamètre interne 1 μ m (spectre de gauche) et 10 μ m (spectre de droite). D'après (Schmidt et al., 2003).

La diminution relative de l'intensité de l'ion [maltoheptaose + Na]⁺ interviendrait aux alentours de 50 nl/min, cette propriété n'est donc pas observée dans les systèmes actuels de couplage nanoLC-nanoESI qui utilisent des colonnes de 75 μ m de diamètre interne pour des débits de l'ordre de 200 nl/min.

Spectres de masse en mode ESI et calcul des masses

A partir d'analytes possédant plusieurs fonctions protonables, l'ionisation électrospray produit des ions moléculaires multichargés notés [M + zH]^{z+} en mode positif. Les spectres MS contiennent alors plusieurs pics, chacun correspondant à une valeur de z précise (**Fig. 3A**).

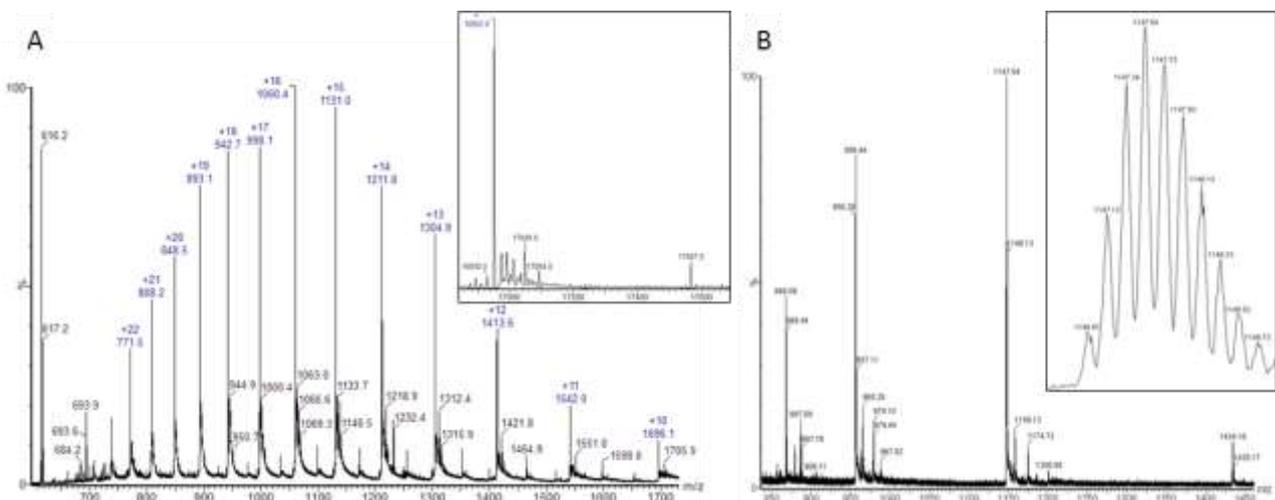


Figure 3. Spectre ESI-MS. (A): myoglobine, 16951 Da), (B); chaîne B de l'insuline, masse mono-isotopique M = 5729.6 Da. En insert, la déconvolution de la protéine et le massif isotopique du peptide 5 fois chargé (mesure réalisée avec un Q-TOF).

Dans le cas d'une protéine comme la myoglobine, l'intensité des ions qui correspondent aux différents niveaux de charge varie de manière régulière. Ici, dans un mélange 0.1 % acide formique 50 % méthanol aqueux, les états de charge +15 et +16 sont les plus intenses.

Deux pics contigus m_1 et m_2 correspondent à deux valeurs z_1 et z_2 avec $z_2 = z_1 - 1$. Il est alors possible de déterminer z puis calculer la masse de la molécule à partir des relations:

$$m_1 = (M + z_1H)/z_1 \quad (1)$$

$$m_2 = (M + z_2H)/z_2 \quad (2)$$

En remplaçant z_2 par $(z_1 - 1)$ dans l'équation (2), les solutions sont :

$$z_1 = (m_2 - 1)/(m_2 - m_1) \text{ et } M = m_1 z_1 - z_1$$

Soit pour le spectre de la myoglobine présenté en **Fig. 3A**, avec $m_1 = 1131.0$ et $m_2 = 1060.4$,

$z_1 = (1060.4 - 1)/(1131.0 - 1060.4) = 15.0$. La masse moyenne mesurée est alors

$$M = 1131 \times 15 - 15 = 16950 \text{ Da.}$$

Dans le cas où la résolution est suffisante pour séparer les pics du massif isotopique, il est possible de déterminer directement z à partir de la lecture du massif. Lorsque $z = 1$, la différence entre deux pics du massif est égale à une unité, Si $z > 1$, alors $z = 1/(m_b - m_a)$ avec m_b et m_a l'apex de deux pics contigus du massif. Pour la chaîne B de l'insuline (**Fig. 3B**), le massif isotopique de l'ion à plus intense à 1147.54 Th est présenté en insert. La résolution de l'instrument permet de séparer les différents pics. L'écart de masse entre 2 pics étant 0.19 (soit 0.20), $z = 5$ et $M = 1146.9 \times 5 - 5 = 5729.5 \text{ Da}$, soit une erreur de 2 ppm par rapport à la masse monoisotopique théorique.

1.1.3 Analyseurs hybrides Q TOF et IT Orbitrap

L'ensemble des analyseurs MS et MS/MS actuels peuvent être couplés à l'ionisation électrospray : Temps de vols (TOF), Trappe Ionique (IT), Quadripole (Q), Orbitrap, et Résonance Ionique Cyclotronique (ICR). Dans le cadre de cette thèse, deux instruments hybrides, c'est-à-dire associant des analyseurs différents, ont été utilisés : le Q-TOF Premier de Waters et l'Orbitrap XL de Thermo Scientific (IT-Orbitrap). Seuls ces deux appareils seront détaillés dans les paragraphes suivants.

Spectrométrie de masse hydride de type Q-TOF

Le spectromètre de masse hybride de type Q-TOF est constitué de 2 analyseurs, un quadripôle et un temps de vol, séparés par une cellule de collision (**Fig. 4**). Un guide d'ion appelé T Wave permet la transmission des ions de la source ESI au quadripôle.

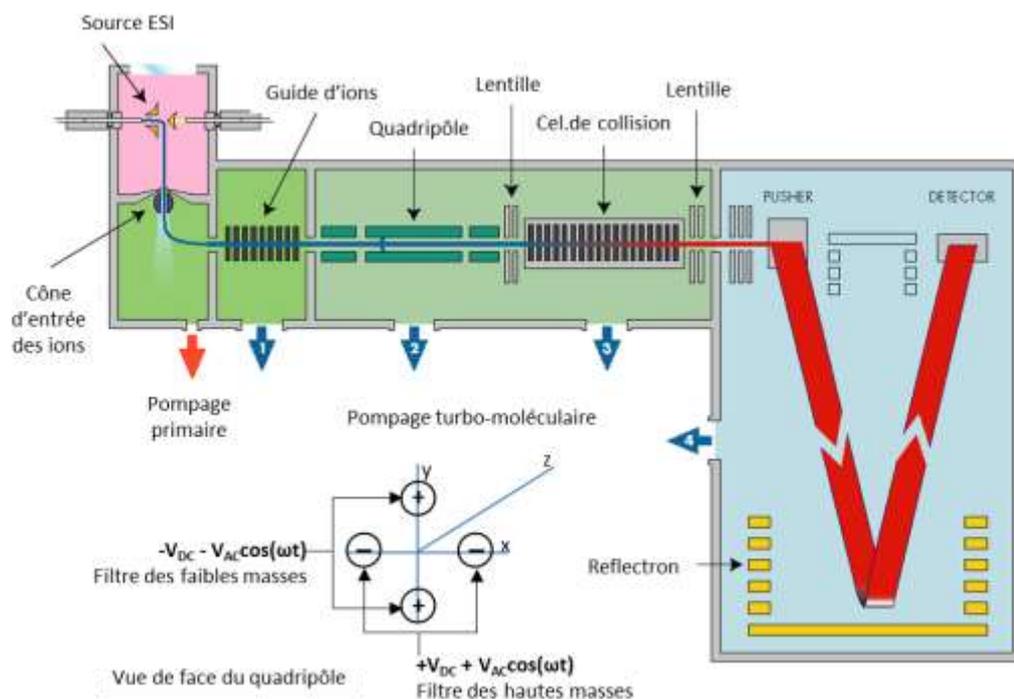


Figure 4. Q TOF Premier. L'appareil fabriqué par Waters utilise une source ESI qui permet d'infuser un étalon (Lock Spray) en même temps que l'analyte. Un guide d'ions (T Wave) permet de focaliser les ions vers le premier quadripôle (Resolving Quadripole). Un second module T-Wave sert de cellule de collision et un tube de temps de vol avec reflectron guide les ions vers un système de détection à galettes multicanaux.

Le quadripôle est un analyseur dont le principe fut publié en 1953 par Helmut Steinwedel et Wolfgang Paul, ce dernier recevant le prix Nobel de physique en 1989 pour cette invention (Paul, 1990). Le quadripôle est constitué de 4 électrodes soumises à un potentiel sinusoïdal ϕ_0 .

$\phi_0 = V_{DC} + V_{AC} \cos(\omega t)$ avec V_{DC} =, tension continue, $V_{AC} \cos(\omega t)$ = courant alternatif de fréquence angulaire ω .

Les potentiels sont appliqués sur les 4 électrodes de telle manière que les 2 électrodes opposées sont soumises aux mêmes tensions (voir fig. 4). Considérant un ion positif, les électrodes de tension $+V_{DC}$ filtrent les ions de haute masse tandis que les électrodes de tension $-V_{DC}$ filtrent les ions de basse masse. En l'absence de tension V_{DC} , le quadripôle fonctionne en radiofréquence uniquement, les électrodes vont inverser leur potentiel à tour de rôle et les ions seront focalisés au centre du faisceau. Le quadripôle installé sur le Q-TOF premier permet en théorie une fenêtre de sélection des ions d'une unité de masse; il est ainsi possible de sélectionner le pic monoisotopique au sein d'un massif et de ne transmettre que lui à la cellule de collision.

La cellule de collision est une chambre dans laquelle un gaz inerte, l'argon, est infusé à la pression de 4×10^{-3} mbar. Cette cellule est à l'origine un quadripôle fonctionnant en radiofréquence seule. Ici, la cellule de collision qui est appelée T wave est constituée d'une série de lentilles.

Chaque lentille est soumise à une tension continue et une tension alternative. Deux lentilles adjacentes sont soumises à des potentiels de phase opposée. La tension continue sur ces deux lentilles permet d'obtenir une barrière de potentiel qui stoppe les ions, cette barrière de potentiel avance pour créer une onde qui emmène les ions jusqu'au tube de vol. La pression d'argon et l'énergie des ions est telle que les ions se fragmentent.

Le tube de vol est placé perpendiculairement à la trajectoire des ions provenant du quadripôle. Il est doté d'un miroir électrostatique qui permet de doubler le parcours des ions dans le tube de vol : les ions décrivent un V du pusher (lentille qui focalise les ions et les propulse dans le tube de vol) jusqu'au détecteur.

Considérant uniquement le temps de vol après l'accélération:

Sachant que $E = \frac{1}{2} mv^2$ (E : énergie cinétique de l'ion accéléré)

$E = qU$ (q : charge électrique de l'ion, U : tension appliquée)

$v = d/t$ (d : distance de vol)

$$\text{Alors } t = \frac{d}{\sqrt{2U}} \sqrt{\frac{m}{q}}$$

En d'autres termes, le temps nécessaire à un ion pour atteindre le détecteur sera proportionnel à la racine carrée de son rapport m/z. Pour l'appareil utilisé au laboratoire, l'analyseur TOF permet une justesse supérieure à 10 ppm et une résolution de l'ordre de 9000.

Les modes d'acquisition permettent l'obtention de spectres MS ou MS/MS. En mode MS, seule une radiofréquence est appliquée au quadripôle, l'ensemble des ions est transmis à la cellule de collision dont l'énergie appliquée est réduite pour empêcher la fragmentation des ions, le TOF mesure l'ensemble des masses qui lui sont transmis. En mode MS/MS, les tensions et la radiofréquence du quadripôle sont ajustées pour ne permettre la transmission que d'un ion de m/z choisi. La sélection avec le quadripôle permet d'isoler un ion à la masse m/z = 1000 +/- 1 avec une intensité correspondant au 3/4 de l'intensité de l'ion en mode MS. L'ion parent est transmis à la cellule de collision dont l'énergie est augmentée afin de permettre une fragmentation CID. Les ions fragments (ions fils) sont finalement transmis au TOF et mesurés. Les temps de mesure pour produire un spectre MS et MS/MS avec un rapport Signal/bruit satisfaisant sont respectivement de l'ordre de 0.5 sec et 2 sec. Il est nécessaire d'augmenter le « temps de scan » pour les spectres MS/MS afin d'accumuler suffisamment de micro-scans et compenser la baisse d'intensité propre à la fragmentation.

La résolution et la précision de mesure du Q TOF sont fonction du TOF. L'association du quadripôle au TOF permet de combiner la robustesse du quadripôle pour la sélection d'ions parent avec la précision de mesure du temps de vol (~ 5 ppm, FWHM $> 10\,000$) (Chernushevich et al., 2001). Cette résolution peut être poussée à 17 000 lorsque le deuxième réflecteur, disposé entre le pusher et le détecteur, est activé et permet aux ions de parcourir alors un trajet en « W ». Malheureusement, la réduction de sensibilité associée au mode W ne le rend utilisable que pour des échantillons non limités en quantité.

Spectrométrie de masse hybride de type IT Orbitrap

Cette géométrie correspond à un appareil commercialisé par Thermo scientific : le LTQ Orbitrap XL (Makarov et al., 2006). L'instrument est constitué d'une trappe ionique linéaire (LTQ pour Linear Trap Quadrupole), d'une cellule de transfert C-Trap, d'une cellule de collision HCD (HCD pour Higher Collision Dissociation) et d'une deuxième trappe ionique de type Orbitrap (**Fig. 5**).

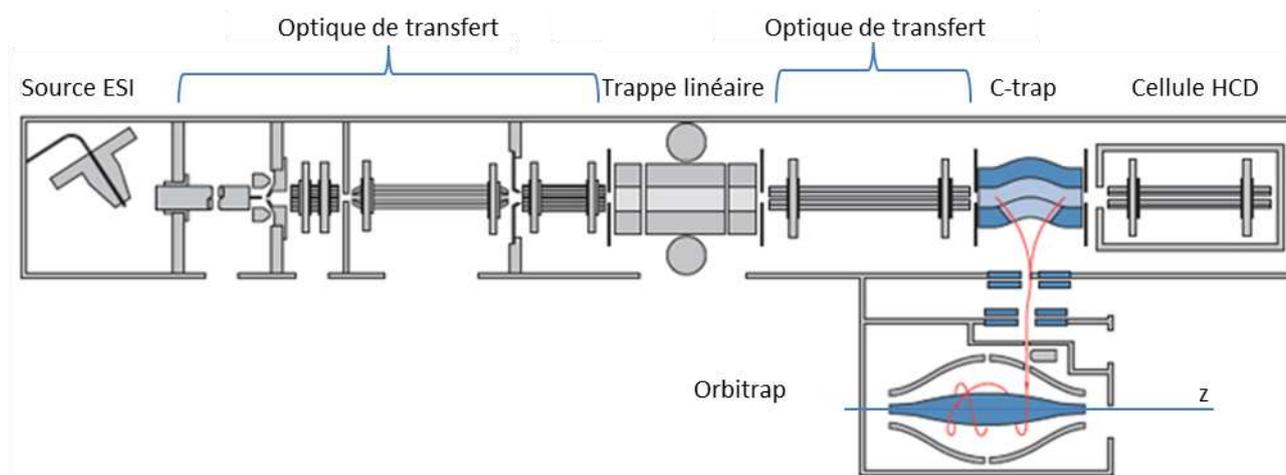


Figure 5. Orbitrap XL. L'Orbitrap XL est un spectromètre de masse hybride composé d'une trappe linéaire, d'une cellule de fragmentation HCD et d'un analyseur Orbitrap. La trappe ionique permet des mesures MS et MS^n avec une résolution voisine de 2000 pour une exactitude de l'ordre de 100 ppm. L'Orbitrap permet des mesures de masse exacte avec une erreur inférieure à 2ppm et une résolution supérieure à 50000.

La trappe linéaire est un quadripôle possédant des électrodes supplémentaires à ses extrémités (end cap). Les électrodes d'extrémité sont des portes électrostatiques qui permettent le confinement des ions dans la dimension axiale (z) avec l'aide d'une pression de gaz inerte (1 mtorr d'Helium). Les ions sont ensuite éjectés vers la C trap. La trappe linéaire possède également son propre système de détection avec une éjection radiale des ions qui permet de produire des scans MS et MS^n .

La C-trap est un piège électrostatique qui permet de focaliser les ions au moyen d'une radiofréquence et d'une pression de gaz inerte. Elle permet d'éjecter les ions orthogonalement vers l'Orbitrap. Cette éjection est pulsée et focalise à haute vitesse les ions qui pénètrent dans l'orbitrap par une fenêtre de 1 x 10 mm (Makarov, 2000).

La cellule HCD est une cellule de collision quadripolaire munie de miroirs électrostatiques à ses extrémités pour permettre de retenir puis expulser en sens inverse les ions qu'elle contient. Elle est utilisée notamment pour palier l'incapacité de la trappe linéaire à piéger des ions fragments de m/z inférieurs à environ 1/3 de la valeur du m/z de l'ion parent.

L'Orbitrap est une trappe ionique orbitale constituée d'une électrode centrale en forme de barre entourée par une électrode externe cylindrique. Les ions sont introduits dans l'Orbitrap de manière tangentielle et tournent autour de l'électrode centrale. Comme les planètes tournent autour du soleil, l'attraction électrostatique dirigée vers l'électrode centrale est compensée par la force centrifuge des ions. Les aspects mathématiques du mouvement des ions dans l'Orbitrap ont été publiés en 2000 par Alexander Makarov (Makarov, 2000). La vitesse angulaire de déplacement d'un ion de long de l'axe z est fonction de sa charge q et de sa masse m :

$$\omega = \sqrt{(q/m)k} \text{ en rad/sec (soit } \omega = \sqrt{k(\frac{e}{m/z})}.)$$

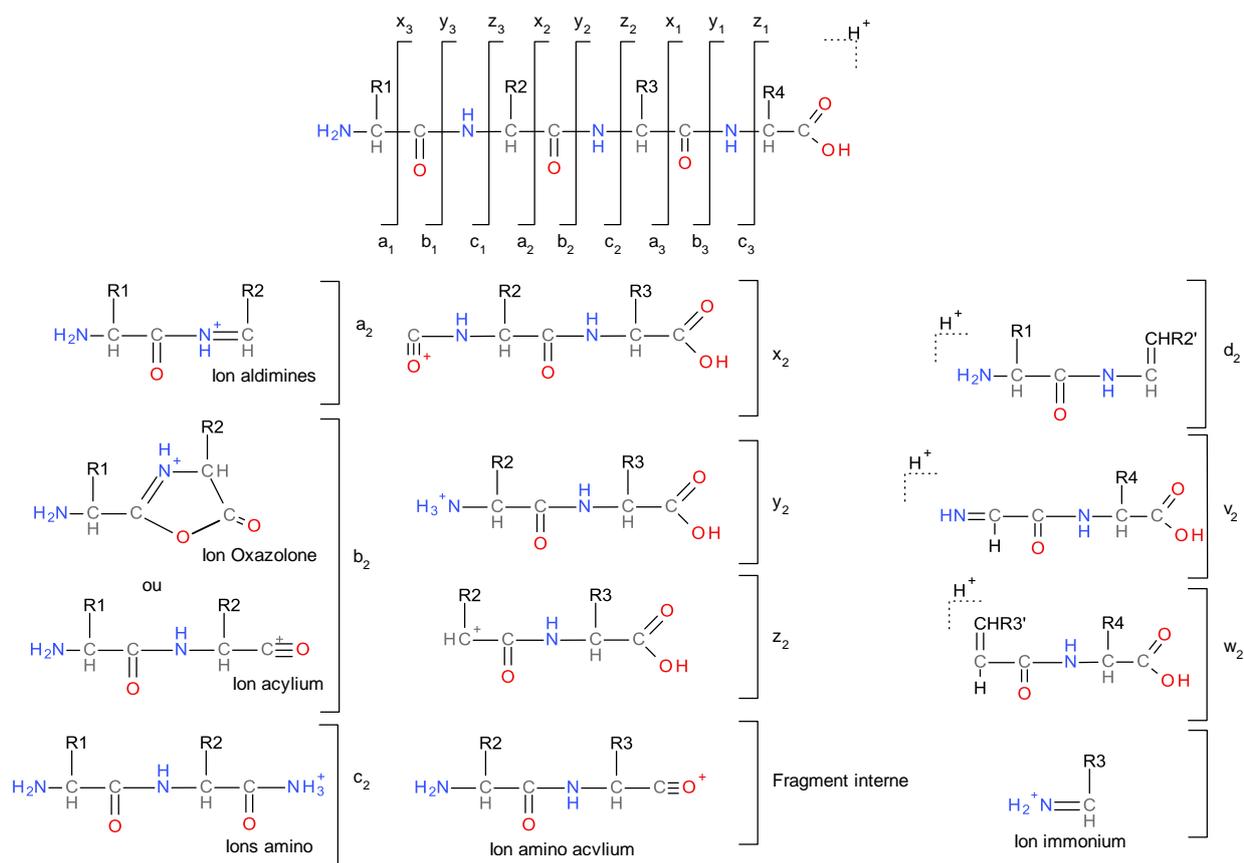
A titre d'exemple, la fréquence d'oscillation f le long de l'axe z pour m/z 524 est de 300 kHz ($f = \omega / 2\pi$) (Makarov et al., 2006). Sa fréquence d'oscillation sur l'axe z crée un courant induit qui est mesuré par l'électrode externe. Dans le cas de l'analyse d'un mélange d'ions, le signal mesuré est alors la combinaison des fréquences des différents ions piégés. Une transformation de Fourier permet de décomposer le signal de synthèse en ses différentes composantes ; des traitements additionnels du signal permettent de reconstituer le spectre de masse.

La résolution et la précision de mesure de l'Orbitrap sont sensiblement supérieures au Q-TOF. A titre d'exemple, sur un temps d'acquisition dans l'Orbitrap de 1.07 s, la résolution FWHM est de 100 000 pour m/z = 400 (Scigelova et al., 2011). Suivant l'étalonnage, l'exactitude de mesure est inférieure en routine à 2 ppm. Lee et collaborateurs ont publié une erreur moyenne de 0.1 ± 1.7 ppm avec un étalonnage interne (Lee et al., 2011). L'utilisation d'ions contaminants présents dans les spectres MS permet également un re-étalonnage dynamique. Dans ce cas, les calibrants sont choisis automatiquement en fonctions de leur intensité et permettent une erreur moyenne de 0.03 ± 0.5 ppm (Zhang et al., 2011).

1.1.4 Fragmentation par CID des peptides en mode positif

Nomenclature

La dissociation induite par collision (CID) est un processus de fragmentation à basse énergie qui transfère aux ions une énergie de quelques dizaines d'électron-volt et qui peut être décomposé en deux étapes. La première étape correspond à une excitation qui est la conséquence des collisions de l'ion avec le gaz de la cellule de collision ; une partie de l'énergie cinétique est convertie en énergie interne. La deuxième étape correspond à une décomposition de l'ion parent en ions fragments. Les fragments observés fournissent une information de séquence lorsque la rupture se produit au niveau des liaisons amide le long du squelette peptidique. Une nomenclature a été établie à la fin des années 80 et propose une classification en fonction de l'extrémité N- ou C-terminale contenue dans le fragment (Biemann, 1990; Roepstorff and Fohlman, 1984) (**Fig. 6**).



Perte d'ammoniac (-17 sur R,H,N,Q) notée a*, b*, y*. Perte d'eau (-18 sur S,T,E,D) notée a⁰, b⁰, y⁰.

Figure 6. Nomenclature des ions obtenus par la fragmentation d'un peptide en mode positif. Les ions fragments a, b, c (ions portant la partie N-terminale) et les ions fragments x, y, z (fragments portant la partie C-terminale) sont produits en fonction du site de clivage de la liaison peptidique. Les autres ions fragments correspondent à une fragmentation des chaînes latérales (ions d, v, w), ou à des doubles clivages conduisant à des ions fragments internes ou à des ions immonium caractéristiques des acides aminés.

Les ions fragments a, b, c contiennent l'extrémité N-terminale alors que les ions fragments x, y, z contiennent l'extrémité C terminale. Les différentes lettres précisent le site de clivage de part et d'autre du carbone et de l'azote de la liaison amide, la numérotation en indice renvoie à la position de l'acide aminé concerné par la coupure en considérant le premier acide aminé comme l'acide aminé N- ou C-terminal contenu dans le fragment. Cette nomenclature envisage la totalité des ions fragments imaginables. Suivant le mode de fragmentation utilisé, il est possible de ne produire que certains types de fragments. Par exemple, la fragmentation CID à basse énergie ne produit pas de fragments c, x, z. La structure des ions b a été initialement décrite comme linéaire (ion acylium), l'étude des voies de fragmentation a démontré par la suite une structure cyclique à 5 éléments de type oxazolone. Plus récemment, une structure cyclique a également été démontrée pour les ions a (Bythell et al., 2010).

Interprétation des spectres MS/MS

La fragmentation CID à basse énergie génère essentiellement des ions b et y. Dans le cas idéal où les séries d'ions b et y sont complètes, la lecture de la séquence consiste à identifier les séries puis calculer les différences entre fragments de la même série. Par exemple, $b_n - b_{n-1}$ permet d'identifier l'acide aminé en position n. Le succès de cette approche dépend grandement de la qualité des séries. Si des ions sont manquants ou peu intenses, ce qui correspond à la majorité des cas, le séquençage *de novo* est alors difficile voire impossible. De nombreux auteurs ont cherché à comprendre les mécanismes qui interviennent lors de la fragmentation afin de déterminer les relations entre la nature des fragments et leur intensité.

Vue d'ensemble des voies de fragmentation

Une vue d'ensemble des voies de fragmentation est présentée dans les revues de Paizs et Suhai (Paizs and Suhai, 2004, 2005) et de Harrison (Harrison, 2009b). Deux grandes voies de fragmentations sont définies en fonction de la mobilité des protons (Dongre et al., 1996) et sont basées soit sur un mécanisme de fragmentation appelé charge dirigée (charge directed, ChD) soit sur un mécanisme de fragmentation appelé charge à distance (charge remote, ChR) (**Fig. 7**).

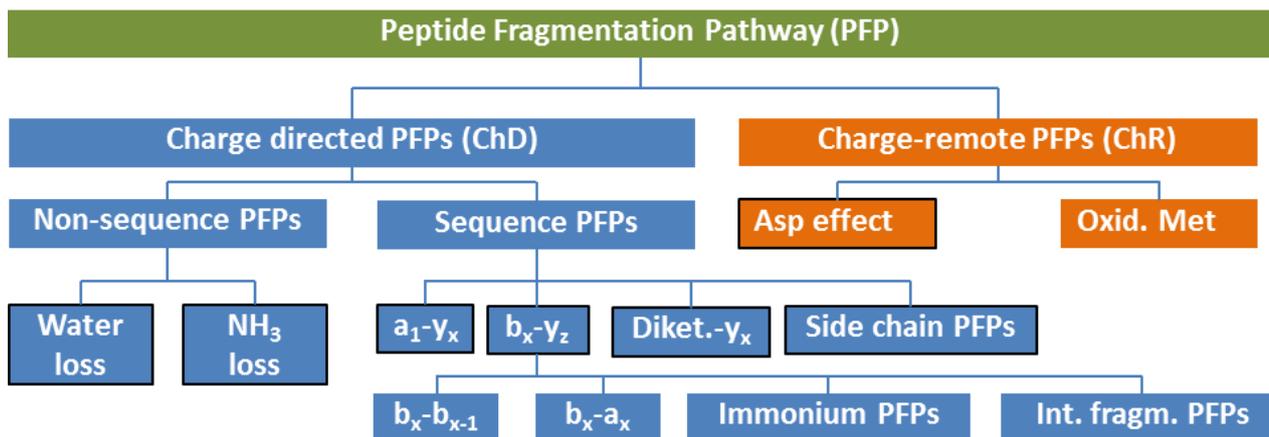


Figure 7. Les différentes voies de fragmentation (connues et supposées) des peptides protonés (Paizs and Suhai, 2005). Une première différence s'opère en fonction de la mobilité du proton. S'il est séquestré par une arginine, la fragmentation sera de type *charge remote* avec l'intervention d'un proton acide (Asp effect) ou du sulfoxide d'une méthionine oxydée (Oxid Met). Si le proton est mobile, la fragmentation sera de type *Charge directed* avec soit une perte de neutre et la production de fragments non informatifs pour ce qui est de la séquence en acides aminés (Non Sequence PFPs) soit une fragmentation des liaisons amide suivant différents schémas réactionnels qui aboutiront *in fine* à la formation d'ions a,b,y, de fragments internes et d'ions immonium. Les mécanismes encadrés en noirs sont détaillés par la suite.

Dans le cas du mécanisme ChD, le proton est mobile, il est associé à une fonction amine présentant une affinité protonique qui permet au proton de se déplacer entre ses différents accepteurs potentiels. La fragmentation peut alors correspondre à une perte de neutre en laissant intact les liaisons peptidiques (*non sequence PFPs*). Cette fragmentation ne permet pas de déterminer la séquence en acides aminés, elle est sans intérêt sauf à signer la présence de fonctions carboxyliques ou de fonction alcool dans le cas d'une perte d'eau (*Water loss*) ou de lysine ou arginine, asparagine et glutamine dans le cas d'une perte d'ammoniac (*NH₃ loss*). La voie b_x-a_x peut être également considérée comme une perte de neutre puisqu'elle correspond à une perte de CO (-28u) des ions b_x . Dans le cas plus intéressant des mécanismes de fragmentation aboutissant à des fragments informatifs en termes de séquences (*Sequence PFPs*), plusieurs voies sont proposées et expliquent la formation des différents fragments observés en CID : a, b, y, ions immonium et ions fragments internes. La première étape de la réaction est une délocalisation de la charge apportée lors de l'ionisation sur l'azote des différentes liaisons amide (azote amide).

Dans le cas du mécanisme ChR, la charge est fixée par un groupement ayant une forte affinité protonique (ammonium quaternaire, arginine) et le proton n'est plus mobile. Le mécanisme de fragmentation implique alors des protons localisés par exemple sur la fonction carboxylique de l'acide aspartique (Asp effect, voir paragraphe suivant).

Fragmentation ChR

Beaucoup de peptides contenant Asp et Glu produisent des fragments b intenses avec Asp ou Glu en position C-terminale. Cette fragmentation orientée est observée essentiellement quand le nombre de protons est égal au nombre d'arginines. Dans ce cas, l'arginine séquestre le proton qui n'est plus mobile, la fonction carboxylique de Asp transfère alors son proton à l'amide voisin coté C-terminal et crée une interaction ionique (salt bridge) avec l'arginine ; la fragmentation intervient alors à droite de l'acide aspartique. Si le nombre de charges est supérieur au nombre d'arginines, le mécanisme fera alors intervenir le mécanisme ChD et le clivage ne sera plus orienté (Gu et al., 2000; Wysocki et al., 2000). Cependant, l'utilisation de l'échange H/D et la dérivation N terminale par une amine quaternaire montre que les mécanismes ChR et ChD peuvent intervenir sur un même ion (Cydzik et al., 2011).

Fragmentation ChD, modèle du proton mobile

Les peptides peuvent être protonés sur différents sites ; amine N terminale, chaîne latérale, azote d'un amide. Lors de l'excitation dans la cellule de collision, l'énergie interne des ions augmente et les sites de protonation énergétiquement moins favorables peuvent alors se protoner par déplacement du ou des proton(s) mobile(s) (Dongre et al., 1996; Wysocki et al., 2000). Le modèle du proton mobile fait donc référence à la possibilité de migration des protons des groupements ayant une affinité protonique élevée vers des groupements qui ont une affinité protonique moindre comme l'azote de la liaison amide. La détermination de l'affinité protonique (PA) des acides aminés et des peptides est détaillée dans la revue de Harrison (Harrison, 1997). Elle est fonction de la basicité en phase gazeuse (GB) tel que $PA = GB + T\Delta S^0$, soit le classement suivant: Arg>Lys>His> α NH₂>>N amide.

La protonation de la fonction amide a pour conséquence de fragiliser la liaison peptidique et transforme le carbone de groupe CO en un site favorable aux attaques nucléophiles des groupements voisins riches en électrons. Deux voies de fragmentation ont alors été abondamment décrites ; la voie « b_x-y_z » et la voie « dikétopipérazine Y_{N-n} ».

Voie b_x-y_z

Dans cette voie de fragmentation, le proton mobile saute de sa position initiale, par exemple N-terminale, à une fonction amide qui est alors fragilisée (**Fig. 8**). Le carbone en position n-1 de l'amine protonée devient un site accepteur d'électrons. L'attaque nucléophile du groupement

cétone de la liaison amide n-1 provoque une cyclisation et la formation d'un intermédiaire oxazolone protoné capable de réarrangement. Selon l'affinité protonique de la partie N-terminale et C-terminale, des ions b ou y seront formés. Dans le cadre de peptides doublement chargés, les deux fragments peuvent être observés conjointement.

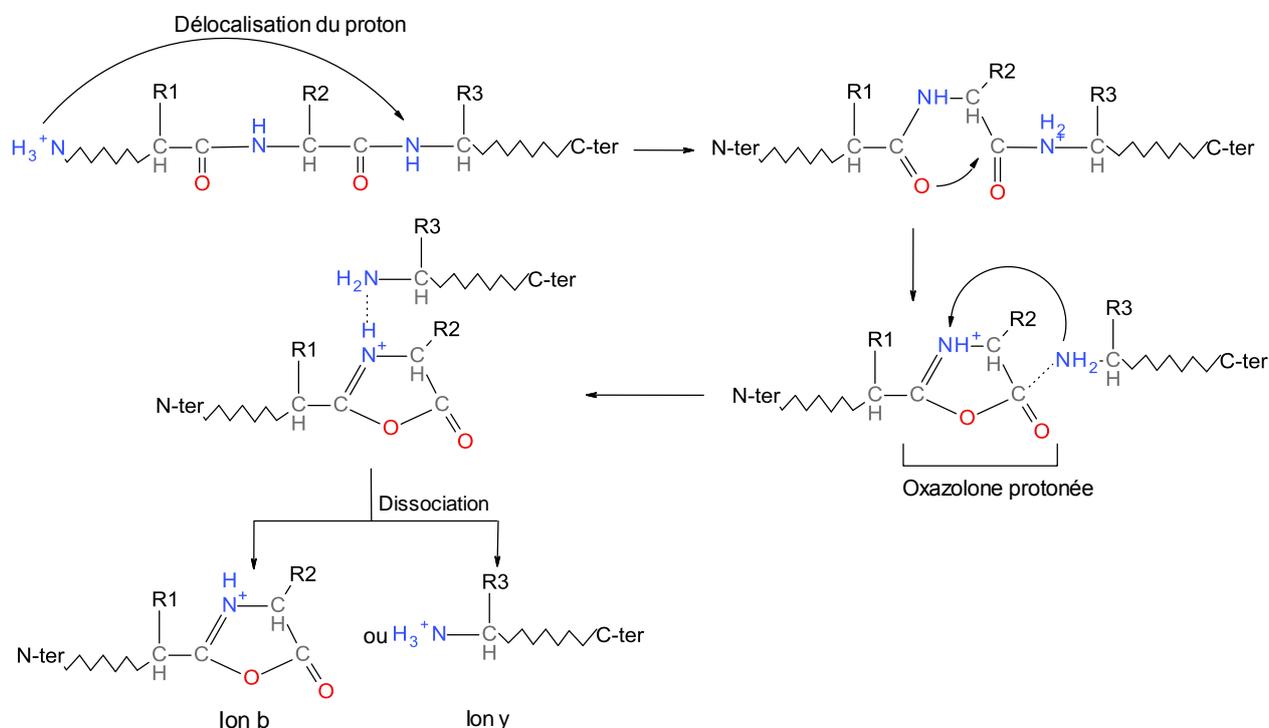


Figure 8. Mécanisme de fragmentation de la voie b_x - yz . La protonation de l'azote de la fonction amide rend le carbone adjacent électrophile. Une cyclisation (formation d'un oxazolone) et un réarrangement conduiront à la formation d'ions b ou y.

A partir des fragments oxazolone b_x formés, de nouveaux fragments oxazolone b_{x-1} et des fragments a_x peuvent être produits. Une conséquence directe de l'étude de ces réactions en phase gazeuse est la description des ions b non plus sous une forme linéaire mais sous une forme cyclique de type oxazolone.

Voie dicétopipérazine Y_{N-n}

La voie dicétopipérazine Y_{N-n} (**Fig. 9**) génère uniquement des ions y. N correspond au nombre d'acides aminés et n désigne l'acide aminé à droite duquel la dissociation a lieu. Dans un premier temps, le proton mobile saute sur le groupement NH d'une fonction amide à l'identique de la voie b_x - yz . Par contre, l'attaque nucléophile du carbone en position n-1 de l'amine protonée est menée ici par l'amine N-terminale, provoquant une cyclisation de la partie N-terminale sur le carbone électrophile. Pour les ions parents mono-chargés, cet intermédiaire dicétopipérazine n'a pas une

affinité protonique suffisante pour se protoner, un fragment y sera systématiquement formé (Paizs and Suhai, 2005). Par contre pour les peptides tryptiques doublement chargés, la formation dicétopipérazine b_n protonée est majoritaire (Savitski et al., 2008).

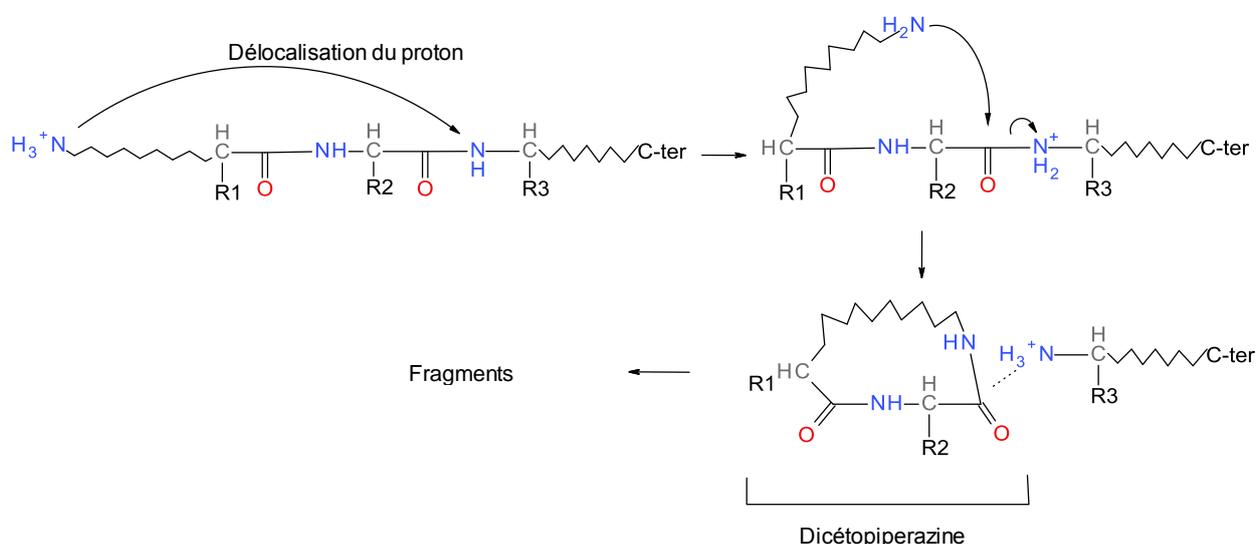


Figure 9. Mécanisme de fragmentation de la voie Y_{N-n} . La protonation de l'azote de la fonction amide rend le carbone adjacent électrophile. Une cyclisation par l'amine N-terminale (formation d'un cycle dicétopipérazine) conduira à la formation d'ion b ou y .

Pour les ions Y_{N-3} , Y_{N-4} , Y_{N-5} , les pertes de neutre sont respectivement des cyclo-tri, cyclo-tetra, cyclo-penta-peptides. L'ion Y_x formé possède à la fois le proton mobile et le proton provenant de l'amine N-terminale. Dans le cas particulier où $n = 2$ (Y_{N-2}), la dicétopipérazine N-terminale doit être précédée d'une isomérisation cis-trans pour que l'amine N-terminale soit placée à proximité du carbone électrophile. L'étude du peptide penta-alanine a montré que cette voie était minoritaire par rapport à la voie b_x-y_z , notamment à cause des contraintes stériques imposées par le repliement de l'amine N terminale sur les liaisons amides (Paizs and Suhai, 2004).

Voie a_1-y_x

La voie a_1-y_x (**Fig. 10**) est initiée par la relocalisation du proton mobile sur l'azote de la liaison amide voisine puis la formation d'un dimère constitué d'une partie aziridone (cycle à 3 éléments) et d'une partie amine. L'aziridone n'est pas stable. Si la charge reste sur le cycle, une perte de CO aboutira à la formation d'un ion a (**Fig. 10**, voie 1). Si la charge est sur l'amine, le fragment y sera alors détecté (**Fig. 10**, voie 2). Ce mécanisme a été décrit par Harrison et col. dans le cadre de l'étude de di-peptides Val-Xxx (Harrison et al., 2000).

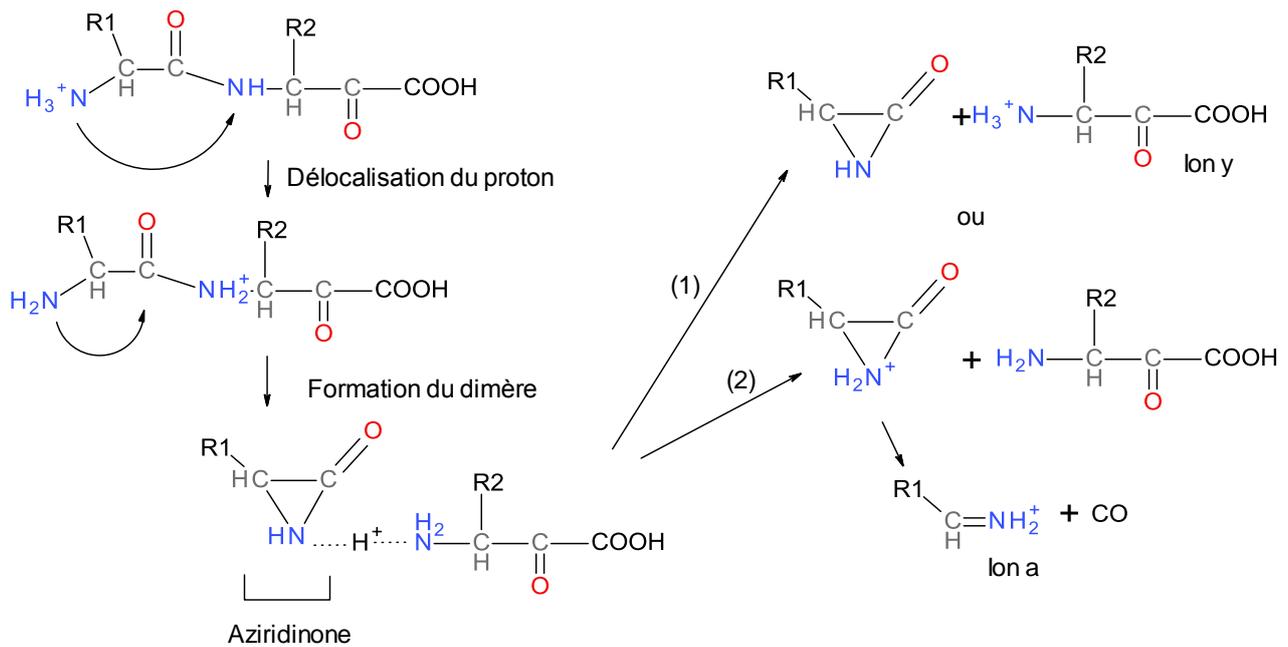


Figure 10. Mécanisme de fragmentation de la voie α_1 - γ_z . Après relocalisation du proton mobile sur l'amide N-terminale puis formation d'un dimère aziridinone-amine, un fragment α_1 ou γ_{N-1} sera produit en fonction de la localisation du proton.

Autres voies de fragmentation

- Protonation de l'oxygène amide

La protonation de l'oxygène de la fonction amide est énergétiquement plus favorable que la protonation de l'azote de la fonction amide, ce qui a amené différents auteurs à proposer des mécanismes de dissociation qui n'impliquent pas l'azote-amide mais l'oxygène-amide (Wysocki et al., 2000). Ici, le mécanisme est initié par la formation d'un groupement carbonyl au niveau de la fonction amide attaquée. Une cyclisation avec transfert de la charge sur l'amine du cycle est ensuite observée. Dans un troisième temps les fragments b ou γ sont formés. Si la protonation de l'oxygène est thermodynamiquement favorable, la forme intermédiaire cyclisée est quant à elle beaucoup moins favorable.

- Effet de la proline

La présence d'une proline entraîne généralement la production d'ion γ_n intenses produits par le clivage du côté N-terminal de la proline. La liaison amide du côté C-terminal de la proline est stabilisée et la forte affinité protonique de la proline favorise l'observation des fragments γ (Bleholder et al., 2011). Cette observation est robuste et peut être utilisée comme critère pour la validation des identifications (Tabb et al., 2003).

- Perte d'eau et d'ammoniac

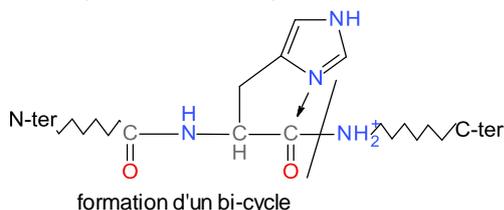
Les pertes d'eau se produisent sur les fragments b et y contenant Asp, Glu, Ser et Thr ainsi que sur la fonction carboxylique C-terminale. Les pertes d'ammoniac se produisent également sur les fragments b et y des peptides contenant Gln, Asn, His, Lys et Arg ainsi que sur l'amine N-terminale. Pour la perte d'eau, les fragments y^0 (i.e. $y - H_2O$) sont plus fréquents et plus intenses que les fragments b^0 mais ils ne sont pas caractéristiques d'acides aminés contenant des oxygènes dans la chaîne latérale. Un cas particulier concerne le glutamate en position N-terminale qui se transforme généralement en pyroglutamate accompagnée par une perte d'eau importante. Pour l'ammoniac, au contraire, les fragments b^* (i.e. $b - NH_3$) sont généralement spécifiques (>95 %) des acides aminés basiques (Savitski et al., 2007).

- Clivage de la liaison amide par attaque nucléophile des chaînes latérales

Ce mécanisme intervient essentiellement suite à une attaque de type *charge remote* de l'histidine, de la lysine et de l'arginine (**Fig. 11**). Ces acides aminés ont leur chaîne latérale protonée. Le transfert de ce proton sur l'azote de la fonction permet une attaque nucléophile du carbonyle en α de l'amine protonée puis la formation d'un cycle (Paizs and Suhai, 2005).

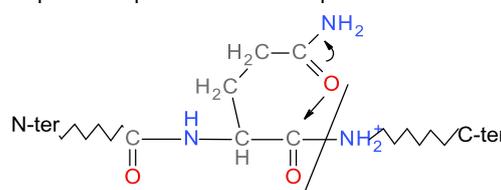
Histidine:

Attaque nucléophile du CO amide par N imidazole



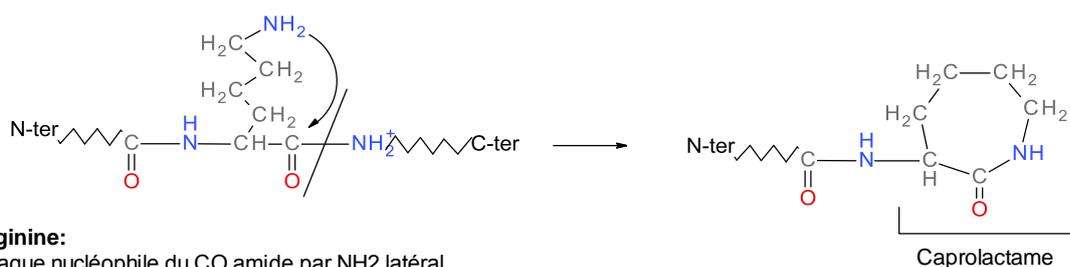
Glutamine, asparagine:

Attaque nucléophile du CO amide par O amide latérale



Lysine:

Attaque nucléophile du CO amide par NH_2 latéral



Arginine:

Attaque nucléophile du CO amide par NH_2 latéral

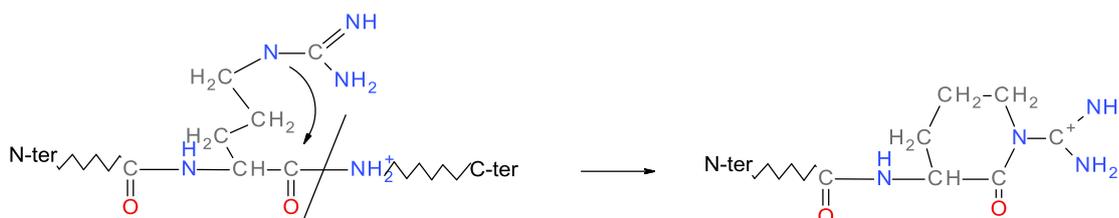


Figure 11. Attaque nucléophile de la liaison amide par les chaînes latérales. D'après (Paizs and Suhai, 2005)

Dans le cas de l'histidine, le groupement imidazole est très probablement protoné et transfère son proton à l'azote de la liaison amide adjacente, coté C-terminale. Dans un deuxième

temps, l'attaque nucléophile de l'oxygène amide par l'azote imidazole crée un bi-cycle qui entraîne le clivage de la liaison amide avec l'observation d'un ion b_x ou y_z (Wysocki et al., 2000). Avec une trappe ionique, l'histidine favorise nettement l'observation des ions b portant l'histidine à leur extrémité N-terminale (Tabb et al., 2003).

Dans le cas de la glutamine ou de l'asparagine, le groupement amide latéral (attaque nucléophile de l'oxygène amide) peut être impliqué dans la rupture de la liaison amide côté C-terminal.

Perte du phosphate en CID

L'analyse en mode CID de peptides contenant des phosphosérines et phosphothréonines montre qu'ils sont fréquemment sujets à la perte d'acide phosphorique (H_3PO_4 , 98 Da). Cette perte de neutre sur les ions précurseurs protonés produit des spectres MS/MS avec un pic $[M + nH - H_3PO_4]^{n+}$ majoritaire. L'observation de cet ion peut être utile pour signer des peptides contenant une phosphosérine ou une phosphothréonine. En contrepartie, les ions b et y qui portent l'information de la séquence sont peu intenses, et la localisation précise du groupement phosphate peut être difficile voire impossible si plusieurs acides aminés sont susceptibles d'être phosphorylés.

Plusieurs mécanismes sont proposés pour expliquer la perte de neutre par élimination d'acide phosphorique (**Fig. 12**). Il était couramment admis que l'élimination de H_3PO_4 en phase gazeuse faisait intervenir un mécanisme de β -élimination (DeGnore and Qin, 1998). D'autres mécanismes ont été décrits par la suite, avec la formation d'aziridine ou d'oxazolone (Reid et al., 2000). De manière générale, ces mécanismes de cyclisation correspondent à une attaque nucléophile de l'acide aminé phosphorylé sur son oxygène carbonyle en position β (Rozman, 2011). Pour des peptides Pro-pSer/pThr, une fragmentation aboutissant à un ion y particulier ($y+10$ amu) a été décrite (Medzihradzky and Trinidad, 2012). Dans ce cas, la formation d'un cycle oxazolone entre le carbone β de l'acide aminé phosphorylé et le carbonyle de la liaison amide $n-1$ produit un fragment y de type oxazolone qui contient un carbone supplémentaire (+12) par rapport à un fragment y classique, mais également une insaturation sur le cycle oxazolone (-2) soit un incrément de 10 amu par rapport à l'ion y habituel.

La perte de neutre peut également correspondre soit à une élimination directe de l'acide phosphorique soit à une élimination en 2 temps, d'abord du groupement phosphoryl (HPO_3) puis de H_2O (Palumbo and Reid, 2008).

Enfin, dans les trappes ioniques où la fragmentation est réalisée sur des ions confinés pendant des temps longs de l'ordre de la dizaine ou centaine de millisecondes, le réarrangement du phosphate (« scrambling ») sur un autre acide aminé hydroxylé a été décrit, indépendamment de la nature de l'acide aminé phosphorylé (Cui and Reid, 2013). L'étude de 33 peptides de synthèse phosphorylés sur sérine, thréonine et tyrosine a montré un réarrangement pour les 3 acides aminés (Palumbo and Reid, 2008). Dans cette étude, 45 % des peptides sont sujets au réarrangement.

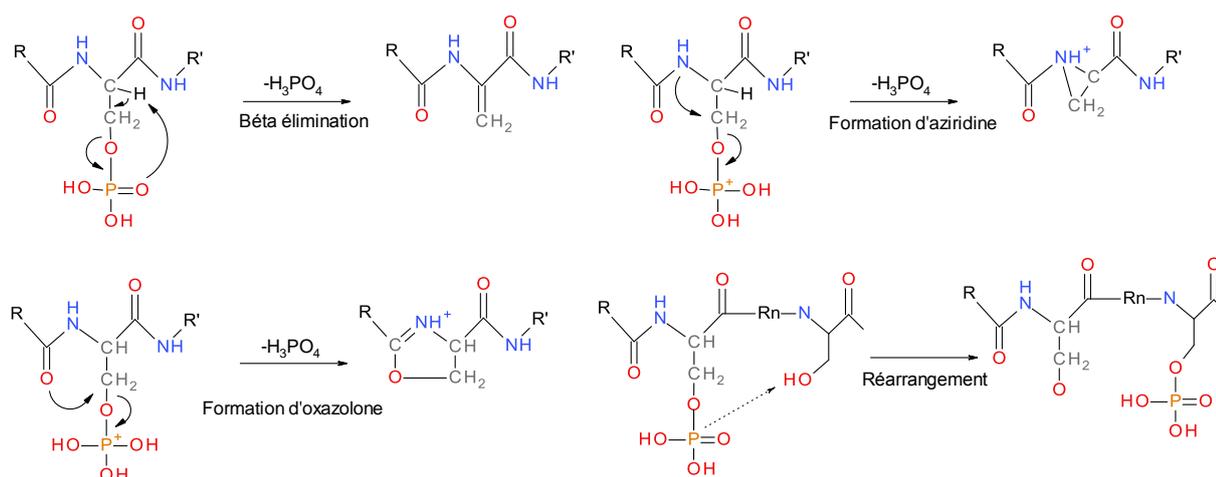


Figure 12. Réaction du phosphate en phase gazeuse. β -élimination, formation de cycle à 3 ou 5 atomes et réarrangements. Cette liste n'est pas exhaustive.

Prévision des intensités des ions fragments

La connaissance des mécanismes de fragmentation et des énergies nécessaires à ces réactions permet d'envisager la construction de modèles prédictifs calculant l'intensité des différents fragments en fonction de la séquence de l'ion moléculaire. Il faut noter que l'ensemble des articles cités précédemment développent cette idée. Différents travaux présentent des relations entre les fragments observés. Une relation linéaire a été établie entre la basicité en phase gazeuse (GB) de tripeptides Gly-Gly-Xxx-OH et $\log(y_1/b_2)$ (Harrison, 1999). Cette relation a été étendue sous la forme $\ln(b_x/y_z) = (PA_{N\text{-ter}} - PA_{C\text{-ter}})RT_{\text{eff}}$ pour prévoir l'intensité des ions b et y produits par la voie $B_x\text{-}Y_z$ (Paizs and Suhai, 2004).

Une approche différente mettant en œuvre une analyse en composante principale (ACP) de la distribution des ions fragments de 15 000 peptides tryptiques a montré que l'ACP permettait de distinguer deux classes de peptides : une classe 1 de peptides courts (8.4 résidus) dont les spectres MS/MS sont dominés par les fragments b_2 et y_{N-2} et une classe 2 de peptides plus longs (12.4 résidus) dont les fragments majoritaires sont Y_{N-4} et Y_{N-5} (Savitski et al., 2008). L'explication de cette distribution bi-modale est basée sur l'hypothèse d'une isomérisation cis-trans plus rapide des

acides aminés N-terminaux pour les peptides les plus courts, permettant une cyclisation et la formation d'un ion b_2 dicétopipérazine. L'étude de peptides $(Ala)_xHis$ avec x allant de 5 à 10 a confirmé la présence de ces 2 classes de peptides dépendant de la longueur de chaîne (Harrison, 2009a). Par des expériences d'échange isotopique hydrogène/deutérium, Bythell et collaborateurs ont infirmé l'hypothèse de formation des ions b_2 de la classe 1 suite à l'isomérisation puis la formation de dicétopipérazine, et indiqué que les fragments b_2 ont une structure habituelle de type oxazolone (Bythell et al., 2009). En tant qu'exception, les peptides courts avec Gln ou Glu en position N-terminale ne forment pas d'ion b_2 intenses à cause de leur facilité à perdre une molécule d'eau mais se cyclisent en imide dont l'affinité protonique est faible ; la charge se relocalise alors sur la partie centrale du peptide (Godugu et al., 2010).

Les règles de fragmentation présentées dans les paragraphes précédents ont été intégrées dans un modèle de prédiction de l'intensité des fragments et appliquées à une bibliothèque de 2605 spectres CID préalablement identifiés. Les scores de similarité entre spectres théoriques et spectres observés augmentent de 0.4 à 0.7 entre la méthode conventionnelle (corrélation sans prendre en compte l'intensité des ions) et la méthode publiée qui détermine l'intensité des différents ions fragments (Zhang, 2004). Ce modèle a été amélioré et prend maintenant en compte les règles de fragmentation propres à différentes modifications ; phosphorylation de S, T, Y oxydation de M, hydroxylation, méthylation et acétylation de K (**Fig. 12.1**) (Zhang, 2011). A titre d'exemple, la figure ci-dessous présente le spectre calculé et le spectre observé pour un peptide phosphorylé sur la sérine.

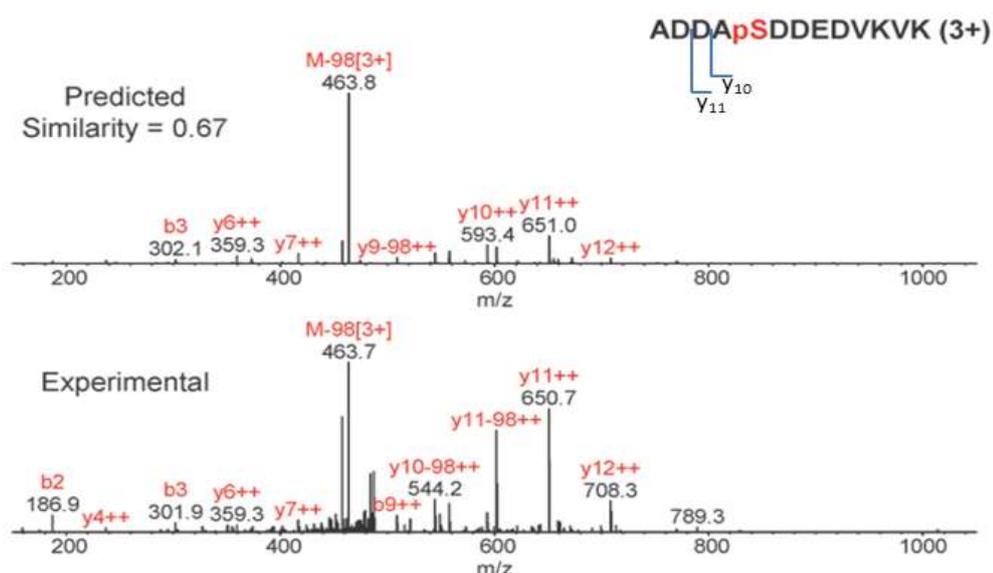


Figure 12.1. Prédiction des intensités des ion fragments en mode CID. Exemple de fragmentation d'un peptide phosphorylé doublement chargé. Le spectre calculé (en haut) est similaire au spectre expérimental (en bas). D'après (Zhang, 2011)

En conclusion, si la meilleure compréhension des mécanismes de fragmentations ne permet pas de prédire un spectre de fragmentation à partir d'une séquence d'acides aminés, les règles qui se dessinent permettent d'envisager une amélioration de la validation des peptides identifiés dans les grandes séries d'identifications produites par l'analyse protéomique, en prévoyant les intensités de certains fragments caractéristiques des voies de fragmentations précédemment décrites.

1.2 L'analyse protéomique

1.2.1 Généralité

Le terme « protéome » est une fusion des mots PROTEines et génOME et désigne l'ensemble des protéines exprimées par un génome. Le mot est certainement né à la fin des années 1990 (Wilkins et al., 1996b). Dans le papier intitulé « From proteins to proteomes: large scale protein identification by two-dimensional electrophoresis and amino acid analysis », Wilkins identifie 20 protéines d'*Escherichia Coli* en séparant les protéines par électrophorèse bidimensionnelle puis en réalisant un transfert sur membrane pour permettre une composition en acides aminés. La connaissance du point isoélectrique, du poids moléculaire et de la composition en acides aminés de chaque protéine permet une identification dans la base de données de séquences protéiques de la bactérie (web.expasy.org/aacompident/). Pour plus de fiabilité, l'analyse est validée par la technique de référence de l'époque, le séquençage d'Edman de l'extrémité N-terminale de la protéine (Wilkins et al., 1996a).

Si le terme « protéomique » est apparu en 1996, la possibilité d'identifier des protéines en grand nombre est antérieure, puisque les premières identifications par spectrométrie de masse ont été réalisées en 1993. Pappin et collaborateurs furent les premiers à présenter une approche systématique utilisant le MALDI TOF et une bases de données protéique (voir paragraphe suivant)(Pappin et al., 1993). Cette approche est une évolution majeure dans l'analyse biochimique des protéines. La caractérisation ne dépend plus de la seule analyse chimique mais utilise les données de séquençage du génome pour faire une corrélation entre mesures expérimentales et valeurs théoriques calculées à partir de séquences. Cette approche protéomique permet d'augmenter considérablement la quantité de protéines identifiées et modifie radicalement la vision du biologiste et du biochimiste sur son modèle.

L'analyse protéomique peut être divisée en une approche qualitative (construction de catalogues généralement associés à une localisation subcellulaire) et une approche quantitative

(**Tab. 2**). Il s'agit alors de comparer le contenu protéique de différents échantillons afin d'en déterminer les différences. Cette comparaison peut s'opérer sur le matériel biologique entier ou fractionné (organes, compartiments cellulaires) et met en jeu différentes techniques de séparations. Dans tous les cas, la mesure qui permet l'identification des protéines est réalisée par spectrométrie de masse.

Tableau 2. Méthode d'analyse du protéome

ANALYSE PROTÉOMIQUE			
Type	Niveau	Séparation	Mesure
<ul style="list-style-type: none"> Qualitative Quantitative 	<ul style="list-style-type: none"> Globale Cellulaire subcellulaire 	<ul style="list-style-type: none"> Electrophorèse Chromatographie Partage 	<ul style="list-style-type: none"> MALDI MS ESI MS/MS

L'analyse du protéome peut porter sur un protéome entier ou sur une fraction du protéome. Cette fraction est spécifique d'un organe ou organisme entier, d'un type cellulaire ou d'un compartiment subcellulaire. Les techniques d'électrophorèse et de chromatographie ou de partage permettent de séparer les constituants (protéines ou peptides) pour une mesure par spectrométrie de masse.

L'analyse protéomique s'est également focalisée sur l'analyse de protéines ayant une caractéristique physicochimique originale (protéines membranaires, protéines de complexes) et sur l'analyse des protéines portant des modifications post-traductionnelles ; les termes de complexome, de phosphoprotéome ou de glycoprotéome sont alors apparus. A travers ses différents axes de recherche, la protéomique a connu une forte croissance de 1995 à 2005 et est aujourd'hui la technique de référence pour la caractérisation d'un contenu protéique. A titre d'exemple, 7292 publications contenaient ce mot « protéome » en 2012 (**Fig. 13**)

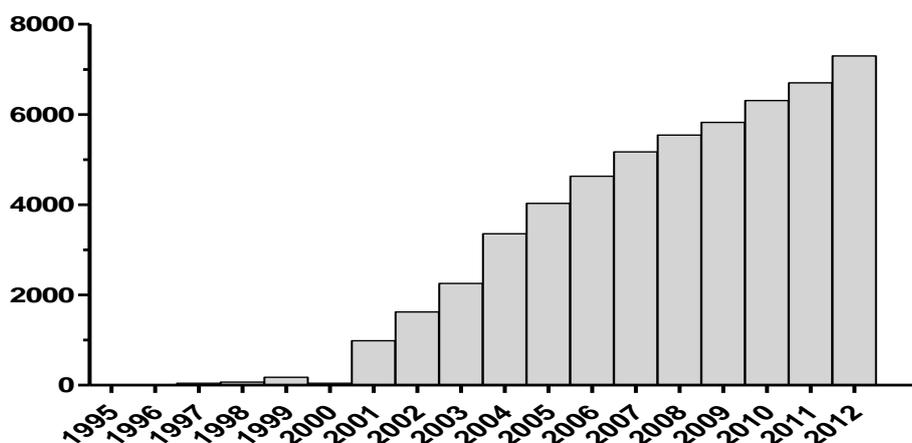


Figure 13. Nombre de publications référencées dans PubMed (<http://www.ncbi.nlm.nih.gov/pubmed>) et contenant les mots clé proteom*.

Il est possible de distinguer deux périodes dans l'évolution des techniques protéomiques, une première période qui correspond à l'utilisation exclusive de l'électrophorèse bidimensionnelle (gel 2D) pour séparer les mélanges protéiques complexes et quantifier les protéines du mélange ; le MALDI TOF est alors généralement utilisé pour identifier les spots des gels 2D. Une deuxième période fut initiée notamment par les travaux de John Yates qui a montré en 1996 qu'un hydrolysats tryptique de levure pouvait être analysé directement par un système LC-ESI-MS/MS et permettre l'identification de plus de 100 protéines en une seule analyse (Link et al., 1999).

Aujourd'hui, la première difficulté d'une analyse protéomique est le choix des techniques à utiliser pour fractionnement des protéines et le fractionnement des peptides (**Fig. 14**). Si on excepte l'approche Top-down, les protéines sont identifiées à partir de l'identification d'un ou plusieurs peptides (Bottom-up). Les protéines peuvent être séparées par électrophorèse ou chromatographie mais cette étape n'est pas indispensable : une hydrolyse enzymatique peut être menée sur l'extrait total ; le mélange de peptides est alors analysé par chromatographie liquide mono- ou multi-dimensionnelle couplée à la spectrométrie de masse MS/MS. Un préfractionnement par IEF en gel ou liquide est parfois réalisé.

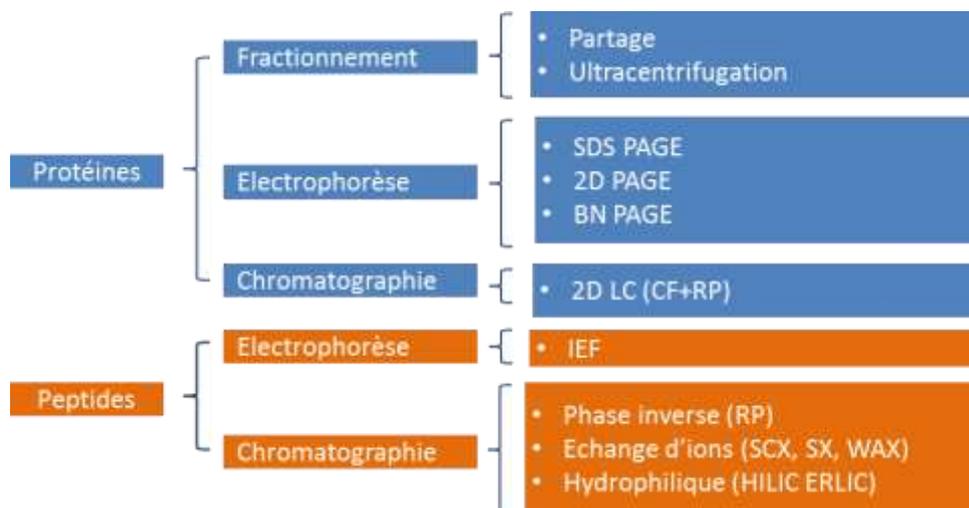


Figure 14. Les techniques de séparation en analyse protéomique. En général, l'électrophorèse est utilisée pour séparer les protéines et la chromatographie liquide pour séparer les peptides. La séparation de mélanges protéiques par chromatographie liquide et l'électrophorèse de peptides sont des techniques peu employées à ce jour.

1.2.2 Analyse protéomique par électrophorèse bidimensionnelle

En 1993, la première technique de protéomique par spectrométrie de masse a mis en œuvre une séparation des protéines par électrophorèse bidimensionnelle et une identification des spots protéiques par MALDI-TOF (Pappin et al., 1993). La technique d'électrophorèse bidimensionnelle a été présentée pour la première fois en 1975 par O'Farrell (O'Farrell, 1975). L'auteur présentait une

séparation en deux dimensions de protéines d'*Escherichia Coli* marqué au ^{14}C , avec en première dimension une iso-électrofocalisation (IEF) et en deuxième dimension une SDS-PAGE. A partir de 10 μg de protéines, O'Farrell a séparé plus de 100 constituants (Fig. 15A). Plusieurs évolutions ont rendu la technique accessible, avec une commercialisation de l'ensemble des matériels et réactifs. Une évolution majeure est l'immobilisation du gradient de pH sur une bandelette IPG (Immobilized pH Gradient) (Giorgio Righetti, 1984). Les bandelettes IPG permettent de charger une plus grande quantité de protéines et la première dimension est plus reproductible. Les techniques de détection ont également évolué : l'utilisation de la fluorescence a permis d'augmenter la sensibilité et le marquage des protéines permet maintenant d'analyser dans le même gel différents échantillons par la méthode DIGE (Differential In Gel Electrophoresis) (Saibil, 2013). Cependant, au regard des électrophorèses bidimensionnelles présentées dans la **Figure 15A**, il n'est pas certain que ces évolutions permettent de produire des électrophorèses plus sensibles et résolutive.

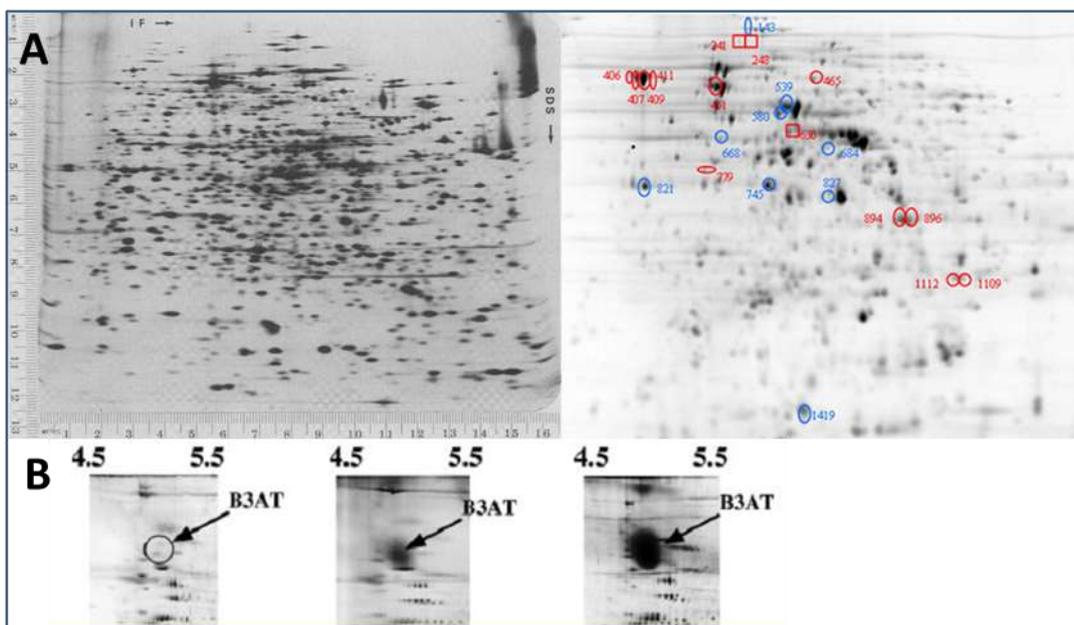


Figure 15. Electrophorèse bidimensionnelle d'E Coli. A gauche, réalisation en 1975 : 10 μg d'extrait total, détection de ^{14}C par film photographique (O'Farrell, 1975). A droite : réalisation en 2013 : 150 μg d'extrait total, détection par fluorescence ; les dessins de couleur désignent des spots discriminants (Sheridan et al., 2013). En B : Electrophorèse bidimensionnelle de la protéine membranaire B3AT. Le tampon de solubilisation utilisé pour l'IEF contient de l'urée, de la thiourée et différents détergents. De gauche à droite : CHAPS, Brij96, dodécyl maltoside, d'après (Rabilloud, 2009).

Un défaut majeur de cette technique est l'impossibilité de faire migrer dans la première dimension des protéines peu solubles comme les protéines membranaires. En effet, l'IEF ou IPG ne tolère que pas ou peu les détergents ioniques de type SDS, qui sont substitués par des détergents moins performants de type non ionique. L'optimisation des tampons de solubilisation a permis d'améliorer pour quelques cas particuliers la séparation des protéines membranaires. Par exemple,

le transporteur d'anion B3AT à 12 domaines transmembranaires peut être ou non détecté en fonction des tampons de solubilisation utilisés pour la première dimension (**Fig. 15B**) (Rabilloud, 2009). La synthèse de nouveaux détergents a également permis d'améliorer l'analyse des protéines membranaires (Santoni et al., 2000). Cependant le cas B3AT est marginal et l'électrophorèse bidimensionnelle reste peu adaptée à l'analyse des protéines membranaires. Une autre limitation de la technique est la fenêtre d'analyse des gels 2D (PM de 10 à 100 kDa et pI entre 4 et 8) qui exclut de la séparation les nombreuses protéines à caractère basique et les protéines de très hauts poids moléculaires (Rogowska-Wrzesinska et al., 2013).

De 1975 à la fin des années 90, les protéines visualisées sur les gels d'électrophorèses bidimensionnelles sont restées pour l'essentiel mystérieuses. Les techniques d'identification demandaient de grosses quantités de protéines, de l'ordre du μg , et l'identification se faisait essentiellement par micro séquençage et par composition en acides aminés (Wilkins et al., 1996a). Les années 90 voient l'essor de l'ionisation MALDI et la commercialisation de spectromètres MALDI-TOF qui permettent de mesurer simplement des molécules entières, non fragmentées par l'ionisation. En 1993, un article précurseur présente un logiciel accessible par le web, MOWSE pour Molecular Web SEarch qui identifie les protéines sur la base de leur carte peptidique (Pappin et al., 1993). La carte peptidique massique (*peptide mass fingerprinting*) est une liste de masses correspondant aux peptides obtenus à partir de la digestion enzymatique d'une protéine purifiée (purifiée par exemple par électrophorèse 2D). Dans un deuxième temps, une analyse informatique compare la liste de masses obtenues par MALDI-TOF avec une base de données indexée qui contient la liste de masses obtenue par digestion *in silico* des protéines. Comme la liste de masses expérimentale est incomplète par rapport à la digestion *in silico* d'une protéine, les auteurs calculent un rang d'identification basé entre autre sur le nombre de peptides expérimentaux présents dans la protéine identifiée et le nombre total de peptides théoriques, de manière à limiter le biais causé par le poids moléculaire de la protéine. Ce logiciel est le premier à être utilisable par l'intermédiaire d'internet et son évolution aboutira au logiciel MASCOT.

Le nombre d'articles qui utilisent l'électrophorèse bidimensionnelle a considérablement décliné. Pour autant, la technique conserve un intérêt majeur, car elle est à ce jour la technique séparative de choix pour mettre en évidence les différentes modifications post-traductionnelles d'une protéine (Rogowska-Wrzesinska et al., 2013).

Electrophorèse SDS-PAGE et LC-MS/MS

Afin de pallier aux limitations inhérentes à l'électrophorèse bidimensionnelle (gamme de PM et pl restreintes), et grâce aux progrès de la spectrométrie MS/MS qui permet d'identifier les protéines en mélange à partir de l'identification des peptides, la simple SDS-PAGE est utilisée pour préfractionner les mélanges protéiques. La digestion en gel permet d'extraire facilement les peptides qui sont ensuite analysés par LC MS/MS. Une comparaison de cette technique par rapport aux techniques off gel (SCX ou off gel IEF des peptides) a été réalisée en gardant un temps d'analyse MS constant, montrant que la SDS-PAGE et digestion en gel permettait d'identifier le plus grand nombre de protéines (Antberg L Fau - Cifani et al., 2012). Par ailleurs, les identifications des protéines séparées par SDS-PAGE sont suffisamment reproductibles pour permettre une quantification relative de plusieurs milliers de protéines (Piersma et al., 2013).

1.2.3 Chromatographie liquide et spectrométrie MS/MS

Deux éléments clés expliquent le succès du couplage LC-MS, à savoir (i) la source ESI est compatible avec tous les débits HPLC, de 20 nl/ml à 1 ml/min, une simple union suffit au couplage, (ii) la qualité des spectres de masse MALDI est fortement dégradée lorsque l'échantillon contient des contaminants (sels, détergents) ou lorsque le mélange est complexe (discrimination spectrale). Le couplage de la chromatographie liquide est facile à mettre en œuvre, permet d'éliminer les éventuels sels non volatils et simplifie le mélange en le fractionnant. La chromatographie utilisée devant le spectromètre de masse est essentiellement une phase inverse à bas pH. Afin d'augmenter le pouvoir de séparation, elle peut être précédée d'une autre chromatographie dont les critères de séparation sont différents, en général un échange d'ions.

Paramètres d'optimisations des colonnes HPLC

Le diamètre de colonne est un paramètre qui influence dramatiquement la sensibilité des systèmes LC-MS. La concentration des analytes en sortie de colonne peut être calculée par la formule suivante :

$$C = \frac{mN^{1/2}}{(2\pi)^{1/2}V_0(1+k)}$$

avec m : quantité chargée sur la colonne, N : efficacité de la colonne, V_0 : volume de colonne, k : facteur de rétention (Abian et al., 1999; Chervet et al., 1996).

Pour une même phase solide, la concentration en sortie de colonne est donc inversement proportionnelle au volume de colonne V_0 , ou encore inversement proportionnelle au carré de son

diamètre ($V = \pi r^2 \times \text{Hauteur}$). Autrement dit, transposer une analyse d'une colonne de 1 mm de diamètre à une colonne de 100 μm permet un gain de sensibilité d'un facteur 100. Sachant que l'efficacité de transmission des ions de la source est inversement proportionnel au débit, ce gain de sensibilité chromatographique est d'autant plus intéressant que l'efficacité de transmission des ions à l'analyseur de masse est amélioré pour des sources nanospray (Oosterkamp et al., 1998).

Si la diminution des diamètres de colonnes permet d'augmenter la sensibilité des analyses LC-MS, elle affecte également les débits et les capacités de chargement des colonnes (**Tab. 3**). Il apparait clairement que les colonnes de petits diamètres sont conçues pour analyser des petites quantités d'échantillon pour des petits volumes injectés.

Tableau3. Influence du diamètre sur les caractéristiques chromatographiques des colonnes.

Diamètre interne	Volume	Débit optimum	Injection	Conc. Relative au détecteur	Capacité
4.6 mm	4.1 ml	1 ml/min	100 µl	1	8469
2 mm	783 µl	0.2 ml/min	20 µl	5.3	1598
1 mm	196 µl	50 µl/min	5 µl	21.2	400
320 µm	20 µl	5 µl/min	500 nl	206	41
50 µm	490 nl	120 nl/min	10 nl	8459	1

Considérant une colonne de 75 µm, diamètre couramment utilisé en analyse protéomique, un volume d'échantillon facilement manipulable, soit 10 µl, demande 50 min pour être chargé sur la colonne qui fonctionne à un débit de 200 nl/min. Il apparaît donc indispensable d'associer au système une cartouche de plus gros diamètre, classiquement 300 µm, pour permettre un chargement rapide de l'échantillon. Ce type de cartouche de préconcentration est maintenant systématiquement utilisé (**Fig. 16**).

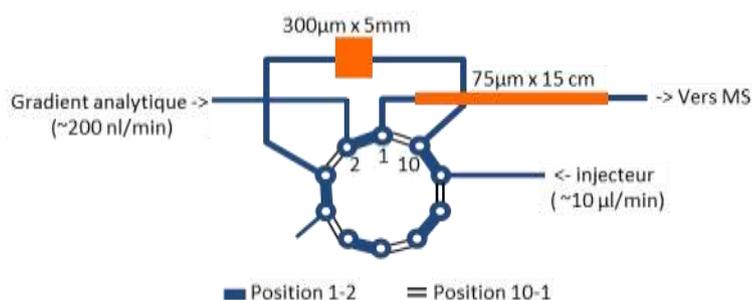


Figure 16. Montage nanoLC avec colonne de préconcentration. La colonne analytique (75 µm I.D.) et la cartouche de préconcentration (300 µm I.D.) sont connectées à une vanne à 2 positions. En position 1-2, l'échantillon provenant de l'injecteur est chargé sur la cartouche à un débit rapide, alors que la colonne analytique n'est reliée qu'au débit analytique (200 nl/min). En position 10-1, cartouche et colonne sont reliées au gradient analytique.

La longueur de colonne est un deuxième paramètre qui permet l'amélioration des séparations chromatographique. Son influence peut être évaluée par le calcul de la capacité en pic (n_c), ce paramètre permettant de déterminer l'efficacité de la séparation (Lan and Jorgenson, 1999).

La relation $n_c = (T_2 - T_1) / \omega$ où T_2 et T_1 sont respectivement les temps de rétention du dernier et premier pic d'analyte et ω la largeur moyenne des pics, indique que la capacité en pics augmente en fonction de la racine carrée de la longueur de la colonne (Medina et al., 2001). En phase inverse C18, la capacité en pic passe de 162 pour une longueur de 7.5 cm à 450 pour une longueur de 60 cm (Wang et al., 2006). La longueur de colonne permet donc d'améliorer la séparation. Mais en contrepartie, le temps d'analyse est augmenté ainsi que la perte de charge, l'utilisation de four à colonne est alors nécessaire pour diminuer la pression.

La taille des particules est un autre paramètre d'optimisation qui permet d'augmenter la capacité de pics en diminuant le temps d'analyse. Hongji Liu montre que n_c augmente de 50 % pour des particules réduites de 3 à 1.7 μm alors qu'elle n'augmente que de 30 % quand la longueur de colonne est multipliée par deux (Liu et al., 2007). Il apparait en effet que :

$$\begin{cases} n_c \sim \sqrt{L} \text{ avec } L \sim t, L: \text{ longueur de colonne} \\ n_c \sim d_p \text{ avec } t \sim d_p^4, d_p: \text{ diametre particules} \end{cases}$$

Soit en pratique avec une pression de 80 MPa en facteur limitant :

$n_c=500$ avec $L = 1\text{m}$, $d_p = 2 \mu\text{m}$ $t=120$ min

$n_c=1000$ avec $L=2\text{m}$ avec $d_p = 3 \mu\text{m}$, $t = 500$ min (Novakova et al., 2013).

Il apparait donc que la réduction de la taille des particules est une solution intéressante pour améliorer l'efficacité des séparations, mais ceci s'accompagne d'une forte augmentation de la contre pression (loi de Darcy).

Le tableau ci-dessous est une synthèse basée sur des calculs théoriques et permet d'illustrer clairement les relations entre t , L , et d_p (Poppe, 1997). Pour un facteur limitant (la pression) fixé à 200 bars, la séparation optimale est obtenue avec 2.3 jours de chromatographie sur une colonne de 35 m et des particules de 17 μm . Des particules de diamètre 10 fois inférieures (1.7 μm) permettent une chromatographie « plus raisonnable » de 20 sec sur une colonne de 35 mm.

Tableau 4. Caractéristiques optimales d' une chromatographie à 200 bars, d'après (Poppe, 1997)

Nb. de plateaux	Temps de rétention	d_p (μm)	L (m)
1000	0.2 sec	0.5	0.0011
10000	20 sec	1.7	0.035
100000	2000 sec	5	1.1
1000000	2.3 jours	17	35

La contre pression est le facteur limitant. Alors que la séparation est améliorée avec une diminution des tailles de particules, la contre pression qui en résulte limite la longueur des colonnes et oblige l'utilisation de systèmes chromatographiques dit « ultra haute pression » (UHPLC) qui peuvent atteindre 100 MPa. Une alternative est le développement de phases polymériques dont les caractéristiques mécaniques permettent de limiter la contre pression. A titre d'exemple, une colonne à polarité de phase inverse polymérique de 3 mètres de long (C18, 100 μm x 3 m) et une séparation de 40 h à 500 nl/min à 20 MPa permettent d'atteindre une efficacité en pic de 1600

(Horie K Fau - Sato et al., 2012). Une situation intermédiaire est l'utilisation de particules de petite taille (2 μm) dont le cœur est non poreux et la périphérie poreuse. Elles permettent de limiter la contre pression et d'obtenir des capacités en pic de l'ordre de 530 pour une séparation de 5h (200 μm x 25 cm) (Schuster Sa Fau - Boyes et al., 2012). Définitivement, les développements de nouvelles phases permettent d'augmenter la capacité en pic (Sandra et al., 2008), les répercussions sur les performances des système LC-MS étant importantes du fait de la corrélation linéaire entre n_c et le nombre de peptides identifiés. A titre d'exemple 14292 peptides peuvent être identifiés dans un seul run LS-MS (Orbitrap) de 8h (C18, 2 μm , 75 μm x 50 cm, 750 bars, $n_c >600$). (Kocher T Fau - Swart et al., 2011).

Chromatographie monodimensionnelle

La chromatographie directement liée à la source ESI est une chromatographie dont les tampons ne contiennent pas de sel. Dans le cadre de l'analyse des peptides, elle est essentiellement de type phase inverse (RP) avec un greffage par des chaînes octadécylsilane (C18). Avec ce type de chromatographie, le choix approprié du contre-ion permet d'augmenter considérablement la rétention des peptides. Les contre-ions les plus performants sont fluorés et l'augmentation de rétention augmente avec leur hydrophobie, acide trifluoroacétique (TFA) < acide pentafluoropropionique (PFPA) < acide heptafluorobutyrique (HFBA) (Shibue M Fau - Mant et al., 2005). Malheureusement ces contre-ions ne peuvent pas être employés en couplage LC-MS à cause de la suppression de signal qu'ils entraînent (Garcia, 2005). Une concentration en TFA diminuée à 0.05 % est cependant utilisée en électrospray (Premstaller A Fau - Oberacher et al., 2001). La même solution a été employée avec l'HFBA (Washburn Mp Fau - Wolters et al., 2001) mais la grande majorité des protocoles utilisent l'acide formique 0.1 % et un gradient linéaire de 0-5 % à 30-50 % (Macek et al., 2009). Dans ces conditions, dans le cas de l'analyse du protéome de levure qui constitue un excellent modèle pour l'évaluation des performances, plus de 4000 protéines ont été identifiées en une seule analyse de 4h avec une colonne UHPLC (C18 1.8 μm , 75 μm x 50 cm)(Nagaraj N Fau - Kulak et al., 2012).

Chromatographie multidimensionnelle

A la manière de l'électrophorèse bidimensionnelle qui a su associer deux principes électrophorétiques et multiplier sa capacité de séparation, la chromatographie multidimensionnelle (généralement bidimensionnelle) s'est rapidement développée. Elle a été aidée par la

commercialisation de systèmes HPLC bidimensionnelles entièrement automatiques. La quasi-totalité des montages chromatographiques utilise une chromatographie en phase inverse en deuxième dimension, devant le spectromètre de masse. La première dimension utilise généralement une phase stationnaire correspondant à un échangeur fort de cations (SCX). Plus récemment, d'autres associations de colonnes ont été publiées. Elles conservent une phase inverse en deuxième dimension mais la première dimension est assurée par une chromatographie SAX (Strong Anion Exchange), une chromatographie RP à pH basique ou une chromatographie HILIC (Hydrophilic Interaction Chromatography) (Xie F Fau - Smith et al., 2012).

MudPit(SCX+RP) pour *Multidimensional protein identification technology* est une version intégrée du couplage SCX –RP (**Fig. 17**). Son montage diminue tous les volumes morts puisque l'aiguille nanospray sert également de colonne analytique, la pointe de l'aiguille contient un fritté qui retient une phase inverse et une phase échangeuse de cations. Des injections par palier de sels volatiles (acétate d'ammonium) transfèrent les peptides par paquets de la partie SCX à la partie phase inverse. En 2001, cette technique permit l'identification de 1484 protéines (5540 peptides) pour une analyse de plus de 20h (Washburn Mp Fau - Wolters et al., 2001).



Figure 17. Multidimensional protein identification technology (MudPit). A gauche : l'aiguille nanospray est remplie par un segment de phase SCX et par un segment de phase inverse. Une union permet d'appliquer d'une part la tension de spray (kV) et d'autre part le débit (gradient). A droite : la phase est retenue par un fritté constitué dans le dernier mm de l'aiguille nanospray.

En 2013, l'analyse a été reproduite par le même groupe avec un échantillon similaire (extrait de protéines totales de *Saccharomyces cerevisiae*). Pour un temps d'analyse de 8h, 3844 protéines ont été identifiées (29952 peptides) (Webb et al., 2013). Il faut noter que les auteurs ont modifié leur montage chromatographique en dissociant la partie SCX de l'aiguille-colonne nanospray. Une cartouche de préconcentration (250 μ m x 5.5 cm) constituée de SCX + RP permet de charger l'échantillon et élimine les sels dirigés vers la poubelle. Dans un deuxième temps, la cartouche est connectée à l'aiguille-colonne C18 (100 μ m I.D.) pour la conduite d'une expérience MudPit semblable à celle de 2001. L'augmentation du nombre de protéines identifiées est essentiellement due au spectromètre de masse (Orbitrap vs. Trappe ionique quadripolaire) : l'identification de près de 60 % des 6617 protéines prédites à partir du génome a ainsi été rendue possible.

L'association SAX+RP est une solution actuellement en développement qui présente la même orthogonalité que la SCX vis-à-vis de la phase inverse. L'utilisation de SAX se justifie notamment par le type de peptides que produit l'hydrolyse trypsique. De par sa spécificité, la trypsine produit majoritairement des peptides acides (Cargile et al., 2004) dont la charge nette à pH élevé leur permet théoriquement d'être retenus sur une colonne échangeuse d'anions. Appliqué à un extrait protéique de culture cellulaire (macrophage de rat), 9469 protéines (126318 peptides uniques) ont été identifiés par ce moyen (Ritorto et al., 2013).

L'association RP+RP a également été testée dans le même article et présente une efficacité d'identification équivalente (8627 protéines pour 99110 peptides différents)(Ritorto et al., 2013). Dans cette configuration, la première phase inverse est réalisée à pH élevé (pH 10) alors que la deuxième phase inverse utilise un pH bas (pH 2.5), le changement de pH conférant une différence de spécificité entre les 2 dimensions (Stephanowitz et al., 2011) qui permet d'atteindre une capacité en pics de 8500, ou 3982 après correction, montrant ainsi que le mode de calcul de la capacité en pic influence grandement le résultat (Donato et al., 2011).

L'association HILIC + RP est une configuration dont la première dimension permet une grande diversité de greffage du support chromatographique : diol, cyano, sulfobétaïne, amide, etc. (Donato et al., 2011). A titre d'exemple, 2554 *protéines de S.cerevisiae* ont été identifiées avec une colonne amide (Zhao et al., 2012). Elle est plus particulièrement adaptée au fractionnement de peptides hydrophiles, par définition peu retenus sur une phase inverse.

1.2.4. Analyse protéomique quantitative

La quantification des protéines par l'intermédiaire des peptides est exclusivement réalisée par LC-MS/MS. Plusieurs stratégies sont envisageables pour comparer l'abondance de peptides entre plusieurs échantillons (**Tab. 5** et revues (Kumar and Mann, 2009) (Rodriguez-Suarez and Whetton, 2013)). Deux grandes approches sont classiquement mises en parallèle : la quantification relative et la quantification absolue. L'approche de quantification absolue est ciblée, car la connaissance des protéines /peptides à doser est un pré requis pour produire des standards marqués. Au contraire, la quantification relative relève plus particulièrement de la démarche protéomique, car elle ne demande pas *d'a priori* et quantifie l'ensemble des peptides mesurés.

Tableau5. Les principales méthodes d'analyse protéomique quantitative

PROTEOMIQUE QUANTITATIVE	
APPROCHE GLOBALE: Quantification relative	APPROCHE CIBLEE: Quantification Absolue
Label-free: <ul style="list-style-type: none"> • Spectral count • Intensité des pics 	Ajout dosé de: <ul style="list-style-type: none"> • Peptides marqués: AQUA • Protéine marquée: PSAQ
Marquage isotopique: <ul style="list-style-type: none"> • ICAT / ITRAQ/TMT • dimethyl • ¹⁸O 	ICAT: Isotope-coded affinity tag ITRAQ : Isobaric Tag for Relative and Absolute Quantitation TMT : Tandem Mass Tags SILAC : Stable Isotope Labeling with Amino acid in Cell Culture
Marquage métabolique: <ul style="list-style-type: none"> • SILAC 	AQUA : Absolute quantification PSAQ : Protein Standard Absolute Quantification

D'un point de vue pratique, les échantillons sont mélangés (sauf dans le cas du Label-free) à différents moments de leur préparation (**Fig. 18**). De ce point de vue, la méthode SILAC présente un avantage certain, puisque le mélange est réalisé au moment de l'échantillonnage. Toute source de variation liée aux manipulations ultérieures affecterait de la même manière les 2 échantillons analysés (Emadali and Gallagher-Gambarelli, 2009). Au contraire, les marquages chimiques sont réalisés après extraction des protéines. À l'exception d'ICAT qui ajoute une étiquette sur les cystéines réduites, les autres étiquettes sont greffées sur les fonctions amines. En ce sens, l'échantillon est préalablement digéré par la trypsine. Toutes les étapes du protocole réalisées avant le marquage sont susceptibles d'introduire des variations artefactuelles.

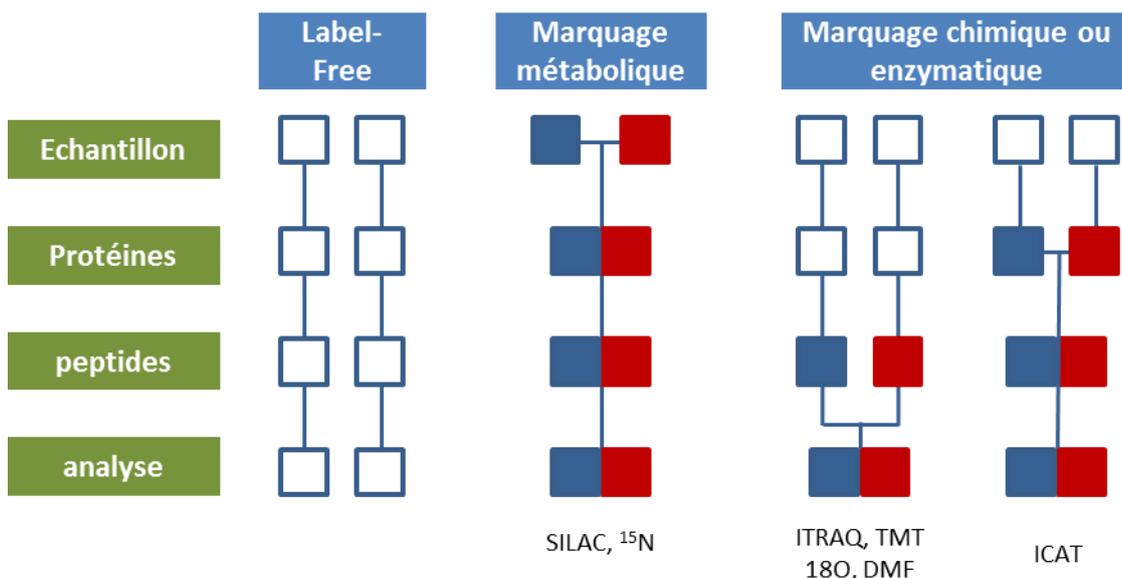


Figure 18. Stratégies d'analyse protéomique quantitative. Suivant la technique, les échantillons sont mélangés (marquage) ou analysés séparément (label free).

Quantification label-free

La quantification label-free repose sur l'observation que l'intensité d'un pic MS est proportionnelle à la concentration du peptide correspondant, y compris dans un mélange complexe (Chelius and Bondarenko, 2002). L'intérêt immédiat de cette méthode est l'absence de traitement chimique sur l'extrait de peptides, l'inconvénient étant que les échantillons sont analysés les uns après les autres. La quantification est donc sensible aux variations inter-analyses, aux instabilités de la source ESI et aux éventuels problèmes de suppression ionique. De nombreux outils logiciels ont été développés pour le traitement des données label-free (Nahnsen et al., 2013). Il existe 2 grandes méthodes de traitement du signal : le *Spectral counting* et la quantification par la mesure des intensités de pics (Neilson et al., 2011; Zhu et al., 2010).

Le ***Spectral Counting*** est une méthode simple de comptage des spectres MS/MS correspondant à une même protéine. Elle est basée sur l'observation que le nombre de spectres MS/MS est corrélé avec l'abondance de la protéine dans le mélange. Le paramétrage de l'instrument et sa méthode d'exclusion permet de favoriser ou non les protéines majoritaires (Zhou et al., 2012). Afin de corriger un biais lié à la longueur de la protéine et au nombre de peptides tryptiques qui en découlent, différentes méthodes de calcul sont proposées : SI (Spectral Index), PAI (Protein Abundance Index), SAF (Spectral Abundance Factor) (McIlwain et al., 2012). La valeur de *spectral counting* peut également être combinée aux TIC (Total Ion Current) des spectres MS/MS pour améliorer la quantification (Freund and Prenni, 2013). A titre d'exemple, autour de 2000 protéines de rhizome de prêle ont été quantifiées, dont 87 surexprimées lors de la croissance du rhizome (Balbuena et al., 2012).

La quantification par l'intensité des pics est l'autre méthode de quantification qui utilise les spectres MS. Le traitement des données nécessite un alignement des temps de rétention entre analyses chromatographiques, une normalisation des intensités, et un traitement du signal pour éliminer le bruit de fond. L'application de cette méthode à des cellules HUVEC avec séparation par 2D-LC (RP+RP) a permis de quantifier 6700 produits de gènes dont 99 surexprimés dans les cellules HUVEC traitées par un médicament (Ding et al., 2013).

Quantification par marquage isotopique

La solution idéale pour éliminer toute variation de quantification due à la variabilité de l'analyse chromatographique est de comparer les intensités des ions au même temps de rétention, dans le même spectre MS/MS. Le marquage isotopique est la réponse, il permet de différencier les

peptides de différents échantillons en incorporant des isotopes rares pour une même composition élémentaire (^2D , ^{13}C , ^{15}N , ^{18}O). Les extraits peuvent alors être mélangés avant l'analyse LC-MS, puisque la mesure de masse permettra de les différencier. Les protéines ou peptides peuvent être marqués par deux méthodes. La première méthode consiste à utiliser le métabolisme du modèle étudié pour incorporer des acides aminés essentiels marqués (modèle animal) ou une source d'azote ^{15}N (modèle végétal). La deuxième technique consiste à extraire les protéines des échantillons et de les marquer par une étiquette (ITRAQ, ICAT, TMT, Diméthyl). Dans ce cas, l'étiquette différenciera les peptides en mode MS (ICAT, ^{18}O , TMT, Diméthyl) ou en mode MS/MS (étiquettes isobares ITRAQ).

Le marquage ^{18}O peut être réalisé lors digestion à la trypsine dans une solution d'eau enrichie en ^{18}O (Heller et al., 2003). Lors de l'hydrolyse, deux molécules d'oxygène sont ajoutées à l'extrémité C-terminale du peptide, conférant une différence de masse de 4 Da par rapport aux peptides hydrolysés dans l'eau ^{16}O . Une revue présente les avantages mais également les inconvénients de la technique ; marquage incomplet et incorporation d'un seul atome ^{18}O , chevauchement des massifs, back-exchange (Fenselau and Yao, 2009). L'échange peut également être réalisé par une incubation dans de l'acide concentré, provoquant alors également un marquage sur D et E (Haaf and Schlosser, 2012).

Le marquage SILAC (Stable Isotope Labelling by Amino Acid in Cell culture) est un marquage métabolique qui consiste à cultiver une population de cellules dans un milieu contenant les acides aminés essentiels Lysine et Arginine marqués par des isotopes ^{14}C et ^{15}N (par exemple $^{13}\text{C}_6^{15}\text{N}_2$ -L-lysine pour la lysine (écart +8) et $^{13}\text{C}_6^{15}\text{N}_4$ -L-arginine pour l'arginine (écart+10) (Emadali and Gallagher-Gambarelli, 2009). Cette technique présente l'avantage de mélanger les extraits dès l'échantillonnage. Appliquée aux phosphopeptides, le marquage SILAC de cellules HeLa a permis d'identifier 6600 sites de phosphorylation, 14 % des sites ayant une variation d'expression supérieure à 2 après stimulation par le facteur de croissance EGF (Olsen et al., 2006). Plus récemment, le marquage SILAC a été également appliqué à des cultures de cellules primaires (Liberski et al., 2013) et à des rongeurs (Konzer et al., 2013). Un développement appelé superSILAC a été proposé par le groupe de Mathias Mann et permet d'utiliser SILAC avec des échantillons non marqués de type biopsie (Geiger et al., 2011). La culture SILAC de lignées représentatives de la tumeur étudiée permet d'utiliser leur pool comme ajout dosé qui sert ensuite de référence à l'ensemble des protéines quantifiées. Avec cette technique et un enrichissement des protéines glycosylées, les sécrétomes de 11 lignées de cancer du poumon ont été comparés à des cultures

primaires de poumons, puis à des prélèvements sanguins : 1398 sites de N-glycosylation ont été identifiés, dont 881 sites quantifiables entre les lignées. Cinq profils d'expression permettent de classer les lignées en fonction des stades cancéreux qu'elles représentent. Certains des biomarqueurs identifiés sont également détectés dans les sérums de patients atteints de cancer du poumon (Boersema et al., 2013). La méthode superSILAC a également appliquée à des extraits de tissus hépatiques murins après stimulation ou non par l'insuline. Elle a permis de vérifier l'activation de kinases appartenant à la voie de signalisation du récepteur de l'insuline (Lundby et al., 2012).

Le marquage Diméthyl permet de comparer 3 échantillons dont les peptides sont modifiés par du formaldéhyde (H_2CO). Trois formes de formaldéhyde sont utilisées : une forme légère H_2C , (ajout de 28 Da), une forme intermédiaire D_2CO (+32 Da), une forme lourde $D_2^{13}CO$ (+36 Da)(Kovanich et al., 2012). Cette approche a été utilisée pour comparer le phosphoprotéome de plants de maïs soumis à un stress hydrique (Bonhomme et al., 2012), le phosphoprotéome de foie cancéreux (Song et al., 2011), ainsi que l'identification des substrats des 2 kinases ; PKA (Protein kinase A) et Mps1 (human monopolar spindle 1) (Henrich et al., 2013).

Les marquages ITRAQ 8plex et TMT⁶ sont basé sur des kits commercialisés respectivement par AB SCIEX et Thermo Scientific. Les deux réactifs utilisent une chimie de couplage basée sur le N-hydroxy-succinimide (NHS) qui réagit sur les amines à pH élevé (>8). Ces marquages sont isobares, c'est à dire que les 6 étiquettes TMT et les 8 étiquettes ITRAQ permettent de marquer respectivement 6 ou 8 échantillons différents avec un incrément de masse identique, indépendant de l'étiquette utilisée. Lors de la fragmentation en mode MS/MS, un fragment spécifique de chaque étiquette est libéré.

Le marquage TMT est illustré en détail dans la **Figure 19**. Le groupe réactif (succinimide) réagit avec les fonctions amines des peptides pour greffer une étiquette (tag) de masse 230 Da. Lors de la fragmentation en mode CID, un fragment caractéristique se forme au site de clivage indiqué sur la figure.

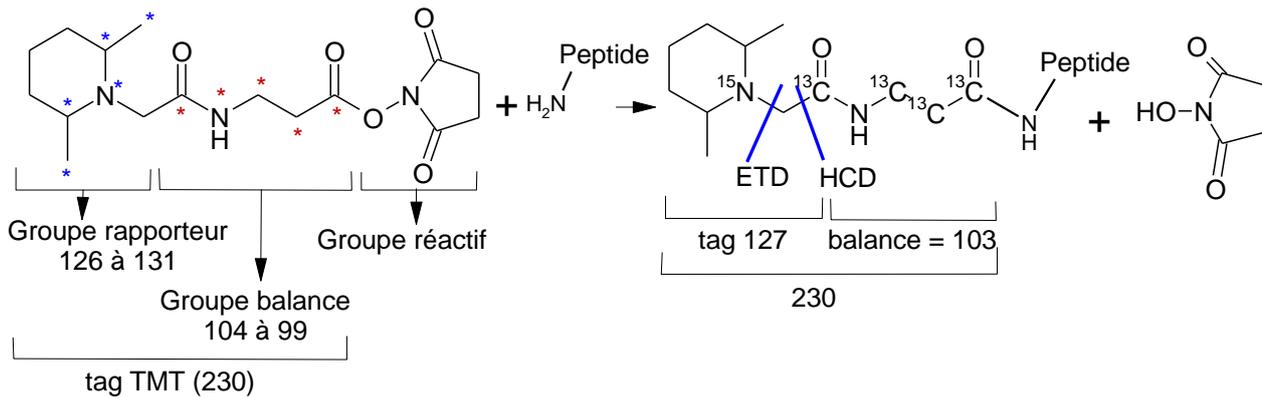


Figure 19. Marquage isobare avec le réactif TMT (Tandem Mass Tag). Les étoiles bleues correspondent aux 5 atomes qui peuvent être échangés avec des isotopes lourds, les étoiles rouges marquent les atomes du groupe balance qui seront ou non marqués afin de conserver une même masse pour l'ensemble du tag.

Grace aux dernières générations de spectromètres de type Orbitrap qui permettent des précisions inférieures à 1 ppm pour une résolution de 100 000, le réactif TMT est également décliné sous la forme de 10 étiquettes différentes. Pour une même formule brute, l'augmentation du nombre de groupes rapporteur est basée sur l'utilisation soit de ^{13}C , soit de ^{15}N . Plusieurs tag TMT existent sous une forme légère (**L**) ou une forme lourde (**H**) dont la différence de 0.006 u est discernable par des instruments à haute résolution. Par exemple, TMT127L ($^{12}\text{C}_8\text{H}_{16}^{15}\text{N}_1^+$) = 127.1247 u, et TMT127H ($^{12}\text{C}_7^{13}\text{C}_1\text{H}_{16}^{14}\text{N}_1^+$) = 127.1310 u sont discernables à partir d'une résolution supérieures à 30 000 (McAlister et al., 2012; Werner et al., 2012).

La comparaison des différentes stratégies de marquage a été réalisée dans quelques rares articles. Les conclusions sont difficiles à tirer puisque l'efficacité de la technique peut être évaluée par différents paramètres, par exemple le nombre de protéines identifiées et quantifiées, mais également la précision de la quantification, la justesse, la linéarité du dosage. Avec un système de calcul de rangs, le classement suivant a été obtenu : Diméthyl LTQ-FTICR > iTRAQ QStar > Label-free LTQ-FTICR > Diméthyl QStar > Label-free QStar (Sjodin et al., 2013). Une autre comparaison portant sur SILAC, label free, TMT et iTRAQ 4plex présente la méthode label-free comme la technique la plus sensible mais la moins précise (Li et al., 2012). Une comparaison restreinte au label-free et iTRAQ 4plex d'une souche de *Chlamydomonas reinhardtii* a permis de quantifier un plus grand nombre de protéines par label free (329 contre 124) (Wang et al., 2012a).

Plus récemment, la technique *label free* a été comparée au marquage TMT pour différentes cultures d'hépatocytes humains. Si l'approche label free permet d'identifier considérablement plus de peptides que l'approche TMT (respectivement 1644 et 721 peptides), le marquage TMT permet une quantification considérablement plus précise (Megger et al., 2013).

1.3 L'analyse des phosphopeptides

1.3.1 Sites de phosphorylation

La phosphorylation des protéines est une modification post-traductionnelle courante qui est impliquée dans la régulation des fonctions cellulaires. Nous ne parlerons ici que de la phosphorylation dans les cellules eucaryotes (dans les cellules procaryotes, la phosphorylation intervient essentiellement sur l'histidine, l'acide glutamique et l'acide aspartique).

Les kinases protéiques (PK) sont les enzymes qui catalysent le transfert d'un groupement phosphoryl (PO_3^{2-}) d'un nucléoside triphosphate, habituellement l'ATP, sur la chaîne latérale d'un acide aminé. Les kinases font partie d'une famille de protéines regroupant plus de 518 gènes putatifs, soit 1.7 % de la totalité des gènes humains (Manning et al., 2002). En termes de spécificité, il existe 2 grandes classes de kinases : 428 kinases (connues ou prédites) phosphorylent la sérine et la thréonine en phosphosérine (pS) et phosphothréonine (pT), alors que les 90 kinases restantes sont spécifiques des tyrosines. Un autre groupe de protéines kinases moins étudié catalyse l'addition de phosphate sur les amines pour former une liaison phosphate amide (Attwood, 2013). La liaison P-N a été particulièrement étudiée pour l'histidine en position 1 ou 3 de l'imidazole (Besant et al., 2003) dans le cadre de l'étude de l'histone H4 dont les niveaux de phosphorylation sont associés à la prolifération cellulaire et aux mécanismes cancéreux (Besant and Attwood, 2012).

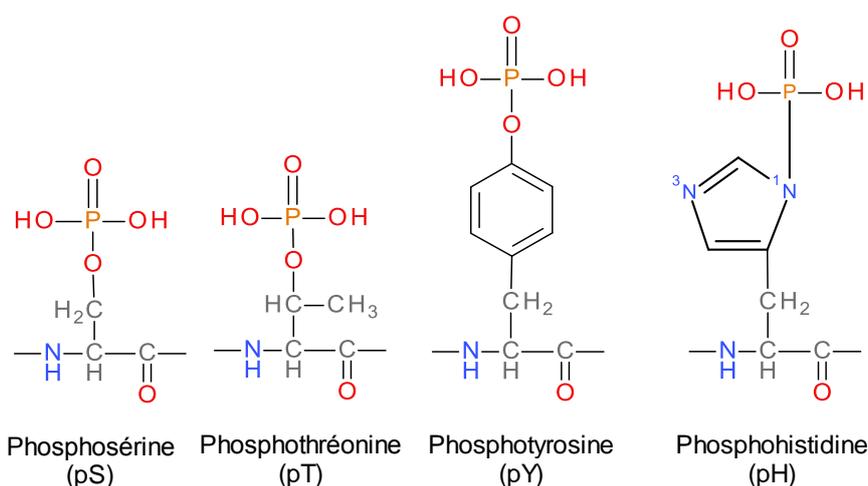


Figure 20. Acides aminés phosphorylés. Le phosphoryl PO_3^{2-} peut former une liaison énol-phosphate (P-O) ou une liaison amide phosphorique (P-N). La phospho-histidine peut être phosphorylée en position 1 ou 3, sa double phosphorylation n'a jamais été observée.

La proportion de pS, pT et pY a été estimée en premier lieu à travers les travaux de Tony Hunter, prix Nobel de Médecine 2005 pour la découverte des tyrosines kinases. Avec des

expériences de marquage radioactif au ^{32}P de cellules de poulet, hydrolyse par HCl 6N et quantification des acides aminés phosphorylés par chromatographie en couche mince et autoradiographie, les rapports pS/pT/ pY ont été estimés à 92.19/7.77/0.03 sans stimulation et 92.48/7.24/0.28 après transfection des cellules par le sarcome de Roux (Hunter and Sefton, 1980). Des articles plus récents ont proposé 1800/200/1 (Gronborg et al., 2002; Mann et al., 2002) et les analyses phosphoprotéomiques ont finalement permis de calculer un ratio sur la base du nombre de pS, pT et pY contenus dans les phosphopeptides identifiés. Avec cette méthode, Olsen et al. ont calculé un ratio 86/12/2 à partir de l'identification de 6600 sites de phosphorylation, la proportion de pY atteignant 6 % lorsque les cellules Hela sont stimulées par le facteur EGF (Olsen et al., 2006). Cette proportion semble également observée dans les tissus ; un ratio 83/15/2 a été calculé à partir de l'étude phosphoprotéomique de neuf organes de la souris (Huttlin et al., 2010). En conclusion, si l'ordre pS>pT>pY est établi et qu'il y a environ 10 fois plus de phosphosérines que de phosphothréonines, la phosphotyrosine est minoritaire avec une proportion fluctuante entre 0.03 % et 6 %.

Par ailleurs, il semble que les kinases aient plus de 17000 substrats (phosphoprotéines) annotés dans la base de données PhosphoSite (Hornbeck et al., 2012). Il y aurait donc environ 1/3 du protéome susceptible d'être phosphorylé. Si plus de 30 % des protéines possèdent un site potentiel de phosphorylation, la quantité de protéines phosphorylées à un instant donné est une donnée qui n'est jamais déterminée par les approches globales. Il n'existe pas en effet de méthode simple de dosage des phosphoprotéines totales : les détections colorimétriques ou immunométriques sont peu spécifiques, et la composition en acide aminés est difficile à mettre en œuvre à cause du caractère labile de la liaison phosphate dans les conditions d'hydrolyse.

1.3.2 Elimination du groupement phosphate

Il est possible d'éliminer le groupement phosphate de 3 manières différentes : par voie enzymatique, par une hydrolyse en milieu acide et par une hydrolyse en milieu basique (β -élimination). L'hydrolyse enzymatique est réalisée par les phosphatases qui constituent une superfamille de 199 protéines (Sacco et al., 2012). Les phosphatases sont utilisées en biochimie, notamment pour confirmer la spécificité des immuno-détections (Western blot), l'action de la phosphatase devant inhiber le signal.

L'hydrolyse en milieu acide à 0°C avec HF-pyridine (70 %) est utilisée de manière plus marginale pour cette même application (Kuyama et al., 2003). L'hydrolyse enzymatique et l'hydrolyse acide se traduisent par une perte de 80 Da correspondant à l'élimination du groupement phosphonyl (PO_3^{2-}), l'acide aminé restant hydroxylé contrairement à l'hydrolyse alcaline.

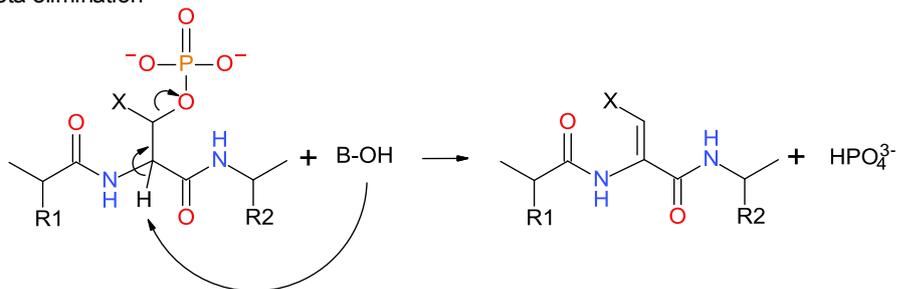
Elimination du groupement phosphate par β -élimination

L'hydrolyse alcaline est une réaction de β -élimination qui entraîne la fonction OH et forme un alcène électrophile qui est souvent utilisé pour réaliser une addition nucléophile. Seules la sérine et la thréonine sont susceptibles de former un intermédiaire insaturé appelé respectivement déhydroalanine et méthyl-déhydroalanine, la phosphotyrosine ne possédant pas de carbone beta disponible pour ce type de réaction. La base utilisée pour l'hydrolyse alcaline est soit NaOH soit $\text{Ba}(\text{OH})_2$. Byford et collaborateurs ont clairement montré que $\text{Ba}(\text{OH})_2$ était plus efficace que NaOH (Byford, 1991) et accélérât d'un facteur 100 la vitesse initiale (k_r) de la β -élimination pour le peptide LRRApSLG (phospho-kemptide). La substitution de pS par pT diminue d'un facteur 20 la vitesse initiale de β -élimination. Malgré ces résultats, la soude est généralement préférée car elle est plus simple d'emploi que $\text{Ba}(\text{OH})_2$ dont la titration est délicate (formation rapide de carbonates insolubles). Une utilisation originale de la β -élimination consiste à éluier les phosphopeptides retenus par les colonnes d'affinité de type IMAC par 100 mM $\text{Ba}(\text{OH})_2$ puis de dériver les acides aminés déshydratés (Thompson et al., 2003).

Addition nucléophile.

Après β -élimination, l'alcène produit est très électrophile et permet facilement une réaction de Michael (**Fig. 21**). Cette réactivité n'empêche pas de conserver les dehydro-analogues en l'état en milieu basique (Rusnak et al., 2004). Cependant, une addition nucléophile est systématiquement réalisée : des dérivés alkyl-thiol ou sulfite ont été utilisés avec succès.

1: Béta élimination



2: Addition nucléophile

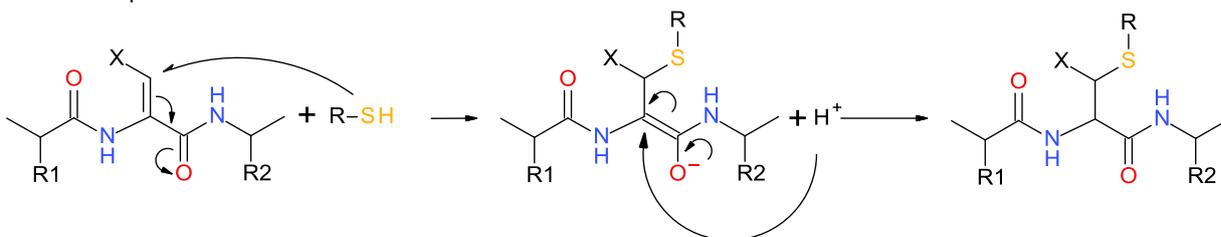


Figure 21. Réaction de β -élimination / addition nucléophile. X = H et CH₃ pour pS et pT. Les intermédiaires insaturés sont appelés déhydroalanine ou β -méthyl-déhydroalanine. Lors de l'addition nucléophile, ils sont convertis en composés soufrés si l'éthanethiol (R=C₂H₅) ou le sulfite (Na₂SO₃) est utilisé comme produit d'addition. (Li et al., 2002).

Le **Tableau 6** passe en revue les principaux articles qui traitent du sujet, des années 1990 à ce jour. Trois types de réactifs nucléophiles ont été utilisés : des amines, des dérivés thiolés et du sulfite de sodium. Le protocole de référence utilise 2 étapes : d'abord élimination du phosphate, puis addition nucléophile (Meyer et al., 1986). Les publications suivantes ont montré la possibilité d'effectuer les 2 étapes en ajoutant simultanément les différents réactifs et les articles plus récents présentent des réactifs nucléophiles qui ajoutent une signature isotopique aux peptides modifiés.

Du point de vue de l'analyse par spectrométrie de masse, l'élimination du phosphate présente en théorie 3 avantages : (i) une meilleure ionisation en mode positif des peptides déphosphorylés par rapport à leurs homologues phosphorylés, (ii) la suppression en MS/MS des problèmes liés à la perte de neutre (-98 u) des peptides qui contiennent des phosphosérines et phosphothréonines, (iii) la suppression des risques de réarrangements de la liaison ester phosphate en phase gaz lors de la fragmentation.

Tableau 6. Revue des principaux articles relatifs à la β -élimination

Auteur	Année	Traitement	Nucléophile (X)	Commentaire
Meyer et al. (Meyer et al., 1986)	1886	+ 50 μ l de (60 μ l 10M X, 200 μ l DMSO, 80 μ l éthanol ; 65 μ l NaOH 5N) 1h 50°C sous N ₂	CH ₃ NH ₂ ou HSCH ₂ CH ₂ OH ou CH ₃ CH ₂ SH	Article de référence. Mise en évidence d'une β -élimination de la sérine lors de la dégradation d'Edman (dT-T-Ser)
Byford (Byford, 1991)	1991	100 μ l peptides en eau + 200 μ l Ba(OH) ₂ saturé (~0.155M) arrêté avec acide acétique 1M	0.2-1M CH ₃ NH ₂ (avec pH>12)	Effet catalytique de Ba ²⁺ : β -élimination des thiols libres bloquée si réduction et alkylation préalable de la cystéine

Jaffe et al. (Jaffe et al., 1998)	1996	Protocole Meyer, 200µl de réactif, 18h 50°C sous Ar	CH ₃ CH ₂ SH	Analyse MS/MS (ion trap) de la caséine
Li et al. (Li et al., 2002)	2002	0.1M NaOH + 0.6M Na ₂ SO ₃ 24h 25°C, arrêt par 50µl acide acétique 0.3M	Na ₂ SO ₃	Rendement pS =90 %, pT = 80 % Utilisation du sulfate car meilleur rendement d'addition et meilleure solubilité (par rapport à l'éthanethiol).
Thompson et al. (Thompson et al., 2003)	2003	1µl Ba(OH) ₂ + 0.4µl X 1h 37°C. Arrêt avec NH ₄ ⁺ sulfate	(CH ₃) ₂ NCH ₂ CH ₂ SH.HCl	β-élimination sur colonne IMAC Fe ³⁺ => élution des peptides de la colonne.
Li et al. (Li et al., 2003)	2003	0.1M NaOH 0.6M Na ₂ SO ₃ 37°C, 24h ou 0.52M NaOH 1.36M éthanethiol 50°C 1-18h	Na ₂ SO ₃ ou CH ₃ CH ₂ SH	Observation de β-élimination sur S et T quand réaction conduite à plus de 25°C.
Rusnak et al. (Rusnak et al., 2004)	2004	0.18 M Ba(OH) ₂ + 0.2M X, 1h 25°C	(CH ₃) ₂ NCH ₂ CH ₂ SH.HCl	Addition nucléophile simultanée ou séquentielle : résultats équivalents.
Ahn et al. (Ahn et al., 2004)	2004	Phosphopeptides sur Fe ³⁺ IMAC +10µl de 2 % NaOH 1 % MEA 15min-2h 45°C, arrêt avec TFA puis ZipTip	HSCH ₂ CH ₂ OH	ΔM = -20 u observé en MALDI MS.
Chen et al. (Chen et al., 2010)	2010	12.5 µl peptide + 9.5µl DMAS + 3µl ethanol + 12.5µl Ba(OH) ₂ saturée, 3h RT sous N ₂ . Arrêt avec 5µl 10 % TFA puis ZipTip, puis 50µl 0.5M X dans NaOH 0.5M, 5h 50°C. Arrêt avec 5µl 10 % TFA	Thiocholine	Thiocholine = > ajout d'une charge + favorable à l'ionisation. Marquage isotopique de la thiocholine permettant d'analyse différentielle
Kim et al. (Kim et al., 2011)	2011	10µl peptides + 10µl 0.5M BrX en ACN ou ethanol + 10µl 0.15M Ba(OH) ₂ , 18h temp. ambiante. Arrêt par TFA 10 %	p-bromobenzyl mercaptan	Signature isotopique du Br (⁷⁹ Br = 51 % et ⁸¹ Br = 49 %).

Greffage des phosphopeptides sur support solide

La réaction de β-élimination peut être utilisée pour ajouter un groupement fonctionnel permettant la fixation du phosphopeptide à une phase solide. Par exemple, après oxydation des cystéines, l'addition d'éthanethiol permet de rajouter un groupement SH utilisé dans un deuxième temps pour greffer une biotine maléimide. Le remplacement du phosphate par la biotine permet alors la purification des peptides biotinylés par une colonne d'avidine (Oda et al., 2001). Plus simplement, le thiol libre peut être utilisé pour la fixation sur une phase sépharose -thiol (McLachlin and Chait, 2003). D'autres articles ont présenté des méthodes analogues en ajoutant une possibilité de quantification par l'utilisation de réactifs marqués (Goshe et al., 2002; Qian et al., 2003). Les

sondes utilisées peuvent comporter 4 fonctionnalités : une partie réactive vis-à-vis du phosphopeptide, une partie contenant une étiquette isotopique, une partie permettant la purification par chromatographie d'affinité et une partie contenant une liaison labile permettant l'élution de la colonne d'affinité (van der Veken et al., 2005).

Réactions parasites. Trois acides aminés sont susceptibles de subir une β -élimination en milieu basique : la cystéine, la sérine et la thréonine. La cystéine produit le même intermédiaire déhydroalanine que la phosphosérine. Cette β -élimination indésirable est évitée en oxydant, ou en réduisant puis alkylant les cystéines. De manière plus problématique, la β -élimination des sérines non phosphorylées a été également décrite (Li et al., 2003) (McLachlin and Chait, 2003), et aucune alternative satisfaisante n'a été présentée à ce jour. On peut noter la rareté des publications sur ce point. Le travail de Li et collaborateurs montre en 2 articles le piège de cette approche. Le premier article présente une méthode optimisée de β -élimination en utilisant uniquement des phosphopeptides de synthèse (Li et al., 2002). Un an plus tard, les mêmes auteurs publient une nouvelle étude en étudiant la β -élimination des phosphopeptides mais aussi des formes non phosphorylées (Li et al., 2003). Ils montrent alors que les traitements alcalins à plus de 25°C entraînent une β -élimination des groupements hydroxyles de la sérine et de la thréonine. Un autre travail utilisant la β -élimination pour greffer les phosphopeptides sur une phase solide fonctionnalisée par des thiols détermine que 1.7 % des peptides non phosphorylés sont greffés à la colonne (McLachlin and Chait, 2003). Ce pourcentage de conversion non spécifique des sérines et thréonines est un problème majeur dans le cadre des approches phosphoprotéomiques. Dans un extrait protéique total, la proportion des formes phosphorylées n'est pas estimée, elle ne représenterait que quelques pourcents. En d'autre terme, l'utilisation de la β -élimination pour la phosphoprotéomique peut aboutir à l'identification de sites de phosphorylation artéfactuels.

1.3.3 Techniques d'enrichissement des phosphopeptides.

Les premières techniques d'enrichissement ont été mises au point pour les phosphoprotéines. L'implication de la spectrométrie de masse et sa contrainte de gamme de masse pour l'analyse MS/MS a orienté la technique vers l'enrichissement des phosphopeptides. Il existe plusieurs méthodes pour purifier les phosphopeptides : précipitation sélective, chromatographie d'affinité, ou chimie de greffage après élimination du groupement phosphate. De très loin, la

chromatographie d'affinité sur métal chélaté ou sur oxyde de métal sont les deux techniques les plus employées.

Précipitation des phosphopeptides

Sachant que le phosphate précipite sous forme de sel de calcium $\text{Ca}_3(\text{HPO}_4)_2$ lorsque CaCl_2 est ajouté à la solution, Zhang et collaborateurs montrent qu'il est possible de précipiter les phosphopeptides de la même manière (Zhang et al., 2007). La précipitation suivie d'une purification IMAC d'un extrait d'embryon de riz a ainsi permis d'identifier 242 phosphopeptides. Dans le cadre de la maladie d'Alzheimer, une analyse LC-MS/MS après précipitation des phosphopeptides issus d'extraits de cerveau a permis l'identification de 466 phosphopeptides (Xia et al., 2008). Plus récemment, les sels de lanthanide ont été décrits comme pouvant précipiter les phosphopeptides de caséine ou ovalbumine, ErCl_3 étant le plus efficace (Guzel et al., 2013). Cette méthode d'enrichissement reste exceptionnelle par rapport à la chromatographie sur colonne d'affinité.

Chromatographie d'affinité par métal chélaté (IMAC).

Les Phases chélatrices. Immobilized Metal Affinity Chromatography (IMAC) désigne les phases solides sur lesquelles sont greffés des groupements chélateurs de métaux. Deux ligands sont couramment employés : Imidoacétate (IDA) et nitriloacétate (NTA). Ils chélatent les ions métalliques par l'intermédiaire de leurs fonctions carboxyliques et de l'azote de l'amine secondaire ou tertiaire (**Fig. 22**). Plus récemment, une phase polymérique portant des groupements phosphate a également été utilisée pour immobiliser des ions métalliques Ti^{4+} . (Zhou et al., 2008).

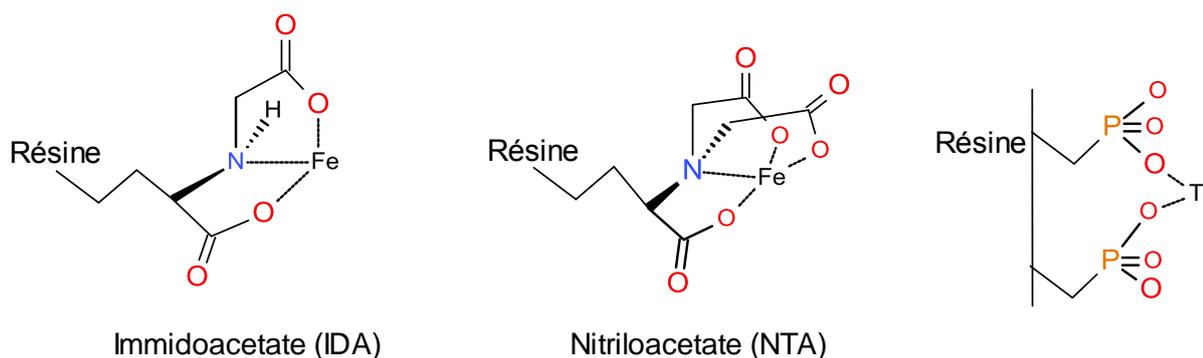


Figure 22. Résines chélatrices de métaux. Résine IDA (Imidoacetate), NTA (nitriloacetate), et phosphate pour la fixation des cations

La comparaison des résines NTA et IDA a été réalisée par le suivi de phosphopeptides marqués ^{32}P dans les différentes fractions de la chromatographie IMAC. Si la capacité de fixation semble équivalente entre les deux résines, l'élution à pH élevé (ammoniaque pH 10.5 ou 50 mM

phosphate d'ammonium) permettra une meilleur rendement avec la résine IDA (Nuhse et al., 2003).

Choix de l'ion métallique. L'ion métallique le plus souvent utilisé est Fe^{3+} sous forme de chlorure ferrique. Sur colonne IDA, les rendements de fixation et d'élution des phosphopeptides ont été déterminés pour différents métaux chélatés grâce à un mélange de phosphopeptides marqués au ^{32}P . (Posewitz and Tempst, 1999). Les auteurs différencient d'un part les rendements de fixation des phosphopeptides ($\text{Ga}^{3+} = \text{Fe}^{3+} = \text{Zr}^{3+} > \text{Al}^{3+}$) et d'autre part les rendements d'élution ($\text{Ga}^{3+} \geq \text{Fe}^{3+} > \text{Zr}^{3+} > \text{Al}^{3+}$). L'intérêt du Gallium a été confirmé par Aryal et collaborateurs (Aryal et al., 2008). Avec une phase chélatrice de type polyphosphate, Ti^{4+} et Zr^{4+} immobilisés permettraient de purifier un plus grand nombre de phosphopeptides que Fe^{3+} (Yu et al., 2009; Zhou et al., 2008).

Méthylation des fonctions acides. Les peptides riches en acide aspartique (D) et en acide glutamique (E) se fixent également aux colonnes IMAC par leur fonction carboxylate. La méthylation permet d'éviter cette interaction non spécifique. La réaction la plus couramment utilisée est une méthylation par du chlorure de méthyl en solution méthanolique. Elle aboutit à l'ajout d'un méthyle sur les fonctions carboxyliques et sur les cystéines S carboxyméthylées, soit un incrément de masse de 14 Da (Ficarro et al., 2002; Ficarro et al., 2005). Il apparait cependant que l'estérification est partielle et complexifie le mélange (Xu et al., 2005). Cette approche n'est plus utilisée dans les articles les plus récents.

Tampon de charge. La rétention des phosphopeptides sur les colonnes d'affinité est réalisée à un pH entre 3 et 5, généralement avec 0.1-0.2M d'acide acétique. Il faut noter que les peptides sont généralement produits par trypsinolyse à pH basique et que les sels d'ammonium inhibent la rétention des phosphopeptides (Tsai et al., 2008). A défaut d'un dessalage des peptides, seul 25 mM de triéthylammonium-bicarbonate ou -formate sont utilisables avec une colonne NTA- Fe^{3+} . L'ajout de solvant organique permet d'augmenter la solubilité des peptides et renforce l'interaction entre les phosphopeptides et le métal chélaté. Un mélange acétonitrile/méthanol/eau (1/1/1) a été proposé (Lee et al., 2007a; Ndassa et al., 2006; Ye et al., 2010) et des auteurs proposent jusqu'à 80 % d'acétonitrile dans le tampon de charge (Alpert, 2008; Nawrocki et al., 2004).

Tampon de charge, Importance du pH. Les acides aspartiques et glutamiques ont des pKa respectifs de 3.65 et 4.25 alors que les résidus phosphorylés ont un pKa moyen de 1.52. Il est possible de limiter l'interaction ionique des fonctions carboxyliques avec le métal en utilisant un tampon de charge dont le pH est suffisamment bas pour protoner ces fonctions carboxyliques ; le

chargement en acide acétique de pH 3.5 à 2.5 permet de baisser la liaison non spécifique d'une colonne Fe³⁺-NTA de 40 % à 4 % (Tsai et al., 2008).

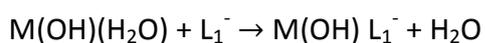
Tampon de charge, importance de l'acide. Avec un contrôle précis du pH et de la concentration d'acide dans le tampon de charge, Tsai et collaborateurs montrent que l'augmentation concentration de 3 à 6 % d'acide acétique pH 3 permet d'augmenter la spécificité de NTA-Fe³⁺ de 85 % à 96 %. (Tsai et al., 2008). Plus récemment, des conditions extrêmes utilisées pour des colonnes IMAC-Ti⁴⁺ (6 % TFA, 14 % acide acétique, 80 % ACN) ont été utilisées avec succès (Nawrocki et al., 2004).

Tampon d'élution, sélectivité. Pour une analyse MALDI-TOF, l'élution des phosphopeptides par la matrice de cristallisation acidifiée (2,5DHB à saturation dans 50 % ACN, 1 % ac. phosphorique) est une solution plus efficace qu'une élution en tampon basique (Gan et al., 2008). Pour une analyse LC-MS, l'acide phosphorique n'est pas utilisé, l'élution à pH élevé est généralement préférée. Lee et collaborateurs ont montré que l'élution en milieu basique (50 mM Na₂HPO₄ pH 9.2 était plus efficace que l'élution en milieu acide (5 % TFA) (Lee et al., 2007b). L'utilisation de phosphate d'ammonium et l'ammoniaque est aujourd'hui la plus fréquente. La possibilité d'éluer les phosphopeptides à pH bas et élevé a été utilisée pour une élution séquentielle des phosphopeptides appelée SIMAC pour Sequential Elution from IMAC (Thingholm et al., 2008). Une colonne NTA-Fe³⁺ est utilisée pour retenir les phosphopeptides. D'une part la fraction non retenue subit une autre chromatographie d'affinité sur colonne TiO₂, d'autre part la première élution de NTA-Fe³⁺ par 1 % TFA 20 % ACN permet d'éluer les peptides monophosphorylés, moins affins pour le métal, et 25 % NH₄OH pH 11.3 permet d'éluer les peptides multiphosphorylés retenus plus fortement.

Complémentarité des méthodes IMAC. Sachant que différents métaux (Ga, Fe, Ti) permettent une purification efficace des phosphopeptides, il paraît naturel de vouloir comparer les phosphopeptides purifiés avec chacun de ces métaux. Aucune étude de ce type n'a été publiée à ce jour. Récemment Chuan et collaborateurs ont montré que IMAC-Ti⁴⁺ et IMAC-Fe²⁺ étaient complémentaires, le recouvrement des phosphopeptides purifiés par l'une ou l'autre méthode ne dépassant pas 10 % du total des séquences (2629 peptides identifiés) (Nawrocki et al., 2004). Il n'est cependant pas sûr (ceci étant un avis personnel) que la différence de population soit représentative d'une complémentarité dans la mesure où les conditions d'élution utilisées étaient radicalement différentes : IMAC-Fe²⁺ élué par NH₄H₂PO₄ pH 4.4 et IMAC-Ti⁴⁺ élué par 25 % NH₄OH.

Chromatographie d'affinité sur oxyde de métal (MOAC)

Caractéristiques de l'oxyde (IV) de titane. La chromatographie sur colonne d'oxyde de métal a été introduite plus récemment que la chromatographie IMAC. L'intérêt de cette phase solide est sa stabilité thermique et chimique pour des conditions qui conduiraient à l'élution des sels métalliques des colonnes IMAC. Différents oxydes tels que ZrO_2 et Al_2O_3 ont été utilisés, mais le dioxyde de titane (TiO_2) est le plus couramment employé pour la purification des phosphopeptides. D'autres oxydes de titane (TiO , Ti_2O_3 , Ti_3O_5) seraient également utilisables, ainsi que le titane pur (Sano and Nakamura, 2007). Le dioxyde de titane existe sous 3 formes ; anatase, rutile et brookite, dont les différences résident en l'organisation du cristal. Il sert essentiellement sous sa forme anatase de pigment blanc obtenu après calcination dans les peintures, les plastiques, les dentifrices ou les aliments (colorant E171) et accessoirement de support chromatographique. En chromatographie, les particules poreuses sont obtenues par polymérisation en présence de colloïdes. (Gan et al., 2008). Des particules sphériques, ayant un diamètre de 3.5 μm et des pores 32 nm ont ainsi été décrites. De manière générale, la surface des particules d'oxyde de titane est hétérogène et porte des groupes OH et des ions Ti^{4+} (Nawrocki et al., 2004). En fonction du pH, les titanols sont sous formes $TiOH_2^{(2/3)+}$, $TiOH$ ou $TiO^{(2/3)-}$ (pI \approx 5). La capacité d'échange de ligand est due à la présence d'acide de Lewis (Ti^{4+}) et la présence de molécules d'eau liées par coordination. Une série d'équations décrit les équilibres entre l'oxyde de métal, les molécules d'eau et les bases de Lewis (Nawrocki et al., 1993). A bas pH, les deux équations suivantes illustrent les échanges de ligands :



$M(OH)L_1^- + L_2^- \rightarrow M(OH) L_2^- + L_1^-$ ou M correspond à l'oxyde de métal, L_1 et L_2 les bases de Lewis dans l'éluant et le soluté. La force de l'interaction entre la base de Lewis L_1 et les sites de liaison de l'oxyde de métal suit un ordre décroissant phosphate > fluorure > citrate > sulfate > acétate > formate > nitrate > chlorure (Nawrocki et al., 2004). Le phosphate est donc un ligand fort des oxydes de métaux et cette propriété est utilisée pour la purification des phosphopeptides.

Chromatographie en colonne. Les premières utilisations de TiO_2 comme phase stationnaire ont été réalisées en système HPLC (Nawrocki et al., 2004; Zarei et al., 2011). La rétention/élution des phosphopeptides sur la phase est réalisée par 0.1 % TFA / borate pH8 (Sano and Nakamura, 2004), 0.1 % TFA / 100 mM phosphate de potassium pH8 (Zarei et al., 2011) ou 0.1 M acide acétique / NH_4HCO_3 pH9 (Pinkse et al., 2004).

Chromatographie en pointes jetables. Le protocole de fabrication de micro-colonnes de TiO₂ dans des pointes de pipettes de type « gel loader » a été décrit en détail (Thingholm et al., 2006). Dans cette configuration, les auteurs ont testés différents tampons de chargement et d'élution des phosphopeptides. Un tampon de charge constitué par une solution aqueuse à 300 mg/ml DHB, 80 % ACN, 0.1 % TFA et une élution par NH₄OH pH 10.5 ont été proposés pour diminuer la liaison non spécifique des peptides (Larsen et al., 2005).

Tampon de charge, rôle de l'acide. De nombreux articles ont présenté des études dont l'objectif était d'empêcher la liaison des peptides riches en acides aminés D et E. S'il est clair que l'oxyde de titane lie également les peptides riches en D et E, il est par contre moins évident de tirer une conclusion limpide des articles publiés. Certains sont contradictoires, notamment en ce qui concerne l'utilisation du DHB pour réduire la liaison non spécifique. Différents acides ont été testés pour inhiber la liaison des peptides non phosphorylés à l'oxyde de titane. Une comparaison systématique de différents acides (glycolique, phtalique, lactique) dans le tampon de charge à 80 % ACN a montré que l'acide lactique (300mg/ml), au contraire du DHB dont l'intérêt a été contesté par ailleurs ((Bodenmiller et al., 2007), permettait de supprimer l'interférence des peptides non phosphorylés sans affecter la liaison des phosphopeptides (Sugiyama et al., 2007). L'acide glutamique (solution aqueuse 2 % TFA, 65 % ACN saturée en acide glutamique) a également été proposé comme additif (Wu et al., 2007). Des tampons de charge constitués de mélanges (20 % acide acétique, 420 mM acide 1-octanesulfonique, 50 mg/ml DHB) ont également été proposés (Mazanek et al., 2010). A ce jour, il n'y a pas de protocole de référence pour la purification des phosphopeptides par MOAC ; néanmoins, il est clair que l'utilisation de concentrations élevées d'acide (>100 mM) permet de limiter les liaisons non spécifiques. Très souvent, 2 à 5 % TFA et 80 % ACN sont utilisés dans le tampon de charge, souvent en association avec un autre acide (Cheng et al., 2013; Han et al., 2012; Urbaniak et al., 2013).

Tampon d'élution. De la même manière que pour la chromatographie IMAC, L'élution sélective permet d'augmenter le nombre de peptides identifiés. L'utilisation d' éluants de force croissante, dans l'ordre ; 0.5 % NH₄OH, 0.5 % pipéridine, 0.5 % pyrrolidine, 5 % NH₄OH, 5 % pipéridine, et 5 % pyrrolidine, permet d'éluer des fractions de phosphopeptides dont la nature physicochimique (index GRAVY, proportion d'acides aminés D et E, sites de phosphorylation) est nettement différente (Kyono et al., 2008).

Autres chromatographies

Les deux méthodes présentées ici sont extrêmement marginales. Elles ont été utilisées ponctuellement.

Résine échangeuse d'anion fort (SAX). La chromatographie SAX permet de retenir efficacement les peptides acides à pH élevé. L'élution est généralement réalisée par un gradient salin. L'élution par un gradient de pH est plus délicate mais permet d'éviter l'utilisation de concentration en sels importantes. Le couplage avec la LC-MS est alors possible. Jie Dai et collaborateurs ont proposé un système *on line* de chromatographie SAX avec une élution par pH décroissant (pH 8.2 à pH 2) des peptides vers une deuxième dimension classique (phase inverse) (Dai et al., 2009). A partir de 500 µg de peptides (culture de cellules HeLa), 1833 phosphopeptides sont élués tout au long du gradient avec un pic de peptides multiphosphorylés dans les dernières fractions issues de la colonne SAX.

L'hydroxyapatite. L'hydroxyapatite est un polymère de phosphate de calcium hydroxylé $\text{Ca}_{10}(\text{PO}_4)_6(\text{OH})_2$. Ce support chromatographique est utilisé notamment comme dernière étape de chromatographie (polishing) lors de la purification d'anticorps. L'hydroxyapatite a été utilisé pour isoler des phosphopeptides à partir d'un extrait de cerveau. Avec une approche MudPit, il a ainsi été possible d'identifier et quantifier 3069 phosphopeptides (Fonslow et al., 2012).

Chromatographie multidimensionnelle appliquée aux phosphopeptides.

Une méthode efficace pour augmenter le nombre de phosphopeptides identifiés est l'utilisation d'une dimension chromatographique supplémentaire, chacune des fractions produites par la première dimension étant enrichie en phosphopeptides par IMAC ou MOAC, ou à nouveau fractionnée par une autre chromatographie. Finalement, les fractions produites sont analysées par RPLC-MS.

Première dimension par échange d'ions. En 2003, une analyse *off line* par chromatographie d'échange d'anions SAX (POROS SAX, 25 mM NH_4HCO_3 30 % ACN, pH7, élution entre 80 et 500 mM NaCl) puis IMAC (Fe^{3+} -IDA) a permis d'identifier plus de 200 phosphopeptides d'une fraction membranaire (Nuhse et al., 2003). L'association de SCX+SAX a été mise en œuvre par la suite. En récupérant la fraction non retenue de SCX et en la séparant par SAX, les 23 fractions obtenues ont permis d'identifier 849 phosphopeptides sur un total de 14105 peptides provenant d'un extrait de foie de souris (Dai et al., 2007). Bien que les phosphopeptides, de nature acide, soient en théorie désignés pour être séparés par échange d'anions, les chromatographies échangeuses de cation ont

été par la suite plus largement utilisées. La combinaison de SCX (polySULFOETHYL A) avec IMAC a permis d'identifier 5500 et 13000 phosphopeptides d'extrait de foie et de drosophile respectivement (Villen and Gygi, 2008). Nie et collaborateurs ont comparés l'enrichissement par TiO_2 avec SAX- TiO_2 ou SCX- TiO_2 (Nie et al., 2012). A partir d'extraits de cellules HeLa soumis à une protéolyse par la trypsine, 929 phosphopeptides sont identifiés par TiO_2 seule avec 4 h d'analyse, 1686 phosphopeptides pour SCX- TiO_2 et 40 h d'analyse, la meilleure combinaison étant une combinaison SAX-SCX- TiO_2 qui permet d'identifier 2705 phosphopeptides pour 80 h d'analyse.

Première dimension ERLIC. Si une colonne d'échange d'ions est utilisée avec une phase mobile de type organique, par exemple 70 % ACN, les solutés sont alors retenus par une interaction hydrophile même si leur charge est identique à celle de la phase. Andrew J. Alpert a utilisé cette combinaison appelé ERLIC (Electrostatic Repulsion Hydrophilic Interaction) pour la séparation de phosphopeptides (Alpert, 2008). Pour des cellules stimulées par EGF, facteur de croissance qui stimule la phosphorylation, la méthodologie ERLIC permet d'identifier 926 phosphopeptides, SCX 194 phosphopeptides, mais SCX-IMAC 1315 phosphopeptides (Gan et al., 2008). Logiquement, le mode ERLIC a donc été associé à une purification IMAC ou TiO_2 . Pour des cellules embryonnaires de poulet, une méthodologie ERLIC ou IMAC seule identifie moins de 20 phosphopeptides alors que la combinaison ERLIC+IMAC permet d'identifier 428 phosphopeptides (Chien et al., 2011).

Etudes comparatives. Face à la multitude de techniques, la question du choix se pose naturellement. La comparaison des modes ERLIC- TiO_2 , SCX- TiO_2 , et HILIC- TiO_2 pour un extrait de cellule HeLa a montré que SCX- TiO_2 permettait d'identifier le plus grand nombre de phosphopeptides ; cependant ERLIC- TiO_2 permet d'identifier plus de peptides multiphosphorylés, cette classe de phosphopeptides n'étant pas ou peu retenue par SCX (Zarei et al., 2011). Cette complémentarité a été vérifiée dans une étude comparant différentes combinaisons entre ERLIC, SCX, TiO_2 . A partir d'un extrait de cellules HeLa, l'association SCX- TiO_2 a permis l'identification de 4373 phosphopeptides alors que l'association ERLIC- TiO_2 avec une chromatographie SCX- TiO_2 de la fraction non retenue de ERLIC permet d'identifier 6433 phosphopeptides (Zarei et al., 2012).

Association IMAC et TiO_2 . Le développement en parallèle de deux chromatographies d'affinité pour les phosphopeptides a naturellement provoqué des comparaisons. Bodenmiller et collaborateurs montrent que dans leurs mains seuls 35 % des 921 phosphopeptides identifiés par IMAC- Fe^{3+} (555 phosphopeptides) ou TiO_2 (366 phosphopeptides) sont retrouvés dans les 2 enrichissements (Bodenmiller et al., 2007). Cette complémentarité a été mise en pratique dans la méthodologie SIMAC présenté dans le paragraphe précédent. Les peptides non retenus sur une

colonne IMAC-Fe³⁺ peuvent être retenus sur une colonne TiO₂ (Thingholm et al., 2008). Cette association a été revisitée par les mêmes auteurs avec l'approche TiSH (TiO₂-S₁IMAC-H₁LIC) (Engholm-Keller et al., 2012). A partir de 300 µg de peptides hydrolysés par l'endoprotéinase lys-C et provenant de culture de cellules pancréatiques, la première purification est réalisée par TiO₂, puis une chromatographie IMAC-Fe³⁺ avec une élution séquentielle en milieu acide puis basique produit 2 fractions qui sont elles-mêmes séparées par HILIC. Cette technique permet d'identifier environ 6600 phosphopeptides uniques.

1.3.4. Dosage des acides aminés phosphorylés

Généralités

Au laboratoire, les premières étapes de l'analyse protéomique sont la préparation de l'échantillon et l'extraction des protéines. Il est donc rigoureux de quantifier les protéines dans l'échantillon de départ et dans l'extrait afin de déterminer un rendement d'extraction. Pour ce qui est de l'analyse phosphoprotéomique, l'objet central de l'étude, les phosphopeptides, ne sont pas quantifiés. Aucun article de phosphoprotéomique ne détermine les quantités absolues de phosphopeptides. Une détermination de ce type lèverait pourtant un certain nombre d'interrogations :

- **Quel est le niveau global de phosphorylation de l'échantillon?** Les articles de développement présentent des méthodes de purifications de phosphopeptides à partir d'échantillons divers : mélange de protéines, culture de cellules, organes, rendant impossible la comparaison des méthodes entre elles. La connaissance des quantités absolues d'acides aminés phosphorylés avant et après purification permettrait de calculer les rendements des purifications et de définir le facteur d'enrichissement entre l'échantillon de départ et la fraction purifiée.

- **Quelles sont les proportions de pS, pT et pY dans l'échantillon?** Les ratios 2000/100/1 (pS/pT/pY) présentés dans les années 2000 ont évolué. Actuellement, les analyses phosphoprotéomiques présentent un ratio 100/10/1 sur la base du comptage des sites de phosphorylation. Dans la mesure où les rendements de purification de chacune des 3 classes de phosphopeptides ne sont pas déterminés, il est possible que le ratio 100/10/1 soit également erroné.

A ce jour, l'analyse des acides aminés phosphorylés est basée sur les techniques de compositions en acides aminés, mais le caractère labile de l'orthophosphate nécessite des

ajustements des méthodes d'hydrolyse. Le dosage des acides aminés d'une protéine ou d'un peptide consiste dans un premier temps à hydrolyser les liaisons peptidiques, puis dans un deuxième temps à doser les acides aminés libres. Il convient de noter le remplacement de la chromatographie d'échange d'ions par la chromatographie phase inverse après dérivation chimique des acides aminés. En effet, les méthodes de référence utilisent une dérivation des acides aminés qui améliore la séparation en RPLC et augmente la sensibilité. Une revue complète sur l'hydrolyse et la composition des acides aminés a été publiée en 1998 (Fountoulakis and Lahm, 1998). Brièvement, Il existe 3 méthodes d'hydrolyse : l'hydrolyse en milieu acide, l'hydrolyse en milieu basique, et l'hydrolyse enzymatique. L'hydrolyse par un acide aqueux est la méthode la plus répandue ; elle utilise généralement HCl 6N, pendant 20 ou 24h, à 110°C sous vide. Dans ces conditions, l'asparagine et la glutamine sont converties respectivement en acide aspartique et acide glutamique, le tryptophane est détruit, la cystéine est modifiée, la tyrosine, la sérine et la thréonine sont partiellement détruites dans des proportions plus importantes pour la sérine (Darragh et al. 1996). Des additifs tels que le phénol ou l'acide thiocholique sont ajoutés au milieu d'hydrolyse pour limiter l'oxydation,. Les évolutions les plus significatives des méthodes d'hydrolyse portent sur la diminution des temps d'hydrolyse grâce à l'utilisation de micro-ondes et/ou l'addition de TFA, et l'augmentation de la sensibilité de détection des acides aminés.

Dérivation et chromatographie des acides aminés

La composition en acides aminés a longtemps utilisé une séparation sur échange d'ions suivie d'une dérivation à la ninhydrine (Moore and Stein, 1954). Cette technique est de moins en moins employée et d'autres méthodes utilisant une dérivation stable sont maintenant couramment employées. Les acides aminés libérés par l'hydrolyse peuvent être dérivés avec des réactifs chromophores ou fluorescents puis détectés et quantifiés par HPLC. Les réactifs utilisés sont entre autres le phényl isothiocyanate (PITC) l'ortho-phthalaldéhyde (OPA), le 9-fluorométhyl chloroformate (FMOC) et le 6-aminoquinolyl-N-hydroxy-succinimidyl carbamate (ACQ) (**Fig. 23**). En effectuant une recherche par mot clé, il apparait que les deux dérivations les plus employées sont la dérivation au FMOC (52 %) et OPA (23 %) (Molnar-Perl, 2011). Le 6-aminoquinolyl-N-hydroxy-succinimidyl carbamate a été publié en 1993 (Cohen and Michaud, 1993) puis est devenu un kit commercialisé par Waters. A ce titre, il est peu référencé dans la base PubMed NCBI bien qu'il soit utilisé dans de nombreux laboratoires d'analyse.

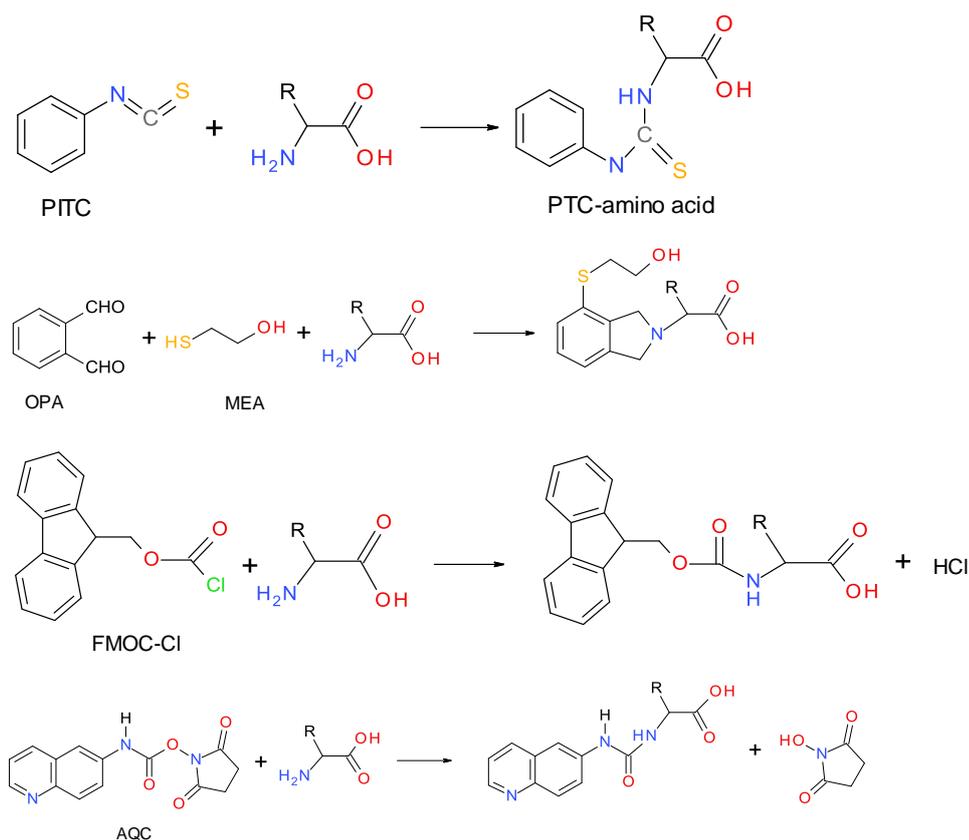


Figure 23. Principales réactions de dérivation utilisées pour l'analyse de composition en acides aminés. Le PhenylthioCyanate (PITC) réagit avec les amines pour former un dérivé phenylthiourée détecté par UV (268-270 nm). L'ortho-phthalaldehyde (OPA) réagit avec les groupements amine en présence d'un agent réducteur comme le β-mercaptoéthanol (MEA) pour former un dérivé isoindole détecté par UV (338 nm) ou par fluorescence (Ex/Em : 340/450 nm). Le 9-fluorométhyl chloroformate (FMOC-Cl) réagit avec les amines pour former un dérivé détecté par UV (265 nm) ou par fluorescence (266/305 nm). Enfin, le 6-aminoquinolyl-N-hydroxy-succinimidyl carbamate (AQC) réagit avec les amines pour former un composé fluorescent (250/395 nm) également détectable en UV (260 nm).

La dérivation avec la ninhydrine et le PITC permettent un dosage avec une limite de quantification de l'ordre de la picomole d'acides aminés, alors que l'OPA et ses dérivés présentent des sensibilités supérieures. Cependant la mise en œuvre de la dérivation à l'OPA est délicate car les réactifs et les dérivés formés sont instables (Molnar-Perl, 2001).

Le FMOC est plus simple d'emploi, car le réactif et les conjugués formés sont stables. Deux difficultés sont cependant associées à l'utilisation du FMOC : (1) l'excédent de FMOC s'hydrolyse pour former un FMOC hydroxylé (FMOC-OH) qui est élué au milieu du chromatogramme (2) le marquage de His et Tyr est hétérogène et produit des dérivés mono- ou di-substitués qui compliquent leur quantification (Malmer and Schroeder, 1990). Par ailleurs, le pic majoritaire de FMOC-OH peut être éliminé par une extraction au pentane (Bank et al., 1996).

A titre d'exemple, deux chromatogrammes de standards d'acides aminés sont présentés (**Fig. 24**), dérivés soit par FMOC (à gauche) soit par AQC (à droite). Les deux séparations sont réalisées

avec des colonnes RP C18, Chaque acide aminé dérivé produit un pic résolu qui permet sa quantification. Dans le cas de la dérivation AQC, l'utilisation de l'UPLC permet de réduire le temps d'analyse à 10 min (30 min pour une HPLC standard).

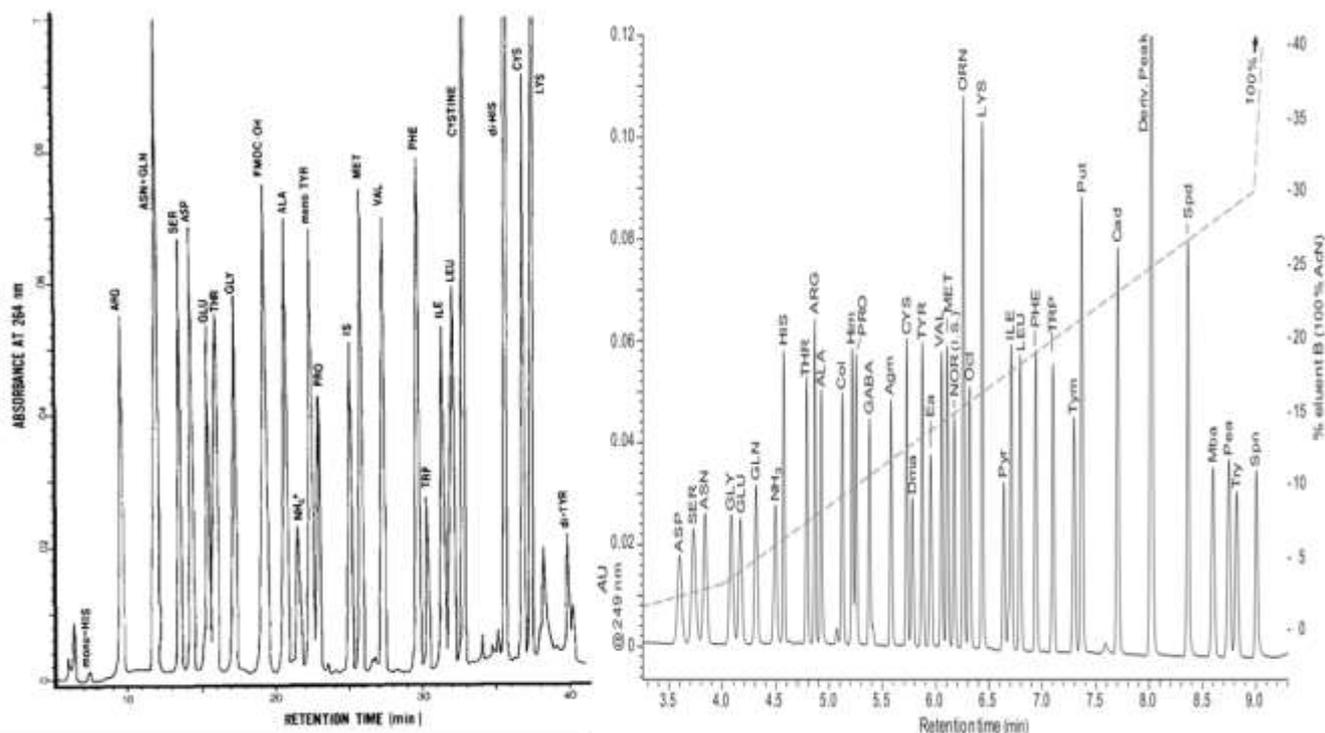


Figure 24. Séparation d'acides aminés standards. Dérivation par le Fmoc (à gauche)(Malmer and Schroeder, 1990) ou dérivation par l'AQC (à droite) (Fiechter et al., 2012). Waters™ commercialise le réactif AQC sous forme d'un kit, ici analysé en UPLC.

Dans le cadre du dosage des acides aminés libres, des protocoles ne nécessitant pas de dérivation ont été proposés de manière ponctuelle (Armstrong et al., 2007; Qu et al., 2002). Dans ces articles, les auteurs utilisent une séparation sur colonne C18 et des contre-ions particuliers, acide perfluoroheptanoïque ou acide tricaéfluoroheptanoïque, qui augmentent la rétention des acides aminés sur la phase inverse. Il faut cependant noter que certains acides aminés sont peu retenus sur la colonne et sont élués dans les premières minutes de chromatographie dans l'ordre d'éluion suivant : Asx, Ser, Glx, Gly, Cys, Thr, les acides aminés les plus retenus étant les acides aminés hydrophobes et les acides aminés basiques.

Dans une logique de simplification, une quantification par ESI-MS en infusion a également été proposée (Louwagie et al., 2012). Cette méthode n'est appliquée qu'à quelques acides aminés pour la quantification de protéines purifiées et le problème de suppression ionique dû au mélange d'acides aminés n'a pas été abordé.

Dosage des acides aminés phosphorylés

Le dosage des acides aminés phosphorylés ajoute un niveau de difficulté supplémentaire aux techniques standard d'analyse des acides aminés. En effet, la liaison phospho-ester est extrêmement sensible aux pH extrêmes ; les acides aminés phosphorylés libèrent leur phosphate en milieu basique (β -élimination) et acide (hydrolyse de la liaison ester). La solution couramment employée est une hydrolyse par HCl 6 N 110 °C pendant un temps réduit, de 1 à 4 h maximum (Capony and Demaille, 1983; Carlomagno et al., 1985; Malencik et al., 1990; Morrice and Aitken, 1985; Niedbalski and Ringer, 1986). Seul un article récent propose un dosage des acides aminés phosphorylés après une hydrolyse de 24 h (Fu et al., 2011), montrant par là l'oubli des données anciennes.

Peu de résultats ont été produits sur la stabilité des acides aminés phosphorylés. Yan et collaborateurs ont montré que sous vapeur d'HCl 5.7 N à 110 °C pendant 4 h, pS et pY sont stables alors que de 25 % de pT est dégradé (Yan et al., 1997). Après 6 h d'hydrolyse, 50 % de pS et 75 % de pT sont dégradés alors que pY reste intact, Ces données sont à prendre avec précaution, car Ringer a présenté en 1991 des résultats sensiblement différents (Ringer, 1991) ; dans des conditions d'hydrolyse équivalentes (HCl 6 N 110°C), les temps d'hydrolyse sont limitées à 3 h pour pS et pT alors que pY est décrit ici comme plus sensible, nécessitant une hydrolyse plus courte, de 1 h à 1.5 h maximum.

On peut s'étonner de l'absence d'une analyse systématique de la stabilité des acides aminés phosphorylés (Yan et al., 1998). Par ailleurs, dans ces conditions de temps d'hydrolyse court, il est probable que l'hydrolyse de la liaison peptidique soit partielle ; le rendement d'hydrolyse dépendra alors de la sensibilité des liaisons peptidiques à l'hydrolyse, variable en fonction des acides aminés adjacents, et de la dégradation des acides aminés phosphorylés, elle-même variable. Dans ces conditions, la robustesse des méthodes présentées prête à discussion. D'autre part, les acides aminés phosphorylés sont minoritaires par rapport aux autres acides aminés et par rapport à leurs précurseurs non phosphorylés. Un effet matrice peut également fausser leur quantification. Murthy et collaborateurs ont contrôlé qu'un excès 100 molaire des acides aminés non phosphorylés ne modifiait pas leur méthode de quantification (dérivation au PITC) (Murthy and Iqbal, 1991). Avec une dérivation au N-hydroxysuccinimidyl fluorescein-o-acétate (SIFA) et une séparation sur colonne C8, Deng et collaborateurs ont également confirmé que les autres acides aminés n'interféraient pas dans la mesure où ils étaient élués plus tardivement dans le gradient (Deng et al., 2007).

A ce jour, les dosages d'acides aminés phosphorylés n'ont jamais été réalisés pour caractériser les fractions issues de purification par colonne d'affinité.

CHAPITRE 2 :
Analyse protéomique de cellules coliques cancéreuses traitées par
thérapie génique CD/5FC.

2.1 Présentation du sujet

2.1.1 Epidémiologie des cancers

Au sein des Nations Unies, l'Organisation Mondiale de la Santé (OMS) est l'autorité directrice et coordonnatrice des travaux internationaux dans le domaine de la santé. Son agence de recherche sur le cancer, le Centre International de Recherche sur le Cancer (CIRC), a publié une étude mondiale sur l'incidence et la mortalité des cancers (IARC, 2008) qui présente le cancer comme une cause majeure de mortalité (13 %) après les maladies infectieuses (29.5 %). Les principaux cancers par les nombres de décès par an sont dans l'ordre décroissant :

- Le cancer du poumon (1,37 million de décès)
- Le cancer de l'estomac (736 000 décès)
- Le cancer du foie (695 000 décès)
- Le cancer colorectal (608 000 décès)
- Le cancer du sein (458 000 décès)

En France, l'institut national du Cancer (INCA), a publié récemment une étude sur l'évolution de la mortalité et l'incidence des cancers en France (Binder-Foucard et al., 2013). En 2012, 355 000 nouveaux cas ont été diagnostiqués (200 000 chez l'homme et 155 000 chez la femme), soit un accroissement d'environ 110 % par rapport à 1982. Cette augmentation d'un facteur 2 est due à l'accroissement de la population (31 %), au vieillissement de la population (34 %) et à l'accroissement du risque lui-même (43 %) (InVS and INCa, 2013). En termes de mortalité, chez l'homme, les trois cancers les plus mortels en France sont le cancer du poumon (21300 décès en 2012), le cancer colorectal (9200 décès) et le cancer de la prostate (8900 décès). Chez la femme, le cancer du sein (11900 décès) est suivi du cancer du poumon (8600 décès) et du cancer colorectal (8400 décès).

Les risques de cancer colorectal sont associés à une dimension socio-économique. Ils sont augmentés par une alimentation riche en viande et graisses animales ainsi que par l'obésité (Bardou et al., 2013). Si l'effet protecteur des fruits et légumes est discuté (Key, 2011), l'effet pro-cancéreux de l'alcool est clairement établi au-delà d'un verre par jour (Fedirko et al., 2011).

2.1.2 Mécanismes moléculaires du cancer colorectal

Mutations multiples. En 1990, Eric R. Fearon et Bert Vogelstein ont proposé un modèle génétique du cancer colorectal (Fearon and Vogelstein, 1990). Il est basé sur la nature monoclonale de la tumeur. La néoplasie colorectale est la conséquence de plusieurs mutations de proto-oncogènes et de gènes suppresseurs de tumeurs. Plusieurs mutations sont nécessaires au développement d'une tumeur maligne, notamment la mutation du gène suppresseur de tumeur Adenomatous Polyposis Coli (*APC*) et l'activation de l'oncogène v-Ki-ras2 Kirsten Rat Sarcoma Viral oncogene (*KRAS*). Dans certains cas, une mutation dans un gène suppresseur de tumeurs, dont *TP53*, a un caractère dominant négatif (la mutation d'un seul allèle suffit pour l'expression du phénotype). Ces mutations sont associées à des altérations de plusieurs régions chromosomiques (délétion de 1p, 5, 8p, 17p, 18q, 22q, insertion dans 8q, 1q, 20q). Les mutations impliquées sur les chromosomes 17 et 18 sont considérées comme importantes pour le caractère métastatique de la tumeur (Gonzalez-Gonzalez et al., 2013).

Le récepteur EGF et sa voie de signalisation. Le facteur de croissance épidermique (Epidermal Growth Factor EGF) est une protéine soluble de 53 acides aminés sous sa forme mature. La liaison de l'EGF à son récepteur membranaire (EGFR) entraîne une cascade d'évènements qui aboutissent à une prolifération cellulaire, une résistance à l'apoptose et à une angiogenèse (**Fig. 25**). La mutation *KRAS* identifiée précédemment (35-45 % des tumeurs colorectales) concerne une des kinases de la voie de signalisation de l'EGFR. Les récepteurs de l'EGF sont surexprimés dans la plupart des tissus cancéreux. L'utilisation d'anticorps anti-EGFR permet de bloquer la voie de signalisation et la prolifération cellulaire (Rodeck et al., 1987) et l'utilisation d'anticorps anti-EGF (panitumumab, cetuximab) est proposée comme nouveau traitement du cancer colorectal métastatique. Cependant, les mutations *KRAS*, *PIK3* et *PTEN* sont des biomarqueurs prédictifs de la résistance aux thérapies ciblées sur le récepteur EGF (Pritchard and Grady, 2011; Siena et al., 2009).

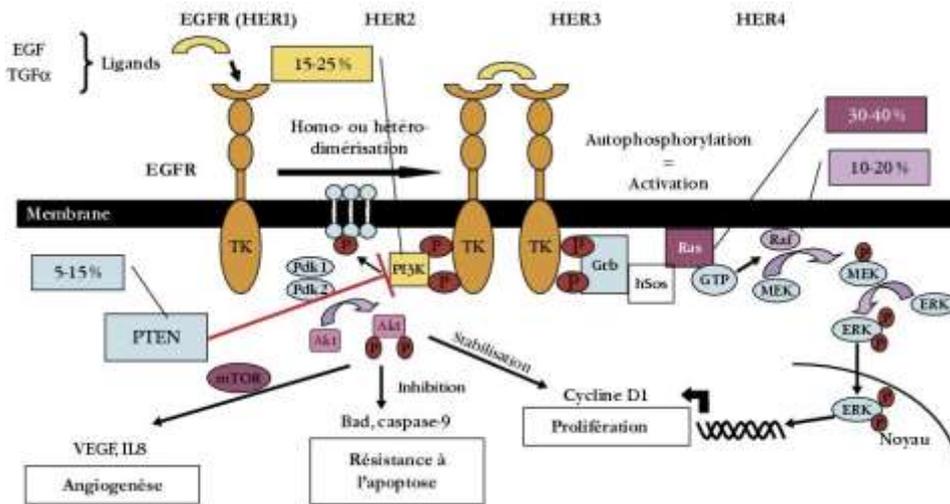


Figure 25. Voie de signalisation du récepteur à l'EGF (EGFR). Les ligands de la famille des EGFRs entraînent une dimérisation du récepteur et la phosphorylation de sa partie intracellulaire. Cette activation provoque une phosphorylation en cascade de kinases (Ras, Akt, mTOR, PI3K, ERK) qui activeront des facteurs de transcription impliqués notamment dans la prolifération. Illustration tirée de (Laurent-Puig et al., 2008)

2.1.3 Traitement du cancer colorectal par le 5-fluorouracile (5FU)

Le 5FU est un analogue fluoré de l'uracile. Il peut être métabolisé comme l'uracile et est incorporé alors dans l'ADN sous sa forme désoxyribose ou dans l'ARN sous sa forme ribose (**Fig. 26**). La cible principale du 5FU est la thymidilate synthase (TS) qui est indispensable au métabolisme des bases pyrimidines. La thymidilate synthase catalyse la conversion de déoxyuridine monophosphate (dUMP) en déoxythymidine monophosphate (dTMP). Prenant la place du dUMP, le 5-FdUMP forme un complexe stable avec la TS, bloquant ainsi la production de dTMP nécessaire à la synthèse de la thymidine. Le déficit en thymidine qui en découle est compensé par une incorporation d'uridine dans l'ADN, cette substitution de T en U provoquant l'apoptose (Longley et al., 2003).

Le 5FU est métabolisé à plus de 80 % par la dihydropyrimidine déshydrogénase (DPD) au niveau hépatique. (Heggie et al., 1987). Cette enzyme intervient dans le catabolisme des bases pyrimidiques et hydroxyle le 5FU, première étape de son catabolisme. Le 5FU est finalement éliminé sous forme de β-alanine fluorée. Une administration directement au niveau de la tumeur permettrait d'éviter la dégradation du médicament par le foie et d'augmenter sa concentration au niveau de la tumeur : la thérapie génique est une des possibilités d'administration localisée du médicament.

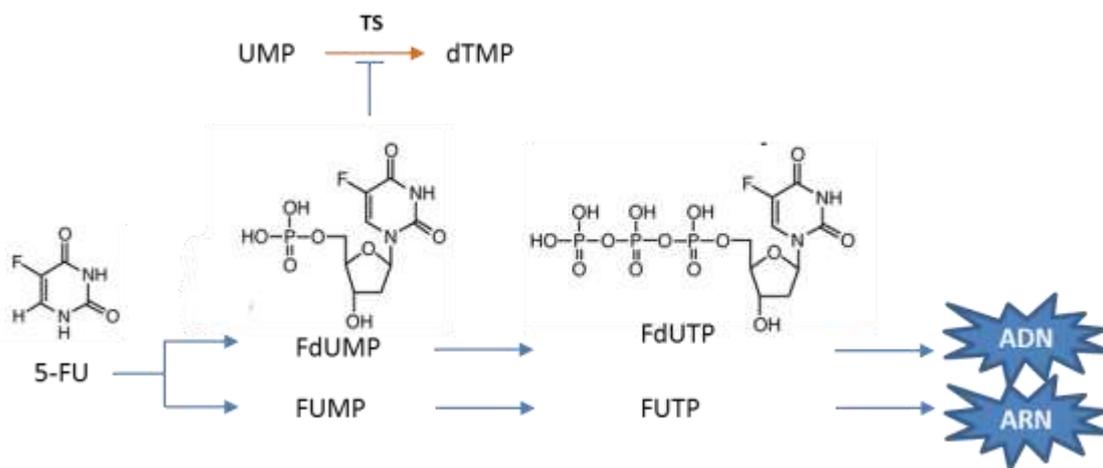


Figure 26. Métabolisme du 5-fluorouracile. Incorporation à l'ADN. Le 5FU est transformé en 5-fluorodésoxyuridine monophosphate (FdUMP) par la thymidine kinase puis en 5-fluorouridine triphosphate (5-FdUTP) par deux pyrimidine kinases. Le 5FdUTP est ensuite incorporé à l'ADN et inhibe son métabolisme. Le 5-FdUMP est également un inhibiteur de la thymidylate synthase (TS) nécessaire au métabolisme des pyrimidines. D'après (Lecomte et al., 2006)

2.1.4 Thérapie génique par la 5-Fluorocytosine (5FC)

Le 5FU est hautement toxique et son efficacité dépend fortement des variabilités génétiques des patients (Lecomte et al., 2006). Il est possible de contourner les problèmes liés à une administration systémique en utilisant une approche de thérapie génique. Le principe est basé sur l'administration de la prodrogue, la 5-fluorocytosine (5FC), qui est ensuite métabolisée en 5-fluorouracile (5FU) par une enzyme, la cytosine désaminase (CD). La CD étant absente chez l'homme, son gène est apporté par un vecteur viral ou des cellules souches transformées avec le gène *CD* (Cihova et al., 2011). L'utilisation de cellules souches mésenchymateuses (MSCs) présentant des affinités particulières pour les tumeurs permettrait de cibler la production de 5FU au niveau des tissus cancéreux. La preuve de principe a été démontrée par l'utilisation de xélogreffes de tumeurs de colon humain chez la souris. 89 % des tumeurs injectées avec des MSCs transformées disparaissent après 24 jours de traitement (n=9) alors que les contrôles non traités développent tous des tumeurs (Kucerova et al., 2008). Du fait des différences de DL_{50} entre 5FC et 5FU, respectivement 15000 mg/kg et 115mg/kg par voie orale chez la souris, l'administration de 5FC peut être réalisée à des doses importantes, soit 500mg/kg chez la souris.

2.1.5 Objectif de l'étude

Les mécanismes moléculaires de la thérapie génique CD/5FC sur des cellules cancéreuses de colon de rat ont été étudiés sur la lignée cellulaire, DHD/K12/PROb (PROb), qui est transfectée par le gène *CD* (PRObCD). Après vérification que le traitement par 5FC induisait une mort cellulaire

spécifique de la métabolisation du 5FC en 5FU par le gène *CD* transfecté, le protéome des cellules traitées par la 5FC a été comparé au protéome des cellules non traitées (CT) dans l'objectif d'identifier des niveaux d'expression protéique affectés par la thérapie génique.

2.2 Choix analytiques

Electrophorèse bidimensionnelle. La séparation des extraits protéiques a été réalisée par électrophorèse bidimensionnelle suivant des conditions standards : première dimension sur bandelettes de gradient de pH (IPG) entre pH 3 à 10, deuxième dimension en SDS PAGE sur gel à 10 % d'acrylamide après réduction des bandelettes IPG par le Dithiothréitol et alkylation par l'iodoacétamide. Les protéines ont été détectées par une coloration au bleu de Coomassie colloïdal puis les gels ont été scannés et analysés par un logiciel dédié à l'électrophorèse bidimensionnelle (ImageMaster 2D Platinum, Amersham Bioscience).

Identification par MALDI-TOF. Les spots dont l'intensité était significativement modifiée par le traitement 5FC ont été identifiés par MALDI-TOF (Voyager DE PRO, PerSeptive Biosystems) après digestion à la trypsine et cristallisation dans une matrice α -Cyano-4-hydroxycinnamic acid (CHCA) sur un support de nitrocellulose. L'identification des protéines à partir de l'empreinte peptidique massique obtenue a été réalisée avec le logiciel Aldente (Expasy, base Swissprot).

Analyse par Western Blot. Les modifications post-traductionnelles de l'HSP90 ont été étudiées par Western Blot avec des anticorps anti-acétyl-lysine, anti-phosphosérine et anti-phosphotyrosine.

Identification par LC-ESI-IT. Un système LC-trappe ionique (Surveyor - LCQ XP, ThermoFinnigan) a été utilisé pour l'identification des modifications post-traductionnelles. Une colonne Hypersil C18 (200 μ m x 10 cm) a été utilisée avec une source électrospray à 5 μ l/min. Un mode d'acquisition classique pour ce genre d'instrument a été utilisé (1 scan MS suivi de 3 scans MS/MS des 3 ions les plus intenses du scan MS). L'identification des peptides a été réalisée par le logiciel Bioworks 3.1 (Thermo Scientific).

2.3 Résultats

2.3.1 Publication

Ce travail a fait l'objet d'une publication : « Treatment of colon cancer cells using cytosine deaminase/5-fluorocytosine suicide system induces apoptosis, modulation of the proteome, and HSP90 phosphorylation ».

Les résultats ont été obtenus à partir d'une culture de cellule cancéreuse de colon transformée par le gène *CD*. Deux conditions de culture ont été comparées ; la première condition (notée FC) correspond aux cellules transformées et traitées par le 5FC ; la deuxième condition (notée CT) est le contrôle et correspond aux cellules transformées mais non traitées par le 5FC. Sans traitement par le 5FC, les cellules poussent rapidement et adhèrent à la boîte de Pétri. L'expression du marqueur annexine V et l'incorporation d'iodure de propidium ont été mesurés par cryométrie de flux. Lorsque le 5FC est ajouté au milieu, le nombre de cellules adhérentes au plastique diminue rapidement et les cellules flottantes expriment les marqueurs de la mort cellulaire ; annexine V (apoptose) et iodure de propidium (nécrose). Si le 5FC est ajouté au milieu de culture des cellules non transformées, aucune mort cellulaire n'est observée.

Les différentes conditions de cultures testées ont permis de montrer que le gène *CD* seul comme le précurseur 5FC seul ne présentent pas de toxicité intrinsèque. Seules les cellules transfectées par le gène *CD* sont sensibles au 5FC. La partie protéomique de l'article présente une analyse par électrophorèse bidimensionnelle et une identification de protéines par LC-ESI-MS. Afin de permettre une analyse quantitative, les électrophorèses bidimensionnelles ont été répétées et les images ont été quantifiées. Dix-neuf protéines sont significativement affectées par le traitement CD/5FC. Parmi elles, la protéine HSP90 est identifiée dans un spot dont le point isoélectrique varie en fonction du traitement, de 5.5 à 5.1 pour respectivement la condition contrôle et la condition traitement.

L'analyse par spectrométrie de masse a permis la caractérisation d'un site de phosphorylation sur la sérine 254. La thérapie génique CD/5FC appliquée à des cellules cancéreuse de colons active probablement l'expression de kinases dont un substrat est l'HSP90.

Treatment of colon cancer cells using the cytosine deaminase/5-fluorocytosine suicide system induces apoptosis, modulation of the proteome, and Hsp90 β phosphorylation

Luc Negroni,^{1,2} Michel Samson,^{1,3}
Jean-Marie Guignonis,^{1,4} Bernard Rossi,³
Valérie Pierrefite-Carle,³ and Christian Baudoin³

¹Plateforme Protéomique IFR50, ²Centre National de la Recherche Scientifique and ³Institut National de la Santé et de la Recherche Médicale Unité 638, ⁴Faculté de Médecine, Université de Nice Sophia-Antipolis, Nice Cedex 2, France

Abstract

The bacterial cytosine deaminase (*CD*) gene, associated with the 5-fluorocytosine (5FC) prodrug, is one of the most widely used suicide systems in gene therapy. Introduction of the *CD* gene within a tumor induces, after 5FC treatment of the animal, a local production of 5-fluorouracil resulting in intratumor chemotherapy. Destruction of the gene-modified tumor is then followed by the triggering of an antitumor immune reaction resulting in the regression of distant wild-type metastasis. The global effects of 5FC on colorectal adenocarcinoma cells expressing the *CD* gene were analyzed using the proteomic method. Application of 5FC induced apoptosis and 19 proteins showed a significant change in 5FC-treated cells compared with control cells. The up-regulated and down-regulated proteins include cytoskeletal proteins, chaperones, and proteins involved in protein synthesis, the antioxidative network, and detoxification. Most of these proteins are involved in resistance to anticancer drugs and resistance to apoptosis. In addition, we show that the heat shock protein Hsp90 β is phosphorylated on serine 254 upon 5FC treatment. Our results suggest that activation of Hsp90 β by phosphorylation might contribute to tumor regression and tumor immunogenicity. Our findings bring new insights into the mechanism of the anticancer effects induced by CD/5FC treatment. [Mol Cancer Ther 2007; 6(10):2747–56]

Received 1/17/07; revised 6/11/07; accepted 9/4/07.

Grant support: Institut National de la Santé et de la Recherche Médicale.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked advertisement in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Requests for reprints: Christian Baudoin, Faculté de Médecine, Unité INSERM U638, Université de Nice Sophia-Antipolis, Avenue de Valombrose, 06107 Nice Cedex 2, France. Phone: 33-49337-7704; Fax: 33-49381-9456. E-mail: baudoin@unice.fr

Copyright © 2007 American Association for Cancer Research.

doi:10.1158/1535-7163.MCT-07-0040

Introduction

Suicide gene therapy consists of the transfer into tumor cells of a “killer gene,” which converts a nontoxic compound into a lethal drug. This nontoxic compound can be systemically administered without any side effect. This strategy has proven to be effective in several tumor models in animals (1, 2). The bacterial cytosine deaminase (*CD*) gene, which is not expressed in eukaryotic cells, encodes a protein capable of converting cytosine into uracil. When expressed in mammalian cells, this protein transforms the nontoxic antifungal agent 5-fluorocytosine (5FC) into the widely used chemotherapeutic drug 5-fluorouracil (5FU), and expression of the *CD* gene within cancer cells results in a tumor-targeted chemotherapy. Although bypassing the toxic side effects of a systemic chemotherapy is in itself an interesting topic, the key effects of this strategy lie in the existence of two associated bystander effects. The first one, called the local bystander effect, is known to induce a complete tumor regression even if only a small proportion of the tumor cells express the suicide gene (3, 4). This effect is due to the toxicity of 5FU in cells expressing the suicide gene and to the diffusion of 5FU to the neighboring cells. The second effect, the distant bystander effect, results in the regression of distant wild-type tumors, mediated by the triggering of an immune reaction after the destruction of the tumor expressing the suicide gene (5–8).

In previous studies, we have developed a CD/5FC-based gene therapy model against colon cancer (9). A plasmid vector containing the *Escherichia coli CD* gene has been introduced into a BDIX rat colon carcinoma cell line (PROb). Intrahepatic injection of the modified cells (PRObCD) into the liver of syngeneic animals generates a single experimental “suicide tumor.” Treatment of animals with 5FC induced the regression of CD-expressing tumors and rendered the treated animals resistant to challenge with wild-type tumor cells (PROb). Immunohistologic analysis of experimental tumors indicates that natural killer (NK) cells are the major immune components involved in this antitumor effect (7–9).

To dissect the mechanisms involved in the regression of the tumor in our gene therapy model, we have examined changes in protein expression that occurred during cell death in the PRObCD cells exposed to 5FC *in vitro*. Using Annexin V staining, we first show that 5FC induced apoptosis in CD-expressing cells. Proteomic analysis of 5FC-treated cells identified several proteins involved in protein chaperoning, detoxification, and the antioxidative network. Attentive examination of the function of the

identified proteins revealed that most of them are involved in the resistance to anticancer drugs and apoptosis. In addition, Hsp60 and Hsp90 β , which have been described to be involved in the activation of NK cells by stressed apoptotic tumoral cells and in antigen presentation are, respectively, up-regulated and phosphorylated on serine 254 (10–14). Our data corroborate previous works indicating that the functions of Hsp90 are regulated by reversible posttranslational modifications (15, 16) and suggest that activation of Hsp90 β by phosphorylation contributes to tumor regression and activation of the immune system.

Materials and Methods

Cell Culture

DHD/K12/PROb (PROb) cells are a colon carcinoma cell line originating from a chemically induced colon cancer in BDIX rats (17). These cells are poorly immunogenic and induce progressive and metastatic tumors in syngeneic hosts. The cells were maintained in DMEM (Bio Whittaker) supplemented with 10% FCS. PRObCD cells, generated by the transfection of pCD β geo plasmid expressing the CD gene (8) in the PROb cell line, were maintained in the same medium in the presence of G418 at 100 μ g/mL. 5FC was dissolved in DMEM (15 mg/mL) and added in the culture medium at a concentration of 500 μ g/mL.

Annexin V-FITC Assay

Following 12, 24, 48, and 72 h of 5FC treatment, culture media containing any floating cells were collected from the drug-treated or control cells and centrifuged. The remaining monolayer cells were detached using trypsin-EDTA. The floating and adherent cells were washed once with cold PBS and centrifuged. The two cell pellets (adherent and floating cells) were stained with FITC-labeled Annexin V (AnV) and propidium iodide (PI) according to the recommendations of the manufacturer. Ten thousand events were collected using a FACScan flow cytometer (Becton Dickinson). Flow cytometry data were analyzed using Cell Quest software. The distribution of cells was placed into the following three groups: viable (AnV–/PI–), apoptotic (AnV+/PI– and AnV+/PI+), and necrotic (AnV–/PI+).

Reagents

5FC was from ICN. The following antibodies were used: clone 4G10 antiphosphotyrosine antibody (Upstate Biotechnology), clone PY20 antiphosphotyrosine antibody (Zymed Laboratories), clone SPA-845 anti-Hsp90 monoclonal antibody (StressGen Biotechnologies), clone Ac-K-103 antiacetylated lysine monoclonal antibody (Cell Signalling Technology), antiphosphoserine polyclonal antibodies (Zymed Laboratories), and goat anti-mouse and goat anti-rabbit polyclonal antibodies conjugated to horseradish peroxidase (Dako). Trypsin, proteinase K, and endoprotease GluC were from Promega.

Western Blot Analysis

Following two-dimensional gel electrophoresis, proteins were transferred to polyvinylidene difluoride membranes

(Millipore Corporation). Immediately after transfer, membranes were soaked in amido black (0.1% naphthol blue black, 10% methanol, 2% acetic acid), washed in 50% methanol and 7% acetic acid, and then scanned. Membranes were blocked with 5% (w/v) bovine serum albumin in TBST [20 mmol/L Tris-HCl (pH 7.5), 150 mmol/L NaCl, and 0.1% Tween 20] for 1 h at room temperature and incubated with primary antibodies overnight at 4°C. The primary antibody concentrations used were 1:1,000 for antiphosphotyrosine clone 4G10 and clone PY20, 1:1,000 for the monoclonal antiacetylated lysine, 1:250 for the polyclonal antiphosphoserine, and 1:1,000 for the monoclonal anti-Hsp90. The membranes were washed thrice for 5 min with TBST and incubated in TBST containing the appropriate secondary antibody conjugated to peroxidase for 1 h at room temperature. The membranes were washed as described above, and the labeled proteins were detected using an enhanced chemiluminescence kit (Amersham International) according to the manufacturer's instructions. Detection was done with autoradiography films (BioMax Light Film; Eastman Kodak).

Sample Preparation for Proteome Analysis

Treated and nontreated cells were rinsed thrice with an ice-cold sucrose solution [250 mmol/L sucrose, 10 mmol/L Hepes (pH 7.5), 0.2 mmol/L CaCl₂]. Rehydration buffer (urea 7 mol/L, thiourea 2 mol/L, CHAPS 4%, and Triton X-100 0.24%) was added to the cell monolayer and incubated for 15 min at room temperature. To remove genomic DNA, spermidine (10 mmol/L) was added for an additional 15 min. After centrifugation (22,000 \times g, 30 min, 15°C), the supernatant and the pellet were separated. The supernatant was precipitated with the Ready Prep 2D clean-up kit (Bio-Rad Laboratories). Proteins were resuspended with rehydration buffer and the protein concentration was determined by the Bradford assay (Bio-Rad Laboratories) using bovine serum albumin in rehydration buffer as standard curve.

Two-Dimensional Gel Electrophoresis and Protein Staining

Each sample was thawed, adjusted to a protein concentration of 2.3 mg/mL, and supplemented with 100 mmol/L of 2-hydroxyethyl disulfide (Acros Organics). The samples (700 μ g) were loaded onto IPG ReadyStrips (pH 3–10; 17 cm, linear gradient; Bio-Rad), and rehydrated for 7 h at 20°C. Isoelectric focusing on the Protean IEF Cell from Bio-Rad was carried out as follows: 15 min at 250 V, a slow (15 h) voltage ramping to 10,000 V and a final step at 10,000 V up to 500,000 Vh. Immediately after focusing, immobilized pH gradient strips were wrapped in plastic trays (rehydration trays; Bio-Rad) and stored at –80°C. Prior to SDS-PAGE, proteins in the immobilized pH gradient were reduced in 4% SDS in 1.5 mol/L Tris (pH 8.8) containing 130 mmol/L of DTT and 30% glycerol for 15 min, alkylated in 4% SDS in 1.5 mol/L of Tris (pH 8.8) containing 130 mmol/L of iodoacetamide and 30% glycerol for 20 min. Finally, strips were equilibrated in 4% SDS, 1 mol/L of Tris (pH 6.8) and 30% glycerol for 90 min. The second dimension separation was run overnight at room

temperature on a 1-mm-thick 10% polyacrylamide gel at a constant voltage up to 1,300 Vh (Protean II XL cell). Immediately after electrophoresis, gels were fixed in ethanol/phosphoric acid (40%:3%) for 2 h, rinsed with deionized water thrice for 30 min, and then stained for 16 h with Bio-Safe Coomassie Stain (Bio-Rad). Gels were scanned using a GS-800 calibrated densitometer (Bio-Rad). Gel patterns were compared using ImageMaster 2D Platinum (version 5.0, Amersham Bioscience). Relative quantification of the detected spots was made as a percentage of the total spot volume for every gel (% vol). After these normalizations, based on the total quantity of valid spots on each gel, heuristic clustering was used to blindly classify similar gels in order to confirm the two classes (control and 5FC-treated cells). Lastly, statistical differences in expression between spots from control and 5FC-treated cells were assessed through Kolmogorov and Wilcoxon tests. The selected spots were subsequently treated for identification by mass spectrometry.

In-gel Tryptic Digestion and Protein Identification by Mass Spectrometry

The protein spots were manually excised from the two-dimensional gel and digested according to a standard protocol. Briefly, after a washing step with 50% ethanol, proteins were reduced and alkylated with 10 mmol/L of DTT and 50 mmol/L of iodoacetic acid, respectively. The spots were washed twice with 50% ethanol, shrunk with acetonitrile and then subsequently digested with 100 ng of enzyme in 50 mmol/L of ammonium bicarbonate and 15% acetonitrile. The peptides were extracted with 0.5% trifluoroacetic acid, 50% acetonitrile, and 100% acetonitrile. Matrix-assisted laser desorption ionization-time of flight mass spectrometry used a Voyager DE-PRO (PerSeptive Biosystems) set in positive reflectron mode. Samples were spotted on a thin layer of nitrocellulose (5 mg/mL solution in acetone/isopropanol, 1:1), washed with 2 μ L of cold water, and then overlaid with α -cyano-4-hydroxycinnamic acid (5 mg/mL solution in 50% acetonitrile and 0.3% trifluoroacetic acid). Spectra were obtained with standard settings and external calibration. Protein identification was done with Aldente software on the SwissProt web site.⁵ For the posttranslational modification analysis of Hsp90, LC-MS/MS mass spectrometry was done with a Surveyor system coupled with a LCQ DECA XP ion trap mass spectrometer (ThermoQuest). Peptides were separated on a C18 column (200 μ m \times 10 cm, Hypersil; ThermoQuest) at 5 μ L/min with an acetonitrile gradient from 1% to 40% acetonitrile in 0.1% formic acid. A 3 kV voltage was applied on the micro-ESI needle. The automatic acquisition was set as previously described (18). Data processing was done with Bioworks 3.3. The main search variables were methionine oxidation, serine, threonine, and tyrosine phosphorylation or lysine acetylation as differential modification, no enzyme as cleavage specificity. Peptide

identification was considered significant when $X_{corr} > 1.7$, 2.2, and 3.3 for mono-, di- and tri-charged peptides, respectively; $\delta Cn > 0.1$, $Resp < 3$, and peptide probability $< 5E-3$.

Results and Discussion

Before performing the proteomic studies, we verified that 5FC induced apoptosis in our CD-expressing colorectal adenocarcinoma cell line in a time-dependent manner. 5FC (500 μ g/mL) was added to actively growing PRObCD cells. As shown in Fig. 1A, the number of PRObCD cells rapidly declined and 25% of the cells remained attached to the plastic vessel after 72 h of treatment, whereas nontreated PRObCD cells proliferated rapidly. To assess whether 5FC induced apoptosis, PRObCD-detached cells were stained with Annexin V and propidium iodide. Annexin V binds to phosphatidylserine, which shifts to the cell surface from the inside of the cell membrane during the early stage of apoptosis, whereas propidium iodide penetrates only in cells which are in the late stage of apoptosis and intercalates into genomic DNA (19). As shown in Fig. 1B, ~90% of the detached cells were AnV+, and 10% were AnV+PI+ at 12, 24, 48, and 72 h of treatment, which shows that 5FC induces apoptosis in CD-expressing cells. The number of detached PRObCD cells without treatment was not significant (<0.01% compared with treated PRObCD cells).

To determine the apoptosis stage of cells which were still attached after 5FC treatment, we stained cells with Annexin V and propidium iodide. We detected three populations in the treated and untreated PRObCD cells: AnV+, AnV+PI+, and PI+, which corresponded to the early stage of apoptosis, late stage of apoptosis, and necrosis, respectively. Unstained cells were considered as living cells. As shown in Fig. 1D, AnV+ cells accounted for 23% and 16% of total cells in the presence and in the absence of 5FC, respectively. The absence of staining in 80% of cells and the small difference between treated and untreated cells for AnV+ and AnV+PI+ stainings suggest that detachment of treated PRObCD cells from the plastic vessel precedes the appearance of phosphatidyl serine in the outer leaflet of the plasma membrane. Finally, we show that 5FC, after being transformed into 5FU by CD, induces apoptosis in the colon adenocarcinoma cell line PRObCD.

We then analyzed PRObCD's proteome modification which was induced after 48 h of 5FC treatment. Two-dimensional gel electrophoresis of total proteins extracted from treated and untreated PRObCD cells was carried out to identify differentially expressed proteins.

Two additional controls were done: (a) PROb cells were treated with 5FC. We did not observe any cellular toxicity, apoptosis, and modification of the proteome (data not shown), and (b) PRObCD cells were treated with 5FU for 48 h. This treatment induced the apoptosis of PRObCD cells and modified the proteome in a similar extent compared with CD/5FC. This result suggests that the two killing mechanisms (i.e., CD/5FC and 5FU) are the same.

⁵ <http://www.expasy.org/tools/aldente/>

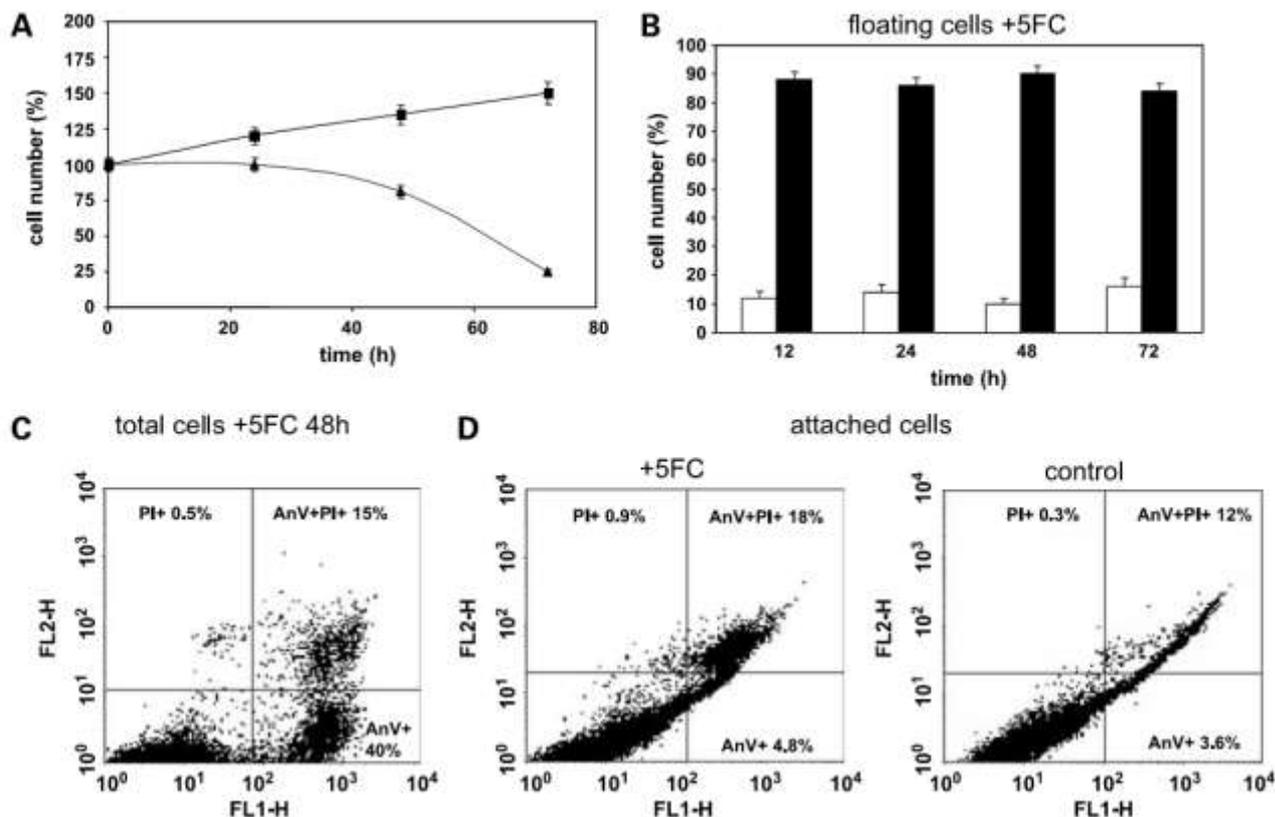


Figure 1. Distribution of cell viability during 5FC exposure. **A**, actively growing PRObCD cells were exposed (▲), or not exposed (■) to 5FC. Attached cells were counted after 0, 24, 48, and 72 h of treatment. The number of treated PRObCD cells rapidly declined and <25% of cells remained attached to the plastic vessel after 72 h of treatment. PRObCD cells were treated for 12, 24, 48, and 72 h with 5FC. Floating cells (**B**) and total cells (**C**) were removed from culture and stained with Annexin V (AnV) and propidium iodide (PI). Apoptosis was determined by flow cytometry as described in Materials and Methods. Filled columns, AnV+ cells; open columns, AnV+PI+ cells ($n = 3$). **D**, PRObCD cells were treated or not for 48 h with 5FC and attached cells were detached from the plastic vessel with trypsin and stained with Annexin V and propidium iodide. Numbers, mean percentages of early apoptotic (bottom right), late apoptotic (top right), and necrotic (top left) cells ($n = 3$).

However, only major proteins can be observed and analyzed on two-dimensional gels. We do not exclude that minor differences occurred between the two treatments.

A representative Coomassie blue-stained gel of proteins from 5FC-treated PRObCD cells is depicted in Fig. 2A. Replicates were carried out for both treated (5FC, $n = 3$) and nontreated cells (CT, $n = 2$). Only reproducible spots were considered as significant. Image analysis with ImageMaster Platinum identified 353 spot groups (i.e., the same spot on different gels) and Kolmogorov/Wilcoxon tests were used to discriminate the up-regulated and down-regulated proteins. Mass spectrometry analysis identified the corresponding proteins. Fourteen up-regulated proteins, six down-regulated proteins, and one protein exhibiting a significant shift in its isoelectric point were reproducibly detected from treated PRObCD cells compared with nontreated cells.

Matrix-assisted laser desorption ionization-time of flight mass spectrometry analysis of the tryptic fragments and a data search allowed for the identification of these proteins, as shown in Table 1. The up-regulated and down-

regulated proteins include cytoskeletal proteins, chaperone proteins, and proteins involved in protein synthesis, protein folding, the antioxidative network, and detoxification. The increase of proteins involved in the antioxidative network and detoxification are likely due to the conversion of 5FC into 5FU by the CD. PRObCD cells need to eliminate 5FU and its metabolic derivatives, FdUMP, 5-FUTP, and 5-FdUTP, which are responsible for the cytotoxic effects in cancer cells (20). The thioredoxin-like 2 protein is involved in the mitochondrial thioredoxin system, which controls redox balance, cell growth, and apoptosis. When thioredoxin levels are elevated, there is increased cell growth and resistance to the normal mechanism of programmed cell death (21). In addition, an increase in the thioredoxin-like 2 protein level seen in many human primary cancers compared with normal tissue seems to contribute to increased cancer cell growth and resistance to chemotherapy (22). The mitochondrial aldehyde dehydrogenase has been implicated in drug catabolism and has been identified as a major actor in the resistance to the anticancer molecule cyclophosphamide.

This has been shown in the medulloblastoma cell line (23), and particularly, in leukemic cells in which an elevated level of aldehyde dehydrogenase is found (24). The genetic variability of aldehyde dehydrogenase is also linked to resistance in cyclophosphamide (25). The dihydropyrimidinase-related protein 2 belongs to the dihydropyrimidinase family, which is involved in the pyrimidine metabolism and in the catabolism of 5FU (26). Patients with a partial dihydropyrimidine deficiency proved to be at risk of developing severe toxicity after 5FU administration (27).

Besides proteins which are directly involved in detoxification, we also found that three cytoskeletal proteins involved in cell survival and drug resistance were increased. α -Actinin 4 binds to the actin cytoskeleton and integrins whose roles in matrix adhesion induced-cell survival are largely documented (28). α -Actinin is also up-regulated by many cytotoxic chemicals in embryonic stem cells and is now used as a biomarker of toxicity in *in vitro* tests (29). In addition, α -actinin is up-regulated in osteo-

sarcoma cell lines (30). Ezrin is a cytoskeletal protein which has a major role in cell polarization through actin binding and the role of ezrin in metastasis behavior has been well documented (31). Ezrin binds to Fas and contributes to the Fas-related resistance of tumors to apoptosis (32). Furthermore, ezrin is connected to P-glycoprotein, which regulates the efflux pumps that are responsible for some of the multidrug resistance mechanisms of tumors (33). The glial fibrillary acidic protein which belongs to the intermediate filament family also binds to the P-glycoprotein, and therefore, may have a role in 5FU resistance (34).

We also found that three proteins involved in protein synthesis were up-regulated after 5FC treatment. These proteins (eIF2 β , eIF4I, and eIF1G) belong to the initiation and elongation factor families. They are not just housekeeping proteins. For instance, eIF4E, which is structurally very close to eIFA1 and eIF2 β , plays an important regulatory role in the control of cellular

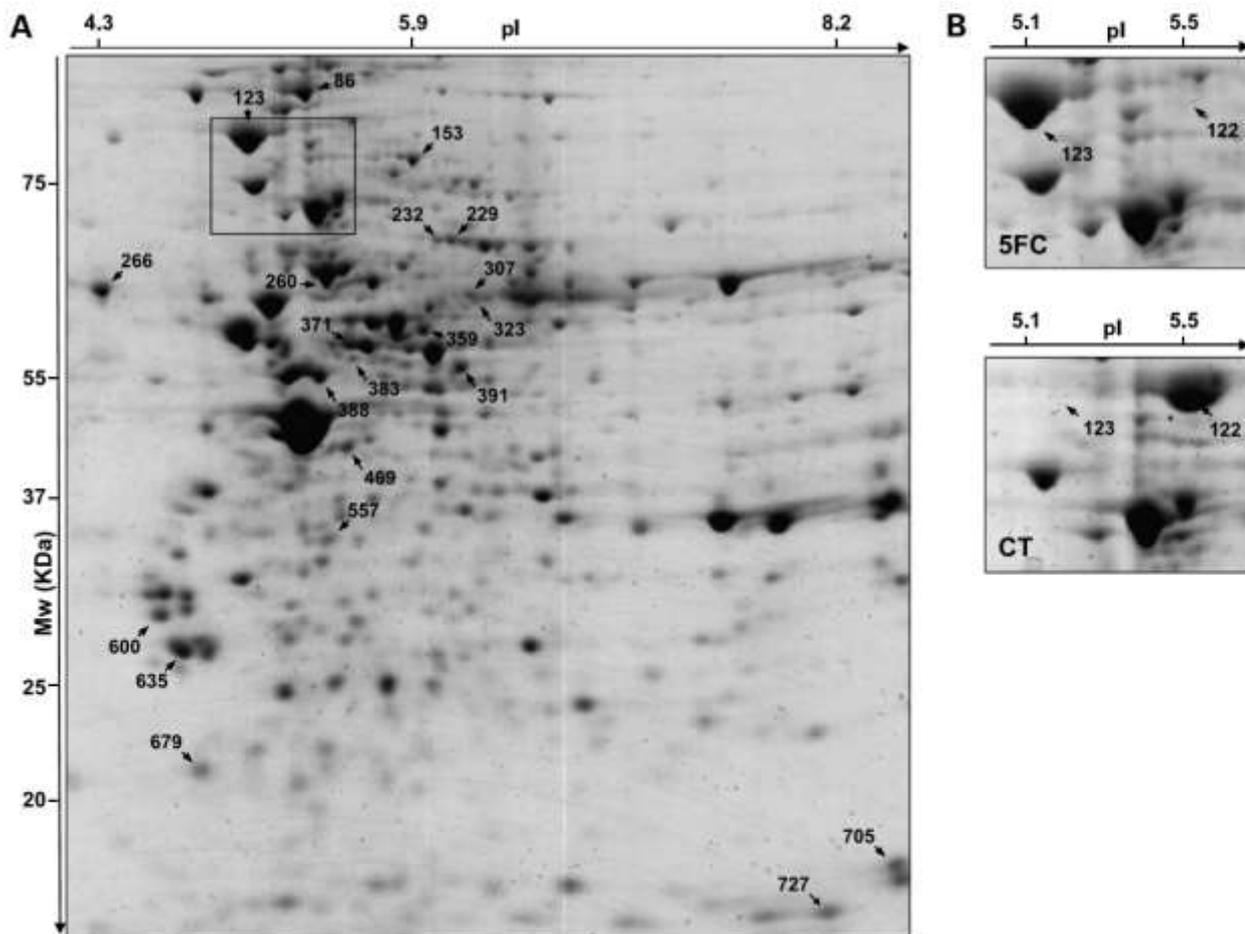


Figure 2. Proteome of 5FC-treated colon adenocarcinoma cells. **A**, two-dimensional PAGE of PRObCD cells treated with 5FC. Gels were done with 700 μ g of total protein extract loaded onto IPG ReadyStrips (pH 3–10). After migration, gels were stained with G250 Coomassie blue. The labeled spots are listed in Table 1. **B**, a section of the two-dimensional gel containing the protein Hsp90 β . Comparison of 5FC-treated cells with control cells shows the shift of the isoelectric point of the protein from 5.5 to 5.1. In the CT condition, Hsp90 β is identified in spot 122 at pH 5.5. After 5FC treatment, Hsp90 β is identified in spot 123 at pH 5.1.

Table 1. Up-regulated and down-regulated protein spots by treatment with 5FC in PRObCD cells

Spot no.*	Theoretical pI/MW [†]	Experimental pI/MW [‡]	P [§]	Accession no.¶	Description	Volume 5FC [¶]	Volume CT	Ratio FC/CT
Chaperones								
122	5/83	5.5/89	7.E-12	P08238	Hsp90 β (basic spot)	—	2.30	
123	5/83	5.1/89	1.E-11	P08238	Hsp90 β (acidic spot)	2.26	—	
260	5.3/58	5.4/61	4E-12	P63039	Hsp60	1.77	1.11	1.60
600	4.6/29	4.59/30	4.E-10	P62260	14-3-3 protein ϵ	0.71	0.39	1.83
266	4.3/46	4.3/63	4.E-08	P18418	Calreticulin	0.57	0.36	1.58
679	4.8/19	4.8/21	9.E-10	P63029	Translationally controlled tumor protein	0.43	0.31	1.40
Protein folding								
229	6.3/61	6.1/66	4.E-12	P80318	T-complex protein 1 subunit γ	0.18	0.30	0.61
727	8.3/18	7.9/15	8.E-11	Q5BK98	Peptidylprolyl isomerase A	0.36	0.68	0.53
Cytoskeletal proteins								
557	5.3/33	5.4/35	4.E-14	Q5RKN9	Capping protein (actin filament) muscle	0.30	0.12	2.44
86	5.2/105	5.3/102	1.E-19	P57780	α -Actinin-4 (basic spot)	1.19	0.53	2.27
371	5.4/50	5.5/53	5.E-07	P14136	Glial fibrillary acidic protein, astrocytes	0.34	0.21	1.65
153	5.9/69	5.1/89	2.E-14	P15311	Ezrin	0.27	0.37	0.74
705	8.2/19	8.2/17	3.E-10	Q544Y7	Cofilin 1	0.50	0.72	0.69
Metabolism								
307	6.3/56	6.11/59	1.E-11	O08651	D-3-phosphoglycerate dehydrogenase	0.15	0.24	0.62
Detoxification								
359	5.8/54	5.9/54	3.E-13	P81178	Aldehyde dehydrogenase	0.37	0.24	1.57
232	6/62	6/66	2.E-13	P47942	Dihydropyrimidinase-related protein 2	0.15	0.21	0.70
Protein synthesis and degradation								
323	6.6/53	6.18/58	4.E-10	Q9CPY7	Cytosol aminopeptidase	0.05	—	
383	5.6/38	5.5/52	2.E-12	Q99L45	eIF2 β	0.12	0.06	2.11
391	6.3/50	6.1/50	2.E-10	Q4FZZ5	eEF1G	0.40	0.30	1.33
388	5.3/46	5.3/50	3.E-12	P60843	eIF4A-I	0.87	0.66	1.32
635	4.7/26	4.7/28	3.E-13	Q9Z2U1	Proteasome subunit α type 5	0.19	0.16	1.17
Protein related to antioxidative network								
469	5.5/38	5.5/43	1.E-07	Q9JLZ1	Thioredoxin-like 2 protein	0.27	0.16	1.73

The proteins are classified according to their function. Data are presented as mean of normalized spot volumes from 5FC and nontreated (CT) cells.

*Spot no. refers to the corresponding spot number in Fig. 2.

[†]Theoretical pI/MW were obtained from SwissProt-computed pI/MW.

[‡]Experimental pI/MW were calculated by ImageMaster.

[§]P value is the probability value calculated by the Aldente peptide mass fingerprinting tool.

[¶]Description corresponds to the protein name given by SwissProt database.

[¶]Volume was expressed as a percentage of total spot volume.

stress-induced apoptosis (35). Similarly, eEF1A2, which is structurally close to eEF1G, is antiapoptotic and has been identified as a tumorigenic oncogene regulating cell growth (36).

In addition, we found four up-regulated proteins which belong to the huge family of chaperone proteins. Chaperones limit aggregation, facilitate protein refolding, and are involved in signal transduction. Under conditions of cellular stress such as chemotherapy, intracellular chaperone levels increase in order to provide cellular protection and maintain homeostasis (37). The 14-3-3 ϵ protein belongs to a highly conserved cellular protein family that can interact with >200 target proteins. 14-3-3 proteins control cell cycle, cell growth, survival, and apoptosis (38, 39). In addition, 14-3-3 σ has been identified as a marker of potential clinical interest for breast cancer detection and contributes to drug resistance in human breast cancer cells (40). We also observed that two chaperones, Hsp60 and Hsp90 β , which are heat shock proteins (37), are up-

regulated and subject to posttranslational modification, respectively. These two proteins have roles in protein chaperoning and in signal transduction during tumorigenesis (41). In addition, it has been described that Hsp60 and Hsp90 have roles in tumor immunogenicity and antigen presentation, respectively (10–12), and thus, might contribute to the activation of the NK cells that we observed in our gene therapy model.

Hsp90 β was further analyzed in order to identify the posttranslational modification responsible for the acidification of the isoelectric point which shifted from 5.5 to 5.1 (Fig. 2B). Two major posttranslational modifications of Hsp90 have been described: phosphorylation and lysine acetylation. It has been shown that phosphorylation of Hsp90 on serine was linked to its chaperoning function. Serine phosphorylation triggers the release of the chaperone from the target protein (16). In addition, Hsp90 is phosphorylated in tyrosines during mammalian sperm capacitation, the process by which spermatozoa gain the

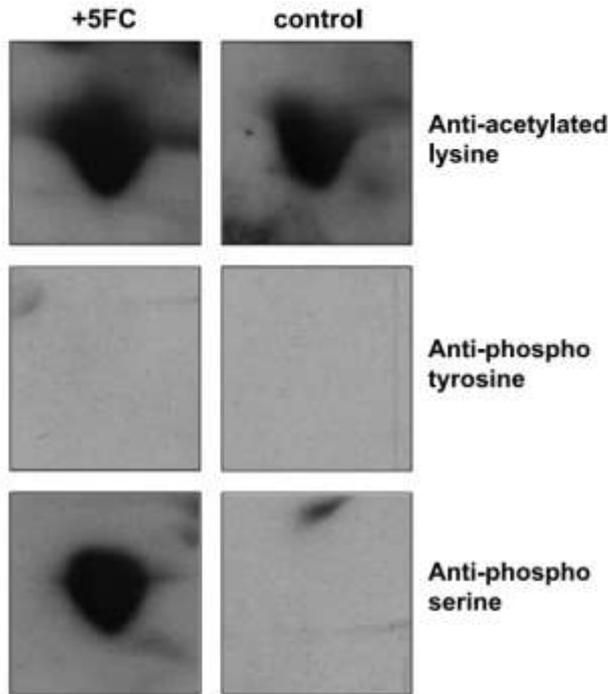


Figure 3. Western blot analysis of Hsp90 β posttranslational modifications. Protein extracts from 5FC-treated or nontreated PRObCD cells were separated with two-dimensional gel electrophoresis and transferred to polyvinylidene difluoride membranes. Areas corresponding to Hsp90 β spots were hybridized with antiacetylated lysine, antiphosphoserine, and antiphosphotyrosine antibodies. Forty-eight hours of 5FC treatment induces serine but not tyrosine phosphorylation of Hsp90 β . In contrast, Hsp90 β is acetylated in 5FC-treated and nontreated PRObCD cells.

ability to fertilize the oocyte (42). Finally, acetylation of Hsp90 regulates chaperone-dependent activation of the glucocorticoid receptor, lysine acetylation of Hsp90, leading to the loss of chaperone activity (15).

We did Western blot analyses on two-dimensional gel electrophoresis for the detection of phosphorylation and acetylation of Hsp90 β . After transfer to a polyvinylidene difluoride membrane, the area corresponding to Hsp90 β spots were hybridized with antibodies specific for phosphotyrosine, phosphoserine, and acetylated lysine. The

membranes were also hybridized with anti-Hsp90 monoclonal antibodies as controls (data not shown). As shown in Fig. 3, we did not detect any signal for phosphorylated tyrosine on Hsp90 β from 5FC-treated and nontreated cell extracts. Hybridizations were done with two antiphosphotyrosine monoclonal antibodies, clones 4G10 and PY20, which are the most widely used antiphosphotyrosine antibodies. In contrast, phosphorylation of the acidic form of Hsp90 β was detected with antiphosphoserine antibodies (Fig. 3). Finally, hybridization with antiacetylated lysine antibodies revealed the acetylation of the basic and acidic forms of Hsp90 β . In conclusion, we show that 5FC treatment induces serine phosphorylation of Hsp90 β and that Hsp90 β is constitutively acetylated in PRObCD cells.

A large proteomics survey previously revealed that Hsp90 β is acetylated on lysine 623 and phosphorylated on serines 254 and 260 in the nuclear protein fractions of HeLa cells (43, 44). Phosphorylation of Hsp90 β on serine 254 is found in the mitotic spindle during division of HeLa cells (45), in synaptosomes from human cerebral cortex (46), and in the inner medulla of rat kidneys (47). In addition, Hsp90 β is phosphorylated *in vitro* by casein kinase II on serines 254 and 225 (48). Finally, Hsp90 β is phosphorylated on tyrosine 483 in Jurkat cells (49). Complementary mass spectrometry analysis with LC-MS/MS (ion trap) was done to determine the acetylation and phosphorylation sites of Hsp90 β . In order to increase the coverage of the protein sequence, three different digestions were done: trypsin, gluC, and proteinase K. This allowed for the identification of 71% of the protein sequence (Fig. 4). Tandem mass spectrometry identified a phosphorylation on serine 254 and showed that serine 260 was not modified (Fig. 5). Tyrosine 483 was also identified as unmodified, confirming the negative results obtained with antiphosphotyrosine antibodies.

Regarding acetylation, Western blot analysis showed that the two forms of Hsp90 β were acetylated. One acetylation has been previously described on lysine 623 but our MS/MS data clearly shows that this lysine was not modified. Two different peptides containing lysine 623 without acetylation were identified after trypsin and proteinase K digestion. The absence of detection of acetylated lysine with MS/MS, despite the positive signal obtained in

1	<u>PEEVHGGEEVETFAFQAEIAQLMSLIINTFYSNKEIFLRELISNASDALDKIRYESLTDPSKLDGKELKIDIIIPNPQERTLTLVDVTGIGMTKADLNN</u>	100
101	<u>LGTIAKSGTKAFMEALQAGADISMIQQFVGVGYFAYSAYLVAEKVVVITKHNDDEQYAWESSAGGSFTVRAHGEPIGRGTVILHLKEDQTEYLEERRVKEV</u>	200
201	VKK <u>HSQFIGYPITL</u> YLEKEREK <u>IS</u> DDAEAEKGEKE <u>EDKED</u> EEKPRI <u>EDVGS</u> DEEDD <u>SG</u> DKKKKTKKIKEKYIDQEELNKTPIWTR <u>MPDDITQEEY</u>	300
301	GEFYKSLTNDWEDHLAVKHFSVEGQLEFRALLFIPRAPPDLFENKKNKNIKLYVRRVPIIMSDCELIPEYLNPIRGVVDS <u>EDLPLNISREMLQOSKIL</u>	400
401	KVIRKNIIVKRCLEL <u>FS</u> EL <u>LA</u> EDKENYK <u>KFF</u> YAFPSKNLKLGIHEDSTNERRLSELLEFRYHTSQSGDEMTSLSEYVSRMKETQKSI <u>IT</u> ITGESKEQVANS <u>AFVE</u>	500
501	<u>RVRKRGFEVVMTEPIDEYCVQQLKEFDGKSLVSVTK<u>EGLELPEDEEKKK</u>MESKAKFENLCKLMKEILDKKVEKVTISNRLVSSPCCIVTSTYGTWTAN</u>	600
601	MERIMKAQALRDNSTMGYMAK <u>HL</u> LEINPDHPIVETLRQKAEADKNDKAVKDLVLLPETALLSSGFSLEDPQTHSNRIYRMIKLGLGIDEDVTAERP	700
701	<u>AAVPDEIPPLEGGEDASRMREVD</u>	

Figure 4. Sequence recovery of Hsp90 β using LC-MS/MS. Hsp90 β , gi|51243733|, was identified with 71% coverage after proteolysis with trypsin (underlined), proteinase K, or gluC (**boldface**). Open boxes, amino acids with posttranslational modifications as described in the literature (i.e., phosphorylation on serines 225, 254, and 260, tyrosine 483, and acetylation on lysine 623). pS, the phosphorylated serine 254 identified in Fig. 5.

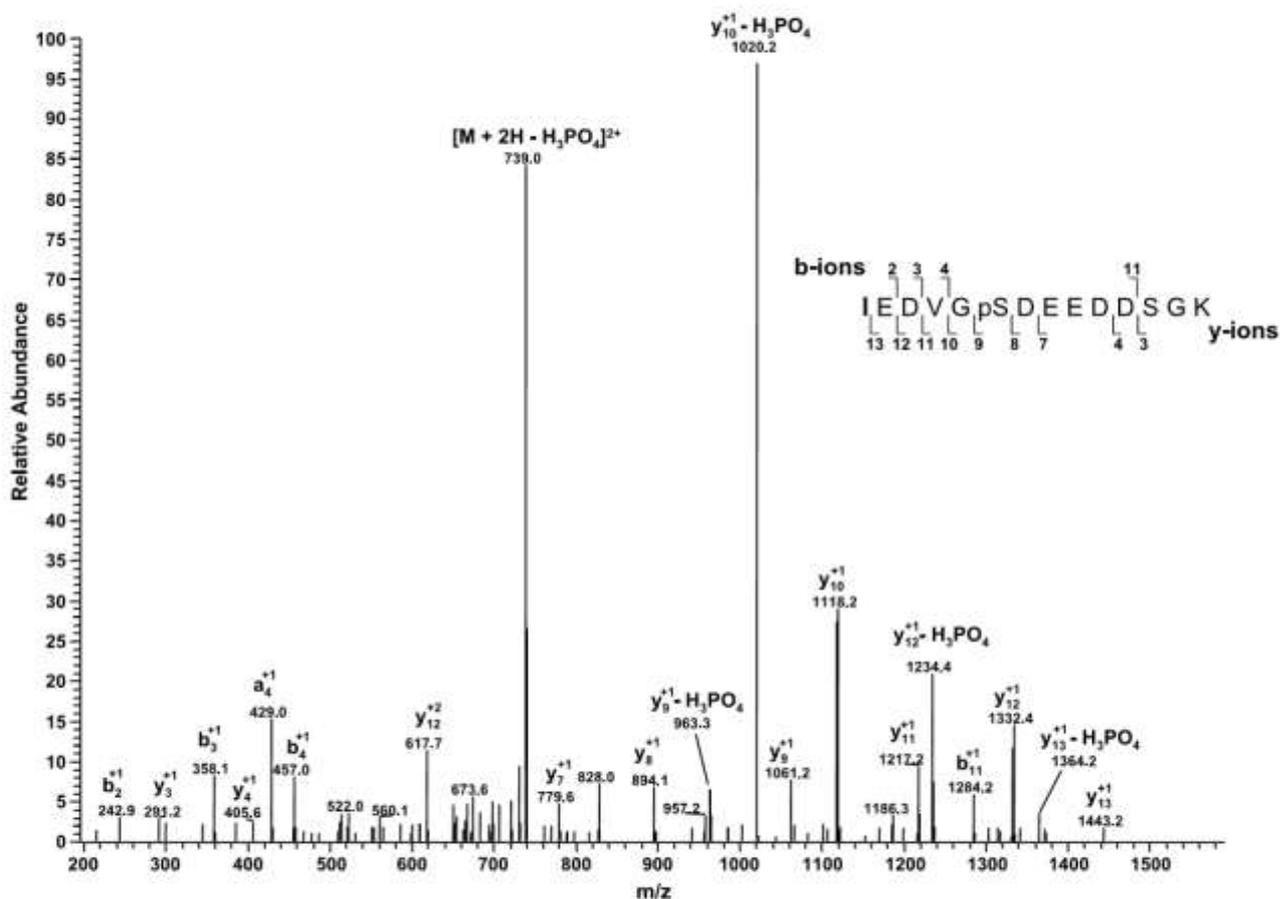


Figure 5. MS/MS spectrum of the doubly charged precursor ion $(M + 2H)^{2+}$ at m/z 787.8 corresponding to the phosphopeptide IEDVGP SDEEDDSGK. The phosphorylation site is unambiguously identified at serine on position 6 (serine 254) with a clear γ -ions serial from γ_{13} to γ_7 . Serine on position 12 (serine 260) is not phosphorylated.

Western blot analysis, suggests that acetylation was uncovered with mass spectroscopy (Fig. 4).

Finally, serine 225 was not covered with our MS/MS analysis. However, it is possible that this serine is also phosphorylated during 5FC treatment. We cannot exclude that in addition to phosphorylation on serine 254, one additional serine phosphorylation occurred on Hsp90 β during 5FC treatment for the following reasons: (a) the pH shift that we found for Hsp90 β is 0.4 pH units, whereas one phosphorylation usually induces an isoelectric point shift of 0.2 pH units on two-dimensional gels (50). (b) It has been shown that three phosphorylations on Hsp27 and one phosphorylation on Hsp70 induce an isoelectric point shift of 0.6 and 0.15 pH units, respectively (50).

The assessment of protein expression profiles by proteomic analysis after treatment with chemotherapeutic agents has the potential to identify novel signaling pathways involved in mediating the downstream response to these therapies, and could greatly facilitate the discovery of novel potential therapeutic targets and/or markers of chemoresistance. Our results show that treatment of PRObCD cells with 5FC induces the up-regulation and down-

regulation of several proteins, most of them having a role in resistance to anticancer drugs and apoptosis. In this context, determination of the expression level of these proteins in different tumors might be of interest. It is possible that one of these proteins will constitute a very useful tool to predict resistance to CD/5FC and/or 5FU treatments.

In addition, we show the phosphorylation of Hsp90 β on serine 254 during 5FC treatment. This protein may have a role in apoptosis which occurred in our gene therapy model. In this light, it has been described that Hsp90 has roles in protein chaperoning and resistance to chemotherapeutic drug-induced apoptosis (41). The resistance to apoptosis is mediated via the binding of Hsp90 to prosurvival client proteins such as Bcr-ABL, c-Raf1, and AKT, which promote their proper folding, assembly, and transportation across different cellular compartments. Failure of Hsp90 activity leads to misfolding of client proteins, which leads to ubiquitination and proteasome degradation (15). Therefore, the chaperone activity of Hsp90 β should be tidily regulated. Although it has been known for some time that Hsp90 is a phosphoprotein, the

significance of this phosphorylation event in the function of Hsp90 is entirely unknown. In this study, we have shown for the first time that Hsp90 β is phosphorylated on serine 254 after anticancer treatment. The phosphorylation of serine 254, very likely, has an important role in the regulation of Hsp90 β activity. It is possible that this phosphorylation induces a modification of the affinity of Hsp90 β toward its client proteins. The phosphorylation of serine 254 might induce the release of client proteins, or inversely, the binding to new client proteins. In this context, it would be interesting to identify proteins bound to Hsp90 β before and after 5FC treatment. Ablation of serine 254 using site-directed mutagenesis would also be interesting to dissect the role of this phosphorylation in its resistance to chemotherapeutic agents.

Besides its role in protein chaperoning and apoptosis, it has been described that Hsp90 has a role in antigen presentation (10–12). Therefore, it is possible that Hsp90 β has a role in NK cell activation, which occurred in our gene therapy model (7–9). Another protein which may have a role in NK cell activation is Hsp60, which has been described to induce tumor immunogenicity (10–12). We have observed that this protein was up-regulated during 5FC treatment. Furthermore, stressed apoptotic tumor cells are capable of providing the necessary danger signals by increasing the surface expression of Hsp60, resulting in the activation/maturation of dendritic cells and the generation of potent antitumor T cell responses (12). Concerning Hsp90, it has been shown that this protein is able to present antigens to T lymphocytes via MHC class II molecules (13). In addition, surface expression of Hsp90 promotes $\gamma\delta$ T cell proliferation in EBV-infected B cells (14). It is possible that phosphorylation of serine 254 occurring during 5FC treatment capacitates Hsp90 β to stimulate NK cells after being released from dying PRObCD cells. Phosphorylation of serine 254 might also capacitate Hsp90 β to present antigenic peptides from apoptotic PRObCD cells to T lymphocytes. To check our assumption, it would be interesting to evaluate the capacity of Hsp90 β purified from 5FC-treated and untreated PRObCD cells to activate dendritic cells.

Acknowledgments

We thank Bernard Rossi, head of the laboratory, who died in May 2006. We also thank Alexandra Charlesworth and Guerrino Meneguzzi for a critical reading of the manuscript.

References

- Lai S, Lauer UM, Niethammer D, Beck JF, Schlegel PG. Suicide genes: past, present and future perspectives. *Immunol Today* 2000;21:48–54.
- Bateman AR, Harrington KJ, Melcher AA, Vile RG. Cancer gene therapy: developments to 2000. *Expert Opin Investig Drugs* 2000;9:2799–813.
- Huber BE, Austin EA, Richards CA, Davis ST, Good SS. Metabolism of 5-fluorocytosine to 5-fluorouracil in human colorectal tumor cells transduced with the cytosine deaminase gene: significant antitumor effects when only a small percentage of tumor cells express cytosine deaminase. *Proc Natl Acad Sci U S A* 1994;91:8302–6.
- Freeman SM, Abboud CN, Whartenby KA, et al. The “bystander effect”: tumor regression when a fraction of the tumor mass is genetically modified. *Cancer Res* 1993;53:5274–83.
- Crittenden MR, Thanarajasingam U, Vile RG, Gough MJ. Intratumoral immunotherapy: using the tumor against itself. *Immunology* 2005;114:11–22.
- Kianmanesh AR, Perrin H, Panis Y, et al. A “distant” bystander effect of suicide gene therapy: regression of nontransduced tumors together with a distant transduced tumor. *Hum Gene Ther* 1997;8:1807–14.
- Pierrefite-Carle V, Baque P, Gavelli A, et al. Regression of experimental liver tumor after distant intra-hepatic injection of cytosine deaminase-expressing tumor cells and 5-fluorocytosine treatment. *Int J Mol Med* 2000;5:275–8.
- Pierrefite-Carle V, Gavelli A, Brossette N, et al. Re: Cytosine deaminase/5-fluorocytosine-based vaccination against liver tumors: evidence of distant bystander effect. *J Natl Cancer Inst* 2000;92:494–5.
- Pierrefite-Carle V, Baque P, Gavelli A, et al. Subcutaneous or intrahepatic injection of suicide gene modified tumour cells induces a systemic antitumour response in a metastatic model of colon carcinoma in rats. *Gut* 2002;50:387–91.
- Multhoff G. Heat shock proteins in immunity. *Handb Exp Pharmacol* 2006;279–304.
- Srivastava P. Roles of heat-shock proteins in innate and adaptive immunity. *Nat Rev Immunol* 2002;2:185–94.
- Feng H, Zeng Y, Graner MW, Katsanis E. Stressed apoptotic tumor cells stimulate dendritic cells and induce specific cytotoxic T cells. *Blood* 2002;100:4108–15.
- Rajagopal D, Bal V, Mayor S, George A, Rath S. A role for the Hsp90 molecular chaperone family in antigen presentation to T lymphocytes via major histocompatibility complex class II molecules. *Eur J Immunol* 2006;36:828–41.
- Kotsioprifitis M, Tanner JE, Alfieri C. Heat shock protein 90 expression in Epstein-Barr virus-infected B cells promotes $\gamma\delta$ T-cell proliferation *in vitro*. *J Virol* 2005;79:7255–61.
- Aoyagi S, Archer TK. Modulating molecular chaperone Hsp90 functions through reversible acetylation. *Trends Cell Biol* 2005;15:565–7.
- Zhao YG, Gilmore R, Leone G, Coffey MC, Weber B, Lee PW. Hsp90 phosphorylation is linked to its chaperoning function. Assembly of the reovirus cell attachment protein. *J Biol Chem* 2001;276:32822–7.
- Martin F, Caignard A, Jeannin JF, Leclerc A, Martin M. Selection by trypsin of two sublines of rat colon cancer cells forming progressive or regressive tumors. *Int J Cancer* 1983;32:623–7.
- Valot B, Negroni L, Zivy M, Gianinazzi S, Dumas-Gaudot E. A mass spectrometric approach to identify arbuscular mycorrhiza-related proteins in root plasma membrane fractions. *Proteomics* 2006;6 Suppl 1: S145–55.
- Castedo M, Hirsch T, Susin SA, et al. Sequential acquisition of mitochondrial and plasma membrane alterations during early lymphocyte apoptosis. *J Immunol* 1996;157:512–21.
- Heidelberger C, Danenberg PV, Moran RG. Fluorinated pyrimidines and their nucleosides. *Adv Enzymol Relat Areas Mol Biol* 1983;54:58–119.
- Patenaude A, Ven Murthy MR, Mirault ME. Mitochondrial thioredoxin system: effects of TrxR2 overexpression on redox balance, cell growth, and apoptosis. *J Biol Chem* 2004;279:27302–14.
- Powis G, Mustacich D, Coon A. The role of the redox protein thioredoxin in cell growth and cancer. *Free Radic Biol Med* 2000;29:312–22.
- Friedman HS, Johnson SP, Colvin OM. Cellular mechanisms of cyclophosphamide resistance: model studies in human medulloblastoma cell lines. *Cancer Treat Res* 2002;112:199–209.
- Sladek NE. Leukemic cell insensitivity to cyclophosphamide and other oxazaphosphorines mediated by aldehyde dehydrogenase(s). *Cancer Treat Res* 2002;112:161–75.
- Di Paolo A, Danesi R, Del Tacca M. Pharmacogenetics of neoplastic diseases: new trends. *Pharmacol Res* 2004;49:331–42.
- Nyhan WL. Disorders of purine and pyrimidine metabolism. *Mol Genet Metab* 2005;86:25–33.
- van Kuilenburg AB, Meinsma R, van Gennip AH. Pyrimidine degradation defects and severe 5-fluorouracil toxicity. *Nucleosides Nucleotides Nucleic Acids* 2004;23:1371–6.
- Reddig PJ, Juliano RL. Clinging to life: cell to matrix adhesion and cell survival. *Cancer Metastasis Rev* 2005;24:425–39.
- Seiler A, Visan A, Buesen R, Genschow E, Spielmann H. Improvement

- of an *in vitro* stem cell assay for developmental toxicity: the use of molecular endpoints in the embryonic stem cell test. *Reprod Toxicol* 2004; 18:231–40.
30. Fellenberg J, Dechant MJ, Ewerbeck V, Mau H. Identification of drug-regulated genes in osteosarcoma cells. *Int J Cancer* 2003;105: 636–43.
31. Fais S. A role for ezrin in a neglected metastatic tumor function. *Trends Mol Med* 2004;10:249–50.
32. Parlato S, Giammaroli AM, Logozzi M, et al. CD95 (APO-1/Fas) linkage to the actin cytoskeleton through ezrin in human T lymphocytes: a novel regulatory mechanism of the CD95 apoptotic pathway. *EMBO J* 2000;19:5123–34.
33. Luciani F, Molinari A, Lozupone F, et al. P-glycoprotein-actin association through ERM family proteins: a role in P-glycoprotein function in human cells of lymphoid origin. *Blood* 2002;99:641–8.
34. Golden PL, Pardridge WM. Brain microvascular P-glycoprotein and a revised model of multidrug resistance in brain. *Cell Mol Neurobiol* 2000; 20:165–81.
35. Holcik M, Sonenberg N. Translational control in stress and apoptosis. *Nat Rev Mol Cell Biol* 2005;6:318–27.
36. Thornton S, Anand N, Purcell D, Lee J. Not just for housekeeping: protein initiation and elongation factors in cell growth and tumorigenesis. *J Mol Med* 2003;81:536–48.
37. Neckers L, Ivy SP. Heat shock protein 90. *Curr Opin Oncol* 2003;15: 419–24.
38. Mhawech P. 14-3-3 proteins—an update. *Cell Res* 2005;15:228–36.
39. Tzivion G, Gupta VS, Kaplun L, Balan V. 14-3-3 proteins as potential oncogenes. *Semin Cancer Biol* 2006;16:203–13.
40. Liu Y, Liu H, Han B, Zhang J. Identification of 14-3-3 as a contributor to drug resistance in human breast cancer cells using functional proteomic analysis. *Cancer Res* 2006;66:3248–55.
41. Whitesell L, Lindquist SL. HSP90 and the chaperoning of cancer. *Nat Rev Cancer* 2005;5:761–72.
42. Ecroyd H, Jones RC, Aitken RJ. Tyrosine phosphorylation of HSP-90 during mammalian sperm capacitation. *Biol Reprod* 2003;69:1801–7.
43. Kim SC, Sprung R, Chen Y, et al. Substrate and functional diversity of lysine acetylation revealed by a proteomics survey. *Mol Cell* 2006;23: 607–18.
44. Beausoleil SA, Jedrychowski M, Schwartz D, et al. Large-scale characterization of HeLa cell nuclear phosphoproteins. *Proc Natl Acad Sci U S A* 2004;101:12130–5.
45. Nousiainen M, Sillje HH, Sauer G, Nigg EA, Korner R. Phosphoproteome analysis of the human mitotic spindle. *Proc Natl Acad Sci U S A* 2006;103:5391–6.
46. DeGiorgis JA, Jaffe H, Moreira JE, et al. Phosphoproteomic analysis of synaptosomes from human cerebral cortex. *J Proteome Res* 2005;4: 306–15.
47. Hoffert JD, Pisitkun T, Wang G, Shen RF, Knepper MA. Quantitative phosphoproteomics of vasopressin-sensitive renal cells: regulation of aquaporin-2 phosphorylation at two sites. *Proc Natl Acad Sci U S A* 2006;103:7159–64.
48. Lees-Miller SP, Anderson CW. Two human 90-kDa heat shock proteins are phosphorylated *in vivo* at conserved serines that are phosphorylated *in vitro* by casein kinase II. *J Biol Chem* 1989;264:2431–7.
49. Rush J, Moritz A, Lee KA, et al. Immunoaffinity profiling of tyrosine phosphorylation in cancer cells. *Nat Biotechnol* 2005;23:94–101.
50. Kumar Y, Khachane A, Belwal M, Das S, Somsundaram K, Tatu U. ProteoMod: a new tool to quantitate protein post-translational modifications. *Proteomics* 2004;4:1672–83.

2.3.2 Résultats complémentaires

Electrophorèse bidimensionnelle

Variabilité des images. Après traitement 5FC, la forme phosphorylée de l'HSP90 (spot 123) est un spot majoritaire sur le gel 2D (voir article et **Figure 27**). Son caractère remarquable de présence/absence en fonction du traitement rend son analyse quantitative excessivement simple. Cette situation est exceptionnelle, car la plupart des spots présente des variations discrètes entre les deux conditions, avec des rapports de volume 5FC/CT allant de 0.5 à 2.5. Dans ce cas, les valeurs moyennes obtenues à partir de plusieurs gels sont utilisées après étude de la variabilité des données. En pratique, les variations peuvent être étudiées avec un même échantillon mesuré plusieurs fois (répétitions techniques pour déterminer la répétabilité) ou avec des échantillons différents (répétition biologique pour déterminer la reproductibilité). A titre d'exemple, les répétitions biologiques de la condition 5FC pour la zone du gel contenant les spots de l'HSP90 sont présentées dans la **Figure 27**.

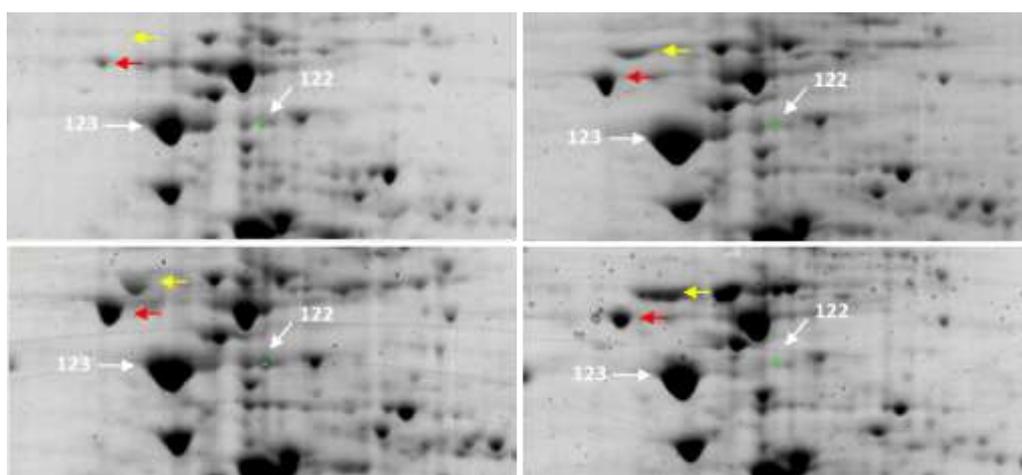


Figure 27. Quatre répétitions biologiques de la condition 5FC. Les 4 répétitions biologiques présentées ici montrent des différences marquées par des flèches jaune ou rouge. Les spots 122 et 123 correspondent respectivement à la forme non phosphorylée et la forme phosphorylée de l'HSP90.

Malgré une bonne reproductibilité globale des gels, certains spots apparaissent ou disparaissent au gré des échantillons, indépendamment du traitement. Les spots pointés par une flèche rouge et une flèche jaune sur la **Figure 27** illustrent cette observation. En effet, alors que les spots environnants sont identiques pour les différents échantillons, ces deux spots ont des comportements particuliers, avec une variabilité importante du deuxième spot (flèche jaune), allant de l'absence à la présence sous forme de spots, de trainée ou de chapelet. Ces variations ne sont pas observées lors de répétitions techniques, elles sont propres aux échantillons. Il pourrait

sembler surprenant qu'une culture de cellules monoclonales puisse exprimer des protéines d'une manière différente pour une même condition de culture. Il est possible que ces variations dépendent de la densité de cellules dans la boîte de Petri.

Effet de la densité de cellules. Une condition supplémentaire a été rajoutée et correspond aux cellules ProbCD cultivées jusqu'à confluence. (**Fig. 28**).

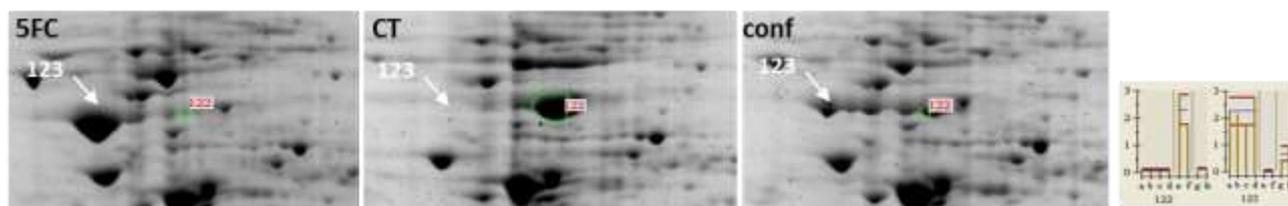


Figure 28. Détection de l'HP90 en fonction du traitement. Les spots 122 (pI=5.1) et 123 (pI = 5.5) correspondent à l'HSP90. Conditions 5FC, CT et CT à confluence (conf). Les cellules sont soit présentes en petits amas (condition CT), soit sont en contact sur l'ensemble de la boîte (condition conf). Les histogrammes présentent les volumes des spots 122 et 123 dans les 3 conditions.

La condition 5FC présente un spot majeur en position 123 alors que le spot 122 est absent. Inversement la condition CT présente un spot majeur en position 122 alors que le spot 123 est absent. Une situation intermédiaire est obtenue avec les cellules non traitées qui atteignent la confluence : les spots 122 et 123 sont présents en quantité équivalente, ainsi que 3 spots intermédiaires (**Fig.28, gel de droite**). Les différents spots intercalés entre le spot 122 et 123 correspondent aussi à l'HSP90, mais leur plus faible quantité par rapport aux spots 122 et 123 n'a pas permis d'identifier les modifications post-traductionnelles correspondantes. Leur décalage vers la gauche (pI plus bas) par rapport au spot 122 suggère que des fonctions basiques sont neutralisées ou que des fonctions acides sont ajoutées.

Retraitement des données de spectrométrie de masse

Une évolution majeure des logiciels d'interrogation des bases de données est l'utilisation d'un nouveau paramètre pour établir la validité de l'identification ; le taux de faux positifs (FDR pour False Discovery Rate). Les différences observées entre l'article et les résultats présentés ci-dessous sont la conséquence de l'utilisation d'un logiciel d'identification plus récent (Proteome Discoverer® 1.4). Cette nouvelle version utilise le même algorithme de corrélation, Sequest®, développé par Jimmy Eng (Eng et al., 1994) mais Proteome Discoverer® utilise une base dite reverse pour calculer le taux de faux positifs (FDR) et réajuster les seuils des scores de validation. Les résultats suivants sont établis avec un FDR < 1 % ; 1 % des séquences présentées ici sont donc identifiées à tort.

Identification de l'HSP90

L'HSP90 existe sous plusieurs isoformes. Chez le rat (*Rattus norvegicus*), deux isoformes sont décrites dans les bases de données protéiques :

- P82995 : l'HSP90 α ou Heat shock 86 kDa(HSP86), nom du gène Hsp90aa1, chrom 6
- P34058 : l'HSP90 β ou Heat shock 84 kDa (HSP84) nom du gène Hsp90ab1, chrom 9

Les deux séquences présentent 85 % d'homologie et sont fortement conservées, elles possèdent plus de 99 % d'homologie avec leurs homologues dans le génome d'autres mammifères séquencés (souris, bovin, singe, humain etc...).

L'analyse du spot 123 par LC-trappe ionique a permis d'identifier distinctement les deux isoformes (**Tableau 7**). Douze peptides sont communs aux deux protéines, 34 sont spécifiques de l'HSP90 β et 23 sont spécifiques de l'HSP90 α . La présence de ces deux protéines dans un même spot est certainement due au fait que les deux isoformes ont à la fois un point isoélectrique presque identiques (5.01 et 5.03) et une masse moléculaire très proche (84 815 et 83 281 Da).

Tableau 7. Liste des peptides identifiés dans le spot 123

P34058 - Heat shock protein HSP 90B		P82995 - Heat shock protein HSP 90A		P34058 & P82995	
Séquence	Modification	Séquence	Modification	Séquence	Modification
IEDVG\$DEEDDSGK	S6(+80)	VVD\$EDLPLNISR	S4(+80)	TLTIVDTGIGMTK	M11(+16)
IEDVG\$DEEDDSGKDK	S6(+80)	ESDDKPEIEDVG\$DEEEEEK	S13(+80)	HNDDEQYAWESSAGGSFTVR	
IEDVG\$DEEDDSGKDKK	S6(+80)	HGLEVIYMIPIDEYCVQQLK	M8(+16); C16(+57)	YIDQEELNK	
EKEI\$DDEAEEEK	S5(+80)	YYSASGDEMVSLLK	M10(+16)	EDQTEYLEER	
LVSSPCCIVTSTYGWTANMER	P5(+16); C6(+57); C7(+57); M19(+16)	LVTSPCCIVTSTYGWTANMER	C6(+57); C7(+57); M19(+16)	EKYIDQEELNK	
LVSSPCCIVTSTYGWTANMER	C6(+57); C7(+57); M19(+16)	HGLEVIYMIPIDEYCVQQLK	C16(+57)	ELHINLIPNKQDR	
RGFEVVMTEPIDEYCVQQLK	M8(+16); C16(+57)	CLELFTELAEDKENYK	C1(+57)	GVVDSDELPLNISR	
GFEVVMTEPIDEYCVQQLK	M7(+16); C15(+57)	CLELFTELAEDKENYKK	C1(+57)	EGLELPEDEEEKKK	
VFIMDSDELPIPEYLNFR	M4(+16); C7(+57)	LIINTFYNSK		HFSVEGQLEFR	
YHTSQSGDEMSTLSEYVSR	M10(+16)	NPDDITNEEYGEFYK		VILHLKEDQTEYLEER	
TKPIWTRNPDDITQEEYGEFYK	K2(Acetyl)	HLEINPDHSIIETLR		YESLTDPSK	
AKFENLCK	C7(+57)	ADLNNLGTIAK		YESLTDPSKLDLSDGK	
KCLELFSLEAEDKENYK	C2(+57)	RAPFDLFENR			
GFEVVMTEPIDEYCVQQLK	C15(+57)	HSQFIGYPITLFEVK			
CLELFSLEAEDK	C1(+57)	DQVANSFAVER			
CLELFSLEAEDKENYK	C1(+57)	APFDLFENR			
CLELFSLEAEDKENYKK	C1(+57)	ELISNSSDALDK			
APFDLFENK		FYEQFSK			
EEEDKEDEEKPK		HIYFITGETK			
ELISNASDALDK		IINTFYNSK			
EQVANSFAVER		LGIHEDSQNR			
FYEAFSK		SLLIINTFYNSK			
GEKEEEDKEDEEKPK		SLTNDWEEHLAVK			
HLEINPDHPIVETLR					
HSQFIGYPITLYLEK					
KHLEINPDHPIVETLR					
LGIHEDSTNR					
NPDDITQEEYGEFYK					
RAPFDLFENK					
RAPFDLFENKK					
RLSELLR					
SIYYITGESK					
SLTNDWEDHLAVK					
YHTSQSGDEMSTLSEYVSR					

Alors qu'une seule phosphorylation était identifiée lors de l'écriture de l'article, le retraitement des données par Proteome Discoverer® 1.4 permet d'identifier 4 phosphorylations sur sérine, une acétylation sur lysine et une hydroxylation de Proline.

Identification de nouvelles modifications post-traductionnelles.

Des Western blots ont été réalisés avec des anticorps anti-acétyllysine, anti-phosphotyrosine et anti-phosphosérine pour les conditions 5FC et CT. La détection de l'acétyllysine est positive pour les 2 spots de l'HSP90 et aucune phosphotyrosine n'est détectée. Par contre, l'anticorps anti-phosphosérine détecte spécifiquement le spot 123 induit par la condition 5FC (article, Fig. 3).

Suite à ce premier résultat, une purification sur phase TiO₂ (GL Science, Japon) a été réalisée en batch (incubation en Eppendorf de l'hydrolysate dessalé avec de la phase TiO₂). Le tampon de charge et de rinçage (0.1 % TFA, 30 % ACN) n'a pas permis d'éliminer les peptides riches en D et E, seuls peptides identifiés par l'élution 0.5M NH₄OH.

Face à cet échec, plusieurs spots 123 provenant de différents gels ont été excisés et réunis, puis des clivages protéolytiques à la trypsine, à la protéase V8 (GluC) et la protéinase K ont été effectués. Les digestions à la protéinase K et GluC ont produit peu de peptides identifiés par MS/MS, alors que la digestion à la trypsine a permis l'identification de 4 sites de phosphorylation, un peptide acétylé et un peptide hydroxylé sur une proline (**Tableau 7**). Ces modifications sont attribuées sans ambiguïté à l'une ou l'autre HSP90 mais une interrogation Blast des séquences identifiées par MS montre que les sites de phosphorylation, d'acétylation et d'hydroxylation sont présents sur les deux HSP90 (**Tableau 8**).

Tableau 8 : Modifications pos-traductionnelles observées sur l'HSP90 par MS et Blast.

P34058 - Heat shock protein HSP 90 bêta		P82995 - Heat shock protein HSP 90 alpha	
Séquence	Modification	Séquence	Modification
(250) IEDVG S DEEDDSGK(263)	S255 (phospho) MS	(258) IEDVG S DEEEEEKK(271)	S263 (phospho) MS
(222) EKEI S DEAEAEK(234)	S226 (phospho) MS	(227) DKEV S DEAEAEK(238)	Blast
(584) LVSPCCIVT S TYGWTANMER(604)	P588 (hydroxyl) MS	(593) LVSPCCIVT S TYGWTANMER(613)	Blast
(285) TKPIWTRNPDDIT Q EEYGEFYK(306)	K286 (Acetyl) MS	(294) TKPIWTRNPDDIT N EEYGEFYK(315)	Blast
(380) VV S EDLPLNISR(392)	S383 (phospho) Blast	(389) VV S EDLPLNISR(401)	MS

En rouge les homologues de séquences entre les deux isoformes.

La phosphorylation de la sérine en position 255 de l'HSP90 β est caractérisée par l'identification du peptide tryptique phosphorylé avec 0, 1 et 2 miscleavages, soit les séquences respectives IEDVGSDEEDDSGK, IEDVGSDEEDDSGKDK, IEDVGSDEEDDSGKDKK (**Fig. 29**). Pour les 3 peptides, la perte du phosphate produit un ion intense $[M + 2H - H_3PO_4]^{2+}$ et dans le cas du peptide

IEDVGSDEEDDSGK**D**KK, la perte de neutre produit un pic majoritaire qui masque les ions de séquence. Malgré cette situation défavorable, le spectre MS/MS permet tout de même l'identification du peptide dans la base de données complète (27251 séquences). La présence de miscleavages est rarement observée. Ici, cinq des six peptides qui contiennent les modifications post traductionnelles ont des lysines internes. Cette digestion partielle est fortuite, elle est la conséquence de regroupement de plusieurs spots 2D dans un même tube. La digestion trypsique a été réalisée sur des volumes de gel plus importants que la normale, sans augmenter suffisamment la quantité de trypsine.

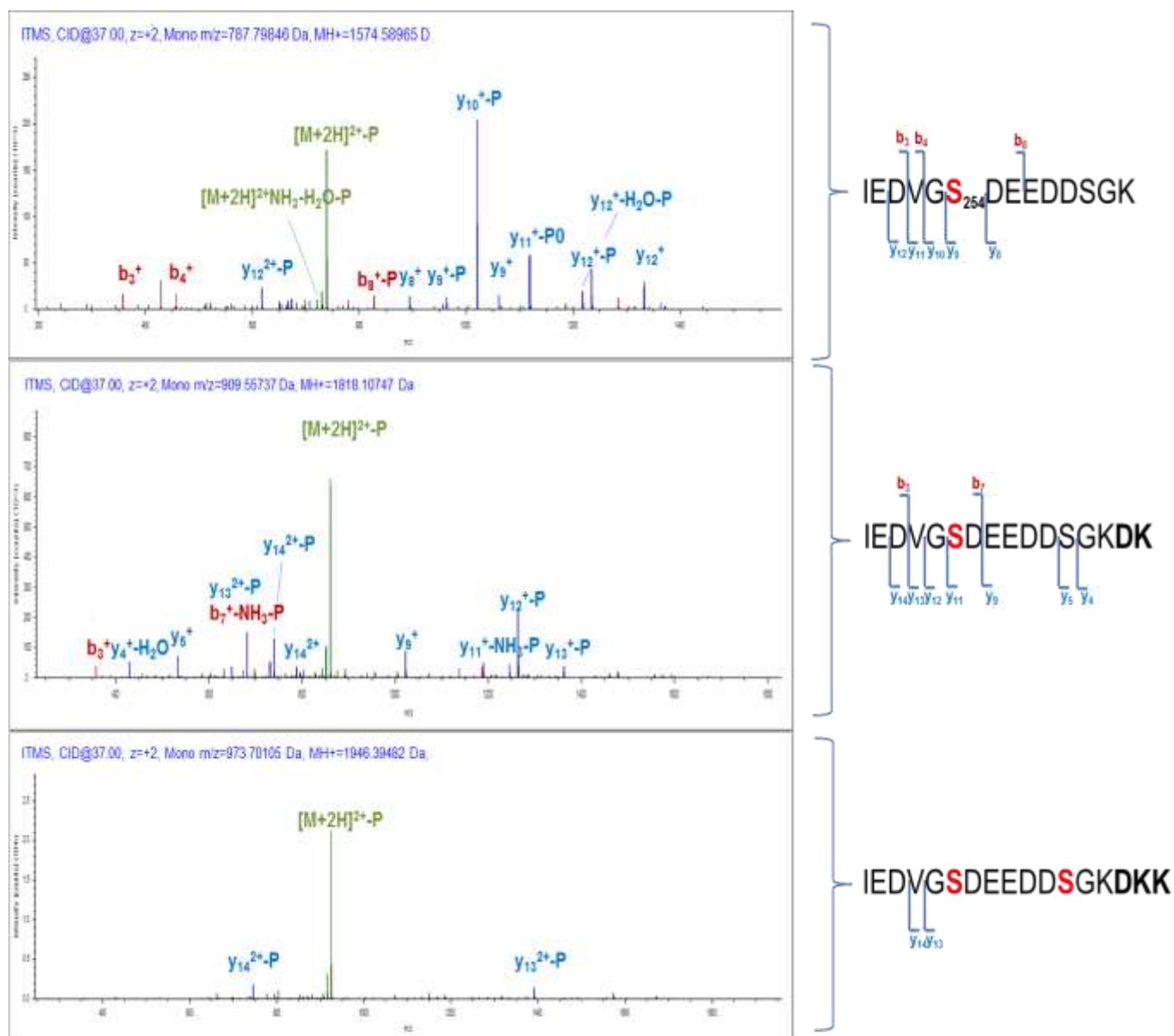


Figure 29 : spectres MS/MS annotés des peptides portant S₂₅₄ phosphorylée. De haut en bas, IEDVGSDEEDDSGK, IEDVGSDEEDDSGKDK, IEDVGSDEEDDSGKDKK.

Le chromatogramme du pic de base (BP pour Base Peak) pour un hydrolysats trypsique du spot 123 est présenté sur la **Figure 30** et les chromatogrammes correspondant aux trois peptides phosphorylés sont présentés en dessous. Ces ions n'apparaissent pas dans le tracé BP, l'intensité des ions phosphopeptides étant trop faible. Ils apparaissent par contre clairement lorsque le signal correspondant à leur m/z est extrait. On peut constater qu'ils sont élués en début de gradient et que l'addition de lysine (K) et l'addition d'acide aspartique (D) diminue la rétention du peptide.

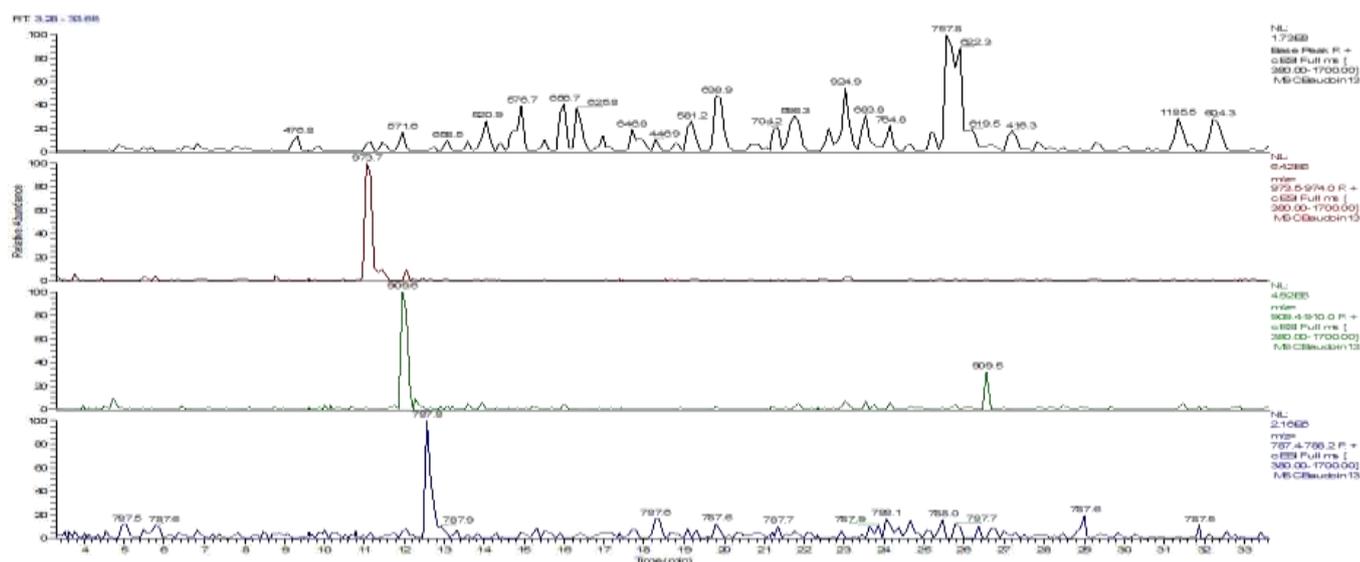


Figure 30. LC-MS de l'hydrolysats trypsique de l'HSP90 phosphorylée, spot 123. De haut en bas : chromatogrammes du pic de base, de l'ion m/z = 973.7 (IEDVGSDEEDDSGKDKK), de l'ion m/z = 909.5 (IEDVGSDEEDDSGKDK), de l'ion m/z = 787.6 (IEDVGSDEEDDSGK).

L'hydrolyse partielle du spot a également permis d'identifier une acétylation en position 2 sur le peptide TKPIWTRNPDDITQEEYGEFYK. Le spectre MS/MS correspondant est présenté **Figure 31**, il montre une série d'ions γ majoritaires qui est complète de l'ion γ_2 à l'ion γ_{13} et qui permet un séquençage *de novo*. Ce peptide n'avait pas été identifié lors de la publication de l'article malgré la qualité de sa fragmentation. Sa valeur de Xcorr (valeur de corrélation entre les spectres théoriques et expérimentaux) est égale à 2.94 dans l'article alors que le seuil de sélection avait été fixé à 3.3. Il a donc été automatiquement exclu de la liste de peptides identifiés. Lors du retraitement des données, la valeurs de XCorr pour les ions tri-chargés a été ajustée automatiquement à une valeur inférieure à 2.9 pour permettre un FDR < 1 % ; le peptide acétylé a donc été sélectionné.

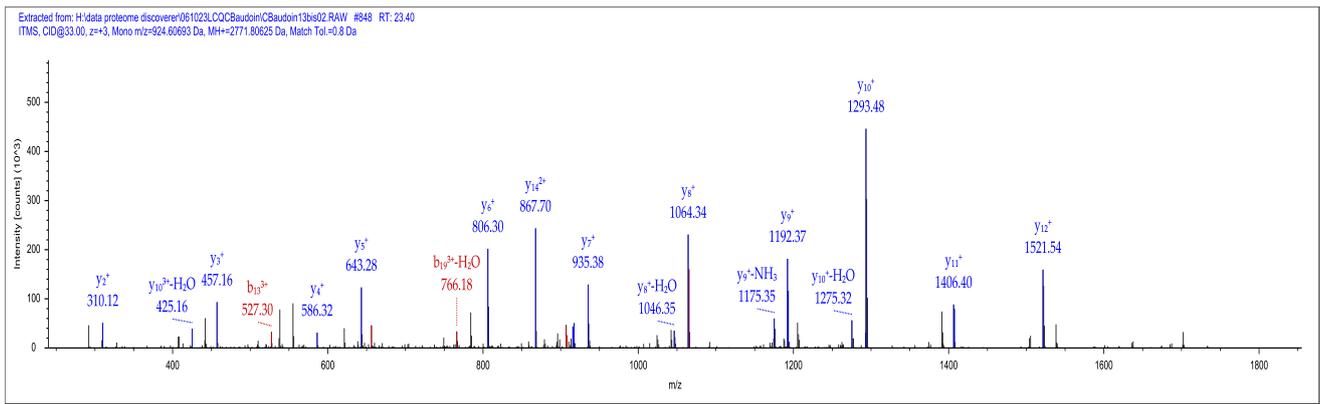


Figure 31. Spectre MS/MS du peptide acétylé en position 286. $m/z = 924.6$, $z=3$. La séquence identifiée (TK*PIWTRNPDDITQEEYGEFYK). Paramètres de validation Sequest : XCorr = 2.94, rank = 1.

2.4 Discussion

Les années passées entre la publication de l'article et le retraitement des données avec de nouveaux outils logiciels ont mis à jour un certain nombre de données supplémentaires par rapport à l'article publié dans *Molecular Cancer Therapeutics*. L'HSP90 β identifiée dans le spot 123 induits par le traitement 5FC avait pour seule modification la phosphorylation S255. Le retraitement des données a permis d'identifier la présence concomitante de l'HSP90 α . Cette observation modifie sensiblement la conclusion de l'article puisque la variation induite par le traitement 5FC est du type « on/off » non seulement pour l'HSP90 β mais aussi pour l'HSP90 α . Par ailleurs, de nouvelles modifications pos-traductionnelles ont été identifiées (2 phosphorylations sur S₂₂₆ et S₃₈₃, une hydroxylation sur P₅₈₈ et une acétylation sur K₂₈₆) (**Tableau 8**). L'acétylation avait été caractérisée par Western Blot dans les deux spots 122 ($pI = 5.5$) et 123 ($pI = 5.1$) ; elle n'explique donc pas le déplacement du spot de pI 5.5 à pI 5.1. Par contre, l'analyse par Western blot avait également permis d'identifier une phosphorylation spécifique du spot 123 et la sérine 255 de l'HSP90 β avait été identifiée comme phosphorylée. Ici, le retraitement des données a permis d'identifier deux autres sérines phosphorylées en position 226 et 383.

Sachant que (1) le spot 122 contient l'HSP90 α et l'HSP90 β , (2) ce spot disparaît entièrement au profit du spot 123 lors du traitement 5FC, (3) le spot 123 contient l'HSP90 α et l'HSP90 β , Il est probable que les modifications post-traductionnelles responsables du changement de point isoélectrique concernent à la fois l'HSP90 α et l'HSP90 β bien que le logiciel d'identification n'attribue les peptides phosphorylés qu'à une seule des HSPs. La recherche de séquences homologues par BLAST montre que les sites de phosphorylations sont présents sur les deux HSPs (**Tab. 8**), il y a donc bien une phosphorylation des deux HSPs. Cette observation pointe du doigt le

problème des identifications basées sur l'approche *Bottom-up* ; les protéines sont identifiées à partir de peptides ne couvrant pas la totalité de la séquence protéique et l'identification de peptides spécifiques d'une isoforme ne permet de conclure qu'à la seule présence de l'isoforme en question. Dans le cas présent, les données d'électrophorèse et de spectrométrie de masse permettent de conclure, contrairement à ce que suggère le logiciel d'identification, que les sérines phosphorylées sont présentes à la fois sur l'HSP90 α et l'HSP90 β .

Paradoxalement, le retraitement récent des données diminue la couverture de séquences de l'HSP90. L'article présentait 71 % de couverture, il n'y a plus maintenant que 50 % de la séquence identifiée. Cette différence provient assurément de la gestion des faux positifs qui sont maintenant ajustés à moins de 1 % alors que leur détermination n'était pas réalisée auparavant. Les seuils appliqués précédemment étaient fonction d'un score (X_{corr} , sp) dont le calcul est sensible au nombre de charges z . Les seuils de X_{corr} fixés pour les différentes valeurs de z étaient fonction de l'évaluation manuelle *a posteriori* des spectres MS/MS et de leur corrélation avec les fragments théoriques. Cette évaluation visuelle de l'identification reposait sur deux critères principaux ; (i) l'affectation des ions les plus intenses du spectre expérimental à des ions théoriques, (ii) la présence de petites séries d'ions dans le spectre expérimentale. L'utilisation d'un FDR fixé à 1 % et la diminution du nombre de peptides identifiés qui en résulte montre que l'évaluation manuelle a abouti à une sélection erronée de peptide. Paradoxalement, de nouveaux peptides sont identifiés, à l'image du peptide IEDVGSDEEDDSGKDK dont le spectre MS/MS ne permettrait pas une validation manuelle (**Fig. 29**).

L'électrophorèse bidimensionnelle est aujourd'hui d'une utilisation de moins en moins fréquente et l'analyse quantitative des hydrolysats est devenue la technique de référence. On peut cependant se demander quel résultats auraient produit une analyse quantitative par LC-MS ? L'étude quantitative des peptides totaux n'aurait pas permis de mettre en évidence l'HSP90, puisque sa quantité totale ne varie pas. Par contre, l'analyse des peptides après enrichissement des peptides phosphorylés aurait permis d'identifier la phosphorylation de l'HSP90 sans pour autant identifier les autres protéines non modifiées et discutées dans l'article. Il apparaît donc important d'analyser dans le cadre d'une approche phosphoprotéomique par LC-MS à la fois les phosphopeptides et les peptides totaux.

CHAPITRE 3 :
Etude Comparative des chromatographies d'affinité IMAC et MOAC

3.1 Introduction

Le chapitre précédent a montré qu'une protéine, l'HSP90, avait un niveau d'expression constant mais que sa phosphorylation était caractéristique du traitement appliqué. Comme l'illustre cet exemple, la plupart des activités cellulaires telles que le métabolisme, le cycle cellulaire, mobilité cellulaire, etc. ne peuvent être caractérisées à partir de l'identification des seules protéines sans prendre en compte les modifications post-traductionnelles qui activent, inhibent ou modulent l'activité des protéines. A ce titre, la phosphorylation prend une place particulière dans le monde des modifications post-traductionnelles, car elle correspond au système de régulation le plus étudié à ce jour.

D'un point de vue biochimique, l'identification des phosphoprotéines et la caractérisation des sites de phosphorylation sont confrontés à la difficulté de purifier une fraction minoritaire et hétérogène, les phosphopeptides, dans un mélange complexe de peptides tout aussi hétérogènes en termes de pI, de masse ou d'hydrophobie. Les phosphopeptides sont des peptides rares, à la fois du point de vue quantitatif (faible abondance relative) et qualitatif (ils représentent un faible pourcentage de la composition d'une protéine). L'analyse de l'ensemble des phosphopeptides nécessite donc l'emploi d'une technique de purification qui a été envisagée à l'aide de deux méthodes. La première méthode consiste à utiliser la réactivité du groupement phosphate pour réaliser une chimie de couplage et/ou de marquage. Cette solution n'est pas ou peu employée à ce jour pour l'analyse phosphoprotéomique. La deuxième méthode consiste à utiliser des phases de chromatographie plus ou moins sélectives pour enrichir la fraction phosphorylée. Deux techniques chromatographiques spécifiques sont utilisées actuellement : la chromatographie d'affinité sur métal chélaté (Immobilized Metal Affinity Chromatography, IMAC) et plus récemment la chromatographie sur oxyde de métal (Metal Oxide Affinity Chromatography, MOAC). Ce chapitre présente une comparaison des deux techniques dans le but d'identifier l'approche la plus favorable pour purifier de manière reproductible des phosphopeptides provenant d'extraits de foie.

3 Choix méthodologiques

3.2.1 Chromatographie en colonne

Les supports d'affinité sont souvent utilisés en *batch*. Dans ce cas, les échantillons sont incubés dans des tubes (Eppendorf) avec la phase solide et une centrifugation permet de récupérer

le surnageant. Cette méthode est simple à mettre en œuvre, mais présente des inconvénients qui peuvent nuire à la reproductibilité des expériences. Lorsque de nombreux échantillons sont traités en même temps, il est difficile de maintenir constant les temps d'incubation avec la phase solide et les temps de rinçage, et des différences importantes peuvent apparaître entre le premier et le dernier échantillon. L'élimination lors de lavages ou la récupération finale des surnageants est également délicate, puisque de la phase solide peut être emportée accidentellement à ces étapes. Enfin, les volumes de phase stationnaire peuvent également varier entre les échantillons dans la mesure où cette dernière sédimente très rapidement. Une solution qui permet de pallier ces problèmes est l'utilisation d'un système HPLC pour contrôler précisément les volumes d'échantillons injectés ainsi que les débits et les conditions de rinçage et d'élution. Cette solution a été choisie et mise en œuvre pour deux phases solides ; POROS 20MC (Applied Biosystems) saturée en Fe^{3+} (POROS- Fe^{3+}) et TiO_2 (GL Science).

3.2.2 Choix des échantillons

Certains articles présentent des protocoles de purification de phosphopeptides avec une mise au point des conditions chromatographiques sur un hydrolysate de phosphoprotéines purifiées, caséines et ovalbumine notamment (Aryal et al., 2008; Aryal and Ross, 2009; Guzel et al., 2013; Kuroda et al., 2004; Xue et al., 2009). Dans le cadre de l'étude de l'HSP90, une vérification avait été réalisée de cette manière en utilisant de la phase TiO_2 et un hydrolysate tryptique de caséines. Malgré des résultats de purification satisfaisants pour les caséines, la mise en œuvre du protocole pour l'HSP90 phosphorylée a été un échec, la chromatographie MOAC ne permettant la purification et l'identification que des peptides riches en D et E.

Ce résultat préliminaire est un exemple de l'utilisation d'un modèle trop simple qui n'est pas représentatif de l'échantillon réel. Afin de limiter ce biais, la mise au point a été réalisée avec un hydrolysate tryptique de caséines diluée au 1/10 dans un hydrolysate d'albumine. Les caséines sont des protéines du lait hautement phosphorylées ; elles sont diluées dans une matrice de peptides d'albumine afin de rendre l'échantillon plus complexe et un peu plus conforme à la réalité. Cette méthode a été utilisée avec succès dans plusieurs articles de références (Larsen et al., 2005; Thingholm et al., 2008). Dans notre cas, la validation a été complétée avec un échantillon biologique, un extrait de protéines de foie de souris. Les quantités de foie frais (100 mg) utilisées

pour la mise au point ont été choisies pour correspondre aux quantités attendues à partir de biopsies hépatiques humaines.

3.2.3 Préparation des hydrolysats tryptiques

Extrait protéique de lait. Le lait de vache contient 3.5 g/l de protéines dont 80 % correspond aux caséines (les caséines α_{s1} , α_{s2} , β sont phosphorylées) (Alais and Linden, 1987). Une préparation de caséines a été obtenue à partir de 100 mL de lait écrémé précipité à pH 4.6 par de l'acide acétique 1M. Le surnageant qui correspond au lactosérum est éliminé et le culot rincé 3 fois par de l'acide acétique 50 mM, puis 3 fois par un mélange 90 % éthanol/10 % d'eau afin de permettre un séchage rapide au SpeedVac (Thermo Scientific). Le culot obtenu (3g) est d'apparence blanc pâle et correspond bien à la quantité de caséines attendue à partir de 100 mL de lait.

Extrait protéique de foie. Le foie d'une souris a été prélevé entier, divisé en morceaux d'environ 100 mg, puis congelé dans l'azote liquide et conservé à -20°C. L'extrait est préparé en broyant à l'aide d'un piston plastique 100 mg de tissu dans un tube de 1.5 mL (Eppendorf) après l'ajout de 0.5 mL de tampon de lyse contenant 0.1 M TEAB pH 9.3, 0.15 M NaCl, 0.1 % SDS, 5 mM NaF, 2 mM Na_3VO_4 et 10 mM DTT. Après 10 min à 100°C, l'échantillon est refroidi dans la glace puis traité pendant 45 min à température ambiante par de l'iodoacétamide (concentration finale 50mM). Après dilution au demi dans 0.1 M TEAB, l'extrait est soumis à une protéolyse par la trypsine.

Protéolyse par la trypsine. La trypsinolyse est réalisée avec une solution stock de trypsine TPCK (Sigma-Aldrich) à 10 mg/mL dans HCl 2mM. La protéolyse est menée à 37°C avec un rapport enzyme/substrat de 1/100 (poids/poids). Pour les protéines purifiées, le ratio est ajusté en fonction de la pesée de la protéine. Pour le foie, le ratio est estimé en considérant 10 % de protéines dans l'extrait frais de foie. Après 1h d'incubation, la même quantité de trypsine est rajoutée. La protéolyse est réalisée dans un Thermomixer (Eppendorf) pendant 10 h sous agitation pulsée ; les échantillons sont ensuite refroidis, puis la digestion est stoppée par ajout de TFA (concentration finale 1 %).

Dessalage des échantillons. Le dessalage des échantillons est réalisé au moyen d'une phase Sep Pak tC18 (Waters) transférée dans une colonne HPLC (5 mm x 4.5 cm). La colonne est montée sur un système HPLC (PF2D, Beckman Coulter) muni d'une boucle d'injection de 500 μL et d'un détecteur UV (280nm). A un débit de 0.9 mL/min, les peptides sont retenus en tête de colonne à

0.5 % d'éluant B puis élués à 60 % de B, la colonne étant régénérée par 90 % de B (A : 0.1 % acide formique, B : 100 % acétonitrile). La fraction éluée à 60 % B est séchée au Speedvac (Thermo Scientific). Pour les échantillons le permettant, les tubes de collecte sont préalablement tarés afin de permettre une pesée de l'extrait sec.

3.2.4 Chromatographie d'affinité en colonne

Les échantillons secs après dessalage sont mis en solution (10 mg/mL) dans le tampon de charge puis injectés (500 µL) par le système PF2D sur la colonne d'affinité. Une seule voie de l'HPLC est utilisée pour rincer la colonne à un débit constant pendant 20 min après chaque injection. Les fractions non retenues (FT) et les fractions éluées sont collectées automatiquement, les tubes servant à la collecte de la fraction éluée étant préalablement garnis par 25 µL d'acide formique pur afin de neutraliser le tampon d'élution (0.7M NH₄OH). Deux phases ont été testées : POROS 20 MC (Applied Biosystems) et TiO₂ (GL Science).

HPLC avec colonne POROS-Fe³⁺. Une colonne (4.6 mm x 4 cm) a été remplie de phase POROS 20 MC (suspension dans l'acétonitrile pur) et utilisée dans les conditions suivantes : débit de 0.5 mL/min en tampon A (50 mM acide acétique, 20 % ACN). Le passeur d'échantillons permet d'injecter l'échantillon (400 µL) ainsi que les solutions nécessaires aux différentes étapes de la chromatographie IMAC : activation par l'éluant A contenant 100 mM FeCl₃, rinçage par 0.5 M NaCl en Tampon A, élution par 0.7 M NH₄OH, régénération par un mélange EDTA 0.1 M, NaCl 0.5M et NH₄OH 0.7 M. Entre les utilisations, la colonne est stockée en 50 % ACN.

HPLC avec colonne TiO₂. Une colonne (2 mm x 5 cm) a été remplie de phase TiO₂ (suspension dans 100 % ACN) et utilisée dans les conditions suivantes : débit de 0.2 mL/min en tampon A (50 mM acide acétique, 20 % ACN) ou en tampon B (50 mM TFA en 20 % ACN). L'élution de la fraction retenue est réalisée dans les mêmes conditions que pour la colonne POROS-Fe³⁺ par NH₄OH 0.7 M.

3.2.5 Analyse LC-MS/MS

Analyses Q-TOF. L'ensemble des analyses de mélanges simples (caséines-albumines) a été produite avec un système ESI-Q-TOF premier (Waters) couplé à une HPLC CapLC (Waters). La séparation des peptides a été réalisée avec une colonne capillaire silice (250 µm x 15 cm) remplie de phase C18 Stability 5µm (C.I.L.). Les peptides sont élués par un gradient de 0 à 30 % B en 30 min (Tampon A : 20mM ac. formique, 2 mM TFA, tampon B : 100 % acétonitrile). Pour l'identification

des peptides, le mode MS Survey a été utilisé (1 MS suivi de 3 MS/MS des 3 ions les plus intenses du spectre MS). Pour la quantification, Les données ont été produites avec le mode *MS Scan* (1 acquisition MS/sec).

Analyse LTQ-Orbitrap. L'analyse des extraits peptidiques de foie a été réalisée avec un système LTQ Orbitrap XL couplé à une HPLC nano (Thermo Scientific). Les échantillons (10µl) sont chargés dans un premier temps sur une cartouche de préconcentration (PepMap C18, 300 µm x 0.5 cm). La cartouche est ensuite reliée par un système de vannes à la colonne analytique (PepMap C18, 75 µm x 15 cm). Les peptides sont élués par un gradient d'acétonitrile de 5 à 40 % B en 105 min. (A : 0.1 % ac. formique en eau, B : 100 % ACN, 0.1 % ac. formique). L'acquisition met en œuvre une méthode « TOP10 » qui consiste en 1 MS suivie de 10 MS/MS des 10 ions les plus intenses du spectre MS).

3.3 Résultats

Ce travail a fait l'objet d'une publication (*Comparison of IMAC and MOAC for phosphopeptides enrichment by column chromatography, J Chromatogr B 891-892, 2012, 109-112*) présentée plus loin.

Dans un premier temps, l'étude a été réalisée avec un hydrolysats tryptique de caséines dilué à 10 % (poids/poids) dans un hydrolysats d'albumine. Deux acides différents ont été utilisés dans le tampon de charge : soit l'acide acétique (pKa = 4.73), soit le TFA (pKa = 0.3). Quel que soit l'acide utilisé, les deux phases permettent de purifier efficacement les phosphopeptides de caséines, mais des contaminants (peptides non phosphorylés d'albumine) sont également présents dans la fraction purifiée, notamment le peptide DAIPENLPPLTADFAEDK (m/z = 978.49, **Fig. 1 de l'article**) en fonction de l'acide utilisé dans le tampon de charge. Le TFA permet de réduire significativement la contamination du peptide DAIPENLPPLTADFAEDK dans les fractions éluées des colonnes POROS-Fe³⁺ et TiO₂. Cependant pour la colonne POROS-Fe³⁺, le TFA 0.1 M diminue de moitié les quantités de phosphopeptides retenus par rapport à l'acide acétique 0.1 M (**Fig. 1 de l'article**).

Dans un deuxième temps, le même protocole a été appliqué avec un extrait de foie de souris. Les contaminants, essentiellement des peptides riches en D et E, correspondent à l'essentiel des peptides purifiés par les phases POROS-Fe³⁺ et TiO₂ lorsque l'acide acétique 0.1M est utilisé. La situation est inversée lorsque l'acide acétique est remplacé par le TFA ; les phosphopeptides correspondent alors à la majorité des peptides identifiés (**Fig. 2 de l'article**).

Constatant que le nombre de peptides identifiés (peptides phosphorylés et non phosphorylés) diminuait dramatiquement lorsque la chromatographie IMAC était réalisée avec 0.1M TFA dans le tampon de charge, la fraction non retenue lors du chargement de l'échantillon sur la colonne d'affinité a été réinjectée sur la colonne d'affinité. Alors que peu de phosphopeptides ont été détectés au deuxième passage de l'extrait sur la colonne TiO₂, un nombre important de phosphopeptides a été identifié au deuxième passage sur colonne IMAC. La colonne IMAC ne fixe donc que partiellement les phosphopeptides lorsque le tampon de charge contient 0.1M TFA.

Finalement, le couplage d'un spectrophotomètre en sortie de colonne IMAC a permis de montrer que le fer était élué de la colonne IMAC lorsque le TFA était utilisé dans le tampon de charge. Par ailleurs, l'utilisation répétée de la colonne IMAC entraîne la formation d'oxyde de fer au court du temps. A cause de l'évolution du support, les caractéristiques chromatographiques des phases IMAC en colonne ne peuvent donc pas être garanties.



Contents lists available at SciVerse ScienceDirect

Journal of Chromatography B

journal homepage: www.elsevier.com/locate/chromb

Short communication

Comparison of IMAC and MOAC for phosphopeptide enrichment by column chromatography

Luc Negroni^{a,*}, Stephane Claverol^b, Jean Rosenbaum^c, Eric Chevet^{c,d}, Marc Bonneau^b, Jean-Marie Schmitter^a

^a Université de Bordeaux, UMR 5248, CNRS – UBx1 – IPB, Centre de Génomique Fonctionnelle, BP 68, 146 rue Léo Saignat, 33076 Bordeaux, France

^b Centre de Génomique Fonctionnelle, BP 68, 146 rue Léo Saignat, 33076 Bordeaux, France

^c Inserm U1053, Université Bordeaux 2, 146 rue Léo Saignat, 33076 Bordeaux, France

^d Inserm U1053, Avenir, Université Bordeaux 2, 146 rue Léo Saignat, 33076 Bordeaux, France

ARTICLE INFO

Article history:

Received 14 October 2011

Accepted 16 February 2012

Available online 23 February 2012

Keywords:

Phosphopeptide enrichment

IMAC

TiO₂

LC-MS/MS

ABSTRACT

Automated phosphopeptide enrichment prior to MS analysis by means of Immobilized Metal Affinity Chromatography (IMAC) and Metal Oxide Affinity Chromatography (MOAC) has been probed with packed columns. We compared POROS-Fe³⁺ and TiO₂ (respectively IMAC and MOAC media), using a simple mixture of peptides from casein–albumin and a complex mixture of peptides isolated from mouse liver. With these samples, selectivity of POROS-Fe³⁺ and TiO₂ were pH dependant. In the case of liver extract, selectivity increased from 12–18% to 58–60% when loading buffer contained 0.1 M acetic acid or 0.1 M trifluoroacetic acid, respectively. However, with POROS-Fe³⁺ column, the number of identifications decreased from 356 phosphopeptides with 0.1 M acetic acid to 119 phosphopeptides with 0.1 M TFA. This decrease of binding capacity of POROS-Fe³⁺ was associated with strong Fe³⁺ leaching. Furthermore, repetitive use of IMAC-Fe³⁺ with the 0.5 M NH₄OH solution required for phosphopeptide elution induced Fe₂O₃ accumulation in the column. By comparison, MOAC columns packed with TiO₂ support do not present any problem of stability in the same conditions and provide a reliable solution for packed column phosphopeptide enrichment.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

Protein phosphorylation has been extensively studied since the 1990s and the development of tandem mass spectrometry and affinity media for phosphopeptide enrichment provided effective tools to investigate phosphoproteomes. IMAC was the first widely used affinity medium for phosphopeptide enrichment [1,2]. Later, MOAC chromatography [3,4], essentially with TiO₂ as support, appeared as an interesting alternative and the sequential use of the two methods was proposed to increase the panel of purified phosphopeptides [5,6]. Logically, solutions offered by two-dimensional

chromatography with combinations such as SAX-MOAC [7], SCX-IMAC [8], IMAC-HILIC [9], or IMAC-IMAC [10] tend to be preferred for the identification of phosphopeptides from complex peptide mixtures. However, screening experiments aimed at biomarker discovery require the analysis of large sample sets and are not compatible with the multiplication of fractions induced by multidimensional chromatography. Therefore, a single step purification of phosphopeptides still remains a hot topic in phosphoproteomics. From another point of view, repeatability and reproducibility are not the strongest points of phosphopeptide enrichment methods that employed supports packed in pipette-tips or in-vial batch processes. The aim of this work was to develop a method with affinity supports packed in columns and using an HPLC system for process automation and column regeneration. Phosphopeptide enrichment with IMAC and MOAC was tested in this configuration. The method was first evaluated with a tryptic hydrolysate of a bovine casein and albumin mixture. The results obtained with this simple mixture were then validated by the analysis of a complex sample (i.e. a mouse liver extract). Superior results were obtained with MOAC operated at very low pH, and these tests highlighted a specific issue of IMAC column chromatography.

Abbreviations: AA, acetic acid; ACN, acetonitrile; IAA, iodoacetamide; DTT, dithiothreitol; IMAC, immobilized metal ion affinity chromatography; LC-MS/MS, liquid chromatography tandem mass spectrometry; MOAC, Metal Oxide Affinity Chromatography; TEAB, triethylammonium hydrogen carbonate; TFA, trifluoroacetic acid.

* Corresponding author. Tel.: +33 557 57 5623.

E-mail address: lnegroni@cbrmm.u-bordeaux.fr (L. Negroni).

1570-0232/\$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.jchromb.2012.02.028

2. Materials and methods

2.1. Chemicals

Acetonitrile (ACN), trifluoroacetic acid (TFA), acetic acid (AA), dithiothreitol (DTT), iodoacetamide (IAA), triethylammonium hydrogen carbonate (TEAB), sodium orthovanadate (Na_3VO_4), sodium fluoride, bovine serum albumin, bovine casein, Glu 1-fibrinopeptide B and TPCK treated trypsin were obtained from Sigma (France). 15.0 Mohms cm water was produced by a Pure Lab Option system (Veolia, France).

2.2. Sample preparation

Liver tissues was processed as previously described [8], with some modification. Briefly 100 mg liver from BALB/c adult mice were homogenized with a piston pellet (Eppendorf, France) in 0.5 ml of buffer containing 0.1 M TEAB pH 9.3, 0.15 M NaCl, 0.1% SDS, 5 mM NaF, 2 mM Na_3VO_4 and 10 mM DTT. The sample was heated 10 min at 100 °C then cooled, and 50 mM IAA was added for 45 min at RT. Soluble proteins (10 mg/100 mg crude extract) were quantified by Bradford assay and digested by trypsin at 37 °C for 18 h (1:100 as E:S ratio). The solution was acidified with TFA and tryptic peptides were purified with a C18 cartridge (washing buffer: 0.1% TFA, elution buffer 60% ACN) then dried with a vacuum centrifuge (Speed Vac, Thermo Scientific, France). Ten mg of dried extract were obtained from 100 mg tissue. Casein mixed with albumin in a 1:10 (w/w) ratio was digested with trypsin, like liver tissue, but with two minor modifications; 0.1% SDS was replaced with 2 M urea and the DTT reduction step was carried out for 1 h at room temperature.

2.3. Chromatographic supports

The IMAC support (POROS 20 MC, Applied Biosystems, France) and the MOAC support (TiO_2 , 10 μm , GL Science, Japan) were packed in a 2 mm \times 5 cm column and mounted in a HPLC system (Beckman Coulter, France) with a 500 μl sample loop, a column oven set at 35 °C and UV detector set at 280 nm. Sample (1 mg/ml, 400 μl), elution buffer (400 μl) and IMAC buffers (400 μl) were injected with the sample loop. Three different loading buffers were tested with POROS 20 MC and TiO_2 , namely buffer A: 0.1 M acetic acid, buffer B: 0.1 M TFA and buffer C: 0.7 M (5%) TFA. Elution buffer was 0.5 M NH_4OH . All buffers were prepared in 20% aqueous ACN. Columns were flushed at 0.2 ml/min for 15 min with 10 mM acetic acid between each runs. IMAC was activated with FeCl_3 and regenerated with EDTA as previously described [11].

2.4. LC-MS/MS and data processing

The collected fractions from casein–albumin tryptic digest were analyzed with a capLC system coupled with a Q-ToF Premier (Waters, France). Peptides were eluted from a 300 μm C18 column with a 0.1–30% ACN gradient for 30 min. Glu 1-fibrinopeptide B was added as internal control. Full MS scans were acquired from 350 to 1700 m/z and quantification was carried out with QuanLynx (Waters).

The eluted fractions from liver digests were analyzed with a nanoLC system (Dionex, France) coupled with a LTQ Orbitrap XL (Thermo Scientific). A 75 μm C18 analytical column was used with a 300 μm \times 0.5 mm pre-concentration cartridge (both from Dionex). A gradient of 5–40% ACN was applied for 105 min. Acquisition was a "TOP 10" method (an Orbitrap MS survey scan followed by ion trap MS/MS scans of the 10 most intense precursor ions). Data processing and peptide identifications were performed with Proteome

Discoverer 1.2 (Thermo Scientific); the different thresholds were set to obtain a False Discovery Rate (FDR) < 1%. Excel (Microsoft) was used to remove the redundant peptides and to count the phosphorylated peptides.

3. Results

3.1. Comparison of media and loading buffers with the casein–albumin mixture

The purification methods were evaluated with tryptic peptides from the 1/10 casein–albumin mixture for each column and condition. Three loading buffers were used for POROS- Fe^{3+} and TiO_2 chromatography: 0.1 M acetic acid (pK_a : 4.76), 0.1 M trifluoroacetic acid (pK_a : 0.3) and 0.7 M trifluoroacetic acid. Indeed, the use of 0.1 M acid concentration was mainly reported for IMAC, while TiO_2 has been used with higher concentrations of TFA, up to 0.7 M, in order to decrease non-specific binding [12]. The eluted fractions were analyzed by LC-MS and a recovery yield was determined for the peptides corresponding to three major ions (see Fig. 1). Two of them are monophosphorylated (VPQLEIVP-NpSAEER, m/z = 830.88 and FQpSEEQQTEDELQDK, m/z = 1031.89 from alpha-s1- and beta-casein, respectively) and one is a contaminant peptide from serum albumin (DAIPENLPLLTADFAEDK, m/z = 978.46). With a TiO_2 column, the decrease of pH has no effect on phosphopeptide recovery but decreases the binding of the contaminant peptides. With POROS- Fe^{3+} the decrease of pH lowers the recovery of phosphopeptides too.

3.2. Comparison of media and loading buffers with liver sample

We confirmed the results obtained with peptides from the casein–albumin mixture with a more complex mixture of peptides from mouse liver. A diagram showing the numbers of identified phosphopeptides and non-phosphorylated peptides for the different methods is presented in Fig. 2. It appears clearly that the use of trifluoroacetic acid as loading buffer increased from 18 to 60% and from 12 to 58% the enrichment specificity of IMAC and MOAC, respectively. However, trifluoroacetic acid dramatically decreased the number of identified phosphopeptides purified with the IMAC POROS- Fe^{3+} column, from 356 to 119 phosphopeptides. For this reason, we verified the content of the flow-through fraction by reloading it on the column (i.e. a second pass of the depleted sample). Whatever the affinity support and the buffer, the number of non-phosphorylated peptides was similar in the first and the second pass. However, phosphorylated peptides were mainly obtained in the first pass, excepted with POROS- Fe^{3+} /TFA for which 119 and 107 phosphopeptides were identified in pass 1 and 2, respectively. This clearly pointed out a problem of low binding capacity of POROS- Fe^{3+} IMAC with 0.1 M TFA loading buffer.

3.3. Leaching of Fe^{3+} from IMAC support and rust formation

Upon replacement of 0.1 M acetic acid loading solvent by 0.1 M TFA for the POROS- Fe^{3+} column, the number of identified peptides strongly decreased. Moreover, we identified 50% phosphopeptides in the flow-through. Therefore, we checked a potential Fe^{3+} release from the IMAC support by monitoring UV absorbance (see Fig. 3). When no FeCl_3 was used to activate the POROS IMAC support, the flow through produced almost no signal. On the contrary, after activation of POROS with FeCl_3 , a peak corresponding to Fe^{3+} leaching was produced by the injection of the loading solvents. This peak was moderate with 0.1 M acetic acid (pH 2.9), but saturated the detector when 0.1 M TFA (pH 1.3) was used as loading solvent. When the column was unpacked after 10 separations, we also observed

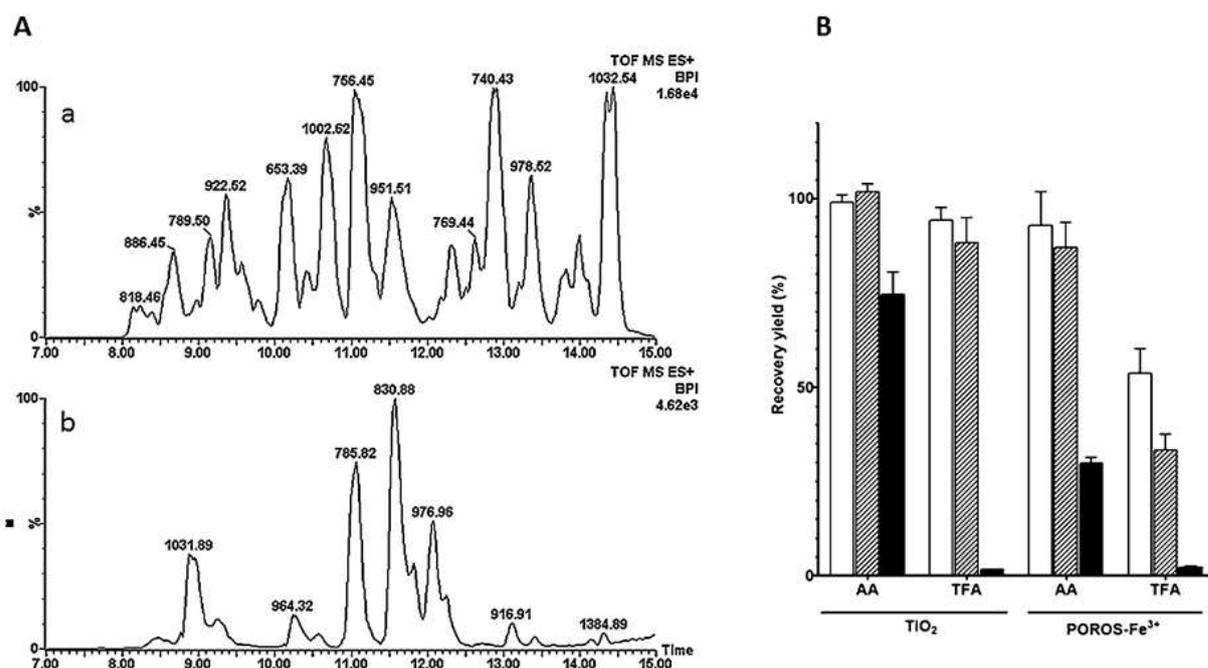


Fig. 1. Enrichment of casein phosphopeptides with POROS-Fe³⁺ and TiO₂ columns. Panel A: base peak chromatogram of casein–albumin sample before enrichment (a) and after enrichment (b) on TiO₂ column (peaks are labeled with m/z values, ion at m/z = 785.84 corresponds to spiked Glu 1-fibrinopeptide B). Panel B: recovery yield of peptides after TiO₂ and POROS-Fe³⁺ enrichment. Peptide VPQLEIVPNpSAEER (m/z = 830.88, white bar), FQpSEEQQTEDELQDK (m/z = 1031.89, striped bar) are phosphopeptides from alpha-s1-casein and beta-casein, respectively. Peptide DAIPENLPPLTADFAEDK (m/z = 978.49, black bar) is a non-phosphorylated peptide (i.e. a contaminant) from serum albumin. AA and TFA correspond to acetic acid and trifluoroacetic acid used in the loading buffer.

that the POROS support had taken an orange-brown color (see Supplementary Material). An extensive EDTA treatment did not wash the color away, because Fe³⁺ had formed a precipitate. This corresponds to rust formation: the elution buffer (0.5 M NH₄OH) led

to the formation of Fe(OH)₃ which spontaneously formed a brown precipitate of Fe₂O₃·3H₂O.

4. Discussion

IMAC-LC and mass spectrometry for phosphopeptides analysis was first reported by Watts et al. [2]. These authors used

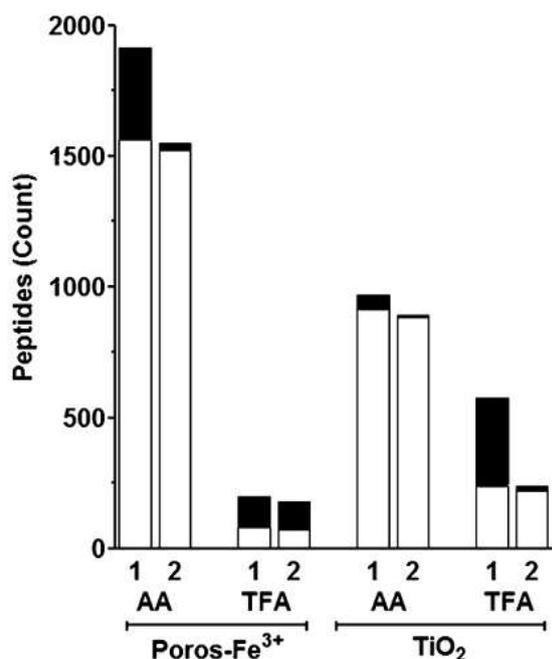


Fig. 2. Number of peptides identified after POROS-Fe³⁺ and TiO₂ purification. (A) Phosphopeptides (striped bar) and unphosphorylated peptides (open bar) are presented for two loading buffers: acetic acid (AA) and trifluoroacetic acid (TFA). Numbers 1 and 2 correspond to the first and the second pass of the sample onto column.

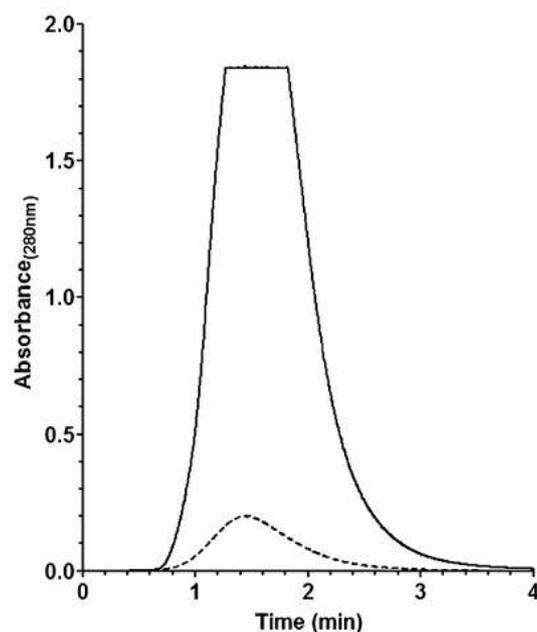


Fig. 3. (A) UV trace (280 nm) for POROS-Fe³⁺ after injection of 0.1 M acetic acid (AA; solid line) and 0.1 M trifluoroacetic acid (TFA; dashed line).

chelating-Sepharose with Fe^{3+} and 0.1 M acetic acid as loading buffer. It quickly appeared that contamination with peptides containing aspartic and glutamic acids was a major weakness of IMAC when it was applied to the purification of phosphopeptides from complex mixtures. Methyl esterification of carboxyl groups has been used to decrease the binding of mono-acidic peptides on IMAC supports [13], but side products (amine esterification) and incomplete esterification are known to increase the complexity of phosphopeptide fractions [14,15]. An alternative consists in lowering the pH of the loading buffer below 3, where D and E are mostly protonated (pK_a : 3.65 and 4.25 respectively) while phosphorylated residues still bear a charge and are able to bind to the IMAC- Fe^{3+} support [16]. However, the high selectivity obtained with a loading buffer at low pH also correlates with a low recovery of phosphopeptides [17,18]. The observation of Fe^{3+} leaching provides an explanation for the balance between IMAC selectivity (low pH) and recovery of phosphopeptides. Indeed, the use of a lower pH increases the salting-out of iron during the loading step of sample onto the column (see Fig. 3). This phenomenon then induces a decrease of phosphopeptide recovery (see Fig. 1), and concomitant occurrence of phosphopeptides in the flow-through fraction (see Fig. 2). Furthermore, we point out an important drawback of the use of Fe^{3+} IMAC enrichment of phosphopeptides by means of a packed, re-usable column: upon reaction with aqueous ammonia, Fe^{3+} precipitates during the elution step of phosphopeptides and accumulates in the column. If re-used, a column packed with IMAC- Fe^{3+} support may contain an unknown amount of Fe^{3+} and Fe_2O_3 , and such a "rusty IMAC" support, featuring a mixed IMAC and MOAC mode, may lead to irreproducible binding of phosphopeptides. To our knowledge, this effect has not been described in previous publications, including those involving an online IMAC- Fe^{3+} system [15,18–20]. Thus, it is highly recommendable to use only disposable IMAC columns and to avoid storing Fe^{3+} activated IMAC over a long period. By comparison, MOAC performed with TiO_2 packed columns does not present this drawback, since these columns may be used for tens of runs without loss of affinity even with very acidic loading solvents such as 0.7 M trifluoroacetic acid.

3.3.2 Résultats complémentaires

Stabilité des colonnes IMAC

Suite aux résultats obtenus avec la phase POROS 20 MC, une autre phase IMAC a été testée : le gel d'affinité PHOS-Select (Sigma). Les deux phases sont sensiblement différentes. En effet, la phase POROS 20 MC est constituée de particules de polystyrène - divinylbenzène qui chélatent les ions métalliques par l'intermédiaire d'un groupement imidoacétate, alors que la phase PHOS-Select est un gel d'agarose qui possède un groupement nitriloacétate capable d'interagir avec le métal par une fonction carboxylique supplémentaire. La **Figure 32** présente les chromatogrammes correspondant à l'injection de différents acides (acide acétique, acide formique, TFA) après activation des deux types de phases par FeCl_3 . Quelle que soit la phase utilisée, un pic est observé après injection d'une solution d'acide 0.1M. Son intensité est corrélée avec la force de l'acide

Acknowledgment

This work was supported by the Programme d'Actions Intégrées de Recherche – Carcinome Hépatocellulaire (PAIR-CHC).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jchromb.2012.02.028.

References

- [1] G. Muszynska, G. Dobrowolska, A. Medin, P. Ekman, J.O. Porath, J. Chromatogr. 604 (1992) 19.
- [2] J.D. Watts, M. Affolter, D.L. Krebs, R.L. Wange, L.E. Samelson, R. Aebersold, J. Biol. Chem. 269 (1994) 29520.
- [3] M.W. Pinkse, P.M. Uitto, M.J. Hilhorst, B. Ooms, A.J. Heck, Anal. Chem. 76 (2004) 3935.
- [4] A. Sano, H. Nakamura, Anal. Sci. 20 (2004) 565.
- [5] T.E. Thingholm, O.N. Jensen, P.J. Robinson, M.R. Larsen, Mol. Cell. Proteomics 7 (2008) 661.
- [6] B. Bodenmiller, L.N. Mueller, M. Mueller, B. Domon, R. Aebersold, Nat. Methods 4 (2007) 231.
- [7] T.S. Nuhse, A. Stensballe, O.N. Jensen, S.C. Peck, Mol. Cell. Proteomics 2 (2003) 1234.
- [8] J. Villen, S.A. Beausoleil, S.A. Gerber, S.P. Gygi, Proc. Natl. Acad. Sci. U.S.A. 104 (2007) 1488.
- [9] D.E. McNulty, R.S. Annan, Mol. Cell. Proteomics 7 (2008) 971.
- [10] J. Ye, X. Zhang, C. Young, X. Zhao, Q. Hao, L. Cheng, O.N. Jensen, J. Proteome Res. 9 (2010) 3561.
- [11] K. Moser, F.M. White, J. Proteome Res. 5 (2006) 98.
- [12] S.S. Jensen, M.R. Larsen, Rapid Commun. Mass Spectrom. 21 (2007) 3635.
- [13] S.B. Ficarro, M.L. McClelland, P.T. Stukenberg, D.J. Burke, M.M. Ross, J. Shabanowitz, D.F. Hunt, F.M. White, Nat. Biotechnol. 20 (2002) 301.
- [14] C.F. Xu, Y. Lu, J. Ma, M. Mohammadi, T.A. Neubert, Mol. Cell. Proteomics 4 (2005) 809.
- [15] S.B. Ficarro, A.R. Salomon, L.M. Brill, D.E. Mason, M. Stettler-Gill, A. Brock, E.C. Peters, Rapid Commun. Mass Spectrom. 19 (2005) 57.
- [16] C.F. Tsai, Y.T. Wang, Y.R. Chen, C.Y. Lai, P.Y. Lin, K.T. Pan, J.Y. Chen, K.H. Khoo, Y.J. Chen, J. Proteome Res. 7 (2008) 4058.
- [17] J. Lee, Y. Xu, Y. Chen, R. Sprung, S.C. Kim, S. Xie, Y. Zhao, Mol. Cell. Proteomics 6 (2007) 669.
- [18] Y. Iwase, S. Honma, M. Matsuzaki, Y. Miyakawa, T. Kanno, K. Ishii, N. Furuichi, K. Furukawa, T. Horigome, J. Biochem. 147 (2010) 689.
- [19] J. Wang, Y. Zhang, H. Jiang, Y. Cai, X. Qian, Proteomics 6 (2006) 404.
- [20] S. Feng, C. Pan, X. Jiang, S. Xu, H. Zhou, M. Ye, H. Zou, Proteomics 7 (2007) 351.

utilisé ; l'acide acétique produit un pic peu intense voir négligeable alors que l'acide formique et le TFA produisent des pics intenses qui saturent le spectrophotomètre. Par contre, sans activation des phases par FeCl_3 , les pics produits par les différents acides sont inférieurs à 0.1 DO ; l'absorbance mesurée est donc liée à l'éluion des ions métalliques. Lorsque le tampon de charge contient du TFA, la diminution du nombre de peptides identifiés par chromatographie IMAC s'explique par la perte d'ions Fe^{3+} qui, soit diminue la capacité de la colonne, soit entraîne une partie des phosphopeptides dans la fraction non retenue.

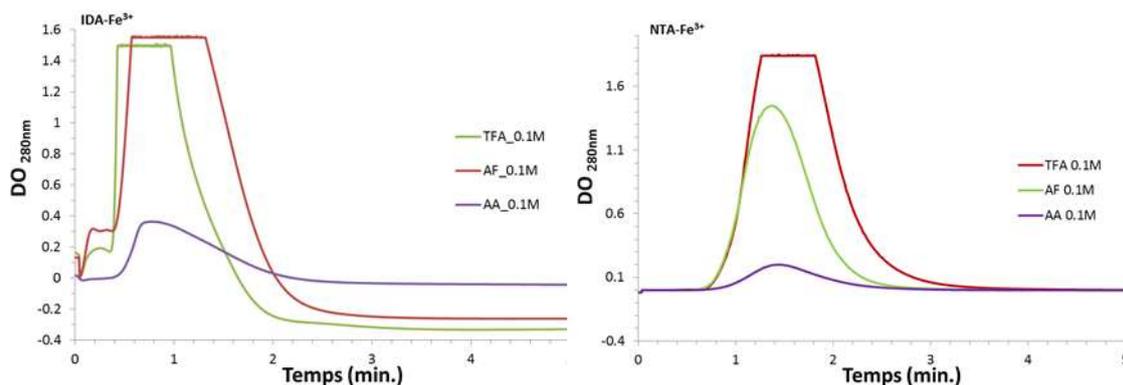


Figure 32. Injection de solution acides sur les colonnes IMAC-Fe³⁺. Les colonnes IDA-Fe³⁺ (POROS 20 MC) et NTA-Fe³⁺ (PHOS-Select) sont chargées en ions Fe^{3+} puis rincées par 50 mM acide acétique. Les chromatogrammes sont obtenus après injection de 0.1M TFA, 0.1 M acide formique (AF), 0.1M acide acétique (AA). L'injection des solutions d'acides sans activation préalable des phases par Fe^{3+} ne produit pas de pic d'absorbance.

L'utilisation des phases IMAC en colonne chromatographique et système HPLC est une source de contamination croisée des échantillons dont le risque est limité par une procédure systématique de régénération de la colonne avant l'injection d'un nouvel échantillon. La régénération est réalisée par l'injection d'EDTA en solution basique, puis la phase est rechargée en Fe^{3+} par une solution de FeCl_3 . Malgré cette régénération, il est apparu que la phase POROS-Fe³⁺ présentait une coloration orangée à brun qui s'intensifiait au cours des utilisations et qui était insensible à la régénération par l'EDTA (**Fig. 33**). Il semblerait donc que l'évolution de la colonne aboutisse à la formation d'oxydes de fer, plus communément appelés « rouille ».

Bien que la chromatographie « IMAC-oxyde de fer » permette la purification des phosphopeptides, cette phase pose un problème évident de reproductibilité. La quantité d'oxyde de fer augmente avec l'utilisation de la colonne ; elle n'est visible qu'après plusieurs cycles de régénération mais il est possible que l'oxyde soit présent dès la première chromatographie. Il faut noter ici que le chlorure de fer utilisé ici était hydraté. Sa nature hygroscopique fait qu'il est difficile de le maintenir anhydre et la présence d'eau favorise certainement son oxydation. La qualité du chlorure de fer a également été mise en avant pour expliquer une contamination plus ou moins importante de la

fraction purifiée (Ye et al., 2010). Prenant en compte d'une part l'impossibilité de garantir une phase solide IMAC homogène et stable dans le temps, et d'autre part les performances supérieures de la phase TiO_2 par rapport à la phase POROS- Fe^{3+} , les développements ultérieurs ont été réalisés exclusivement sur la colonne de TiO_2 .

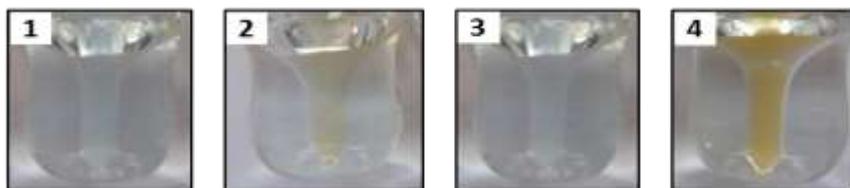


Figure 33. Evolution des phases IMAC au cours du temps. (1) phase POROS 20 MC neuve, (2) phase chargée en Fe^{3+} , (3) phase après une première régénération par l'EDTA, (4) phase après 20 cycles d'activation-régénération.

Influence du tampon de charge pour la chromatographie TiO_2

L'effet de l'acide acétique et du TFA, du chlorure de sodium et de l'acétonitrile dans le tampon de charge a été étudié pour un échantillon modèle contenant 1 mg/mL de peptides tryptiques de caséines et 10 mg/mL de peptides d'albumine. Cet échantillon a été dilué au dixième dans les différents tampons de charge et 500 μL d'échantillon ont été injectés, soit 550 μg peptides. Les quantités de caséines ont été fixées de manière arbitraire à 10 % de l'échantillon afin de permettre une analyse LC-MS quantitative à la fois pour les peptides de caséines et les éventuels peptides d'albumine, sans saturer le détecteur avec les peptides d'albumine. Sauf exception, tous les tampons de charge contiennent 20 % ACN.

Une injection de 500 μL de tampon d'élution (NH_4OH 0.7 M) a été systématiquement réalisée avant de charger l'échantillon sur la colonne de TiO_2 (2 mm x 5 cm). La colonne est rincée en continu par le tampon A (50 mM acide formique, 20 % ACN) et les peptides sont élués par NH_4OH 0.7 M. Par ailleurs, un échantillon « caséines seules » a également été purifié sur colonne de TiO_2 ; les résultats sont présentés dans le **Tableau 9**. Pour cet échantillon de caséines seules, l'intensité du pic d'élution diminue lorsque la concentration d'acide augmente dans le tampon de charge. A concentration d'acide égale, le pic d'élution est plus faible avec le TFA. Enfin, le chlorure de sodium et l'acétonitrile augmentent l'intensité du pic d'élution.

Tableau 9. Absorbance (280 nm) du pic d'éluion par NH₄OH 0.7 M de la colonne de TiO₂

Tampon de charge*	DO (280 nm)		
	Caséines (50 µg)	Caséines + albumine (50 µg + 500 µg)	Différence
AA 0.1M	0.3	>1.8	>1.6
AA 1M	0.1	1.3	1.2
TFA 0.01M	0.3	1.8	1.5
TFA 0.1M	0.1	0.8	0.7
TFA 0.65M (5 %)	0.1	0.5	0.4
AA 0.1M, NaCl 0.1M	0.2	1.8	1.6
TFA 0.1M, NaCl 0.1M	0.2	1.5	1.3
TFA 0.1M dans 80 % ACN	0.2	>1.8	>1.6

AA : acide acétique, TFA : acide trifluoroacétique, ACN : acétonitrile

A partir de ces résultats, deux hypothèses sont possibles : soit le tampon correspondant à la DO la plus faible (TFA 0.65 M) diminue la liaison des contaminants, soit il inhibe la liaison des phosphopeptides. La comparaison de l'échantillon de caséines et de l'échantillon caséines+albumine montre que l'intensité des pics d'éluion est bien supérieure avec le deuxième échantillon, bien que la même quantité de caséines ait été injectée sur la colonne. La détection UV met donc en évidence une contamination importante par des peptides d'albumine de la fraction purifiée. Si l'on compare les données obtenues avec un tampon de charge TFA 0.65 M pour les 2 échantillons, il apparaît que l'absorbance attribuable à la contamination par l'albumine est égale à (0.5 – 0.1) soit 0.4 DO. Cette valeur peut être comparée avec celles des autres conditions (**Tab. 9**, dernière colonne). En considérant que la courbe dose/réponse du spectrophotomètre est linéaire à de telles valeurs de DO, il apparaît que le tampon de charge TFA 0.65 M permet de diminuer d'un facteur 3 la contamination par rapport à l'acide acétique 1 M. Inversement, l'addition de 0.1M NaCl dans le tampon 0.1 M TFA augmente la contamination d'un facteur 2 par rapport au tampon 0.1 M TFA seul. Une forte concentration de solvant aprotique (80 % ACN) augmente également la contamination d'au moins un facteur 3. Ici, Le tampon de charge le plus favorable serait donc 0.65M TFA.

Les différentes fractions purifiées ont été analysées par LC-MS (CapLC-Q-TOF premier). Préalablement à ces analyses, la variation de signal entre analyses LC-MS (45 min) a été évaluée par une série de 30 injections de Glu-1-Fibrinopeptide B (GFP). L'aire sous la courbe moyenne (AUC) pour une injection de 10 µL de GFP à 2.5 µg/mL est égale à 40 ± 1.48 (m ± σ), soit un coefficient de variation de 28 %. Afin de limiter l'effet de la variation entre analyses, du GFP (2.5 µg/ml final) a été ajouté à chaque échantillon et les AUC des différents ions ont été normalisées par l'AUC du GFP.

Chaque échantillon a été analysé 3 fois, la **Figure 34** présente les AUC moyennées de 3 peptides phosphorylés de caséines et de 4 peptides contaminants provenant de l'albumine.

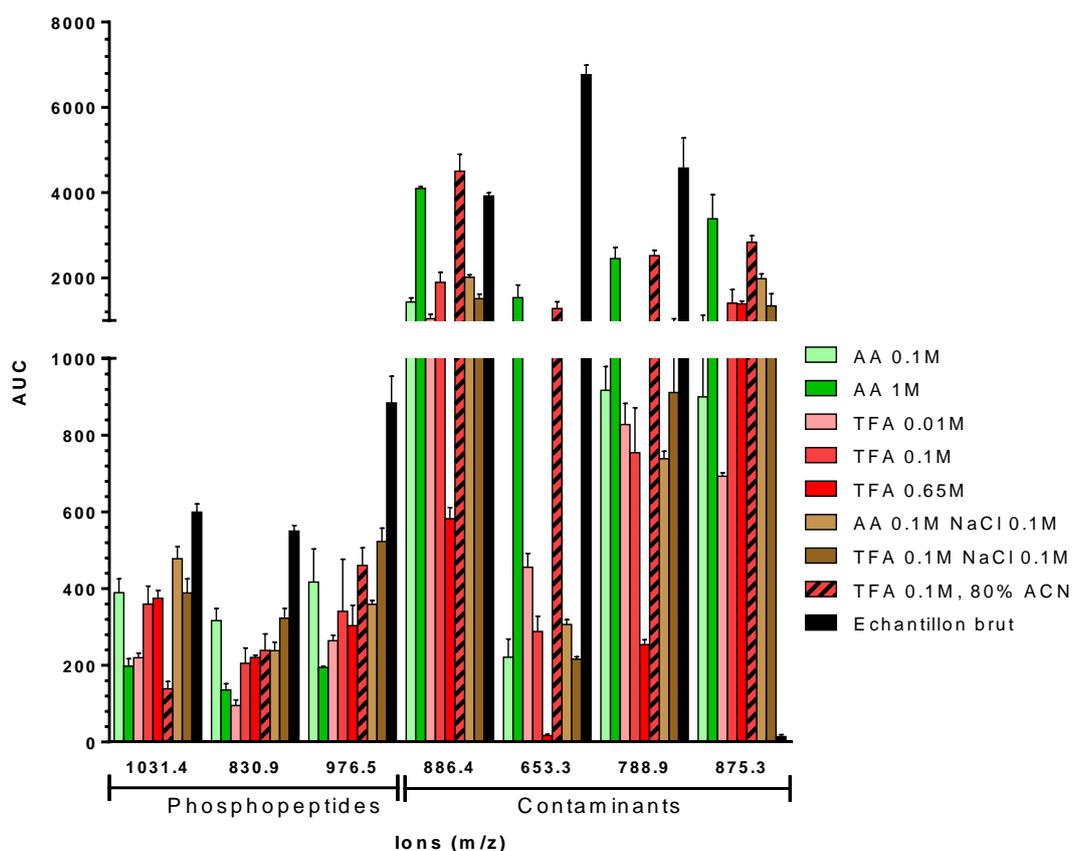


Figure 34. Intensité des peptides phosphorylés de caséines et des peptides contaminants. Les peptides sont notés par leur m/z. Phosphopeptides : FQpSEEQQTEDELQDK (m/z = 1031.4), VPQLEIVPNpSAEER (m/z = 830.9), YKVPQLEIVPNpSAEER (m/z = 976.45). Contaminants : DDSPDLPK (m/z = 886.5), HLVDEPQNLIK (m/z = 653.3), LKPDNTLCDEFK (m/z = 788.8), ECCHGDLLECADDR (m/z = 875.3). Tampon de rinçage : 50 mM AA, 80 % ACN, Elution : 0.7 M NH₄OH.

Pour l'échantillon avant purification (échantillon brut, barre noire), les AUC des peptides de caséines sont plus faibles que celles des peptides d'albumine d'un facteur 10, en accord avec la proportion relative de caséines par rapport à l'albumine. Il faut cependant noter que les AUC dont les valeurs sont supérieures à 4000 correspondent à des ions qui ont saturé le détecteur (intensité > 3x10³). A part ces cas particuliers où la quantification est sous-estimée, les valeurs d'AUC d'un même peptide obtenues pour les différentes conditions peuvent être comparées.

Il apparaît que les peptides contaminants génèrent des AUC supérieures à celle des peptides phosphorylés après purification sur colonne de TiO₂, excepté lorsque le tampon de charge contient TFA 0.65M. Ces résultats sont cohérents avec ceux obtenus à partir de la détection UV. On peut juste noter que la différence entre les tampons de charge TFA 0.01M et 0.1M est plus marquée avec la détection UV que la détection MS. Pour les mesures MS, l'effet sélectif du TFA n'est sensible

qu'à une concentration supérieure à 0.1 M. Il en est de même pour l'effet du chlorure de sodium que l'analyse MS semble sous-estimer par rapport à la détection UV. Par contre, l'utilisation de 80 % d'acétonitrile dans le tampon de charge augmente significativement le signal des contaminants, à la fois en détection MS et UV.

Puisque le TFA permet de réduire significativement la contamination en peptides non phosphorylés, le tampon A (50 mM AA, 20 % ACN) a été remplacé par le tampon B (50 mM TFA, 20 % ACN). Ce tampon est appliqué pendant 20 min après chaque injection. L'AUC des contaminants diminue alors considérablement, mais de manière inégale (**Fig. 35**).

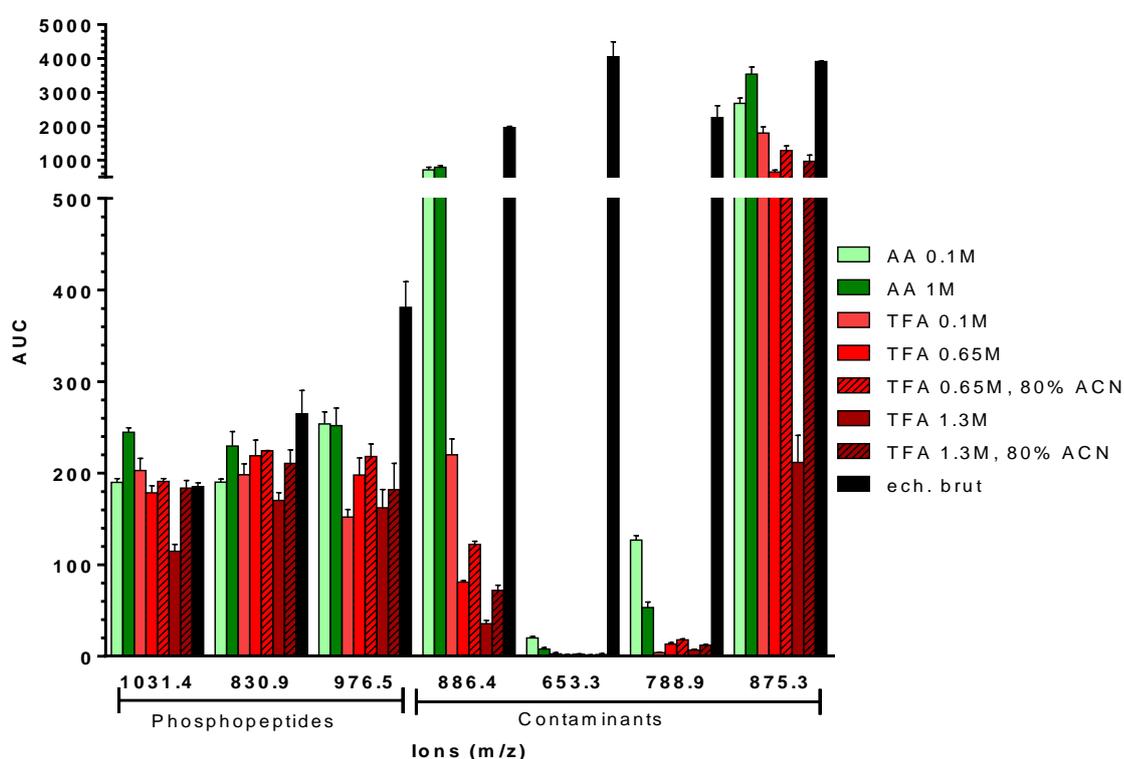


Figure 35. Intensité des ions majoritaires, peptides phosphorylés de caséines et peptides contaminants. Tampon de rinçage B : 50 mM TFA, 80 % ACN, élution : 0.7 M NH₄OH.

La substitution dans le tampon de rinçage de l'acide acétique par le TFA permet d'éliminer complètement une famille de contaminants représentée par l'ion 653.4. D'autres peptides acides contaminent encore l'éluat et nécessitent l'utilisation de TFA concentré (0.65M ou 1.3 M) dans le tampon de charge (ions 886.4 et 788.89). Enfin, certains contaminants représentés par l'ion 875.3 ont une affinité suffisante par la phase TiO₂ pour que le TFA concentré ne suffise pas à les éliminer de la colonne lors du chargement. Ici encore, l'utilisation de 80 % ACN favorise la liaison des contaminants au TiO₂. Ce phénomène est particulièrement visible pour le peptide de m/z 875.3 (ECCHGDLLECADDR) qui contient 5 résidus acides D et E. Alors que seul le TFA 1.3 M permet de

diminuer sensiblement sa liaison au TiO₂, l'ajout de 80 % ACN restaure entièrement la contamination.

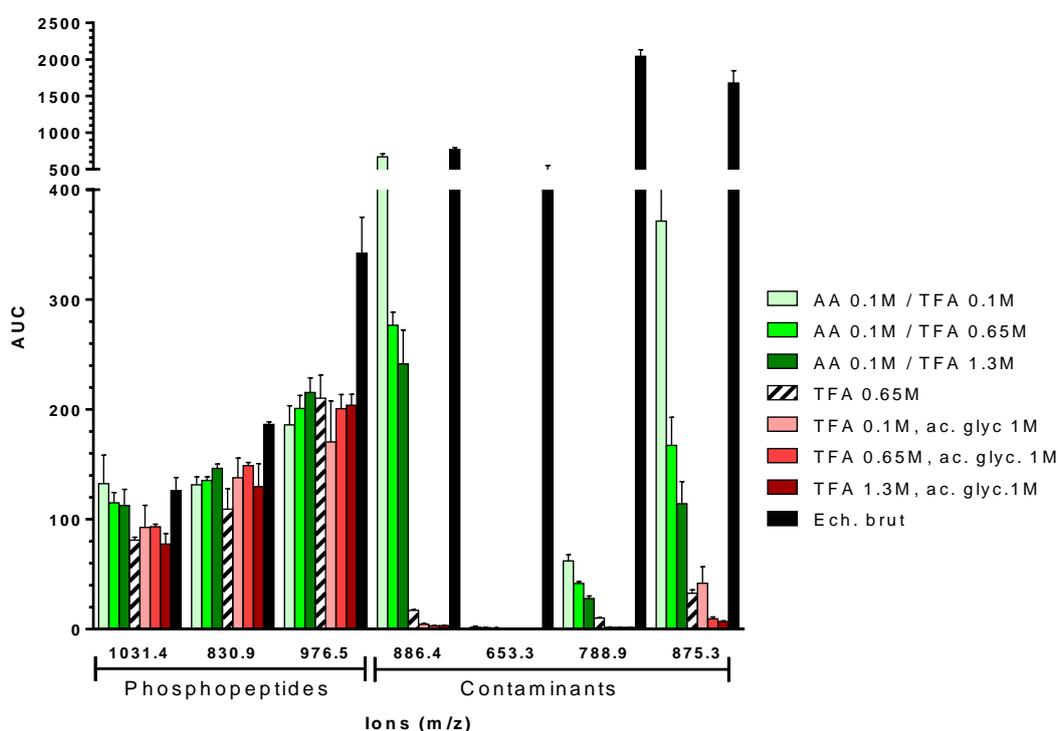


Figure 36. Intensité des ions majoritaires, peptides phosphorylés de caséines et peptides contaminants. Tampon de rinçage B : 50 mM TFA, 20 % ACN, élution : 0.7 M NH₄OH.

Une troisième série d'analyses a été réalisée également avec le tampon B (50 mM TFA, 20 % ACN) avec l'objectif de favoriser la fixation des peptides phosphorylés et de diminuer la contamination par les peptides acides (**Fig. 36**). Pour les trois premières conditions, l'échantillon a été dilué dans une solution d'acide acétique 0.1M et injecté sur la colonne. Puis une deuxième injection de solution très acide (TFA 0.1 M, 0.65 M ou 1.3 M) a été réalisée afin d'éliminer les contaminants. La comparaison des résultats avec une injection de l'échantillon directement dilué dans le TFA 0.65 M (barre hachurée) montre que les contaminants sont beaucoup plus intenses lorsque le chargement est réalisé en acide acétique 0.1M. Apparemment, le TFA ne permet pas d'éliminer les contaminants avec la même efficacité lorsqu'ils sont déjà adsorbés sur la colonne d'affinité. Par ailleurs, l'effet de l'acide glycolique sur la liaison des contaminants a été vérifié pour les trois concentrations de TFA. L'acide glycolique permet une diminution significative du signal des contaminants, le tampon de charge optimal contenant du TFA 0.65 à 1.3M et de l'acide glycolique 1M.

Purification des phosphopeptides, extrait de foie de souris

La purification des phosphopeptides sur colonne de TiO₂ a été validée par l'analyse d'un échantillon de foie de souris. L'extrait de peptides tryptiques (500 µg) a été dilué dans différents tampons de charge contenant du TFA 0.65 M ou 1.3 M et de l'acide glycolique 0.5 M ou 1 M. Les fractions éluées par NH₄OH 0.7M ont été dessalées puis analysées par un système nanoLC-Orbitrap (LTQ Orbitrap XL). Une méthode d'analyse rapide a été mise en œuvre, en utilisant une chromatographie d'une heure avec une méthode d'acquisition appelée « TOP10 » consistant en 1 scan MS suivi de 10 scans CID des 10 ions les plus intenses du scan MS. Suivant les tampons de charge utilisés, le logiciel Proteome Discoverer 1.3 permet d'identifier à partir de la totalité des séquences de la souris (génomme complet de *Mus musculus*, 42882 séquences) entre 237 et 513 peptides avec un taux d'erreur (FDR) inférieur à 1 % (**Tab. 10**).

Tableau 10. Peptides identifiés à partir d'un extrait de foie de souris purifié par TiO₂.

Peptides	Tampon			
	TFA 0.65 M	TFA 0.65 M + ac. glyc. 1 M	TFA 1.3 M	TFA 1.3 M + ac. glyc. 1 M
0 phosphate	99	125	196	113
1 phosphate	121	296	289	303
2 phosphates	14	43	26	34
3 phosphates	2	2	2	2
Phospho	137	341	317	339
Phosphopep. (%)	58	73	62	75

Sans acide glycolique, 58 et 62 % des peptides identifiés sont phosphorylés pour un tampon de charge contenant respectivement 0.65 M et 1.3M TFA. Lorsque l'acide glycolique est ajouté à la solution de TFA, 73 à 75 % des peptides identifiés sont phosphorylés, soit une augmentation de 10 % de la proportion de phosphopeptides, quel que soit la concentration de TFA. Pour un échantillon réel, l'utilisation d'acide glycolique permet de diminuer de 13 à 15 % le nombre de peptides contaminants. En conséquence, un plus grand nombre de phosphopeptides sont analysés et identifiés.

Purification des phosphopeptides, extrait de biopsie humaine

Une vérification a également été réalisée avec une biopsie humaine disponible en quantité importante (**Tableau 11**). Pour des conditions d'acquisition identiques (TOP10), le gradient HPLC a

été rallongé de 1h à 2h, permettant ainsi d'identifier un plus grand nombre de peptides. Les résultats sont comparables à ceux obtenus avec le foie de souris. Avec 0.65M TFA, l'ajout d'acide glycolique permet d'augmenter le nombre de phosphopeptides identifiés ainsi que le pourcentage de phosphopeptides. Avec 1.3M TFA, l'acide glycolique ne permet que de diminuer la contamination ; le pourcentage de phosphopeptides passant de 58 à 65 %. Pour les deux concentrations de TFA, l'acide glycolique permet une augmentation de la pureté (exprimée en % de phosphopeptides) d'environ 10 %.

Par ailleurs la comparaison des pourcentages de phosphopeptides entre foie de souris et foie humain met en évidence une différence significative entre les deux types d'échantillons : le pourcentage de phosphopeptides est systématiquement plus faible pour l'extrait de foie humain.

Tableau 11. Peptides identifiés à partir d'un extrait de biopsie humaine purifié par TiO₂

Peptides	Tampon			
	TFA 0.65M	TFA 0.65M 1M Glyc.	TFA 1.3M	TFA 1.3M 1 M Glyc.
0 Phos.	191	215	442	325
1 Phos.	207	391	555	529
2 Phos.	20	47	60	69
3 Phos.	2	4	4	6
Phosphopeptides	229	442	619	604
phospho. (%)	55	67	58	65

3.4 Discussion

Les chromatographies d'affinité IMAC et MOAC sont les techniques de purification des phosphopeptides les plus répandues. Ce chapitre a permis de vérifier les données précédemment publiées par Larsen et collaborateurs (Larsen et al., 2005) ; le remplacement de l'acide acétique 0.1-0.2 M (pH 2.7-2.9) par le TFA 0.1 % (pH 1.9) permet de diminuer la contamination des peptides riches en acides aminés D et E. Les auteurs expliquent que les fonctions acides carboxyliques des acides aminés D et E sont sous forme -COOH à pH 1.9 (TFA 0.1 %), l'absence de charge négative inhibant leur liaison au TiO₂. Par contre, le phosphoryle lié aux acides aminés hydroxylés a un pKa inférieur à 1.9 et conserve une charge négative qui permet sa liaison au dioxyde de titane.

Ici, nous observons la même augmentation de spécificité lorsque le TFA substitue l'acide acétique mais les résultats obtenus montrent que l'interaction ionique n'est pas la force principale de la liaison TiO_2 -phosphopeptide. Si l'utilisation d'un tampon de rinçage très acide permet effectivement de diminuer la liaison entre le TiO_2 et les peptides acides riches en D et E, il persiste une contamination importante de ces peptides dans la fraction purifiée, y compris lorsque des solutions de TFA 0.65 M et 1.3 M sont utilisées, soit à un pH respectif de 0.4 et 0.2. De même, l'ajout de NaCl 0.1 M est sans effet pour à la fois la liaison des phosphopeptides et la liaison des contaminants. Le caractère de la liaison TiO_2 -phosphopeptides ne semble donc pas (ou peu) être de nature ionique.

La chimie de surface de la phase TiO_2 est extrêmement complexe. En effet, deux types de sites hydroxyle (-OH) sont présents sous différentes formes en fonction du pH, ainsi que 4 types de site Ti^{4+} (acide de Lewis)(Nawrocki et al., 2004). Ces sites permettent à la fois une chromatographie d'échange d'ions et une chromatographie d'échange de ligand dans laquelle les liaisons de coordination entre Ti, (OH), (H_2O) et L^- forment des complexes de type $\text{Ti}(\text{OH})(\text{H}_2\text{O})$, $\text{Ti}(\text{OH})\text{L}^-$, $\text{Ti}(\text{H}_2\text{O})\text{L}^-$, L^- étant une base de Lewis comme par exemple le phosphate des phosphopeptides (Blackwell and Carr, 1992). La chromatographie d'échange de ligands semble effectivement avoir un rôle prépondérant dans les conditions employées pour la purification des phosphopeptides. Elle permet d'expliquer l'interaction forte entre le phosphate des phosphopeptides et la phase TiO_2 à des pH très inférieurs au pKa de l'acide phosphorique, mais le caractère anionique du ligand (i.e. le phosphate du phosphopeptide) ne semble pas nécessaire ici pour permettre les liaisons de coordination.

Les résultats obtenus avec 80 % d'acétonitrile dans le tampon de charge auront également montré qu'une interaction de type hydrophile existe lorsque l'éluant contient une concentration élevée de solvant aprotique. Logiquement, cette interaction hydrophile ne modifie pas ou peu la liaison des phosphopeptides au TiO_2 . Par contre, elle favorise la rétention des peptides acides sur la phase TiO_2 et rétablit une contamination importante de la fraction purifiée malgré l'utilisation de TFA. Lors de l'utilisation de tampon peu acides (tampon A et tampon de charge contenant 0.1M TFA), la chromatographie sur TiO_2 avec 80 % ACN fonctionne dans un mode mixte SAX- HILIC qui favorise particulièrement la contamination par des peptides acides (**Fig. 34**). Lorsque l'acide acétique 50 mM est remplacé par le TFA 50 mM dans le tampon de rinçage, une première population de contaminants dits « faibles » est éliminée. L'augmentation des concentrations de TFA à 0.65 et 1.3M permet d'éluer une deuxième population de peptides acides mais semble également

diminuer la liaison du phosphopeptide FQpSEEQQTEDELQDK (**Fig.35**). Ce phosphopeptide présente un caractère très hydrophile (indice GRAVY = -2.5). Son rendement de purification augmente logiquement lorsque 80 % d'acétonitrile sont ajoutés au tampon de charge. Le mode de chromatographie devient mixte avec une interaction de type HILIC favorisée par le fort pourcentage d'acétonitrile. Cependant l'interaction hydrophile n'est pas spécifique et la contamination des peptides acides est également augmentée, spécialement pour le peptide ECCHGDLLECADDR ($m/z = 875.3$). Plusieurs articles présentent des tampons de charge optimisés contenant 80 % ACN (Han et al., 2012; Jensen and Larsen, 2007; Larsen et al., 2005; Thingholm et al., 2006). Dans ces publications, pour lesquelles la « colonne » est une pointe de pipette type *gel loader* remplie de TiO_2 sur sa partie effilée, l'acétonitrile a été utilisé pour éviter que les phosphopeptides ne se fixent au fritté, qui est constitué d'un morceau de filtre Empore C_8 (3M), (Larsen et al., 2005). Il est par ailleurs intéressant d'utiliser l'acétonitrile pour diminuer considérablement la contre-pression due à la phase (résultats personnels). Dans nos mains, cette solution a entraîné une contamination importante ; il apparaît clairement que la phase TiO_2 fonctionne alors dans un mode HILIC qui favorise une contamination difficile à éliminer.

Dans une logique consistant à favoriser la liaison des phosphopeptides au TiO_2 , l'échantillon (caséines + albumine) a été dilué dans un tampon à faible force éluante (acide acétique 0.1M), puis une deuxième injection de TFA a été réalisée afin d'éliminer les peptides contaminants (**Fig. 36**). La pureté de la fraction enrichie en phosphopeptides est alors bien inférieure à celle obtenue lorsque l'échantillon est injecté directement dans du TFA concentré, sans pour autant que le rendement de purification des phosphopeptides ne soit augmenté significativement. Comme pour le tampon de charge contenant 80 % ACN, les peptides contaminants sont plus difficiles à éliminer lorsqu'ils sont déjà fixés au TiO_2 . Il est possible que l'interaction HILIC oriente les contaminants au contact de la phase solide et favorise dans un deuxième temps une interaction par échange de ligand, dont l'énergie de liaison est suffisante pour rendre les éluants faibles de type acide peu ou pas efficaces. L'éluion n'est alors réalisée que par une base de Lewis, c'est à dire les ions hydroxyle de l'ammoniaque.

Finalement, la capacité de l'acide glycolique à inhiber les liaisons non spécifiques aura été évaluée pour l'échantillon (caséines+ albumine) et un extrait de foie de souris. L'utilisation de cet acide a été motivée par les travaux du groupe de P. Roepstorff sur les *non phosphopeptides excluders* (Larsen et al., 2005), des composés acides dont l'affinité pour TiO_2 permet d'éliminer les contaminants sans affecter la liaison de phosphopeptides. Les *excluders* les plus efficaces sont de

nature aromatique : DHB, acide phtalique, acide salicylique. Les auteurs expliquent que chaque phosphate lie deux atomes de titane alors que les fonctions carboxyliques des contaminants ou des acides aromatiques lient un seul titane par deux liaisons de coordination. Les géométries de ces liaisons de coordination étant différentes, les auteurs supposent que les sites de liaison des contaminants acides et les sites de liaison des phosphates sont différents.

Si les acides aromatiques sont des *excluders* efficaces, ils ne sont utilisés que pour des analyses MALDI-TOF, l'utilisation du DHB prenant un intérêt tout particulier dans la mesure où il est à la fois un *excluser* et une matrice de choix pour l'ionisation MALDI. Malheureusement ces acides aromatiques sont incompatibles avec une analyse LC-MS car ils contaminent la RP-HPLC et l'entrée du spectromètre de masse (Jensen and Larsen, 2007). Les auteurs ont alors proposé l'utilisation de d'acide glycolique 1M ou d'acide lactique 1M, deux acides portant un groupement hydroxyle en alpha du groupe carboxylique. A partir d'un extrait de cellules mésenchymateuses et l'utilisation d'un tampon de charge 5 % TFA, 1 M acide glycolique, 80 % ACN, ils identifient 336 phosphopeptides avec une pureté de 89 % (i.e. 89 % des peptides identifiés sont phosphorylés). Ici un extrait de foie permet d'obtenir 78 % de phosphopeptides identifiés avec un tampon de charge contenant 5 % TFA, 1M acide glycolique, 20 % ACN, soit une augmentation de 15 % par rapport au tampon sans acide glycolique. Il faut noter que l'augmentation de la concentration de TFA de 5 à 10 % dans le tampon de charge augmente la pureté de la fraction purifiée de seulement 5 %. Il semblerait donc que l'acide glycolique permette effectivement la diminution du nombre de peptides contaminants dans l'éluat des colonnes TiO₂, bien que l'utilisation des *excluders* ne soit pas l'élément clef de la purification sur colonne de TiO₂, au contraire du TFA qui permet d'augmenter la pureté de 12 à 58 % lorsqu'il remplace l'acide acétique dans le tampon de charge (**Article, Fig. 2**).

D'un point de vue pratique, ce chapitre avait pour objectif de définir le choix d'une chromatographie pour la purification de phosphopeptides provenant de biopsies hépatiques humaines. Ces dernières seront purifiées par chromatographie sur TiO₂ avec un tampon de charge contenant 5 % TFA, 1M acide glycolique, 20 % ACN. D'un point de vue plus théorique, ce chapitre aura montré que la chromatographie TiO₂ est une chromatographie de phase solide hétérogène qui permet l'échange d'ions, l'échange de ligands et l'interaction hydrophile, le choix des tampons chromatographiques permettant de favoriser certains types d'interactions. Ces propriétés originales présentent un inconvénient puisqu'elles contribuent à la contamination des phosphopeptides par des peptides riches en D et E, mais elles offrent également la possibilité

d'ouvrir le champ d'application des phases TiO_2 à la chromatographie de molécules polaires et hydrophobes.

CHAPITRE 4 :
analyse protéomique et phosphoprotéomique de biopsies hépatiques
dans le cadre du carcinome hépatocellulaire sur foie non fibreux

4.1 Présentation du sujet

4.1.1 Le foie et ses fonctions essentielles

Avec un poids moyen de 1,5 Kg, le foie est le plus volumineux des viscères. Il est enveloppé dans une capsule conjonctive, la capsule de Glisson, qui s'invagine et forme des lobes dont l'organisation est étroitement reliée à son système vasculaire. Son alimentation sanguine se fait par l'artère hépatique et par la veine porte. Cette vascularisation afférente de type veineux est particulière au foie et s'explique par son activité de drainage de l'ensemble des viscères. L'ensemble veine porte - artère hépatique - canal cholédoque constitue le pédicule hépatique. La ramification de ce pédicule permet d'isoler plusieurs régions possédant chacune son système de vascularisation. La connaissance de ces segments permet des ablations sans porter préjudice aux segments avoisinants.

Organisation du lobule hépatique

La coloration par l'acide périodique est utilisée en histologie pour révéler les polysaccharides. Sur les coupes de foie, elle met en évidence une organisation en lobule, le glycogène apparaissant sous forme d'un gradient de concentration décroissant, de la périphérie vers le centre du lobule où est localisée une veine, la veine centrolobulaire. Ce gradient de glycogène suit le flux sanguin qui va de la périphérie du lobule hépatique au centre du lobule où il est collecté. Le lobule hépatique est l'unité fonctionnelle du foie, constitué essentiellement d'hépatocytes organisés en travées entre lesquelles circule le sang veineux vers la veine centrolobulaire. Au sein des travées, dans des canalicules, la bile circule en sens inverse vers le canal biliaire (**Fig. 37**). L'organisation du lobule hépatique est donc déterminée par la disposition des vaisseaux et des voies biliaires. Chaque lobule est séparé de son voisin par des travées de tissu conjonctif. L'espace porte ou espace de Kiernan est situé dans ce tissu conjonctif ; il est caractérisé par une triade veine – artériole - canal biliaire qui permettent respectivement (i) l'apport de nutriments par les branches de la veine porte (ii) l'apport d'oxygène par l'artère hépatique (iii) la collecte des sels biliaires par le canal biliaire.

A partir de l'espace porte se développent donc des ramifications appelées capillaires sinusoidaux qui pénètrent dans le lobule hépatique et convergent vers le centre. Cet ensemble de vaisseaux constitue le système sanguin afférent. Au centre du lobule hépatique, la veine

centrolobulaire collecte le sang des capillaires sinusoidaux. Elle constitue le premier maillon du système veineux efférent qui aboutit aux veines sus-hépatiques puis à la veine cave inférieure.

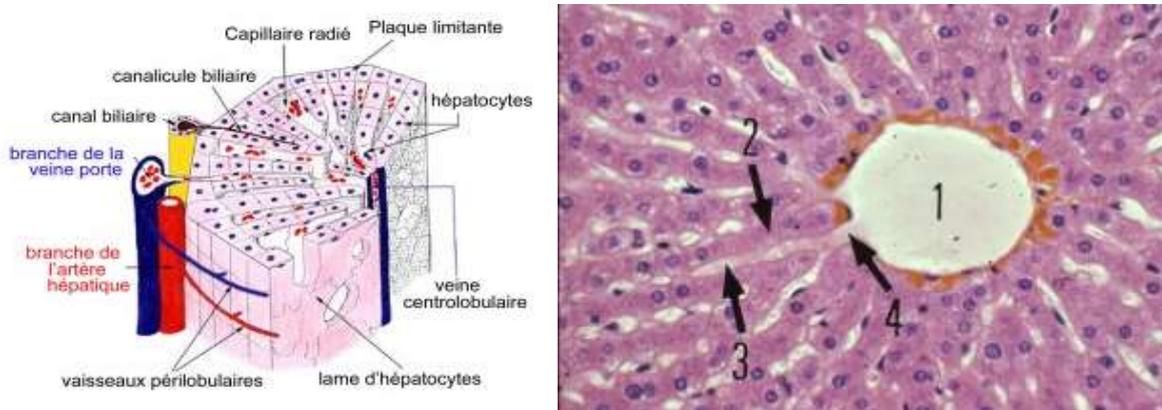


Figure 37. Organisation du lobule hépatique. A gauche : schéma d'un lobule hépatique Les cellules roses représentent les hépatocytes, les points rouges les érythrocytes qui circulent dans les capillaires sinusoides. L'ensemble veine-artériole-canal biliaire caractérise l'espace porte. (source <http://www.transhepate.org>). A droite : microscopie à fond clair d'une coupe de lobule hépatique. La veine centrolobulaire au centre (1) est entourée d'une couche de cellules épithéliales colorées en orange. Les hépatocytes colorés en rose constituent la plus grande partie du volume lobulaire. Ils sont organisés en rangées de cellules appelées travée de Remak (2) qui entourent les cellules épithéliales des canaux sinusoidaux. Source : Faculté de Médecine Pierre & Marie Curie, Paris, <http://www.chups.jussieu.fr/polys/histo/histoP2/foie.html>.

Types cellulaires

Les hépatocytes représentent 80 % des cellules du foie. Les autres types cellulaires sont les cellules des canaux biliaires, les cellules endothéliales, les cellules de Küppfer (macrophages), les cellules étoilées ou cellules de Ito (fonction métabolique de la vitamine A et des lipides, fabrication de la matrice extra-cellulaire autour de cellules endothéliales), les lymphocytes hépatocytaires, et les cellules ovales (cellules pluripotentes, fonction de régénération des hépatocytes et endothéliales).

Fonctions essentielles

Le foie est un organe vital qui assure quatre fonctions essentielles :

- Une fonction de synthèse (i) des protéines plasmatiques : albumine, facteurs de coagulation, protéines du métabolisme du fer (ferritine, transferrine), protéines de l'inflammation (α 1- et α 2-globulines), (ii) des lipides (cholesterol, triglycérides, lipoprotéines) (iii) des glucides (glycogénogenèse), (iv) des corps cétoniques produits lors du catabolisme incomplet des acide gras pendant le jeûne.

- Une fonction de stockage des glucides sous forme de glycogène et des vitamines liposolubles.
- Une fonction de détoxification : conversion de l'ammoniac en urée (cycle de l'urée), glucuroconjugaison de la biliburine et de toxines, dégradation des toxines et médicaments par l'intermédiaire des cytochromes P450.
- Une fonction de production de la bile qui est constituée principalement de sels biliaires (acides dérivés de l'oxydation du cholestérol), de cholestérol, de bilirubine conjuguée.

4.1.2 Le carcinome hépatocellulaire

Introduction

Epidémiologie.

On peut diviser les cancers du foie en deux catégories : les cancers primitifs et les cancers secondaires. Les cancers primitifs du foie, dont fait partie le carcinome hépatocellulaire, peuvent être de type bénin (adénome, hépatome) ou de type malin, carcinome hépatocellulaire (CHC), hépatoblastome (chez le jeune enfant), cholangiocarcinome (cancer de l'épithélium des voies biliaires), angiocarcinome (sarcome des tissus de soutien des vaisseaux). Au niveau mondial, les cancers primitifs du foie, dont 80 à 95 % sont des CHC, correspondent à la troisième cause de mortalité due au cancer, avec approximativement un nombre de décès équivalent au nombre de cas (560 000 cas en 2008) (Nordenstedt et al., 2010). En France, parmi les 355 350 nouveaux cas de cancer estimés en 2012, le cancer du foie représente 8723 cas, 79 % survenant chez l'homme (Binder-Foucard et al., 2013). Les données relatives à la mortalité de ce cancer ne sont pas publiées dans cette dernière étude de l'INCA ; pour l'année 2005, les données brutes sont de 5104 cas diagnostiqués, alors que 5405 décès ont été enregistrés (Belot et al., 2008).

Les facteurs de risques varient en fonction du sexe et de la localisation géographique. En France, l'alcoolisme, l'hépatite B et C, le syndrome métabolique (IMC élevé, taux élevé de triglycérides, taux faible de HDL, glycémie élevée, hypertension artérielle) sont les principaux facteurs de risque. En Europe, une étude épidémiologique a montré que les deux tiers des CHC étaient associés à un des 5 facteurs de risque suivants : tabac (47.8 %), hépatite C (20.9 %), l'obésité (16.1%), Hépatite B (13.2 %), et l'alcoolisme (10.1%). (Trichopoulos et al., 2011). Une autre étude équivalente réalisée en Italie présente des résultats plus communément admis : le risque de

développer un CHC est associé à l'hépatite C (65%), l'alcoolisme (24%), hépatite B (15%), ces trois facteurs représentant 89% des cas (Polesel et al., 2012).

Les cancers secondaires du foie sont des tumeurs liées au développement dans le foie de cellules cancéreuses provenant d'autres organes. A cause de l'effet de drainage par la veine porte, ces cancers secondaires sont souvent associés aux cancers digestifs (estomac, pancréas, intestin, côlon, rectum).

Fibrose et cirrhose hépatique

La fibrose est la formation de tissus fibreux qui désorganisent le tissu hépatique. Lors d'une inflammation, les fibroblastes prolifèrent et produisent de la matrice extracellulaire riche en collagène, formant ainsi un tissu de cicatrisation qui normalement se résorbe progressivement. Dans le cas de la fibrose hépatique qui se développe lors de la destruction des cellules hépatiques, le développement de la matrice extracellulaire désorganise et incapacite les lobules (Friedman, 2008). Son évolution est classée en stades, le stade F0 correspondant à l'absence de fibrose, le stade F4 correspond au premier stade de la cirrhose (**Fig. 38**).

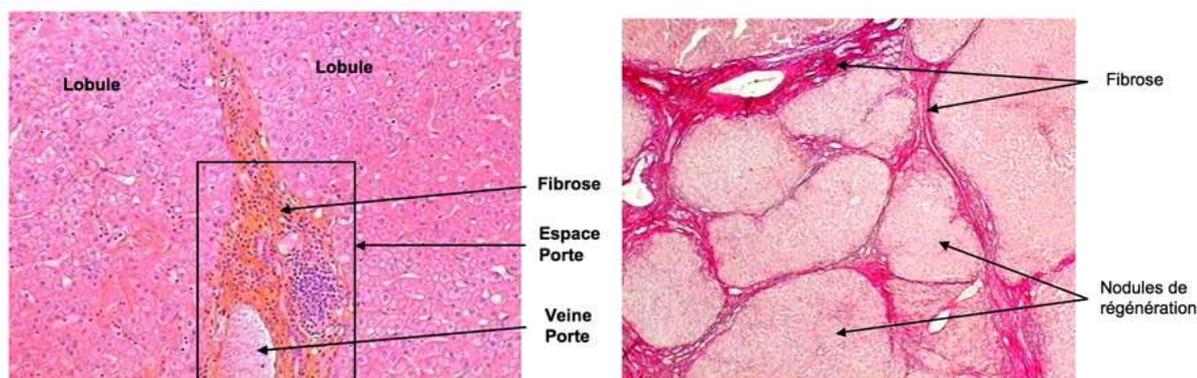


Figure 38. Fibrose hépatique. A droite : fibrose au stade F1 : l'accumulation de matrice extra cellulaire élargit l'espace porte (en encadré). A gauche : fibrose au stade F4 : l'architecture hépatique du lobule a disparu et la fibrose (rose magenta) délimite des nodules de régénération (rose pâle) source : centre Hépatobiliaire Paul Brousse, <http://www.chb.aphp.fr/foieEtMaladies/maladies/cirrhose/index.phtml>.

La cirrhose est un syndrome défini par l'anatomo-pathologie et qui correspond aux stades ultimes de la fibrose. Elle peut être compensée, c'est-à-dire sans symptômes, ou décompensée avec l'apparition de complications (hémorragie, ascite, encéphalopathie...). Dans le monde, 90% des carcinomes hépatocellulaires se déclarent sur des foies cirrhotiques (Nordenstedt et al., 2010).

Mécanismes Moléculaires du carcinome hépatocellulaire

L'hépatocyte normal, sous l'effet des processus d'inflammation et de régénération, peut subir des modifications génétiques qui induisent une dysplasie puis la formation de nodules qui

évolueront en CHC et production de métastases. Dans cette chronologie stéréotypée de l'évolution tumorale, les altérations génétiques s'accumulent et peuvent être présentées au niveau cellulaire en 3 types d'événements ; (i), l'activation de facteurs de croissance et de leurs récepteurs, (ii) l'inactivation de gènes suppresseurs de tumeurs, (iii) l'activation de proto-oncogènes, (iv), La mutation d'autres gènes (**Fig. 39**). Les quatre groupes principaux de gènes qui sont mutés dans le CHC induisent donc une activation des facteurs de transcription ou leur récepteur (IGF-2, TGF- α), inactivation de gènes suppresseurs de tumeurs, (PTEN, P53, RB, RUNX3), l'activation d'oncogènes (MYC, RAS, RAF), la réactivation de voie de signalisation du développement (Hh, Wnt/ β -cat) (Martin and Dufour, 2008; Shiraha et al., 2013).

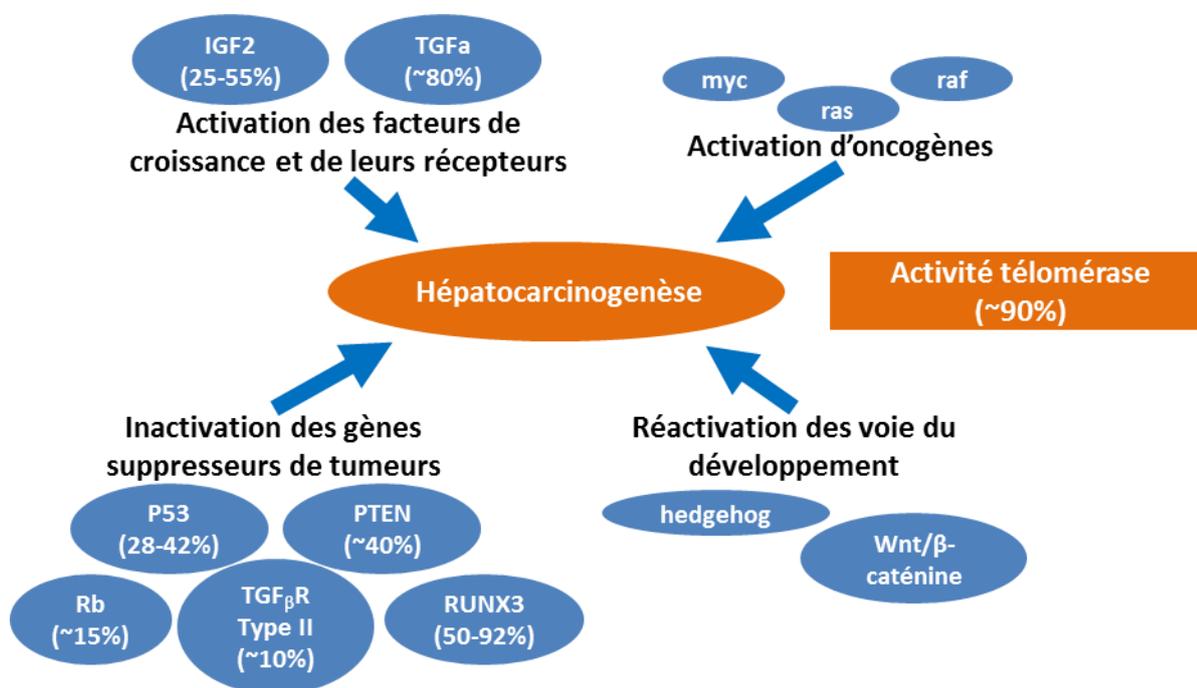


Figure 39. Altération génétiques présentes dans les CHC. IGF2 ; insulin-like growth factor 2, TGF α ; transforming growth factor α , PTEN ; Phosphatase and TENSin homolog, RB ; RetinoBlastoma gene, TGF β R ; Transforming Growth Factor β Receptor, RUNX3 ; Runt-related transcription factor 3. L'activité télomérase est caractéristique des cellules germinales, elle correspond à l'ajout de séquences répétées aux extrémités des chromosomes (télomères). Ces séquences répétées protègent les gènes des extrémités chromosomiques. D'après (Shiraha et al., 2013)

Les analyses récentes du génome de tumeurs hépatiques montrent le caractère cumulatif des mutations. Une étude de 125 biopsies CHC a identifié 135 délétions homozygotes et 994 mutations somatiques (Guichard et al., 2012). Ces mutations sont associées ou non au caractère cancéreux, la discrimination peut être réalisée en analysant conjointement le tissu cancéreux et le tissu périphérique à la tumeur. Cette analyse menée sur 87 biopsies a permis d'identifier 34 mutations spécifiques du CHC. Parmi l'ensemble des mutations qui semble indiquer un caractère cumulatif,

des mutations récurrents ont été identifiées avec un niveau élevé d'ARN: P53 (18%), CTNNB1 (10%), KEAP1 (8%), C16orf62 (8%), MLL4 (7%) and RAC2 (5%) (Cleary et al., 2013).

Analyses protéomiques et phosphoprotéomiques

Analyse chez l'animal. Le protéome du foie de souris a fait l'objet de nombreuses études. Lai et collaborateurs ont montré d'une part que le niveau d'expression protéique était corrélé avec les quantités d'ARN correspondants ($R = 0,62$ $n=5680$ gènes), d'autre part que le protéome du sang contenait un nombre important de protéines du tissu et permettait d'envisager une recherche de bio-marqueurs (Lai et al., 2008). L'analyse par électrophorèse bidimensionnelle de sérums provenant de souris transgéniques qui surexpriment l'EGF et développent des cancers du foie a montré que la réponse anticorps était affectée (disparition des IgK et IgL), et que les protéines *amyloid component P* and *apolipoprotein M* étaient surexprimées d'un facteur 8 à 10 (Gazzana and Borlak, 2008). L'étude du développement de tumeurs hépatiques induites par xénogreffe chez la souris *nude* a montré que le niveau d'expression de la transaldolase était augmentée dans le tissu cancéreux et dans le sang des souris correspondantes (Wang et al., 2011a). Dans la même étude, l'analyse de 72 sérums humains provenant pour moitié de patients atteints de CHC avec métastases a permis de calculer une spécificité et une sensibilité du dosage de respectivement 77.6 et 86.1 %.

Les analyses phosphoprotéomiques de foie de souris ou de rat sont nombreuses. Une série d'articles a utilisé ces modèles pour démontrer l'intérêt de la purification par affinité sur support IMAC (Moser and White, 2006; Villen et al., 2007). Un atlas des sites de phosphorylation a été publié pour 9 organes de la souris et 36000 sites de phosphorylation (soit 6296 protéines) ont été identifiés (Huttlin et al., 2010). Dans cette étude, seulement 3% des phosphopeptides provenaient du foie, alors que 33% des phosphopeptides identifiés provenaient du cerveau. Une étude équivalente sur 14 organes de rat a été réalisée en 2012 (Lundby et al., 2012). Les résultats sont très similaires à ceux obtenus chez la souris : les phosphopeptides sont également moins abondants dans le foie de rat que dans les autres organes.

Analyse de cultures cellulaires. Les études sur cellules permettent de tester plus rapidement l'effet de drogues sur la croissance, la différenciation, et la viabilité des cellules.

Pour des cellules de souris Hepa1-6, Pan et collaborateurs ont montré que l'ajout de pervanadate (inhibiteur de tyrosine phosphatases), de calyculine A et deltaméthrine (inhibiteur de serine et thréonine phosphatases) inhibait la phosphorylation de seulement 28 % des phosphopeptides détectés (Pan et al., 2008).

Pour des cellules humaines (cellules souches mésenchymateuses ,MSCs), deux milieux de culture ont permis d'induire une différenciation des MSCs en hépatocytes. Un premier milieu (supplémentation en facteurs de croissances HFG et FGF-7) induit une inactivation de la voie Wnt/ β -caténine. Un deuxième milieu plus complexe provoque une prolifération anormale des cellules différenciées, associée à la translocation nucléaire de la β -caténine, la surexpression de *cathepsin D precursor*, *adenine phosphoribosyl transferase*, *L-lactate dehydrogenase*, *triosephosphate isomerase*, *inorganic pyrophosphatase or peptidyl prolyl cis-trans isomerase* et la surexpression des gènes Lrp5, Fzd3, myc et P53 déjà cités précédemment (Herencia et al., 2012).

L'étude protéomique d'une lignée cellulaire hépatocytaire résistante au 5-fluoro-uracile (5FU, BEL7402) a montré la surexpression d'annexine 3 chez les cellules résistantes. La surexpression de l'annexine 3 a été constatée pour 15 biopsies sur 20 correspondant à des CHC (Tong et al., 2012).

Analyse de tissus humains

Une étude par électrophorèse bidimensionnelle de 4 biopsies (CHC et tissus adjacents) a permis d'identifier 39 protéines exprimées de manière différentielle entre le tissu cancéreux et le tissu sain, parmi lesquelles *apolipoprotein E* et *chloride intracellular channel 1* ont été validées par une analyse en Western blot (Blanc et al., 2005).

Des niveaux élevés de calréticuline et de fragments de *protein disulphide isomerase A3* ont été détectés dans les sérums de patients CHC, les concentrations mesurées permettant de différencier les CHC des foies cirrhotiques (Chignard et al., 2006).

Une synthèse des données bibliographiques montre que la majorité des protéines surexprimées correspondent à des protéines de stress (HSP) et du métabolisme (Kuramitsu and Nakamura, 2006).

Minagawa et collaborateurs ont étudié 14 biopsies de patients atteints de l'hépatite C (tumeurs et tissus adjacents). L'analyse conjointe du protéome (électrophorèse 2D) et du transcriptome a montré une corrélation particulièrement forte ($r=0,90$) entre les protéines impliquées dans le métabolisme et leur ARN_m. La glutamine synthétase, la vimentine, l'annexine A2 et l'aldo-céto réductase sont surexprimées, alors que l'anhydrase carbonique 1 et 2, l'argininosuccinate synthétase 1, la fructose-1,6-biphosphase et la bétaine-homocystéine méthyltransférase sont sous exprimées dans les tissus CHC (Minagawa et al., 2008). Une analyse protéomique par marquage iTRAQ a également permis d'identifier la sous expression de 5 enzymes du cycle de l'urée dans les tumeurs par rapport aux tissus sains adjacents (Chaerkady et al., 2008).

En 2010, Sun et collaborateurs ont montré une surexpression de *lamin B1* (LMNB1) dans les tissus CHC par rapport aux tissus cirrhotiques et aux tissus sains, le dosage de la protéine dans le sang permettant la détection des stades précoces de CHC (spécificité : 86%, sensibilité 76%) (Sun et al., 2010a). En 2011, la comparaison des données protéomiques de 4 tissus CHC et 4 tissus non CHC a mis en évidence la surexpression de 61 protéines, dont la taline 1, une protéine du cytosquelette dont la présence a été avérée par immunohistochimie, et qui est associée à la réapparition précoce de nodules chez les patients suivis. (Kanamori et al., 2011). En 2012, La surexpression de *heterogeneous nuclear ribonucleoprotein K* (hnRNP K) a été montrée dans des tissus tumoraux par rapport aux tissus cirrhotiques. Le dosage de hnRNP K associé au dosage sérique d'alpha-fœtoprotéine (AFP) permettrait un diagnostic du CHC aux stades précoces (Guo et al., 2012). Plus récemment, des dosages en mode MRM de ApoA1 ont montré une sous-expression de 20% de cette protéine dans des sérums de patients CHC par rapport aux sérums HCV (Mustafa et al., 2013).

Réponse immunitaire associée au CHC. Plusieurs articles ont étudié la réponse des anticorps associée au CHC (Le Naour et al., 2002; Li et al., 2008). Il semblerait que les patients atteints de CHC produisent des auto-anticorps spécifiques. Les antigènes identifiés sont : *DEAD* (Asp-Glu-Ala-Asp) box polypeptide 3, *eukaryotic translation elongation factor 2* (eEF2), *apoptosis-inducing factor* (AIF), *heterogeneous nuclear ribonucleoprotein A2* (hnRNP A2), *prostatic binding protein*, and *triosephosphate isomerase* (TIM). Ces auto-anticorps sont détectés dans 90% des sérums CHC (Li et al., 2008). Le Naour et collaborateurs avaient identifié précédemment la calréticuline comme un antigène spécifique des tumeurs, indépendamment du statut viral (Le Naour et al., 2002).

Les études phosphoprotéomiques chez l'homme sont beaucoup moins nombreuses. L'analyse d'un foie sain, après trypsinolyse et purification sur colonne IMAC-Ti⁴⁺, puis analyse LC-MS a permis d'identifier plus de 2000 sites de phosphorylation, laissant présager des niveaux de phosphorylations dans le foie humain équivalents à ceux observés chez le rat et la souris (Han et al., 2010). Un autre protocole utilisant une chromatographie bidimensionnelle SAX+RP pour séparer les peptides tryptiques a permis d'identifier plus de 1500 phosphopeptides à partir de 500 µg de foie (Wang et al., 2010). Ces auteurs n'ont malheureusement pas appliqué leur protocole à l'étude de tumeurs.

A ce jour, seulement deux publications présentent une analyse phosphoprotéomique de tissus tumoraux. Lee et collaborateurs ont identifiés 39 phosphopeptides dans 3 biopsies CHC et ont validé par Western-blot la diminution d'alpha-HS-glycoprotéine (phospho-ser 138 and 312) dans les tissus tumoraux (Lee et al., 2009). Une autre analyse phosphoprotéomique quantitative a

également été réalisée pour 3 tissus ; un tissu cancéreux et deux tissus sains. Les peptides tryptiques ont été marqués au formaldéhyde puis purifiés par IMAC-Ti⁴⁺ avant d'être séparés par une chromatographie multidimensionnelle SCX-RP. Plus de 1800 sites de phosphorylations ont été identifiés. L'analyse par motif des séquences phosphorylée a montré une surreprésentation des motifs PXpSP, RXXpSP, and pSP dans la tumeur. Ces motifs sont phosphorylés par les ERK ½, kinases impliquées dans la prolifération cellulaire et induites dans différents cancers (Song et al., 2011).

4.1.3 Analyse quantitative par marquage iTRAQ

La première publication en 2004 du réactif iTRAQ (Isobaric Tag for Relative and Absolute Quantitation) présente l'analyse quantitative de mutants de la levure après marquage par un réactif permettant l'analyse conjointe de 4 échantillons (iTRAQ 4Plex) (Ross et al., 2004). Depuis, 900 publications référencées dans PubMed ont utilisé le marquage iTRAQ, dont la dernière version permet de quantifier dans la même analyse 8 échantillons (iTRAQ 8Plex). Le réactif iTRAQ est un produit commercialisé par AB Sciex et réagit en milieu basique avec les fonctions amines par l'intermédiaire d'un groupement succinimide (**Fig. 40**).

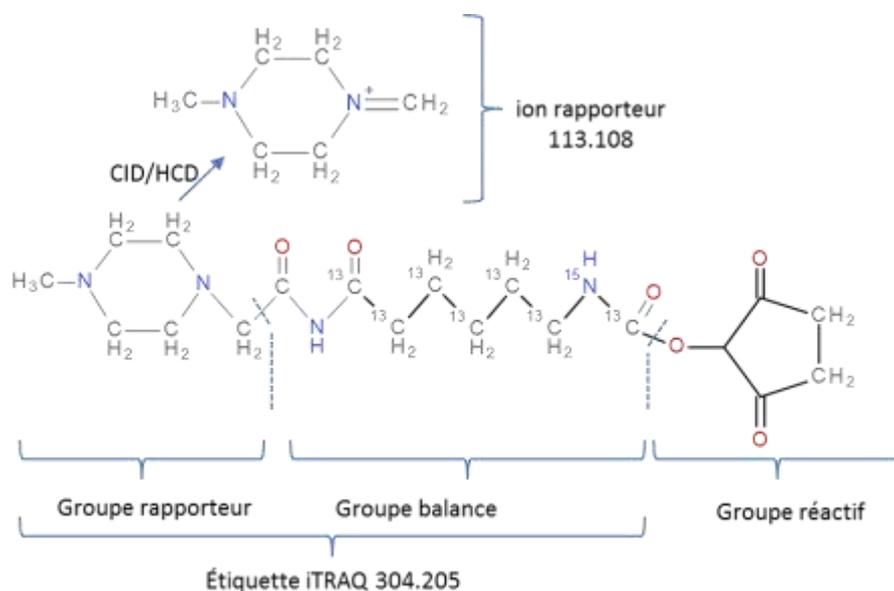


Fig.40. Le réactif iTRAQ 113. Structure supposée du réactif iTRAQ 8plex (non publiée à ce jour). Seuls les ions rapporteurs et le groupement réactif sont connus. Le groupe balance présenté ici est une hypothèse basée sur la composition élémentaire déduite des mesures de masse exacte. Les autres réactifs, iTRAQ 114 à iTRAQ 121, sont obtenus par des marquages ¹³C et ¹⁵N de l'ion rapporteur. En contrepartie, le groupe balance est allégé par l'incorporation de ¹²C et ¹⁴N.

Dans sa version 8Plex, 8 réactifs iTRAQ différents sont utilisés, chacun ayant une masse totale identique de 304.205 Da, mais des ions rapporteurs différents. Ces ions rapporteurs sont observés uniquement dans le spectre MS/MS aux masses 113, 114, 115, 116, 117, 118, 119 et 121 et

correspondent à différentes combinaisons de carbone 14 et azote 14. Le groupe balance permet d'équilibrer la masse moléculaire de l'ensemble du réactif. Après le marquage de chacun des 8 échantillons avec un des 8 réactifs iTRAQ, les échantillons sont regroupés. Lors de l'analyse, les spectres MS ne produisent qu'un seul pic par peptide, quel que soit l'échantillon. Par contre, le spectre MS/MS fragmente les groupements iTRAQ dont les masses sont spécifiques de chaque échantillon. La mesure de l'intensité des ions rapporteurs est fonction de la quantité du peptide dans les différents échantillons.

D'un point de vue expérimental, l'hydrolyse du groupement succinimide augmente avec le pH et la demi-vie du réactif est inférieure à 15 min à des valeurs de pH supérieures à 8. Par contre, elle dépasse 2h pour des valeurs de pH inférieures à 6,5 (Grumbach and Veh, 1991). Inversement, la vitesse de réaction avec les amines augmente avec la valeur de pH. La réaction de couplage est généralement réalisée à un pH proche de 8. Dans un tampon 50 mM phosphate de sodium pH 8,2, le temps de demi-vie de la réaction des peptides avec le réactif iTRAQ est de l'ordre de la minute à 25°C.

Les esters de N-hydroxysuccinimide réagissent spécifiquement avec les amines, mais des réactions parasites sur la tyrosine et la sérine ont également été décrites. Cependant, la O-acylation est réversible avec l'ajout de solution d'hydroxylamine 1 M en tampon borate pH9.2 (Kurosky et al., 1993; Wiktorowicz et al., 2012).

L'analyse MS/MS des peptides dérivés peut être réalisée en MALDI-TOF-TOF (Pflieger et al., 2008; Scheri et al., 2008) ou en ESI-MS/MS. Avec le LTQ-orbitrap, la fragmentation est généralement réalisée dans la cellule HCD (Boja et al., 2009; Pichler et al., 2010; Zhang et al., 2009). Quelques rares publications ont également utilisé une fragmentation ETD (Phanstiel et al., 2008) qui permet la quantification de 5 des 8 réactifs iTRAQ 8plex (Phanstiel et al., 2009).

Si certains articles ont montré que le marquage isobare entraînait une diminution du nombre de peptides identifiés (Li et al., 2012; Patel et al., 2009; Wang et al., 2012a), d'autres publications ont abouti aux conclusions inverses (Putz et al., 2012; Trinh et al., 2013). Par ailleurs, le marquage iTRAQ 8Plex serait moins efficace que le marquage iTRAQ 4Plex et le marquage mTRAQ (Pichler et al., 2010). Enfin, le marquage iTRAQ 4Plex serait plus performant que le marquage mTRAQ (Mertins et al., 2012). Face au caractère contradictoire de ces publications, la seule conclusion raisonnable est la difficulté de comparer les techniques de quantification.

La méthodologie de marquage iTRAQ 8Plex présente l'intérêt de quantifier 8 échantillons différents dans la même analyse. Par contre, le marquage présenterait 3 désavantages : (i) la sous-

estimation de l'amplitude des variations (Evans et al., 2012; Pierce et al., 2008) (ii) l'interférence possible de l'ion ammonium de la phénylalanine pour iTRAQ121 (Ow et al., 2009), (iii) l'augmentation du nombre de charge des peptides dérivés, moins propice à l'identification par les bases de données (Thingholm et al., 2010).

4.2 Matériel et Méthodes

4.2.1 Matériel

Les différentes étapes chromatographiques de dessalage sur support greffé C18 (HPLC Sep Pak tC18) et de chromatographie TiO₂ ont été réalisées sur un système HPLC PF2D (Beckman Coulter) décrit dans le chapitre précédent. Seule la deuxième partie de ce système de chromatographie bidimensionnelle a été utilisée en mode isocratique.

Les analyses par spectrométrie de masse ont été réalisées sur un système nanoLC-LTQ-Orbitrap XL. La nano chromatographie (chaîne Dionex U3000) utilise une colonne de préconcentration C18 (PepMap C18 0.3x5 mm) et une colonne analytique C18 (PepMap C18, 75 µm x 10 cm).

4.2.2 Echantillons

Les biopsies humaines proviennent de patients opérés au Centre Hospitalier de Bordeaux. Elles ont été caractérisées par les anatomo-pathologistes du CHU puis conservées à -80°C. Dix-huit biopsies correspondent à des carcinomes hépatocellulaires sur foie non fibreux, 6 biopsies « normales » correspondent à des tissus qualifiés de normaux après observation des tissus par les anatomopathologistes. Le **Tableau 12** présente les différentes observations, les analyses biochimiques (alpha fœtoprotéine) et les analyses génétiques (P53 et BCAT) des biopsies CHC. La caractéristique commune des 18 biopsies tumorales est le type non fibreux des tissus, les autres caractères étant extrêmement variables. Les patients âgés de 32 ans à 82 ans sont en majorité de sexe masculin. Le taux d'alpha fœtoprotéine n'est supérieur à la valeur seuil (20 ng/mL) que pour 55% des cas. Certaines pathologies associées (hémochromatose, stéatose, péliosis) sont très peu représentées, mais probablement de nature à modifier le protéome des cinq biopsies concernées. La mutation du gène P53 ne concerne qu'une biopsie alors que la mutation bêta-caténine(BCAT) est observée pour la moitié des biopsies. Ces paramètres n'ont pas été associés de manière à créer des sous-groupes ; le panel des biopsies est extrêmement hétérogène.

Tableau 12. Caractéristiques anatomo-pathologiques des biopsies tumorales

Biopsie ⁽¹⁾	1.1	1.2	1.3	1.4	1.5	1.6	2.1	2.2	2.3	2.4	2.5	2.6	3.1	3.2	3.3	3.4	3.5	3.6
Age	77	65	61	69	57	67	71	45	85	72	63	69	32	73	72	82	76	77
Sexe	M	F	M	M	M	M	M	M	M	M	F	M	F	M	M	M	F	M
Alpha foetoprotéine (ng/ml)	<20	<20	<20	>20	<20	>20	<20	>20	NA	>20	>20	NA	<20	>20	<20	<20	<20	<20
Alcoolisme	0	0	1	0	1	1	1	1	1	0	0	1	NA	1	1	1	1	0
Hémochromatose ⁽²⁾	0	0	0	0	0	0	0	0	0	1	0	0	NA	0	0	0	0	0
Diaètre du nodule le plus large	35	45	110	210	190	100	110	100	75	160	45	180	NA	130	95	80	100	50
Invasion vasculaire	1	1	1	1	1	1	1	0	0	1	1	1	NA	0	0	0	1	1
Invasion de la veine portale	1	0	1	1	0	0	1	0	0	1	0	0	NA	0	0	0	0	0
Capsule tumorale (0:abs., 1:intacte, 2:envahie)	0	2	2	0	2	2	2	1	2	2	2	NA	NA	2	0	1	2	1
Nodules satellites	0	0	1	1	0	0	1	0	0	1	1	1	NA	1	0	0	1	0
Invasion de la veine hépatique	0	0	0	1	0	0	0	0	0	1	0	NA	NA	NA	0	NA	0	0
Stéatose tumorale ⁽³⁾	0	1	NA	0	0	0	0	0	3	0	NA	2	NA	0	0	0	1	1
Pélosis ⁽⁴⁾	0	0	NA	0	1	1	0	1	0	0	0	0	NA	0	1	0/1	0/1	0
Stroma fibrotique	0	0	NA	0	0	0	0	0/1	0	0	0	0	NA	0	0	0	0	1
Cholestérase tumorale ⁽⁵⁾	0	0	NA	0/1	0	1	0	1	1	0	1	0	NA	1	0	0	0	0
Inflammation	0	0	NA	0	0	1	0	0	0	1	0	0	NA	0	0	1	0	1
classification d'Edmonson ⁽⁶⁾	I-II	I-II	I-II	III-IV	III-IV	III-IV	I-II	III-IV	I-II	III-IV	III-IV	III-IV	NA	III-IV	I-II	I-II	I-II	III-IV
Mutation P53 ⁽⁷⁾	0	0	1	0	0	0	0	0	0	0	0	0	NA	0	0	0	0	0
Mutation BCATex3 ⁽⁷⁾	0	1	0	0	0	0	0	0	1	0	1	1	1	1	1	1	0	0

(1): la partie entière désigne la série ITRAQ dans laquelle la biopsie a été analysée, la partie décimale désigne le tag ITRAQ utilisé (ex: 2,1 = deuxième série, tag 113)

(2): pathologie provoquée par une accumulation de fer dans le foie

(3): accumulation de graisse (triglycérides) dans les hépatocytes

(4): trouble de la microcirculation sanguine intra-hépatique dans lequel le tissu hépatique est parsemé de cavités remplies de sang

(5): stase de la bile

(6): les grades de la classification d'Edmondson-Steiner sont établis sur la base d'observations histologiques.

Plus le grade est élevé, plus les cellules et le lobule ont des caractéristiques morphologiques propres aux cellules tumorales.

(7): Déterminé d'après les travaux de l'U674 (Dr Zucman-Rossi)

n.d.: non déterminé

Les biopsies "normales" ne présentent pas de lésions macroscopiques et microscopiques

4.2.3. Méthodes

Les différentes étapes du protocole sont les suivantes :

1. Broyage, réduction et alkylation, trypsinolyse
2. HPLC Sep Pak tC18
3. HPLC TiO₂
4. Sep Pak tC18
5. Marquage iTRAQ
6. Sep Pak tC18
7. Analyse LC-MS/MS (Orbitrap)

Broyage, réduction et alkylation, trypsinolyse

Le broyage des biopsies est réalisé par un vibro-broyeur (Retsch) ; 50 à 100 mg de tissu sont placés dans un tube de 2 mL (Eppendorf) contenant 2 billes en acier inoxydable et 1 mL de tampon de lyse (0.1 M TEAB pH 9.3, 0.15 M NaCl, 0.1 % SDS, 5 mM NaF, 2 mM Na₃VO₄ et 10 mM DTT). Deux cycles de 60 sec à 20 Hz permettent d'obtenir une suspension homogène. Le broyage est réalisé à 4°C (les portoirs du vibro-broyeur sont placés à 4°C avant utilisation). Les échantillons sont ensuite transférés dans des tubes de 2 mL. Après 5 min à 100°C, l'échantillon est refroidi puis alkylé à 25°C sous agitation pendant 30 min par 50mM final d'iodoacétamide. La trypsinolyse est réalisée à 37 °C par deux ajouts d'enzyme (à T₀ et T_{1H}) avec un rapport E/S = 1/100 (poids/poids). Le poids de substrat est déterminé par la pesée de la biopsie en supposant 10% de protéines dans l'échantillon. La digestion est réalisée pendant la nuit au moyen d'un mixeur thermostaté (Eppendorf). Après 10h d'incubation à 37°C sous agitation pulsée, l'échantillon est refroidi à 6°C. Le lendemain, les échantillons sont acidifiés par du TFA (1% final). L'acidification de l'extrait provoque le dégagement de CO₂. Après 5 min sous agitation, les échantillons sont centrifugés à 16000 g pendant 15 min. Le surnageant est récupéré et transféré dans un plaque 96 puits adapté au passeur d'échantillon du système HPLC (PF2D, Bekman-Coulter) ; le culot est séché et le résidu sec est pesé.

HPLC Sep Pak tC18

La phase d'une cartouche Sep Pak tC18 (Waters) est utilisée pour remplir une colonne HPLC (5 mm x 4.5 cm). La chromatographie est réalisée avec un système PF2D (Beckman Coulter) disposant d'un passeur d'échantillon, d'un détecteur UV et d'un collecteur (Gilson) à un débit de

0.9 mL/min. Les peptides sont retenus en tête de colonne avec 0.5 % d'éluant B puis élués à 60 % de B, la colonne étant régénérée par 90 % de B (A : 0.1 % acide formique, B : 100 % acétonitrile).

Les injections sont réalisées dans l'ordre suivant :

- Deux Injections d'albumine bovine pour étalonner le système (vérification des temps de rétention de l'intensité du pic de la fraction éluee par 60 % d'acétonitrile)
- Deux injections de tampon 0.1% TFA (vérification d'un éventuel effet mémoire)
- Injections des échantillons

La fraction éluee à 60% B est séchée au SpeedVac (Thermo Scientific). Les tubes de collecte sont préalablement tarés afin de permettre une pesée de l'extrait sec.

HPLC TiO₂

Les échantillons préalablement purifiés par Sep Pak tC18 sont dissous (10 mg/mL) dans une solution 1M TFA, 1M acide glycolique, 20 % ACN. La colonne TiO₂ (2 mm x 5 cm, voir chapitre 3) est utilisée à un débit de 0.2 mL/min en tampon A (50 mM TFA, 20% ACN). Après 15 min. de rinçage par le tampon A, les phosphopeptides sont élués par NH₄OH 0.7 M. Les injections respectent l'ordre suivant :

- Injection de l'échantillon en tampon de charge (1 M TFA, 1M ac. glycolique, 20% ACN)
- Injection du tampon de charge seul
- 2 injections du tampon d'élution (1M NH₄OH)

Les tubes de collecte sont pré-remplis de 25µL d'acide formique pur afin de neutraliser l'ammoniaque. Les fractions éluées sont réduites à un volume d'environ 100 µL puis ajustées à 500 µL par 0.1% TFA avant d'être à nouveau dessalées sur la colonne HPLC Sep Pak tC18 puis séchées. Les phosphopeptides sont remis en solution dans 20 µL d'un mélange eau/ acétonitrile (50/50) puis dosés par la méthode de Lowry.

Dosage colorimétrique des peptides

Le dosage des peptides a été réalisé au moyen d'un kit commercial (DC Protein Assay, Biorad). Le protocole a été modifié comme suit : incubation de 100 µl d'échantillon + 50 µl solution A + 400 µl solution B; mélange vigoureux, repos 15 min et lecture à 750 nm. Une courbe d'étalonnage a été réalisée à partir d'un hydrolysate tryptique d'albumine bovine, la gamme allant de 400 µg/mL à 6.25 µg/mL (6 dilutions sériées au ½).

Marquage ITRAQ

Le marquage ITRAQ 8plex est réalisé en suivant les recommandations du fabricant (AB Sciex) :

- Echantillon entre 50 et 100 µg.
- isopropanol >60%.
- TEAB > 0.1M final.

Chaque extrait est incubé 2h à 25°C. La réaction de couplage est stoppée par 2 µL de tampon Tris 1M pendant 15 min ; les échantillons sont ensuite acidifiés par du TFA (1% final) puis regroupés dans un seul tube. La solution est concentrée à quelques µL, puis ajustée à 500 µL et dessalée par HPLC Sep Pak tC18.

Les réactifs iTRAQ 8plex permettent de marquer 8 échantillons avec une étiquette différente pour chacun alors que l'étude est réalisée sur 24 échantillons. Quatre séries de marquage ont donc été réalisées. Pour chaque série, les étiquettes 119 et 121 correspondent aux mêmes échantillons C1 et C2 (**Tab. 13**). Les échantillons C1 et C2 correspondent à un mélange de tissus non tumoraux, ils servent de référence commune pour le calcul des ratios. Ainsi, les différentes séries pourront être comparées entre elles.

Tableau 13. Séries ITRAQ

Série ITRAQ	Biopsie	ITRAQ	Série ITRAQ	Biopsie	ITRAQ
1	1.1	113	3	3.1	113
1	1.2	114	3	3.2	114
1	1.3	115	3	3.3	115
1	1.4	116	3	3.4	116
1	1.5	117	3	3.5	117
1	1.6	118	3	3.6	118
1	1.C1	119	3	3.C1	119
1	1.C2	121	3	3.C2	121
2	2.1	113	4	4.1	113
2	2.2	114	4	4.2	114
2	2.3	115	4	4.3	115
2	2.4	116	4	4.4	116
2	2.5	117	4	4.5	117
2	2.6	118	4	4.6	118
2	2.C1	119	4	4.C1	119
2	2.C2	121	4	4.C2	121

C1 et C2 correspondent au pool des échantillons de la série 4 (tissus non tumoraux) tels que C1 = 4.1 + 4.2 + 4.3 et C2 = 4.4 + 4.5 + 4.6.

Analyse LC-MS/MS (nanoLC-LTQ-Orbitrap)

L'analyse LC-MS/MS a été réalisée avec un système nanoLC-LTQ-Orbitrap. Pour un temps d'analyse LC-MS/MS de 2h, le spectromètre de masse a alterné une acquisition MS à haute

résolution (60 000) dans l'Orbitrap avec les acquisitions MS/MS des 6 ions multichargés les plus intenses. Six acquisitions MS/MS à basse résolution sont réalisées dans le piège ionique linéaire LTQ et 6 acquisitions à haute résolution (7500) dans l'Orbitrap avec une énergie de fragmentation importante (50%) afin de favoriser l'observation des ions rapporteurs ITRAQ. Une exclusion des ions analysés pendant 30 sec permet d'analyser les ions de plus faibles intensités.

Traitement des données

Proteome Discoverer 1.4

Les spectres MS/MS ont été comparés aux spectres théoriques issus de l'ensemble des séquences protéiques humaines (*complete genome*, UniprotKB), soit plus de 68000 séquences avec les modifications fixes suivantes : +304.205 (ITRAQ 8 plex) sur les amines N-terminales et K, +57.021 (carbamidomethyl) sur les cystéines. Les modifications variables appliquées sont +15.995 (oxydation) sur M et P, +0.984 (déamidation) sur N, +79.966 (phosphorylation) sur S,T,Y. Le module PhosphoRS 3.0 a été utilisé pour calculer une probabilité lors de la présence de plusieurs sites de phosphorylation sur un même peptide. Les peptides identifiés sont filtrés sur la base d'un FDR <0.01, d'un *Peptide Rank* =1 et d'un niveau de charge du précurseur <4. L'erreur sur la masse de l'ion précurseur est fixée à 25 ppm. L'erreur sur la masse des ions fragments est 0.7Da (spectre CID) ou 0.05 Da (spectre HCD).

X-Tandem dans GPM

X-tandem sous sa version intégrée à *GPM manager* a été téléchargé sur le site FPT suivant : <ftp://ftp.thegpm.org/projects/gpm/gpm-xe-installer/>. Installée localement, une interface web permet une identification suivant le même principe de PD 1.4. La même base de séquences humaines a été utilisée ainsi que les mêmes paramètres d'interrogation. X-tandem ne lisant pas le format de fichiers produit par les instruments Thermo Scientific, les données produites par le LTQ-Orbitrap ont été converties en format mzXML par Proteowizard (<http://proteowizard.sourceforge.net/>). *GPM manager* dispose d'une interface qui permet le classement des protéines identifiées par fonction cellulaire, compartiment cellulaire ou fonction moléculaire (d'après la *Gene Ontology*)

Quantification ITRAQ

La quantification ITRAQ a été réalisée par le module *Reporter Ion Quantifier* de PD 1.4. Les données quantitatives brutes ont également été exportées sous forme de tableau Excel et une macro VBA a été développée pour permettre le traitement des données utilisant la moyenne des

ions rapporteurs 119 et 121 comme dénominateur pour le calcul des ratios ITRAQ. Une autre macro VBA a été créée pour aligner les séquences identiques provenant des différentes séries puis extraire de l'alignement les peptides communs aux 4 séries de marquage ITRAQ. L'ensemble de la programmation VBA représente 2300 lignes de code.

Diagramme de Venn

Les diagrammes de Venn ont été réalisés par une application disponible en ligne à l'adresse suivante : <http://www.bioinformatics.lu/venn.php>.

pLogo

pLogo est accessible en ligne à l'adresse suivante <http://plogo.uconn.edu/>. Le logiciel permet d'aligner une séquence sur un acide aminé central et calcule la représentation des acides aminés adjacents. Plus un acide aminé est observé fréquemment à une position donnée, plus la taille de son caractère est importante.

MeV et Genesis

MeV et Genesis sont des logiciels de traitement des données de type microarray. Ils peuvent être téléchargés aux adresses suivantes : <http://www.tm4.org/mev.html> et http://genome.tugraz.at/genesisclient/genesisclient_download.shtml. Mev nécessite l'installation préalable du logiciel R (<http://cran.univ-lyon1.fr/>).

La classification hiérarchique a été réalisée avec Genesis en utilisant les paramètres suivants : *euclidian distance* et *complete linkage clustering*. Les tests statistiques ont été produits par le logiciel MeV.

4.3 Résultats

4.3.1 Préparation des échantillons

Digestion trypsique et dessalage des peptides

La gestion de 24 échantillons pose un problème technique lié au temps de préparation nécessaire à chaque échantillon. Les biopsies devant être utilisées également pour d'autres analyses, elles ont été découpées en 4 à 5 morceaux de 50-100 mg environ (soit 1 à 2 mm de côté). Cette étape réalisée sous hotte à flux laminaire et à 4°C a été la manipulation la plus délicate de l'ensemble du protocole, avec un temps nécessaire à chaque échantillon très différent, de quelques secondes à quelques minutes. Les autres étapes, broyage, réduction, alkylation, digestion par la

trypsine ont été réalisés en une seule fois. Après trypsinolyse, les extraits sont acidifiés et le surnageant obtenu après centrifugation est injecté via le système HPLC dans la colonne de Sep Pak tC18.

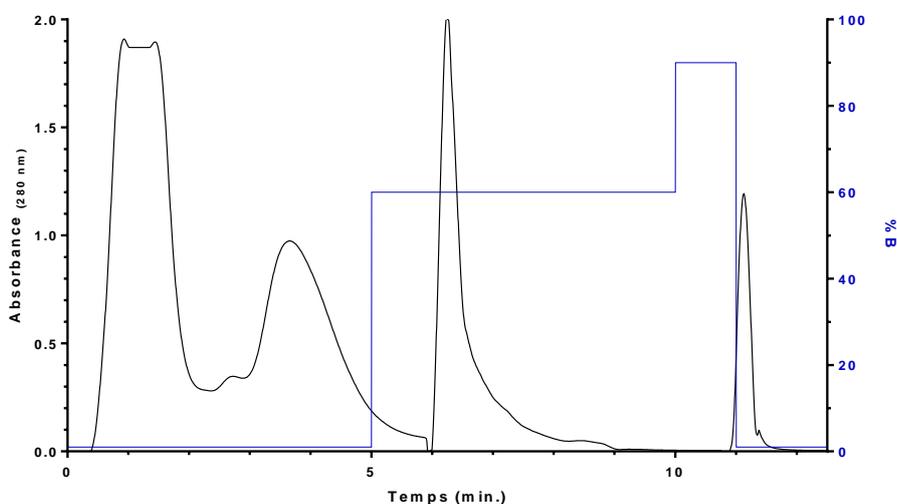


Figure 41. Chromatographie HPLC-tC18 après trypsinolyse des biopsies. La fraction éluée par 60% d'acétonitrile est collectée entre 6 et 8 min.

La Figure 41 présente un chromatogramme type de la première Chromatographie HPLC-tC18. Une injection de 1% TFA produit un petit pic à 1 min (< 0.1 DO), une injection d'échantillon produit plusieurs pics. Seul le pic à 7 min est collecté, il correspond à un palier de 60% d'acétonitrile. Le dernier palier à 90% d'acétonitrile est utilisé pour régénérer la colonne.

Lors des différentes étapes les échantillons ont été pesés afin de vérifier le caractère quantitatif de l'extraction des protéines et de la digestion trypsique (Fig. 42).

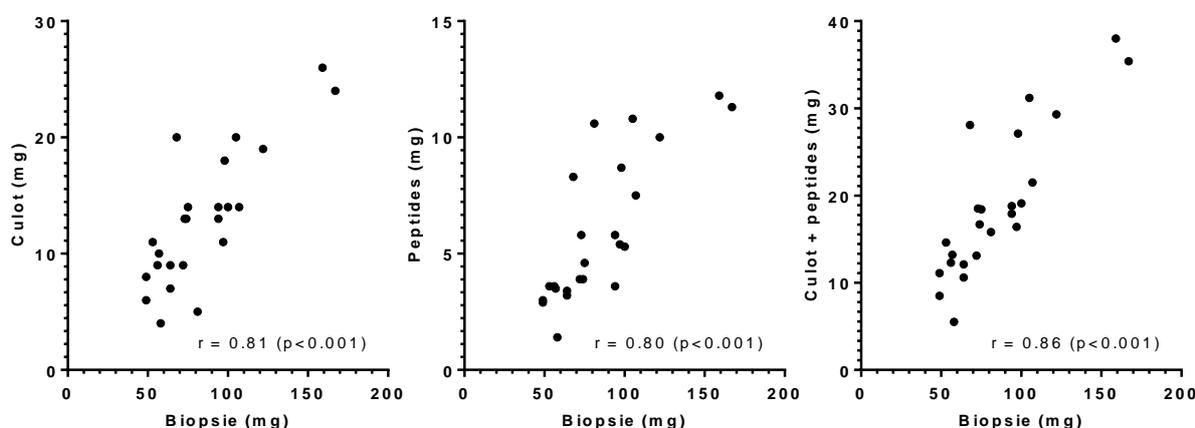


Figure 42. Corrélation entre le poids des échantillons bruts et des extraits après digestion trypsique. L'abscisse correspond à la pesée de la biopsie. Les ordonnées correspondent aux pesées des extraits secs ; culot après trypsinolyse et centrifugation (à gauche), peptides après HPLC-Sep Pak C1 8 (au milieu), somme des 2 fractions (à droite).

La **Figure 42** montre une variabilité importante du poids d'extraits peptidiques par rapport au poids du foie de départ, avec un rendement de $6,8\% \pm 2.4$ ($m \pm SD$) par rapport au foie frais. Si la somme culot + peptides est comparée au poids frais, la corrélation est améliorée ($r=0,86$). L'extraction est donc quantitative et la variabilité observée est certainement à mettre en parallèle avec l'hétérogénéité des échantillons. Cette hétérogénéité est particulièrement visible en fin de broyage des biopsies : la couleur du broyat varie du rouge au vert et la formation d'un anneau de lipide est plus ou moins visible.

Chromatographie TiO_2

La chromatographie TiO_2 a été réalisée à l'aide d'un système HPLC disposant d'un détecteur UV. La **Figure 43** présente les trois chromatogrammes obtenus après injection de l'échantillon, injection du tampon de charge seul, et injection du tampon d'éluion des phosphopeptides. Au vue de l'intensité des pics respectifs, la majorité des peptides ne sont pas retenus par la colonne TiO_2 . La deuxième injection correspond au tampon de charge seule (1M TFA, 1M acide glycolique, 20%ACN) et produit un pic dont l'intensité est essentiellement due à l'acide glycolique. La troisième injection qui correspond à l'éluion des phosphopeptides par NH_4OH 1M produit un pic qui est spécifique des peptides.

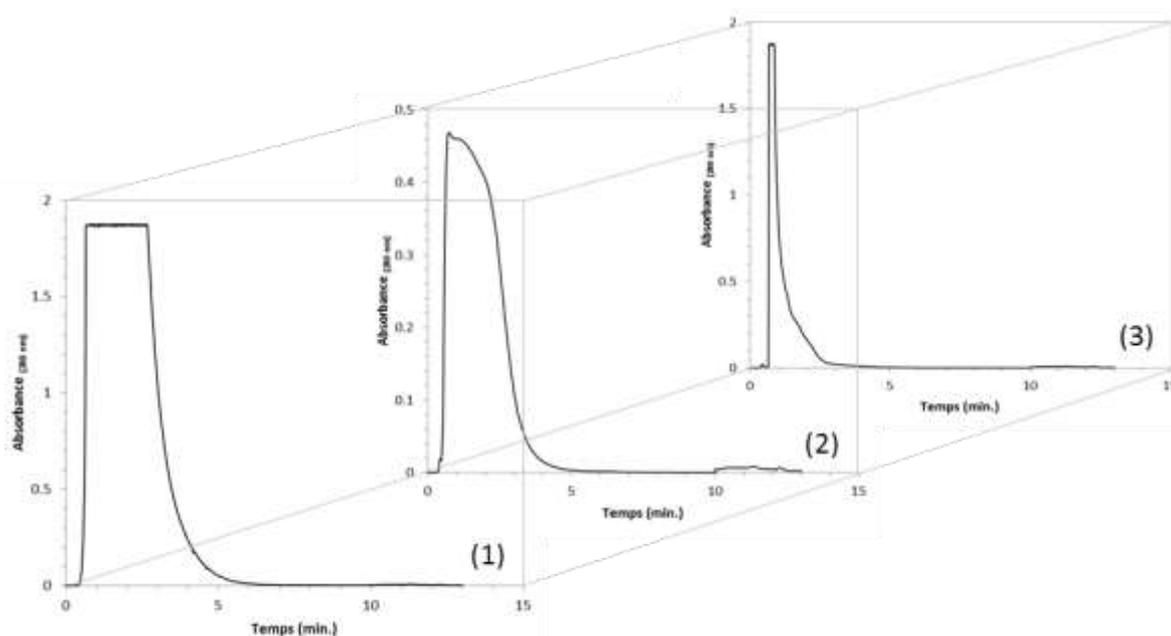


Figure 43. Chromatographie TiO_2 . La purification TiO_2 est réalisée par 3 injections successives : (1) échantillon dans le tampon de charge 1M TFA 1M ac. Glycolique 20% ACN, (2) tampon de charge, (3) tampon d'éluion 1M NH_4OH .

La fraction éluée de la colonne TiO_2 a été ensuite dessalée par Sep Pak tC18 et séchée sous vide (Speed Vac), puis remise en solution dans 20 μL d'une solution acétonitrile / eau (50/50). Le

marquage ITRAQ doit être réalisé sur une quantité de peptides inférieure à 100 µg afin d'assurer un excès de réactif par rapport aux amines N-terminales et aux amines epsilon des lysines. Un dosage colorimétrique a été réalisé afin de connaître la quantité de peptides dans la fraction éluée.

Dosage colorimétrique d'une solution peptidique

Un grand nombre de publications utilise le dosage Bradford commercialisé par BIORAD pour quantifier les échantillons protéiques. Cependant ce dosage est réputé pour être sensible à la nature de la protéine dosée. Deux protéines différentes sont proposées par le fabricant pour réaliser la gamme d'étalonnage ; l'albumine bovine et les gamma-globulines bovines. Pour une absorbance donnée, la concentration de gamma-globuline est 2 fois supérieure à la concentration d'albumine. Si la solution d'albumine est digérée par la trypsine, le signal est alors diminué d'un facteur 5 à 10 (**Fig. 44**). Le dosage Bradford n'est donc pas utilisable pour les mélanges peptidiques, une absorbance pouvant être simplement due à une trace de protéine entière dans l'hydrolysât à doser. Contrairement au dosage Bradford, le dosage Lowry n'est pas ou peu sensible à la nature de la protéine contenue dans l'échantillon; l'albumine ou les globulines produisent une courbe d'étalonnage superposable. Lorsque l'albumine est digérée par la trypsine, la courbe d'étalonnage correspondante est légèrement au-dessus des courbes d'étalonnage des protéines entières. Il est possible que cette différence ne soit pas due à la nature du standard (peptides ou protéine) mais aux différences de pureté de l'albumine de départ et de son hydrolysât trypsique . En effet, l'hydrolysât trypsique a été dessalé par Sep Pak, séché puis pesé. L'étape de chromatographie C18 a peut-être éliminé des contaminants présents dans la solution mère d'albumine.

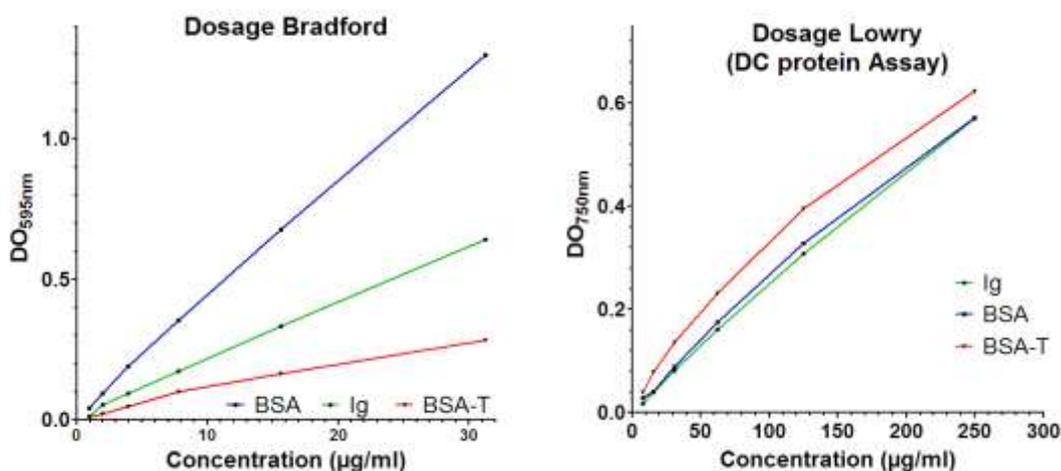


Figure 44. Courbes d'étalonnage. A gauche : le dosage Bradford (Bleu de coomassie), à droite : le DC Protein Assay (dosage Lowry). Courbe bleue : albumine bovine (BSA), courbe verte : gamma-globuline (Ig), courbe rouge : hydrolysât trypsique d'albumine.

L'ensemble des fractions éluées et des fractions non retenues de la colonne TiO_2 ont été dosées au moyen du dosage Lowry. Alors que chaque échantillon chargé sur la colonne de TiO_2 correspond à 1 mg de peptides (déterminé par pesée), la fraction non retenue sur la colonne de TiO_2 correspond à 1.3 ± 0.3 mg ($m \pm \sigma$). Dans les fractions éluées, la quantité de (phospho)peptides est en moyenne de 9.8 ± 4 μg ($m \pm \sigma$). Il faut noter que dans ce cas, les concentrations mesurées correspondent aux premiers points de la gamme d'étalonnage, zone de la courbe où le dosage est le moins précis. Il y aurait donc moins de 10% de phosphopeptides dans un hydrolysate peptidique de foie.

Marquage iTRAQ

Avec ces estimations, 50 μg de peptides non purifiés sur colonne de TiO_2 (fraction peptides totaux) et la totalité des fractions purifiées par chromatographie d'affinité sur TiO_2 (fraction TiO_2) ont été marqués par le réactif iTRAQ (voir **Tableau 12**). Le nombre de biopsies ne permettant pas de réaliser l'analyse par un seul marquage, 4 séries iTRAQ ont été réalisées. Pour les 4 séries, deux échantillons référence sont présents et correspondent aux pools des tissus sains de la série 4 (voir Méthodes).

Les quatre séries de marquage iTRAQ ont été réalisées suivant les préconisations du fabricant en utilisant 1 unité pour les peptides totaux (50 μg) et 1/2 unité pour les phosphopeptides. Après avoir réuni en un tube les groupes de 8 échantillons, les 4 séries iTRAQ de la fraction peptides totaux et les 4 séries iTRAQ de la fraction TiO_2 , l'excès de réactif iTRAQ a été éliminé par une HPLC Sep Pak tC18. La **Figure 45** présente les chromatogrammes obtenus pour les deux types de séries.

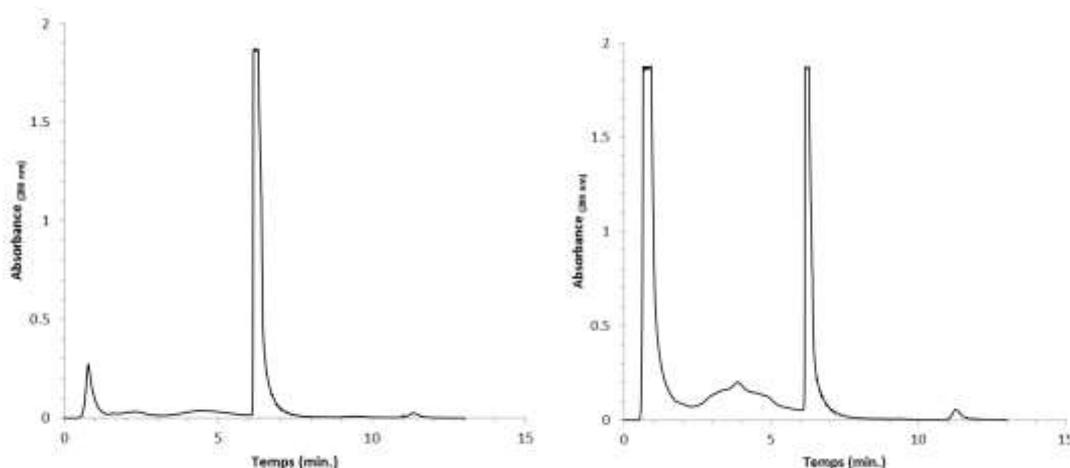


Figure 45. Chromatographie HPLC-tC18 après marquage iTRAQ des échantillons. A gauche : peptides totaux (extrait de peptides avant purification TiO_2), à droite : série phosphopeptides (fraction retenue par la chromatographie TiO_2).

Alors que le pic d'injection observé à 1 min a une intensité inférieure à 0.3 DO pour les séries iTRAQ des peptides totaux, le pic d'injection des séries de phosphopeptides sature le détecteur. L'analyse par spectrométrie de masse du pic d'injection des séries de phosphopeptides a montré exclusivement la présence d'ions monochargés de masse inférieure à 500 m/z. La présence de ce pic intense pour les séries phosphopeptides traduit donc l'excès de réactif utilisé pour le marquage de ces séries. Le pic correspondant à 60 % d'acétonitrile ($T_r = 6.2$ min) a été collecté et concentré au Speed Vac avant analyse LC-MS/MS.

4.3.2 Répétabilité du protocole

Ne disposant pas de quantités suffisantes de biopsies tumorales pour réaliser des répétitions, la variabilité liée à la chromatographie sur support de TiO_2 et au marquage iTRAQ a été étudiée sur une biopsie humaine non tumorale. Trois morceaux de 100, 67 et 83 mg ont été traités en parallèle. Après digestion tryptique et dessalage, la pesée des extraits peptidiques (respectivement 7.85, 4.45 et 5.90 mg) permet de déterminer un rendement d'extraction compris entre 6.6% et 7.85%. Les 3 répétitions ont été remises en solution à 10 mg/mL et 1 mg de chaque extrait a été purifié par chromatographie TiO_2 puis dessalé. Les 3 fractions TiO_2 ont ensuite été quantifiées par le dosage de Lowry, soit respectivement 14, 13 et 14 μ g puis marquées par 0.5 unité iTRAQ chacune avec les étiquettes 115, 116 et 117. Ces fractions sont appelées Rep1 (iTRAQ 115), Rep2 (iTRAQ 116) et Rep3 (iTRAQ 117).

Pour les 3 extraits, les rendements d'extraction de l'ordre de 7% sont identiques au rendement d'extraction des biopsies tumorales et contrôles. La fraction TiO_2 correspond à 1.4% de l'extrait peptidique, soit 0.5% de plus que les fractions TiO_2 des biopsies hépatiques. A titre de comparaison, 1 mg de peptides de caséines permet de purifier 108 μ g de phosphopeptides ($n=2$), soit 11% de l'extrait de départ.

Estimation du taux de faux positifs

L'intensité des ions rapporteurs a été déterminée par PD 1.4 au moyen d'un module dédié : *Reporter Ions quantifier*. Les valeurs brutes ont été exportées avec les données d'identification dans un fichier Excel puis traitées par une macro VBA. Cette macro VBA crée plusieurs feuilles Excel qui correspondent à (1) la liste des peptides sans redondance (2) la listes des peptides non phosphorylés (3) la listes des phosphopeptides (4) la listes des protéines phosphorylées. Elle permet également de récupérer les données de quantification brute (signal iTRAQ mesuré par

l'Orbitrap sans normalisation ou calcul de moyennes). La **Figure 46** présente les intensités des ions rapporteurs en fonction de la nature du peptide ; phosphorylé ou non. Il apparaît que la phosphorylation ne semble pas influencer la fragmentation HCD ; quelle que soit la nature du peptide, Les valeurs se distribuent sur plus de 3 ordres de grandeur, l'intervalle interquartile est identique et les médianes sont comparables.

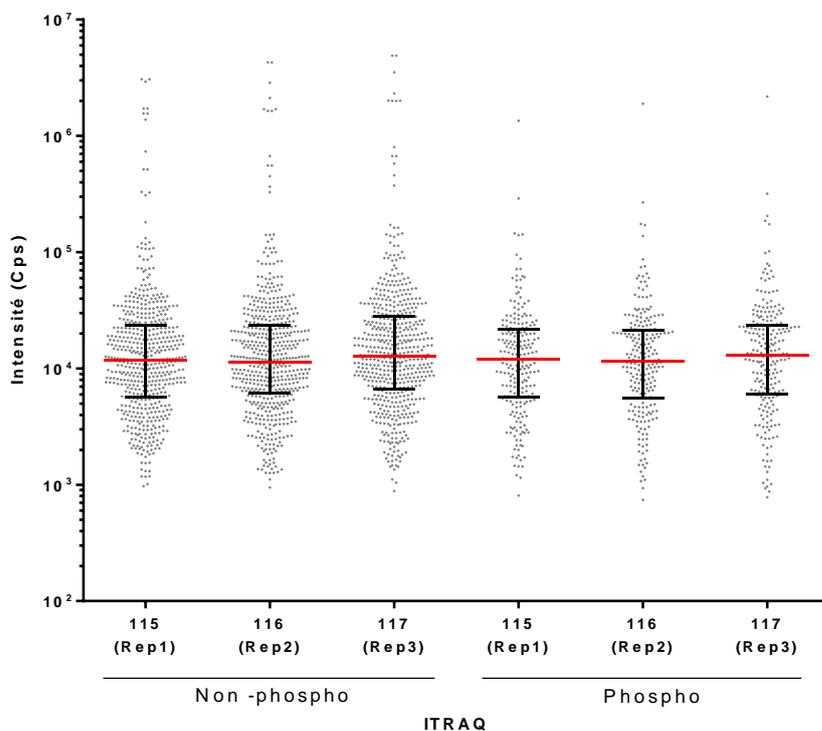


Figure 46. Distribution des données iTRAQ. Distribution des intensités en fonction de la nature du peptide.

Gestion des valeurs manquantes et normalisation

Les valeurs manquantes. Les données relatives aux 3 répétitions contenaient des données manquantes. Elles sont restées manquantes et n'apparaissent donc pas sur les différentes représentations graphiques de la figure ci-dessus. Cependant, le logiciel PD 1.4 propose de remplacer ces valeurs manquantes par la plus basse des valeurs du marqueur concerné. **La Figure 47** présente pour les 3 répétitions la corrélation entre les intensités des ions rapporteurs 115 (Rep1), 116 (Rep2), et 117 (Rep3) par rapport à l'intensité de l'ion 117. Les points entourés correspondent aux valeurs manquantes. Ils forment une ligne horizontale et verticale, sans corrélation avec la deuxième coordonnée. Le remplacement des valeurs manquantes aurait été pertinent si les points concernés étaient groupés en bas de la droite. Mais ici, ils se distribuent sur 3 ordres de grandeurs, générant des faux positifs lors du calcul des ratios. Les données manquantes n'ont donc pas été substituées par la valeur minimale.

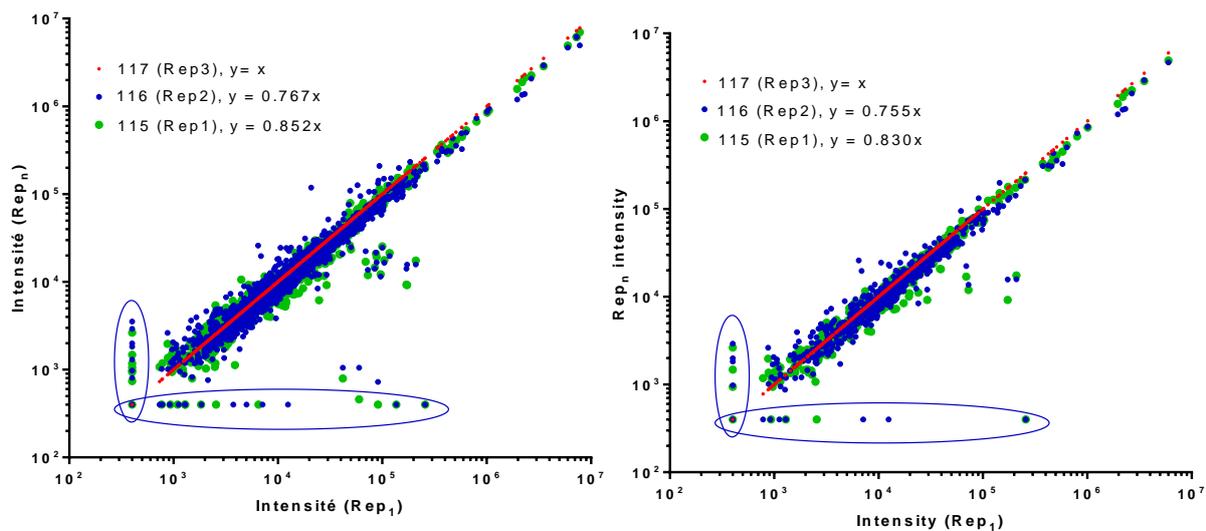


Figure 47. Corrélation entre des différents ions rapporteurs. Les répétitions de préparation et purification TiO₂ d'un foie humain sont marquées par iTRAQ 115, 116, 117. A gauche : corrélation des ions rapporteurs pour chaque scan HCD identifiant un peptide, à droite : corrélation des ions rapporteurs en moyennant les intensités lorsque plusieurs scan HCD sont produit pour le même peptide. Les valeurs manquantes sont entourées.

Valeur moyenne. Un nombre important de peptides sont analysés plusieurs fois. Il est donc possible de moyennner les intensités ITRAQ quand plusieurs scans HCD sont produits pour le même peptide. Cette solution est présentée **Fig. 47**, graphique de droite. La dispersion des points autour de la droite $y=x$ diminue et certains points de données manquantes disparaissent. Ce calcul permet donc d'améliorer sensiblement la qualité des données, il est utilisé par la suite.

Normalisation. La régression linéaire calculée à partir des données ITRAQ brutes définit des droites d'équation $Y_{115} = 0.852X$ et $Y_{116}=0.767X$. En d'autres termes, les signaux des ions rapporteurs 115 et 116 sont globalement moins intenses que ceux des ions 117. Puisque les mêmes quantités de peptides et les mêmes quantités de réactifs ont été utilisées pour chaque échantillon, la somme des intensités pour les différents marqueurs ITRAQ devrait être égale. Une normalisation est donc réalisée pour chaque colonne iTRAQ en divisant les intensités individuelles par la somme de la colonne. De la sorte, chaque colonne de données présente une somme égale à 1.

Calcul des ratios. L'analyse ITRAQ nécessite l'expression des données sous forme de ratio. Généralement, les données sont exprimées en fonction d'un marqueur qui correspond à l'échantillon de référence. De la sorte, si le ratio est par exemple supérieur à 2, il est possible de conclure que le niveau d'expression du peptide est 2 fois supérieur dans l'échantillon par rapport à la référence. En réalisant ce traitement de données pour les 3 répétitions, avec la moyenne des

intensités comme dénominateur, seulement 0.6% des ratios sont en dehors de l'intervalle [0.5-2] communément admis comme une zone de variabilité importante (**Fig. 48**).

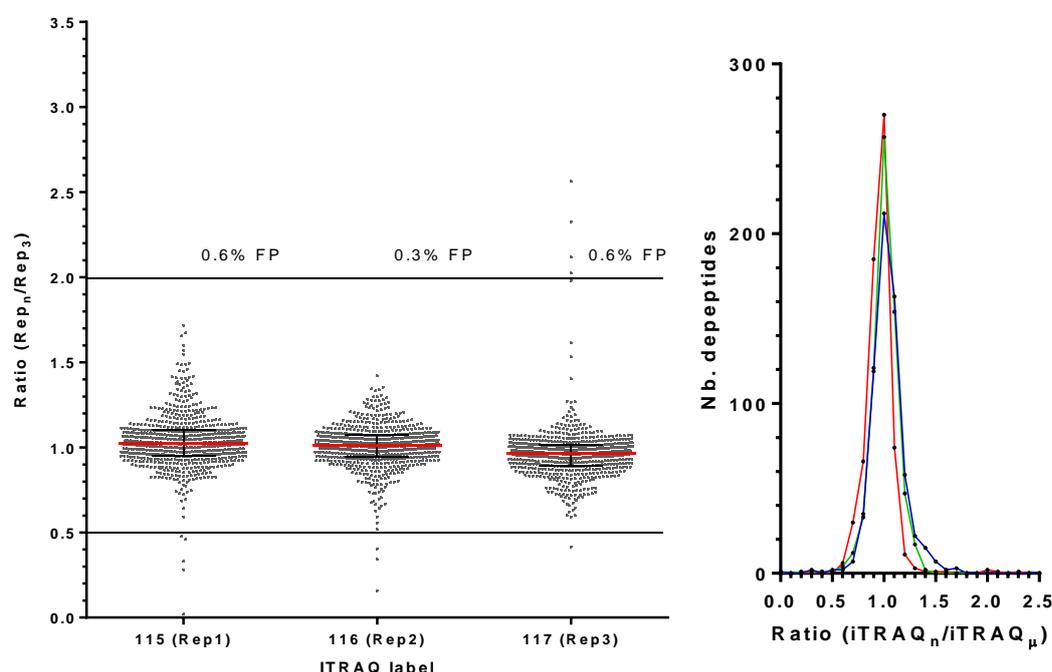


Figure 48. Estimation de taux de faux positif (FP). Analyse en triplicat d'une biopsie humaine et expression des résultats sur forme de ratio. Les répétitions Rep1 (iTRAQ 115), Rep2 (iTRAQ 116), Rep3 (iTRAQ 117) sont exprimées en fonction de leur valeur moyenne μ .

4.3.3 Analyse quantitative de 24 biopsies de tumeurs hépatiques sur foie non fibreux

Analyse LC-MS/MS

Les deux types d'échantillons (peptides totaux et fraction TiO_2) marqués dans 4 séries ITRAQ correspondent à 8 échantillons. Les échantillons peptides-totaux ont été solubilisés dans 150 μL de TFA 0.1%, les fractions TiO_2 ont été reprises par 30 μL , 10 μL de chaque échantillon ont été analysés par LC-MS/MS, soit 26 μg de peptides. Les analyses ont été réalisées sur un LTQ-Orbitrap XL couplé à une nanoLC avec préconcentration. A titre d'exemple, une analyse LC-MS/MS de chaque type d'échantillons est présenté **Figure 49**. En moyenne, les échantillons peptides totaux ont produit 12000 spectres MS/MS alors que les échantillons fraction TiO_2 ont produit 5500 spectres MS/MS. La comparaison des chromatogrammes de courant ionique total (TIC) entre les deux types d'échantillons confirme que les quantités de peptides analysés sont plus importantes pour l'échantillon de peptides totaux. Les ions 480,29 et 406,33 correspondent au réactif ITRAQ non

conjugué à des peptides, ils sont les seuls pics intenses des fractions TiO₂ alors que d'autres pics, des ions multichargés correspondant à des peptides, sont également présents dans les échantillons peptides totaux.

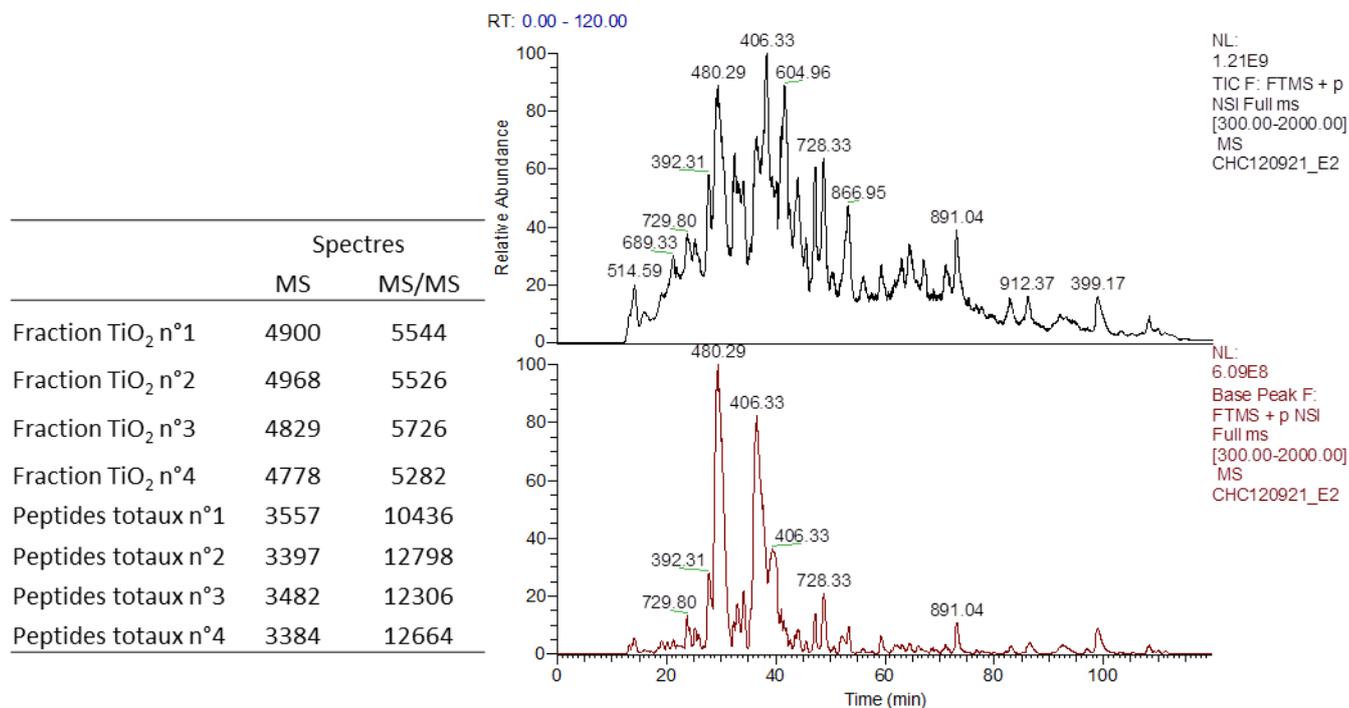


Figure 49. Analyse LC MS/MS des extraits de peptides conjugués au réactif ITRAQ. En haut : intensité du courant ionique Total (TIC : *Total Ion current*), en bas : intensité du pic de base (BP : *Base peak*). Les pics sont annotés de la mesure (*m/z*) de l'ion le plus intense du spectre.

Deux spectres MS/MS sont produits pour chaque ion parent sélectionné. Un spectre MS/MS appelé « spectre CID », est produit dans la trappe linéaire, l'autre spectre MS/MS appelé « spectre HCD » étant produit dans l'Orbitrap après fragmentation de l'ion dans la cellule HCD. Un exemple de spectres CID et HCD est présenté dans la **Figure 50**. Le spectre CID est représentatif des spectres CID de phosphopeptides dont la sérine ou la thréonine sont phosphorylées. La perte de neutre (perte d'acide phosphorique, 98 u) produit un ion majoritaire et les ions fragments b et y sont peu nombreux et peu intenses. Le spectre HCD est réalisé avec une énergie de collision normalisée de 50, soit une augmentation de 15 sur l'échelle de l'instrument par rapport à l'énergie habituellement utilisée pour fragmenter les peptides. Ce réglage particulier permet de favoriser l'observation des ions rapporteurs ITRAQ, au détriment des ions fragments. Les spectres HCD produits dans ces conditions sont donc peu utiles pour l'identification des séquences peptidiques, mais ils permettent une mesure précise des ions rapporteurs. La résolution de l'Orbitrap (7500 pour les spectres HCD) permet une séparation des ions rapporteurs d'éventuels contaminants. Les exemples de spectre CID et HCD de la **Figure 50** correspondent à un ion 2 fois chargé ($m/z = 546.25$). Le spectre CID

présente un ion majoritaire, 49 u en dessous de la masse du précurseur. Cet écart de masse est caractéristique de la perte de neutre qui correspond à l'élimination du phosphate (98 Da), les autres fragments sont alors peu intenses. Le spectre HCD ne présente pas de perte de neutre, les fragments observés sont peu nombreux et les ions majoritaires correspondent aux ions rapporteurs iTRAQ. En insert, le zoom présente les données en mode profil dans l'intervalle 113-121 m/z. Pour ce peptide, les deux références 119 et 121 sont plus intenses que les ions rapporteurs 113 à 118 qui correspondent aux tumeurs. La présence d'un ion de m/z = 120.08 correspond certainement à l'ion immonium d'une phénylalamine présente dans la séquence, alors que l'ion m/z = 113.73 ne correspond à aucun ion immonium. Ici, la présence d'ions contaminants démontre l'intérêt de réaliser la mesure avec une résolution élevée. Sur l'ensemble des spectres MS/MS, des contaminants ont été observés à différents m/z : 113.07, 113.36, 113.39, 113.72, 114.48, 114.55, 114.66, 114.88, 115.09, 115.18, 115.41, 115.46, 115.59, 115.70, 115.87, 115.96, 116.07, 116.25, 116.54, 116.79, 116.98, 117.04, 117.17, 117.97, 118.01, 118.07, 118.35, 118.48, 118.81, 119.26, 119.74, etc. Pour tous ces ions, les intensités représentent entre 1 et 30% de l'intensité de l'ion iTRAQ adjacent.

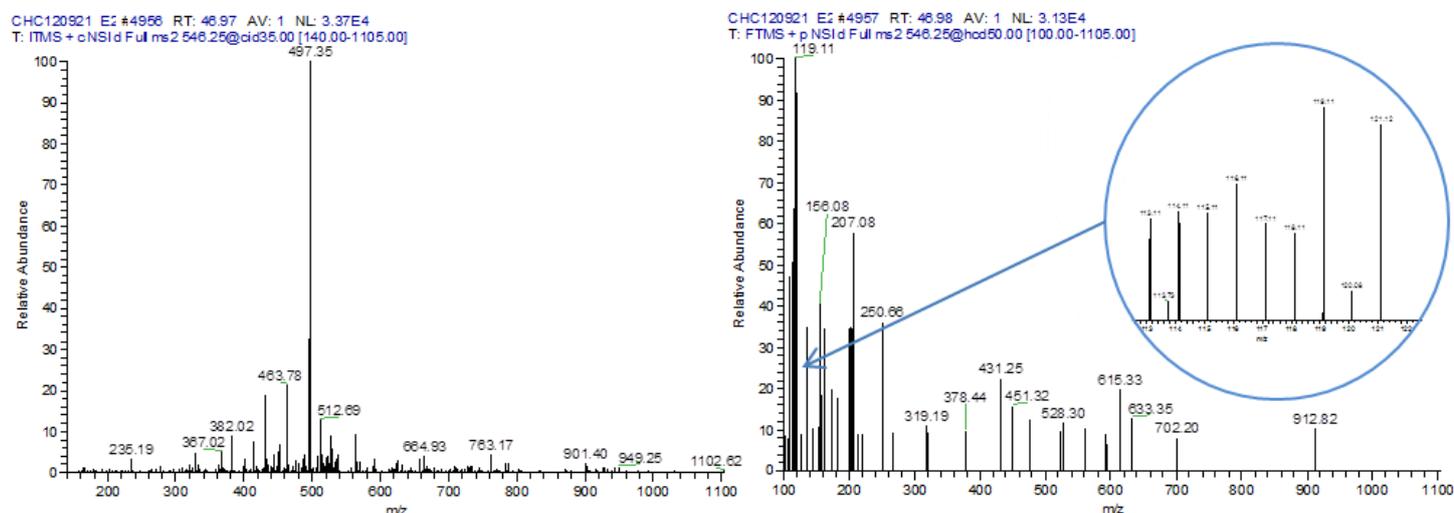


Figure 50. Spectres MS/MS de l'ion 546.26. A gauche : mesure dans la trappe linéaire (spectre CID), à droite : mesure dans l'Orbitrap (spectre HCD). En insert un zoom correspondant aux fragments rapporteurs 113 à 121. Ici les spectres MS/MS CID et HCD ne permettent pas d'identifier le phosphopeptide analysé.

Caractéristiques des peptides identifiés

Identification des peptides par PD 1.4. Les peptides identifiés par Proteome Discoverer 1.4 (voir Matériel et Méthodes) ont été exportés dans Excel et une macro VBA permet d'extraire les principales caractéristiques des peptides identifiés (**Tab.14**) :

- Le pourcentage de phosphopeptides

- Le nombre de peptides avec 1, 2 ou 3 phosphates
- Les sites de phosphorylation S, T,Y
- Le pourcentage de peptides marqués par iTRAQ (Cette option a été utilisée dans un premier temps pour évaluer l'efficacité de marquage)
- Le dernier acide aminé (K, R ou autre).

Tableau 14. Caractéristiques des échantillons fraction TiO₂ et peptides totaux*.

	Echantillons TiO ₂		Echantillons peptides totaux	
	Peptides	Protéines	Peptides	Protéines
0 Phos.	187	84	2024	533
1 Phos.	270	146	2	2
2 Phos.	43	38	0	0
3 Phos.	2	15	0	0
phospho.	315	214	2	2
phospho. (%)	63	72	0	0
pS	321		2	
pT	39		0	
pY	2		0	
Peptides avec 0 ITRAQ	0		0	
Peptides avec 1 ITRAQ	307		967	
Peptides avec 2 ITRAQ	191		1055	
Peptides avec 3 ITRAQ	4		4	
C-ter = K	184		1047	
C-ter =R	279		945	
C-ter (non R/K)	39		34	
Total	502		2026	

* Pour les deux types d'échantillons, les données des 4 séries ITRAQ ont été regroupées. Les valeurs correspondent au nombre de peptides uniques (les doublons ont été éliminés).

Efficacité du marquage ITRAQ (Tab. 14). Une analyse préliminaire par PD 1.4 a été réalisée en fixant la modification ITRAQ comme dynamique (i.e. variable), laissant ainsi au logiciel la possibilité d'identifier des peptides marqués ou non. Avec ces paramètres, aucun peptide non marqué n'a été identifié ; à la totalité des amines N-terminales et des amines epsilon des lysines sont donc couplées au réactif ITRAQ. Aucune arginine n'est conjuguée au réactif. Les interrogations suivantes ont donc été réalisées en considérant le marquage iTRAQ comme statique (i.e. systématique) sur les amines concernées.

Caractéristique des échantillons de peptides totaux (Tab. 14). Pour les échantillons de peptides totaux, aucun enrichissement TiO₂ n'a été réalisé et seuls 2 phosphopeptides sont identifiés sur les 2024 peptides identifiés. Dans les conditions d'analyses utilisées, il apparaît clairement que la complexité du mélange ne permet pas la détection de phosphopeptides sans enrichissement de cette fraction minoritaire. La population de peptides identifiés contient en position C-terminale autant de lysines que d'arginines. Fort logiquement, 50% des peptides sont

marqués par 2 ITRAQ (peptides avec une lysine C-terminale), les 50% restant ne portant qu'un seul réactif ITRAQ (peptides avec une arginine C-terminale). Quelques rares peptides sont marqués par 3 ITRAQ et correspondent à peptides dont la digestion par la trypsine est incomplète ; la lysine supplémentaire est alors marquée par un troisième réactif ITRAQ.

Caractéristiques des échantillons fraction TiO₂ (Tab. 14). La proportion de phosphopeptides dans les échantillons TiO₂ (63%) est inférieure à la proportion de phosphopeptides obtenue à partir d'un foie de souris (75%, voir chapitre 3). Par contre la proportion pS/pt/pY (100/10/1) est identique aux proportions observées précédemment.

Dans ces fraction TiO₂, l'analyse des acides aminés N-terminaux montre que la proportion de peptides comportant une arginine N-terminale est anormalement élevée, avec un ratio K/R = 1.5 alors que K/R = 1 pour la fraction de peptides totaux. (**Tab. 14**). Il faut noter que le ratio élevé est propre aux fractions TiO₂ marquées par ITRAQ, les fractions TiO₂ non marquées ayant une proportion K/R = 1 (**Chap 3, Tab. 11**). Un test du Chi 2 permet de vérifier que la différence de proportion est significative uniquement entre les échantillons TiO₂ + ITRAQ (p<0.001). La surreprésentation des peptides à arginine n'est donc pas spécifique de la purification TiO₂ ou du marquage ITRAQ mais de l'association « purification TiO₂ + marquage ITRAQ ». Fort logiquement, cette proportion K/R anormalement élevée se répercute sur la proportion de peptides marqués avec 1 ou 2 ITRAQ. Cette observation n'est pas expliquée ; une partie des phosphopeptides portant une lysine N-terminale ne sont pas analysés par le spectromètre de masse.

Analyse quantitative des données relatives aux références C1 et C2.

Pour chaque série ITRAQ, les ions reporteurs 119 et 121 correspondent aux échantillons références C1 et C2. Ces échantillons sont chacun des mélanges de 3 tissus normaux analysés auparavant par le groupe du Pr Zucman-Rossi (U674) et dépourvus de mutations associées au carcinome hépatocellulaire. Les ions 119 et 121 devraient donc présenter des corrélations fortes et un ratio 119/121 compris dans l'intervalle [0.5-2]. **La Figure 51** présente la corrélation entre C1 et C2 ainsi que les ratios 119/121 pour l'ensemble des peptides identifiés dans les fractions TiO₂ et les fractions peptides totaux ; 0.6% des ratios sont en dehors de l'intervalle [0.5-2]. Les 40 ratios significativement différents de 1 peuvent correspondre à des peptides exprimés à des niveaux différents entre C1 et C2. Ici, ils sont considérés comme de « vrais » faux positif dus aux erreurs de mesure.

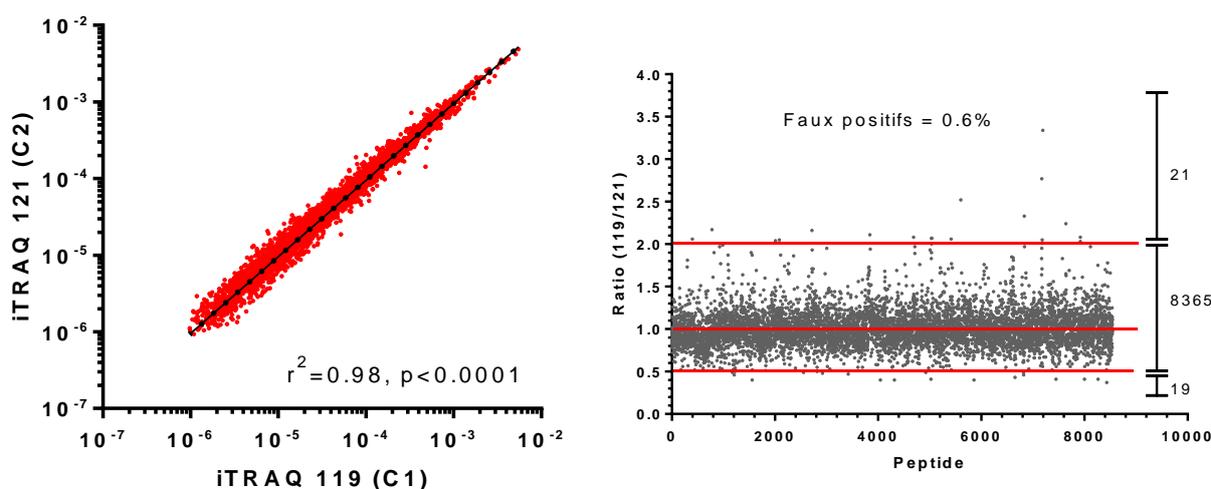


Figure 51. Analyse des variations entre C1 et C2. A droite : corrélation pour l'ensemble des valeurs 119 (C1) et 121 (C2) obtenues avec toutes les fractions TiO_2 et toutes les fractions de peptides totaux. A gauche : ratio 119/121.

Un certain nombre de spectres HCD ne présente un signal que pour l'un ou l'autre des deux ions rapporteurs, bien que le logiciel PD 1.4 propose un remplacement des valeurs manquantes par la valeur minimale de la série. Comme pour le jeu de données précédents, l'utilisation de ce paramètre se traduit sur le graphique de corrélation par la formation d'une flèche pointant vers l'origine, c'est à dire une absence de corrélations entre C1 et C2 pour les intensités faibles.

La qualité de la corrélation entre les intensités 119 (C1) et 121 (C2) a été vérifiée pour chacune des 4 séries de marquage ($r < 0.98$, $p < 0.0001$). Ayant également vérifié que la moyenne des intensités $m_{(119, 121)}$ pour les références C1 et C2 permettait de diminuer la dispersion des points de part et d'autre de la droite de régression linéaire, les ratios ITRAQ ont été calculés avec comme dénominateur la moyenne des intensités iTRAQ 119 et iTRAQ 121.

Détermination des faux positifs à partir de la série 4

La série 4 contient les 6 biopsies normales utilisées en mélange 3 par 3 pour constituer les références C1(119) et C2(121). Les moyennes des ratios individuels pour iTRAQ-113, -114, -115 d'une part, et -116, -117, -118 d'autre part devraient donc être égales aux ratios de iTRAQ-119 et -121 respectivement.

Pour la fraction TiO_2 , 118 phosphopeptides sont identifiés et quantifiés dans la série 4. Sur les 236 ratios calculés, 231 ratios sont compris dans l'intervalle [0.5-2] et 5 ratios sont supérieurs à 2, soit 2% de résultats faux positifs. La **Figure 52** présente un exemple de spectre HCD pour les deux cas de figure rencontrés. A gauche, les signaux iTRAQ sont similaires. Si iTRAQ-118 est moins

intense, son ratio est pondéré par les ratios iTRAQ-116 et -117 et leur moyenne est proche de 1. A droite, il apparait clairement que la moyenne des signaux $m_{(113,114,115)}$ et $m_{(116,117,118)}$ est 3 fois supérieure au signaux 119 et 121. Dans ce cas, la comparaison des ratios aboutira à une conclusion illogique ; mélanger 3 solutions ne peut pas aboutir à un mélange 3 fois moins concentré. Aucune explication ne permet d'expliquer ces 2% de quantifications aberrantes. Ces quantifications aberrantes dans la série 4 ont été utilisées pour éliminer les peptides correspondants dans les séries de tumeurs.

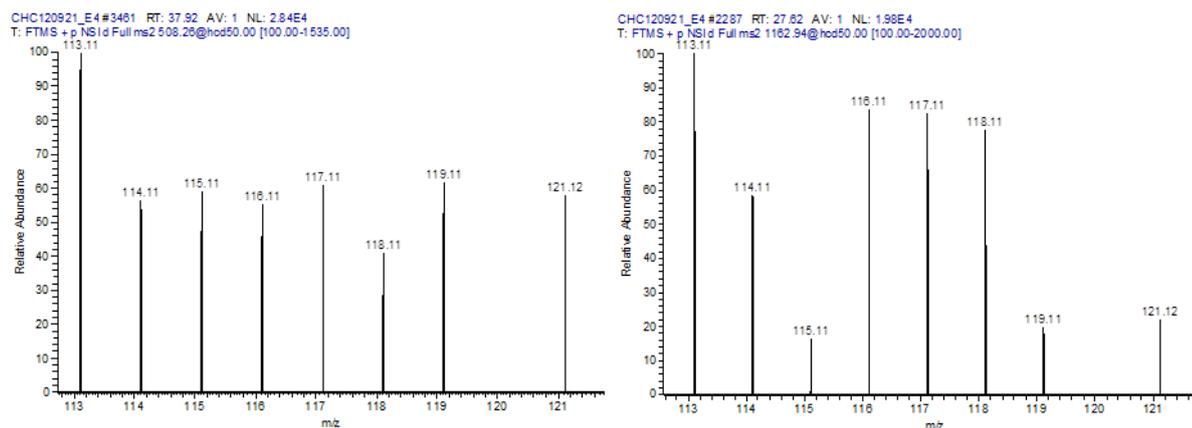


Figure 52. Intensité des ions rapporteurs de la série 4. A gauche : intensités généralement observées, le calcul des ratios permet de vérifier que 119(C1) et 121 (C2) sont respectivement un mélange de 113, 114, 115 et 116, 117, 118, à droite : distribution anormale des intensités.

Identification des peptides communs aux différentes séries

Quatre séries de marquage iTRAQ ont été réalisées, dont trois contiennent les biopsies tumorales. Pour les trois listes de peptides correspondantes, le point commun de ces listes est la présence pour chaque peptide d'au moins un ion iTRAQ 119 ou 121 afin de permettre le calcul des ratios. En d'autre terme, Si un peptide est identifié dans une série, il devrait également être identifié dans les autres séries puisque C1 (119) et C2 (121) sont présents dans toutes les séries. L'analyse de la redondance entre série est illustrée par des diagrammes de Venn et montre un tout autre résultat ; le recouvrement des listes de peptides est faible (**Fig. 53**). Pour la fraction peptide totaux, 402 peptides sont communs aux trois séries de tumeurs. Pour la fraction TiO_2 , l'ensemble des peptides communs aux 3 séries de biopsies tumorales représente 1/3 des peptides identifiés, soit 109 peptides.

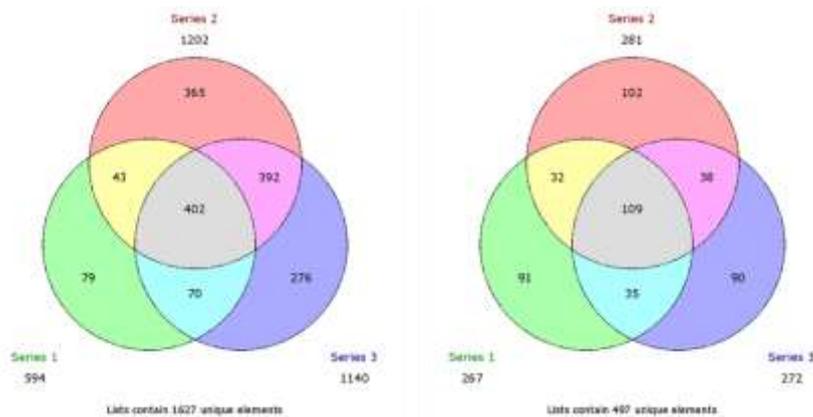


Fig. 53. Diagramme de Venn. A gauche : peptides totaux, à droite : peptide de la fraction phosphorylée.

Caractéristiques générales des protéines identifiées

L'utilisation des annotations de la Gene Ontology (GO) permet de classer les protéines identifiées par compartiment cellulaire, activité cellulaire ou fonction moléculaire. La majorité des protéines identifiées sont d'origine cytoplasmique, mitochondriale, et nucléaires. L'application internet String (<http://string-db.org>) permet de tracer le réseau des interactions connues pour l'ensemble des protéines identifiées (Fig. 54). Avec le jeu de protéines identifiées, des groupes de protéines de dessinent et correspondent aux enzymes du métabolisme (glucides, lipides, protéines), aux enzymes de la détoxification, au ribosome et facteurs associés, aux protéines de choc thermique, aux protéines associées à l'ADN, aux protéines de la matrice, du cytosquelette et aux anticorps.

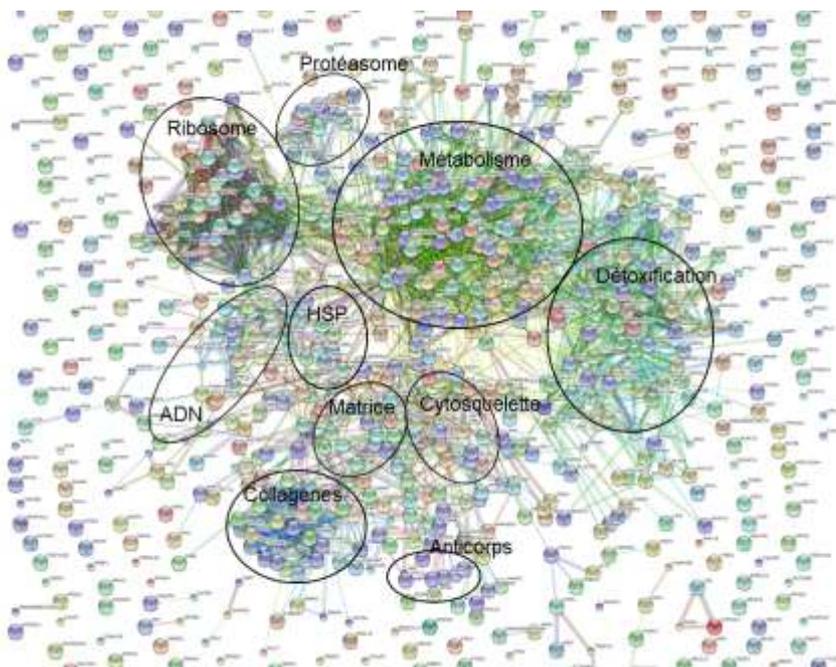


Figure 54. Relations entre les protéines identifiées. Les protéines sont reliées entre elles quand une relation est établie par les bases de données de séquences, d'interactions, de structures, d'articles.

Classification hiérarchique

Vue globale. La classification hiérarchique a été réalisée en considérant chaque peptide comme un gène. Chaque ligne correspond donc à un peptide et chaque colonne à une biopsie. Ici, il ne s'agit pas de déterminer uniquement les ratios significativement différents entre tumeurs (T) et non tumeurs (NT) mais de regrouper les biopsies dont les profils d'expressions sont proches. Seuls les peptides dont les niveaux d'expression sont significativement différents entre T et NT ont été utilisés pour cette classification.

La classification hiérarchique montre une hétérogénéité des biopsies, que ce soit pour les peptides de la fraction TiO_2 et les peptides de la fraction peptides totaux (**Fig. 55**). Il est cependant possible de discerner des clusters de lignes à dominante verte (peptides sous exprimés dans les tumeurs par rapport aux tissus sains) et des clusters de lignes à dominante rouge (peptides surexprimés dans les tumeurs par rapport aux tissus sains).

Pour les peptides totaux comme pour la fraction TiO_2 , la classification hiérarchique des biopsies permet de vérifier que les tissus non tumoraux sont effectivement regroupés dans un cluster unique. Il est intéressant de constater que les contrôles C1 et C2 de chaque série forment deux sous-groupes distincts. Bien que les ratios soient exprimés en fonction de la moyenne de C1 et C2, la classification hiérarchique identifie des différences suffisantes pour reconstituer d'une part un sous-groupe C1, d'autre part un sous-groupe C2. Puis les foies sains s'agrègent sur ces deux sous-groupes à un niveau supérieur de l'arborescence car leur variabilité individuelle les distingue des références C1 et C2.

Pour les peptides totaux, on observe que les biopsies 1.2 et 1.3 ont une signature proche des foies sains ; elles forment un groupe qui s'agrège sur le groupe des biopsies NT (**Fig. 55 A**). Pour les phosphopeptides, les biopsies 1.1 et 1.2 forment également un groupe proche des biopsies NT auquel vient s'ajouter la biopsie 2.1. On peut constater que les phosphopeptides seuls permettent de distinguer plus nettement le groupe des tumeurs du groupe des non tumeurs (**Fig. 55 C**). Par contre, si on ajoute les peptides contaminants de la fraction TiO_2 à la liste des phosphopeptides, un groupe supplémentaire de biopsies tumorales se rapproche des biopsies non tumorales (**Fig. 55 B**).

En comparant les trois classifications hiérarchiques de la **Figure 55**, il apparaît que la distinction entre T et NT est plus nette avec les données provenant de l'analyse des peptides totaux. Il est probable que le plus grand nombre de peptides utilisés pour cette classification améliore la séparation entre tumeurs et non tumeurs. Cependant, les phosphopeptides permettent

un classement similaire (en exceptant la biopsie 2.1) avec un nombre de ligne (i.e. peptides) réduit de deux tiers.

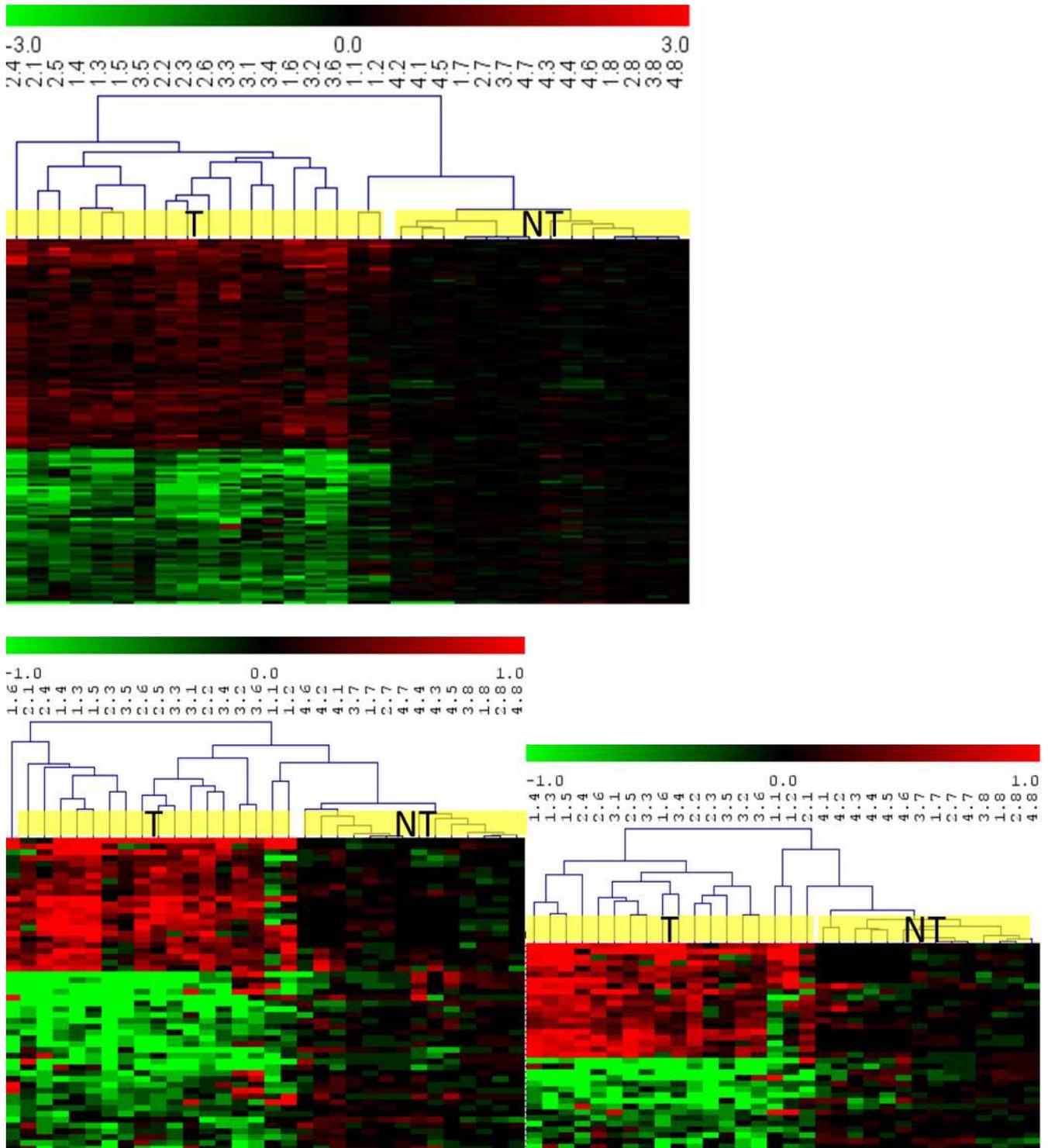


Fig. 55. Classification hiérarchique des biopsies. Chaque ligne correspond à un peptide, chaque colonne à une biopsie. Seuls les peptides significativement différents entre les tumeurs (T) et non tumeurs (NT) sont utilisés pour la classification. **A:** fraction des peptides totaux (106 peptides significatifs entre tumeurs et non tumeurs sur 426 peptides), **B :** fraction TiO₂, phosphopeptides et peptides acides (54 peptides significatifs sur 104 peptides), **C :** fraction TiO₂, phosphopeptides uniquement (36 peptides significatifs sur 66 phosphopeptides).

Cohérence des clusters de peptides et variabilité des échantillons.

Un premier niveau d'analyse permet de constater que des clusters regroupent des peptides appartenant à une même famille de protéines. Par exemple, le cluster présenté dans la **Figure 56** regroupe 11 peptides qui correspondent à des peptides des collagènes COL1A1, COL1A2, COL3A1. Ces peptides ont un profil d'expression très similaire entre eux (comparaison ligne à ligne) mais il apparaît clairement trois groupes de biopsies (comparaison colonne à colonne) ; un groupe sur exprimant les peptides (colonnes de carrés rouges), un groupe sous exprimant les peptides (colonnes de carrés verts) et un groupe dont le niveau d'expression est égal à celui des non tumeurs (colonne de carrés noirs). Malheureusement, pour les biopsies tumorales, ces profils d'expression ne sont pas associés à une caractéristique particulière.

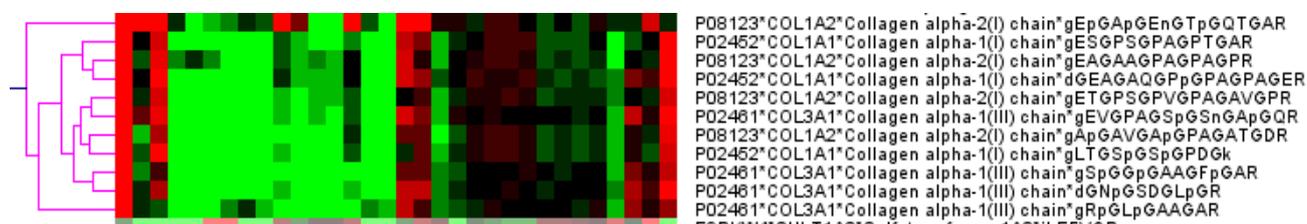


Fig. 56. Regroupement de familles protéiques. Le groupe est constitué de 21 peptides dont 19 correspondent au Collagène 1A1 (P08123), au collagène 1A2 (P02452), Collagène 3A1 (P02461).

De la même manière, un cluster de peptides contient 7 peptides de carboxypeptidase. Treize peptides d'hémoglobines alpha et beta, les deux sous unités de l'hémoglobine sont également regroupées en un seul cluster. Dans tous ces cas de cluster de peptides qui identifient une protéine ou une famille de protéines, le profil d'expression des peptides dépend de la biopsie analysée.

Niveaux d'expression protéique significativement différents entre tumeurs et tissus normaux

Le logiciel Mev (voir Méthodes) a été utilisé pour extraire les données significativement différentes entre les 18 tumeurs et les 6 tissus sains. La liste complète des peptides est présentée en annexe.

Fraction des peptides totaux.

73 peptides sont surexprimés dans les tumeurs, ils correspondent à 49 protéines dont :

- Des protéines de la maturation/dégradation des protéines : HSP90AA1, HSP90B1, HSPD1, HSPA1B, HSPA5, HSPA9, des isomérases (PPIA, P4HB, PDIA6), la valosin-containing protein (VPC), 2 protéines de l'ubiquitination (UBC, UBA1)
- Des protéines du stress oxydatif (PRDX1, TXN)

- Des protéines de la réplication et de la transcription : EEF1G, HNRNPK, HIST1H1C, HIST1H2AH, HIST1H2BK, HIST1H4A
- Des protéines du cytosquelette : ACTN4, CALM1, MYL6, TPM3, TUBA1C, TUBB, VIM
- Des protéines associées à la détoxification : AKR1C4, EPHX1, EPHX1, CES1
- Des enzymes de la glycolyse : ENO1, GAPDG et PGK1

55 peptides sont sous exprimés dans les tumeurs par rapport aux tissus sains et correspondent à 26 protéines parmi lesquelles :

- Des enzymes de la glycolyse : ALDOB, FBP1, GAPDH
- Des enzymes de la détoxification : ADH1A, ADH2A et ADH4, GSTA2, AKR7A3, ALDH1B1
- Des protéines du métabolisme des acides gras : FABP1, ECHS1, ACSL1
- Des protéines sériques : ALB, HBB
- Des enzymes du métabolisme des acides aminés : BHMT, SHMT1
- Des enzymes du cycle de l'urée : CPS1, ASS1

Fraction TiO₂

Les 19 phosphopeptides surexprimés dans les tumeurs par rapport aux tissus sains correspondent à 14 protéines dont :

- Une protéine nucléaire de fonction inconnue : *Neuroblast differentiation-associated protein*(AHNAK)
- La calnexine (CANX)
- Ras GTPase-activating protein-binding protein 1, endoribonucléase liant le domaine SH3 de RASA1 (une protéine activant les Ras GTPase)
- Deux protéines de la transcription des ARN: *Heterogeneous nuclear ribonucleoprotein U-like protein 2* (HNRNPUL2) ; cette protéine lie les pré-mARN dans le noyau et semble modifier entre autre la maturation des ARN, *Transcription elongation factor A protein 1*(TCEA1) ; protéine nécessaire au fonctionnement de l'ARN polymérase II
- Deux protéines associées à la maturation des ARN par le spliceosome (SRRM1 et THRAP3)
- Nuclear casein kinase and cyclin-dependent kinase substrate 1 (NUCKS1); protéine nucléaire de fonction inconnue

Les 15 phosphopeptides sous exprimés dans les tumeurs par rapport aux tissus sains correspondent à 15 protéines dont :

- L'UDP-glucose 6-déhydrogenase qui est impliquée dans la biosynthèse des glycosaminoglycanes, constituant de la matrice extracellulaire
- L'enzyme *2-oxoisovalerate dehydrogenase subunit alpha* (BCKDHA) impliquée dans le catabolisme des acides aminés Leu, Ileu et Val
- *Calcium-regulated heat-stable protein* (CARHSP1); protéine qui inhibe la gluconéogénèse en inhibant notamment G6Pc et PEPCK1 (Fan et al., 2011)
- La protéine *Src substrate cortactin* (CTTN) impliquée dans l'organisation du cytosquelette
- Le facteur *Eukaryotic translation initiation factor 4B* (EIF4B) nécessaire à la liaison de l'ARN au ribosome
- L'*Acétyl-coenzyme A synthétase*, cytoplasmique (ACSS2)
- La *Glutathione S-transférase* (GSTA3) intervenant dans les processus de détoxification des composés électrophiles
- La *Ras GTPase -activating-like protein* (IQGAP2) qui lie CDC42 and RAC1 activés
- NDRG2, protéine qui contribue à la régulation de la voie Wnt
- PRKACB, Protéine kinase AMPc-dépendante
- La *Starch-binding domain-containing protein 1* (STBD1) : protéine probablement membranaire pouvant lier les sucres.

La liste complète des peptides et phosphopeptides dérégulés dans les tumeurs est présentée en annexe.

Analyse des motifs d'acides aminés

Les kinases sont des enzymes spécifiques à la fois pour l'acide aminé phosphorylé et les acides aminés environnants. pLogo (voir Méthodes) est un logiciel en ligne qui permet d'estimer une surreprésentation (ou une sous-représentation) de certains acides aminés dans une séquence. Pour les 19 séquences phosphorylées qui sont significativement surexprimées dans les tumeurs, le motif $-(S/T)P-$ est présent 17 fois. La soumission de ces 19 séquences permet de vérifier que le motif SP est effectivement surreprésenté dans le groupe de phosphopeptides par rapport à

dissocier les événements moléculaires propres au cancer des événements moléculaire propres à la fibrose. On peut cependant noter que les CHC sur foie non cirrhotique (dont la fibrose est aux premiers stades) représentent en France entre 10 et 20% des CHC (Alkofer et al., 2011; Borie et al., 2008) et seraient en augmentation sensible, en association avec l'augmentation du syndrome métabolique dans la population.

La sélection d'un groupe dont la caractéristique (absence de fibrose) est spécifique à seulement ~5% des patients pourrait correspondre à une population homogène de biopsies. Il apparait clairement une hétérogénéité des biopsies qui les définit chacune comme un échantillon unique ; aucun critère anatomo-pathologique n'est commun à deux biopsies. Si certains critères sont significativement surreprésentés (78% d'hommes, 61-67% de patients alcooliques, 67% d'invasion vasculaires, 50% de mutation CTNNB1), les autres critères sont peu fréquents et se distribuent aléatoirement entre les biopsies (stéatoses, stroma fibrotique, mutation P53, nodule satellite, inflammation, etc). Dans ces conditions, il est improbable de distinguer au sein des CHC des sous-groupes. Par contre, une comparaison des 18 nf-CHC avec les 6 tissus sains, par définition plus homogènes car ne possédant pas de signes distinctifs, a été réalisée.

Avantages et inconvénient du marquage iTRAQ. Le réactif iTRAQ permet d'analyser en une seule analyse 8 échantillons. Il semblait donc approprié pour l'analyse de 24 échantillons, à condition d'utiliser une même référence pour les différentes séries de marquage. Quatre séries de marquage iTRAQ ont été réalisées, avec 2 références par série. La série 4 a regroupé les tissus sains dans l'objectif de vérifier que les résultats individuels des tissus sains seraient en accord avec les deux références, chacune correspondant à un mélange de 3 tissus sains. Cette vérification a permis d'identifier 2% de peptides dont les ratios étaient différents d'un facteur 2 par rapport aux références. Ces peptides ont été éliminés de l'analyse. Par ailleurs, chaque série iTRAQ n'a présenté que 1/3 de peptides communs avec les autres séries. Seuls ces peptides communs ont été utilisés pour l'analyse quantitative. Enfin, certains peptides n'ont pas produit de signal pour les ions reporteurs correspondant aux références. Ces peptides ont été également éliminés de l'analyse. En fin de sélection, il restait 57 phosphopeptides et 36 peptides acides dans la fraction éluée de la colonne TiO₂ et 387 peptides dans la fraction peptides totaux. Ce nombre limité de peptides est la conséquence directe des faibles quantités d'échantillon, entre 50 et 100 mg, de la chromatographie minimaliste (simple TiO₂) et du temps d'analyse LC-MS réduit (2h). Les quantités de biopsies n'étaient pas ajustables, par contre un temps d'analyse LC-MS plus long aurait probablement permis d'identifier un plus grand nombre de peptides. La réinjection des échantillons aurait

également permis d'augmenter la proportion de peptides communs aux différentes séries mais les quantités de phosphopeptides purifiés, inférieure à 50 µg, n'ont pas permis d'injections multiples. Enfin, la possibilité d'augmenter le fractionnement en ajoutant par exemple une chromatographie SCX avant l'analyse LC-MS n'a pas été mise en œuvre, toujours en raison de la quantité limitée des phosphopeptides. Ces résultats montrent l'écart important entre la mise en œuvre d'une technique sur un échantillon modèle, souvent abondant et homogène, et la mise en œuvre sur des échantillons cliniques dont les quantités sont limitées.

Aujourd'hui, la nécessité d'un marquage n'est plus impérative. La quantification par *label free* s'est généralisée au point de devenir la technique la plus employée pour l'analyse quantitative globale. Il est certain que cette approche pourrait être employée pour l'analyse des 24 biopsies. Techniquement plus facile, elle permettrait également de s'affranchir des problèmes associés aux résultats sous forme de ratios. Dans le cas d'une poursuite du projet sur un plus grand nombre de biopsies, le dosage *label free* serait certainement utilisé. La technique iTRAQ conserve cependant l'intérêt de disposer d'un contrôle interne (la référence) qui permet dans le cas présent la mise en évidence d'un recouvrement restreint des peptides identifiés entre les séries. La technique *label free* ne disposant d'aucune référence interne, l'estimation des différences de type absence/présence et les différences observées pour les faibles AUC sont difficiles à exploiter, la sensibilité de l'analyse ne pouvant être garantie comme constante tout au long des jours voir des semaines d'acquisition.

Une signature protéomique spécifiques aux nf-CHC. La grande variabilité des tumeurs analysée n'a pas permis d'établir des groupes significatifs parmi les biopsies nf-CHC. La classification hiérarchique permet cependant la différenciation entre les tissus sains et les tumeurs. Dans le groupe des tumeurs, des profils d'expression spécifiques apparaissent mais le nombre d'échantillons restreint ne permet pas d'identifier un sous-groupe particulier. Si on excepte le marquage iTRAQ, le protocole mis en œuvre est simple et robuste. Il serait facilement mis en œuvre avec une approche *label free* pour analyser un panel important de biopsies et établir ainsi un nombre suffisant de profils protéomiques pour définir des groupes dont on peut espérer établir un lien avec le diagnostic ou le pronostic.

Les peptides surexprimés dans les nf-CHC. L'analyse de la fraction peptides totaux permet d'identifier 35 peptides surexprimés et 55 peptides sous exprimés. Des tendances générales se dégagent de ce jeu de données. Sont surexprimés les protéines de choc thermiques (HP10, 60, 70, 90), les protéines liées à l'ADN/ARN (histones), les protéines de détoxification (carboxyestérase,

époxyde hydrolase), les protéines du cytosquelette (actinine, myosine, tubuline), des protéines ou enzymes anti-oxydantes (superoxyde dismutase, thiorédoxine).

Les protéines de choc thermique (HSP) sont une famille de protéines dont l'induction permet aux cellules de survivre dans des conditions de stress. Les HSP sont classées en 6 familles suivant leur poids moléculaires. Ce sont des protéines chaperonnes nécessaires à la bonne conformation des protéines ou à leur élimination dans le cas de dommages irréversibles. Les études protéomiques de CHC ont mis en évidence la surexpression de nombreuses HSP (Khalil et al., 2011) et notamment l'augmentation des concentrations d'HSP60 (Sun et al., 2009), 70 (Zeindl-Eberhart et al., 2004) et 90 (Lim et al., 2002; Sun et al., 2010b).

La cellule cancéreuse subit un stress oxydatif et la thiorédoxine (protéine anti-oxydante) ainsi que la superoxyde dismutase (enzyme anti-oxydante) ont été proposées comme biomarqueurs du CHC (Tamai et al., 2011). Ces deux marqueurs ont effectivement été identifiés dans les biopsies nf-CHC. Concernant les enzymes de détoxification, la carboxyestérase 1 en association avec l'AFP a été récemment présentée comme un biomarqueur performant du CHC (Na et al., 2013). L'expression de l'époxyde hydrolase est également forte à modérée dans les coupes immunohistochimiques de CHC (Coller et al., 2001). La surexpression de la carboxyestérase 1 et de l'époxyde hydrolase ont bien été mises en évidence dans les biopsies nf-CHC, en association avec la ferritine qui peut également traduire une inflammation.

Une augmentation du niveau de la glycolyse dans la production d'énergie système des cellules cancéreuses a été initialement trouvée par le célèbre biochimiste Otto Warburg (Warburg, 1956) et la production des ARN de 8 des 14 gènes de la glycolyse est augmentée dans les cellules CHC (Ganapathy-Kanniappan et al., 2012). Une surexpression de la GAPDH serait également associée à une augmentation de la prolifération et de la chimio-résistance dans le CHC (Ganapathy-Kanniappan et al., 2012). Ici, la GAPDH et l'énolase sont surexprimées dans les nf-CHC. L'énolase est une enzyme de la glycolyse qui agit également comme récepteur de surface du plasminogène. Sa surexpression a également été démontrée dans le cancer du pancréas (Cappello et al., 2009).

Concernant les protéines liant d'ADN ou l'ARN, plusieurs peptides phosphorylés de *Serine arginine repetitive matrix protein* sont surexprimés. Cette protéine est une sous unité du spliceosome qui assure l'excision des introns. Elle serait également induite lors des cancers et sa déplétion par SiRNA permet de réduire la viabilité des lignées cancéreuses par l'activation de l'autophagie (apparition de vacuoles acides) (Quidville et al., 2013). La *Nuclear ubiquitous casein*

and cyclin-dependent kinase substrate est une protéine liant d'ADN et de fonction inconnue. Elle est surexprimée dans 83.5% des cellules cancéreuses du sein (Ziolkowski et al., 2009).

La surexpression des histones H2A et H4 n'est pas documentée dans la littérature. Par contre, la méthylation/acétylation des lysines sont des modifications post-traductionnelles associées aux histones. Elles permettent de limiter l'interaction des histones (chargées positivement) avec l'ADN (chargé négativement) et entraînent une relaxation de la chromatine nécessaire à la transcription. Les déacétylases sont surexprimées dans de nombreux cancers et sont proposées comme cible thérapeutiques (Berry and Janknecht, 2013). Inversement, la surexpression d'histone transacétylase induit l'apoptose de cellules CHC et inhibe la voie de signalisation akt (Zheng et al., 2013). Ici, aucun peptide acétylé ou méthylé n'est identifié dans les nf-CHC mais la surexpression de 8 peptides appartenant aux histones H2A et H4 est observée.

Les peptides sous exprimés dans les nf-CHC. Les peptides sous exprimés dans les nf-CHC correspondent à des protéines dont les fonctions se distinguent des protéines précédentes. L'alcool déshydrogénase est la protéine identifiée avec le plus grand nombre de peptides. La sous-expression de sa forme ADH4 a été proposée comme marqueur pronostique du CHC (Wei et al., 2012). Ici, la sous expression des isoformes ADH1A et ADH2A est également observée.

De la même manière, la sous expression de 5 protéines du cycle de l'urée avait été identifiée par une analyse protéomique d'une biopsie CHC, puis validée par immunohistochimie (Chaerkady et al., 2008). Ici, 2 protéines apparaissent sous-exprimées dans les nf-CHC (Carbamoyl-phosphate synthase, 6 peptides et Argininosuccinate synthase, 1 peptide).

Quelques peptides appartenant à des protéines de la glycolyse indiquent une sous expression de ces dernières dans les CHC. L'analyse des protéines surexprimées montre que cette voie est activée dans le nf-CHC. La sous expression de Fructose-1,6-bisphosphatase et la Fructose-bisphosphate aldolase B identifiée ici est en accord avec la littérature (Liu et al., 2011; Wang et al., 2011b).

De manière remarquable, alors que deux enzymes de de détoxification sont surexprimées (CES1 et EPHX1), deux autres enzymes sont sous exprimées : la glutathion S-transférase A2 et l'*aflatoxin B1 aldehyde reductase*. La GSTA est une enzyme impliquée à la fois dans les processus de détoxifications et le stress oxydant. Une diminution de l'expression de GSTA a été observée dans le cancer du rein, alors qu'une augmentation est observée dans le cancer de la prostate. Un milieu Pro-oxydant pourrait promouvoir la carcinogénèse du rein (Simic et al., 2009).

La sous expression l'UDP-glucose 6-déhydrogénase (G6PD) détectée dans la fraction peptides totaux est également confirmée dans la fraction TiO₂ ; un peptide phosphorylé de la G6PD y est également significativement sous-exprimés. La G6PD est une enzyme cruciale pour la production de NADPH et l'homéostasie du potentiel redox. Il a été montré qu'une déficience de G6PD et de GST prédisposait les cellules à un stress oxydatif et une mort cellulaire (Ho et al., 2013). Dans les nf-CHC analysés, Le système de détoxification est apparemment dérégulé, avec une surexpression des enzymes de la phase 1 et une sous expression des enzymes de la phase 2.

Interprétation des données relatives aux phosphopeptides. En l'état, les variations quantitatives des phosphopeptides ne peuvent pas être interprétées de manière individuelle, elles doivent être vues comme une signature, l'équivalent d'un code barre qui caractérise les biopsies étudiées. En effet, les variations quantitatives des phosphopeptides peuvent correspondre soit à la modification de l'activité kinase/phosphatase, soit à l'augmentation des protéines substrats. Dans un cas, la conclusion portera sur une dérégulation de l'activité kinase/phosphatase, dans l'autre cas, la conclusion sera une dérégulation de l'expression de la protéine (**Fig. 58**). Cette notion a été détaillée dans un papier de référence présentant la technique AQUA (Gillet et al., 2012). Plus récemment, le même groupe dirigé par S.P. Gygi a publié un deuxième article sur la quantification relative à grande échelle des phosphopeptides et de leur protéine correspondante chez la levure (Han et al., 2008). Cette approche a été reproduite pour *Plasmodium falciparum* étudié à trois stades de son développement. Il apparait que les changements de niveau d'expression des protéines ne sont pas corrélés avec leurs niveaux de phosphorylation (Pote et al., 2013).

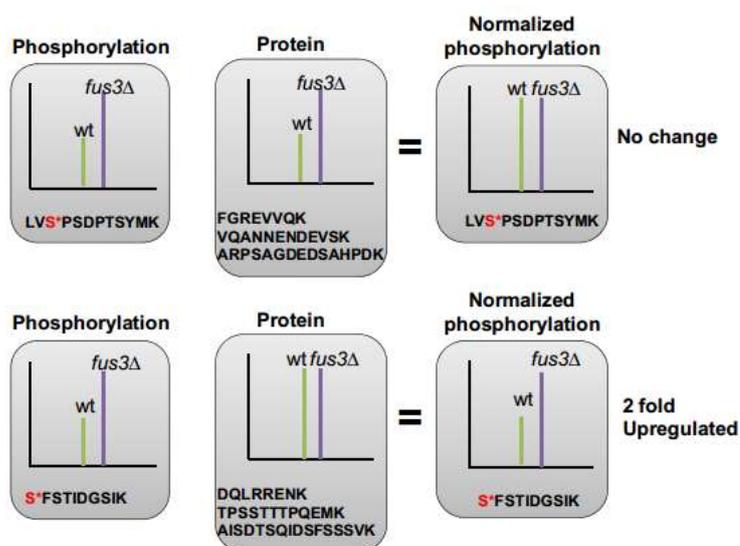


Figure 58. L'interprétation des changements de phosphorylation nécessite une normalisation par l'expression des protéines. En haut : la quantité du phosphopeptide et la quantité de la protéine correspondante sont toutes les deux augmentées d'un facteur 2 : le niveau de phosphorylation est inchangé après normalisation. En bas : le phosphopeptide

est deux fois plus abondant chez le mutant *fus3* et le niveau d'expression de la protéine n'est pas augmentée : le niveau de phosphorylation est augmenté d'un facteur 2 après normalisation. D'après (Han et al., 2008).

La calnexine. La calnexine est une protéine membranaire abondante du réticulum endoplasmique des cellules eucaryotes. Sa partie luminale possède une séquence de liaison aux N-glycoprotéines qui lui conférerait une fonction de repliement des glycoprotéines. La phosphorylation Ser 563 contrôlerait l'activité chaperonne alors que les phosphorylations de Ser 534 et Ser 544 seraient liées à une fonction de transport (Chevet et al., 2010). Par ailleurs, la calnexine interagit avec la calréticuline et ERp57 (Wang et al., 2012b). Ici, la phosphorylation en position 534 n'a pas été identifiée mais les peptide phosphorylés comportant Ser 544 et Ser 563 sont significativement surexprimés dans les tumeurs. L'augmentation de cette activité chaperonne sous contrôle d'une phosphorylation est une illustration supplémentaire de cette activité particulière dans les tumeurs.

Surreprésentation du motif –SP–. Le motif –SP– est le motif le plus fréquemment rencontré dans les identifications de phosphopeptides en grand nombre. L'article le plus significatif utilise un marquage diméthyl pour analyser une tumeur CHC et un tissu sain ; plus de 1800 phosphopeptides sont identifiés en 42 h d'analyse et le motif –SP– apparaît être surreprésenté à la fois dans les 309 peptides surexprimés et les 471 phosphopeptides sous-exprimés (Song et al., 2011). Ici, la surreprésentation de ce site *Proline directed kinase* est spécifique des phosphopeptides surexprimés dans les tumeurs alors que les phosphopeptides sous exprimés ne sont pas enrichis en ce motif. Les *Cyclin dependent Kinases* (CDK), les *Mitogen-activated protein kinases* (MAPK) et les *Glycogen synthase kinase 3* (GSK3 A et B) sont des *Proline directed kinases* au spectre d'action très étendu. Par exemple ERK1/2 catalyse la phosphorylation de centaines de substrats dans le cytoplasme et le noyau et la cascade Ras-Raf-MEK-ERK participe à la régulation de nombreuses activités cellulaires : adhésion cellulaire, migration, différenciation, prolifération et transcription (Roskoski, 2012). La GSK3 est par contre spécifique de la glycogène synthase qui est inhibé après phosphorylation.

Une analyse informatique de la littérature a proposé une liste restreinte de biomarqueurs du CHC. Une protéine significativement augmentée (HSP70) et 4 protéines sous-exprimées (fructose-1,6-bisphosphatase 1, formiminotransférase cyclodéaminase, alcool déshydrogénase et fructose-bisphosphate aldolase B) ont été proposées comme biomarqueurs consensus (Liu et al., 2011). Ici, ces 5 biomarqueurs potentiels sont identifiés mais la liste de protéines candidates est étendue et

des sites de phosphorylation activés dans le cadre du nf-CHC sont proposés. La séquence –(S/T)P- qui est significativement augmentée dans les tumeurs CHC par rapport aux tissus sains fournit une perspective de ciblage des kinases *Prolines directed* dérégulées dans ce type de cancer. En perspective immédiate, la calnexine, la *Thyroid hormone receptor-associated protein 3* et la *Ras GTPase-activating-like protein IQGAP2* ont été choisis parmi les phosphopeptides significativement dérégulés dans les nf-CHC pour tenter de valider les résultats de cette approche quantitative globale par une analyse ciblée par spectrométrie de masse de type MRM.

CONCLUSION GENERALE

L'analyse protéomique est une méthode de choix pour la recherche de biomarqueurs. Sans *a priori*, elle permet d'établir un catalogue de protéines qui sont exprimées dans une cellule, un tissu ou un organisme entier. Cette approche a été mise en œuvre pour une analyse protéomique de cellules coliques cancéreuses et une analyse phosphoprotéomique de biopsies hépatiques.

L'étude de cellules coliques cancéreuses transformées par le gène cytosine désaminase et traitées par la prodrogue 5-fluorocytosine a été réalisée par électrophorèse bidimensionnelle des extraits protéiques. L'analyse d'image a permis la quantification de 353 protéines et isoformes, dont 14 sont surexprimées et 4 sous exprimées lors d'un traitement par la prodrogue. Parmi les protéines dont l'expression est affectée par le traitement, l'HSP90 présente un niveau d'expression constant, mais est identifiée sous deux formes qui diffèrent par leur pI. L'analyse par spectrométrie de masse a permis d'identifier une phosphorylation de la sérine 254 qui pourrait contribuer à la régression tumorale. Cette première étude confirme l'intérêt d'analyser les modifications post-traductionnelles. Les quantités globales d'HSP90 ne varient pas entre les conditions étudiées, Seules les modifications post-traductionnelles permettent de différencier la protéine en question sur la base d'une présence / absence de phosphorylation.

Une étude à caractère technique a défini des conditions d'enrichissement de phosphopeptides qui sont suffisamment robustes pour permettre une analyse quantitative d'échantillons. Les deux chromatographies d'affinité couramment employés pour les analyses phosphoprotéomiques (IMAC et TiO₂) ont été comparées avec comme objectif la sélection d'une seule technique. Les résultats obtenus ont confirmé l'intérêt d'une phase TiO₂ et mis en évidence des caractéristiques originales : (1) les colonnes IMAC utilisée avec du TFA perdent une partie du métal chélaté et présentent un rendement de purification limité, (ii) la chromatographie TiO₂ fonctionne suivant des modes différents en fonction de la teneur en acétonitrile du tampon.

La troisième étude a mis en œuvre le protocole de purification des phosphopeptides établi précédemment. Le réactif iTRAQ 8Plex a été utilisé pour réaliser une analyse protéomique quantitative de tumeurs (biopsies de carcinomes hépatocellulaire sur foie non fibreux) par rapport à du tissu hépatique sain. Les protéines surexprimées dans les tumeurs sont des protéines de choc thermique, des protéines liées à l'ADN/ARN (histones, protéine du spliceosome), des protéines de la phase 1 de la détoxification (carboxyestérase, époxide hydrolase), les protéines du cytosquelette (actinine, myosine, tubuline), des protéines ou enzymes anti-oxydantes (superoxide dismutase, thiorédoxine). Les protéines sous-exprimées sont des protéines du cycle de l'urée, de la détoxification (alcool déshydrogénase) du métabolisme des sucres, des lipides et des acides aminés.

Dans le groupe des peptides phosphorylés, 19 phosphopeptides sont significativement surexprimés et 15 phosphopeptides sont significativement sous exprimés. L'analyse quantitative du phosphoprotéome a également mis en évidence une surreprésentation du motif –(S/T)P- parmi les phosphopeptides surexprimés, indiquant une activation des *proline directed kinases* dans les nf-CHC ou une inhibition des phosphatases correspondantes. Ces peptides/protéines dérégulées sont autant de biomarqueurs potentiels pour le carcinome hépatocellulaire.

Valider les peptides candidats par MRM

L'analyse quantitative globale des biopsies hépatiques a identifié plusieurs peptides appartenant à des protéines significativement surexprimées ou sous-exprimées dans les tumeurs hépatiques. Les contrôles effectués ont permis d'éliminer des résultats de quantifications incohérents et la classification hiérarchique des peptides a permis de grouper des peptides identifiant la même protéine. Cette dernière observation peut faire office de validation pour les peptides concernés, mais il reste une partie importante des données quantitatives qui est associée à des peptides isolés. La pluparts des phosphopeptides significativement sur- ou sous-exprimés sont dans ce cas. Pour ces peptides, une validation est nécessaire. Elle peut être réalisée par Western blot ou par immunohistochimie si les anticorps sont disponibles. La spectrométrie de masse offre également une solution par la technique MRM (Multi Reaction Monitoring). Cette solution a été choisie comme perspective immédiate à l'analyse phosphoprotéomique des biopsies hépatiques. Des peptides Aqua™ comportant des traceurs isotopiques ont été commandés. La quantité de peptide natif pourra alors être déterminée à partir du rapport d'intensité observé entre le signal correspondant au peptide naturel et le signal correspondant au peptide marqué ¹³C ¹⁵N, présent en quantité connue. Les peptides sélectionnés à partir de la liste des peptides significativement différents dans les tumeurs sont deux peptides de calnexine (B4E2T8) : ₅₅₄SDAEEDGGTVpSQEEEDR₅₇₀ et ₅₇₄AEDEILNRpSPR₅₈₅, un peptide de *Ras GTPase-activating-like protein* IQGAP2 (Q13576) : ₁₄YGpSIVDDER₂₂ et un peptide de *Thyroid hormone receptor-associated protein 3* (Q9Y2W1) : ₂₉₂IDISPSTFR₃₀₀. La validation sera réalisée à la fois pour les peptides phosphorylés et les peptides déphosphorylés, afin de déterminer le ratio des deux formes.

De l'échantillon modèle à l'échantillon clinique : une frontière visible.

En 2008, Villen et Gygi ont publié un article de référence pour l'identification des phosphopeptides (Villen and Gygi, 2008). A partir d'un foie de souris, 10 mg de peptides tryptiques

ont été utilisés pour la purification des phosphopeptides pS et pT par SCX (11 fractions) et IMAC. L'analyse par LC-MS (LTQ Orbitrap) a permis d'identifier 4038 phosphopeptides dans les 12 fractions analysées, soit 336 phosphopeptides/h. Cette stratégie d'analyse reste aujourd'hui la référence et fournit les idées majeures pour réussir une analyse phosphoprotéomique : (i) puisque les phosphopeptides correspondent à une fraction minoritaire de la population de peptides, une méthode de purification spécifique et des quantités importantes d'échantillon de départ sont nécessaires. (ii) pour un mélange complexe, le nombre de peptides identifiés sera fonction du temps d'analyse en LC-MS/MS. Une chromatographie multidimensionnelle permet d'augmenter considérablement la capacité de pic de la méthode d'analyse.

Cependant, cette stratégie n'est généralement mise en œuvre que pour un seul échantillon. A ce jour, seuls Bonhomme et collaborateurs ont réussi à concilier une chromatographie SCX-IMAC et l'analyse de 40 échantillons (feuille de maïs), soit 400 analyses LC-MS et 3664 phosphopeptides non redondants identifiés (Bonhomme et al., 2012). Par contre, l'approche SCX-IMAC n'a jamais été appliquée à des séries d'échantillons cliniques dont les quantités sont limitées.

Dans le cadre de la partie 4 de la thèse, la stratégie mise en œuvre a été différente car l'objectif principal n'était pas la performance en termes d'identification, mais la répétabilité du protocole pour permettre l'analyse quantitative d'un grand nombre d'échantillons. Cette condition a justifié de réduire le fractionnement au minimum. La chromatographie multidimensionnelle a été abandonnée au profit d'une simple purification des phosphopeptides par chromatographie TiO_2 . Cet objectif a une conséquence : le nombre de phosphopeptides identifiés est alors réduit. Le deuxième critère qui influence également la performance du protocole est la quantité d'échantillon de départ. Les articles qui présentent des identifications massives de phosphopeptides utilisent des échantillons disponibles en quantité illimitée ; souris, culture cellulaire. Ici, les quantités d'échantillons étaient comprises entre 50 et 100 mg, soit quelques microgrammes de phosphopeptides. Malgré ces limitations, la mise en œuvre du protocole de purification sur du foie humain, sans marquage iTRAQ, a permis d'identifier 619 phosphopeptides non redondants pour 1h d'analyse LC-MS, validant ainsi la capacité de la méthode à identifier des phosphopeptides en nombre avec une préparation minimale des échantillons.

Avec les biopsies hépatiques, le nombre de phosphopeptides identifiés a été considérablement réduit. Après marquage iTRAQ, 100 à 200 phosphopeptides ont été identifiés par biopsie pour une analyse LC-MS de 2h. Cette baisse de performance est certainement due à la méthode d'acquisition utilisée pour les échantillons iTRAQ. L'analyse MS/MS est dédoublée (un spectre CID et un spectre

HCD) pour obtenir respectivement un résultat d'identification et un résultat de quantification. Avec les réglages appliqués à l'instrument (LTQ-Orbitrap XL) les spectres HCD demandent 600 ms, alors que les spectres CID sont produits toute les 200 ms. De ce fait, moins de phosphopeptides ont été mesurés malgré le doublement du temps de la séparation chromatographique. Le nombre de peptides identifiés est donc limité par l'instrument. Il est par contre probable que les nouvelles générations d'instruments Orbitrap qui disposent d'une cellule HCD plus performante et capable de générer des spectres MS/MS à la fois riche en ions de séquences et en ions rapporteurs permettraient d'identifier et quantifier un plus grand nombre de phosphopeptides. Une alternative à tester serait l'approche SWATH (Gillet et al., 2012) qui permet de réaliser avec un spectromètre hybride Q-ToF des spectres MS/MS sur tous les ions précurseurs détectés dans un intervalle restreint de masse (25 u).

Analyser les peptides à phosphotyrosine

Reprenons le fil conducteur de la publication de Villen et collaborateurs (Villen and Gygi, 2008). Les auteurs ont dédié 90 % de leur échantillon pour l'analyse des peptides pY après immunopurification, soit 80 mg de peptides ou environ 1g d'échantillon frais. Grace à la spécificité de la chromatographie d'affinité, 385 phosphopeptides pY non redondants ont été identifiés alors que la purification SCX-IMAC n'avait permis d'identifier que 26 phosphopeptides pY.

Dans le cadre de l'analyse des biopsies, seulement 50 à 100 mg d'échantillon frais étaient disponibles par biopsie, soit 10 fois moins que les quantités utilisées par Villen et al.. L'analyse des phosphopeptides pY n'a donc pas été réalisée, ne disposant pas de quantités d'échantillon suffisantes. Cette limitation est un problème majeur car elle rend inaccessible une population de phosphopeptides impliqués dans des voies de signalisations (PI3K/Akt, MAPK, Jak/Stat, etc) dérégulées dans la plupart des cancers. L'analyse par spectrométrie de masse n'est peut-être pas dans ce cas précis la technique d'analyse la plus appropriée pour produire une analyse quantitative des substrats de kinases ou des kinases elles-mêmes. Les puces à protéines semblent être une approche plus efficace pour analyser des séries d'échantillons cliniques. L'approche RPPA (pour Reverse Phase Protein microArray), une version miniaturisée du dot blot, permettrait de doser des kinases sur des centaines d'échantillons (Kornblau and Coombes, 2011; Malinowsky et al., 2010). Cette approche a été utilisée pour l'étude de l'expression de protéines aujourd'hui inaccessibles par analyse protéomique, notamment PTEN, pAKT, pS6 (un effecteur de la voie PI3K/Akt)(Calderaro et al., 2013). Des (phospho)-peptides de synthèses peuvent également être immobilisés sur les

microarray et la phosphorylation ou déphosphorylation peut être mesurée par une révélation spécifique du phosphate en plus d'une détection par des anticorps spécifiques (Gao et al., 2012; Krishnamoorthy et al., 2013). Les puces à peptides sont maintenant commercialisées, leur mise en œuvre ne demande pas de matériel particulier, et elles correspondent certainement à une technique de criblage qui mériterait d'être évaluée avec des extraits protéiques de biopsies tumorales et des tissus sains.

Développer un approche Top-Down

La première étude présentée dans ce manuscrit a utilisé l'électrophorèse bidimensionnelle pour séparer les protéines, alors que la dernière étude a impliqué l'analyse d'un mélange de peptides sans séparation préalable des protéines. Cette dernière approche est beaucoup plus efficace en terme d'identification puisqu'une seule injection permet d'identifier des milliers voir des dizaines de milliers de peptides pour les plateformes les plus performantes (Kocher T Fau - Swart et al., 2011). Elle présente cependant 2 inconvénients majeurs. D'une part, les peptides identifiés sont interprétés en terme de protéines, bien qu'aucun critère ne permette d'affirmer que les peptides identifiés soient bien le reflet de la protéine entière. D'autre part l'analyse de peptides en mélange provoque des erreurs dans l'identification des isoformes. Cette idée est peu répandue (le point de vue généralement admis consiste à penser qu'un peptide présent uniquement dans une protéine identifie uniquement cette protéine). Le retraitement des données relatives à l'HSP90 (chapitre 2) montre que la situation peut être plus complexe. Dans cet exemple, l'analyse par spectrométrie de masse du spot « acide » correspondant à l'HSP90 a produits deux résultats contradictoires : (i) les deux isoformes HSP90A et l'HSP90B ont été identifiées dans ce spot sur la bases de plusieurs peptides distincts, (ii) dans ce spot, le site de phosphorylation a été identifié uniquement dans un peptide tryptique spécifique de l'HSP90B. Il semble pourtant évident que les deux isoformes, présentant le même décalage de pH et présentes dans le même spot, ont subi la même modification post-traductionnelle. De plus, l'alignement des séquences entre les deux HSP montre que l'HSP90A contient les mêmes sites de phosphorylation que l'HSP90B, et que seuls certains acides aminés, distants du site de phosphorylation, sont différents. En conclusion, lorsque la couverture de séquence des protéines est incomplète, l'absence d'identification d'une modification post-traductionnelle sur une isoforme ne peut pas être interprétée si le peptide non phosphorylé correspondant n'est pas identifié.

L'analyse des modifications post-traductionnelles par l'approche *bottom-up* est confrontée à un deuxième problème d'interprétation. Si on considère les résultats présentés dans le chapitre 4, les phosphorylations identifiées par protéines sont uniques alors que les protéines en question sont généralement modifiées par plusieurs modifications post-traductionnelles. Par exemple PhoshoSitePlus (<http://www.phosphosite.org/>) propose 18 sites de phosphorylation pour la calnexine (seulement 3 sites sont caractérisés par une autre technique que la spectrométrie de masse) alors qu'une seule phosphorylation est identifiée dans les extraits analysés. Si on considère comme exacts les 3 seuls sites vérifiés, 6 combinaisons de phosphorylation sont envisageables. Par ailleurs, la calnexine est également acétylée et ubiquitinée. L'information relative à la combinaison des PTM est perdue dans le mélange de peptides.

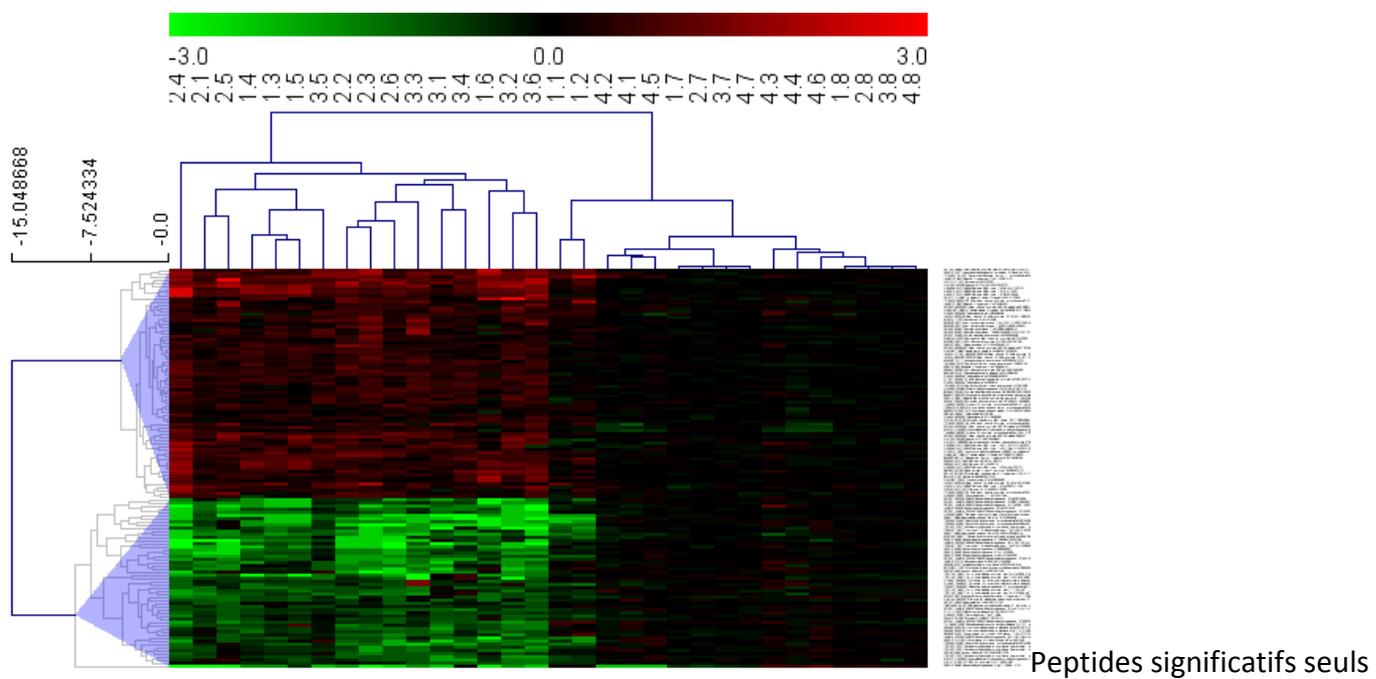
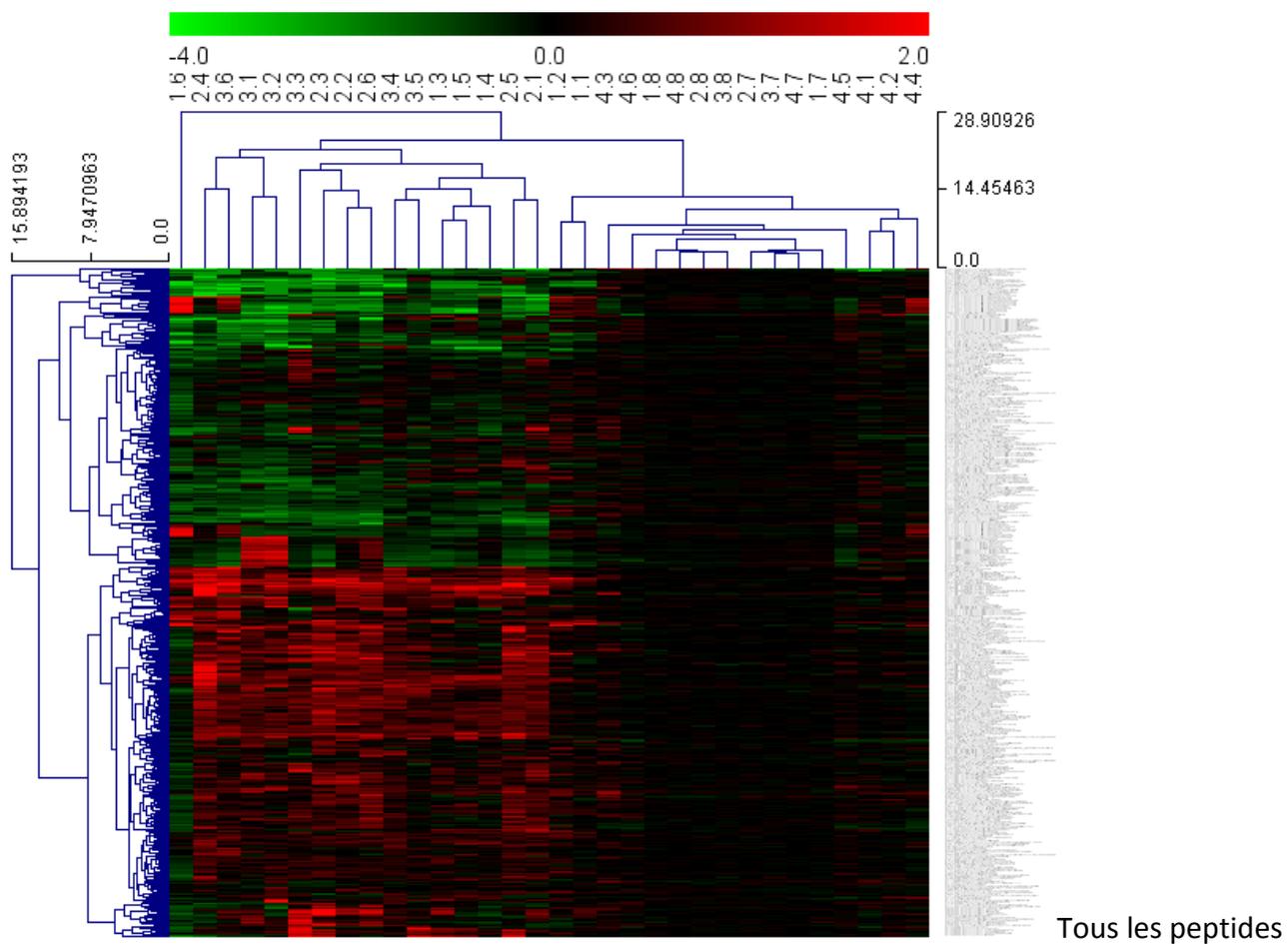
Par contre, l'électrophorèse bidimensionnelle, même si elle présente des limitations, fournit un élément de réponse. En effet, les modifications post-traductionnelles affectent souvent la migration des protéines. Dans le cadre de la première étude, les 2 spots d'HSP90 étaient acétylés mais un seul était à la fois acétylé et phosphorylé lors du traitement par la prodrogue. Ce niveau de connaissance n'est pas accessible avec l'analyse de peptides en mélange mais il serait par contre accessible avec une approche *top-down* utilisant une séparation des protéines et une mesure de masse entière. Cette approche devient possible avec l'exactitude et la résolution de certains instruments (FT-ICR et Orbitrap de dernière génération) et les développements logiciels utilisés pour l'identification à partir des fragmentations de protéines entières (ProsigthPTM, <https://prosigthptm.northwestern.edu/>) (Malinowsky et al., 2010). Cette autre approche de la protéomique est maintenant accessible avec des instruments commerciaux de série. Récemment, une approche *top-down* a identifié 1220 protéines à partir d'une lignée cellulaire H1299 (Catherman et al., 2013). La chromatographie d'affinité TiO₂ de protéines entières pour une approche *top-down* du phosphoprotéome serait un développement méthodologique original.

P21549	Serine-pyruvate aminotransferase (eSPLALAE)GIDNSWR	1.4	1.1	0.9	1.2	0.8	0.7	2.0	1.0	1.1	1.0	1.8	1.4	1.1	0.7	0.8	1.0	1.1	0.9	1.2	0.9	0.9	0.9	0.9	
P21549	Serine-pyruvate aminotransferase (fPVTVA)PAGYDWR	0.7	0.7	1.5	0.8	1.0	0.3	2.0	0.5	0.9	0.6	1.7	0.7	0.8	0.4	0.4	1.1	1.0	1.0	1.5	0.9	1.1	0.8	1.2	0.9
P21549	Serine-pyruvate aminotransferase (QALGLQLFLVK)	0.9	2.1	1.3	1.2	1.3	0.5	2.5	0.6	0.9	1.0	1.9	0.7	1.2	1.1	0.7	0.9	1.1	2.0	1.4	1.0	0.9	0.7	1.1	0.9
F5H0N0	Actin, cytoplasmic 2, N-terminally rAFGAGDDAPR	0.9	0.9	0.9	1.0	0.8	1.9	1.3	1.2	1.2	2.1	1.6	0.9	0.7	1.0	0.7	1.4	1.0	2.9	1.1	1.0	1.0	0.9	1.2	1.0
F5H0N0	Actin, cytoplasmic 2, N-terminally rSYVGDDEAGSK	1.3	1.1	1.2	1.3	1.0	2.7	1.2	1.3	1.7	2.0	1.1	1.2	0.9	0.9	0.9	1.6	1.0	2.8	1.0	0.8	1.0	1.1	1.2	1.1
F5H0N0	Actin, cytoplasmic 2, N-terminally rSYELPDGQVITIGNER	1.0	1.1	1.0	1.2	0.9	1.9	1.4	1.3	1.4	2.4	1.7	1.0	0.9	1.2	0.8	1.7	1.2	2.6	1.0	1.0	1.1	1.0	1.2	1.0
F5H0N0	Actin, cytoplasmic 2, N-terminally rSYELPDGQVITIGNER	0.8	1.1	1.0	1.2	0.9	1.5	1.2	1.2	1.1	2.1	1.3	1.0	1.0	1.2	0.8	1.7	1.1	2.4	1.0	1.0	1.0	1.1	1.1	1.1
F5H0N0	Actin, cytoplasmic 2, N-terminally rWPEEHPVLLTEAPLNPK	0.9	0.9	0.9	0.9	1.0	1.0	1.6	1.1	1.1	2.0	1.6	0.8	0.8	1.0	0.9	1.3	1.2	2.0	1.1	0.9	1.0	0.8	1.2	1.1
P09210	Glutathione S-transferase A2 OS=H aILNYASK	0.7	0.3	0.7	0.7	1.1	0.2	1.3	0.5	0.7	0.3	0.6	0.3	0.2	0.4	0.6	0.5	1.5	0.4	1.3	0.8	1.1	0.9	1.3	1.0
P09210	Glutathione S-transferase A2 OS=H fLQPGSPR	1.0	1.0	1.2	1.0	1.9	0.3	1.4	1.0	1.4	0.7	1.2	0.8	0.4	0.5	0.8	0.7	1.5	0.6	1.0	1.0	0.9	0.8	1.1	0.8
P09210	Glutathione S-transferase A2 OS=H iSNLPTVK	1.1	1.2	1.0	0.9	1.7	0.3	1.3	0.6	0.8	0.2	0.5	0.1	0.5	0.6	0.7	0.7	1.6	0.6	1.1	0.9	0.9	0.8	1.2	0.9
P09210	Glutathione S-transferase A2 OS=H oLNVGLVGNK	0.9	0.9	0.9	1.0	1.0	0.6	1.2	0.7	1.0	0.5	0.7	0.4	0.3	0.4	0.9	0.8	1.4	0.6	1.1	0.8	0.9	1.0	1.1	1.0
P09210	Glutathione S-transferase A2 OS=H wLLAAAGVEFEEK	0.9	1.8	1.2	1.0	1.7	0.6	1.4	1.1	1.3	0.7	0.7	0.5	0.4	0.5	0.6	0.8	0.9	0.8	0.9	0.7	1.4	1.1	1.0	0.8
P09210	Glutathione S-transferase A2 OS=H yPFAFEK	1.0	0.7	1.1	1.0	2.2	0.1	1.4	0.6	0.9	0.2	0.5	0.1	0.2	0.3	0.7	0.7	1.8	0.5	1.3	0.9	1.1	1.1	1.5	1.2
E9PAU8	Liver carboxylesterase 1 OS=Homo aISESGVALTSVLLK	1.3	1.3	1.0	0.6	0.6	0.6	1.5	1.2	2.2	0.6	0.6	1.6	1.8	0.6	3.3	1.3	1.1	1.6	0.8	1.0	0.9	1.0	1.1	1.1
E9PAU8	Liver carboxylesterase 1 OS=Homo eSPLALAE)GIDNSWR	1.5	1.0	1.8	0.9	0.9	0.5	1.8	2.8	4.4	1.4	1.4	2.7	1.3	0.8	2.5	1.5	1.3	1.4	0.9	0.9	1.5	0.9	1.0	1.1
E9PAU8	Liver carboxylesterase 1 OS=Homo eSPLALAE)GIDNSWR	xxx	xxx	1.7	1.1	1.3	xxx	1.0	1.2	2.8	xxx	0.8	1.9	2.5	xxx	3.4	1.4	1.6	1.8	0.7	0.5	0.8	0.6	0.7	0.8
E9PAU8	Liver carboxylesterase 1 OS=Homo eVAFWTLNFAK	1.3	0.9	2.0	0.5	0.7	0.4	2.0	1.9	4.2	0.7	1.2	2.3	1.2	0.5	2.3	0.9	1.1	1.1	0.7	1.2	1.4	1.5	0.9	0.8
E9PAU8	Liver carboxylesterase 1 OS=Homo fTPPQPAEPWFSVK	1.0	0.6	1.1	0.7	0.8	0.5	0.7	1.7	3.2	1.1	1.0	1.9	1.6	1.2	3.1	1.1	1.2	1.2	0.8	1.1	1.2	1.0	1.2	1.2
E9PAU8	Liver carboxylesterase 1 OS=Homo fTPPQPAEPWFSVK	0.9	0.7	1.0	1.1	0.9	0.7	0.6	1.6	3.5	0.9	0.9	2.3	1.4	1.4	1.9	1.7	1.3	1.5	0.9	1.0	1.1	0.8	1.2	1.3
E9PAU8	Liver carboxylesterase 1 OS=Homo gNWGLDQVAALR	1.2	1.0	1.8	1.0	1.3	0.5	1.7	2.2	2.8	1.3	1.5	2.0	1.6	1.1	2.3	1.3	1.1	1.5	0.9	1.2	0.9	0.8	1.0	0.9
E9PAU8	Liver carboxylesterase 1 OS=Homo iPEELQAVK	1.5	1.1	2.5	0.6	0.7	0.4	1.8	2.7	5.7	0.9	1.2	3.5	1.6	0.7	3.4	1.4	1.2	1.9	0.8	1.3	1.0	0.9	1.1	1.1
E9PAU8	Liver carboxylesterase 1 OS=Homo jLGGDDTQK	1.5	1.0	2.3	0.8	0.9	0.7	1.7	2.0	5.2	0.9	0.9	3.3	1.6	0.8	3.4	1.2	1.2	1.7	0.8	1.0	1.0	1.0	1.1	1.1
P00326	Alcohol dehydrogenase 1C OS=HoriNEGFLLR	1.6	1.3	0.6	0.3	0.6	0.2	2.3	0.7	1.5	0.6	0.5	0.8	0.5	0.3	1.0	1.7	0.6	1.3	1.1	1.0	1.2	1.0	0.8	1.0
Q9BQE3	Tubulin alpha-1C chain OS=Homo sAVFDLEPTVIDEVR	1.1	1.1	1.9	1.8	1.8	1.1	1.8	2.8	2.8	3.8	2.2	2.5	1.6	1.8	1.2	1.8	1.9	1.8	1.1	0.9	1.3	0.9	1.1	1.1
Q9BQE3	Tubulin alpha-1C chain OS=Homo sELDLVLR	1.0	0.8	1.6	1.8	1.6	1.1	1.1	1.8	1.6	1.8	1.2	1.3	1.4	1.1	1.1	1.3	1.2	1.1	1.1	0.9	1.3	1.3	1.1	1.1
Q9BQE3	Tubulin alpha-1C chain OS=Homo sFDGALNVLDTEFQTNLPYPR	1.4	1.6	1.1	1.4	1.2	1.1	1.3	1.9	1.9	1.6	1.8	1.6	2.0	2.1	1.1	1.5	1.5	1.5	1.1	1.0	1.5	1.2	1.3	1.2
Q9BQE3	Tubulin alpha-1C chain OS=Homo sILQIVSSITASLR	0.7	1.1	1.0	1.3	1.5	0.8	1.0	2.5	1.5	1.9	1.0	1.3	1.5	1.4	0.7	1.0	1.0	1.3	0.6	0.7	0.7	0.5	0.6	0.5
P69905	Hemoglobin subunit alpha OS=HoriLAVSIVTLTSK	0.8	0.6	0.5	0.9	0.7	0.4	0.6	1.1	0.6	0.6	0.6	1.1	2.2	2.8	0.8	0.4	0.4	0.4	0.7	1.0	0.6	0.6	0.6	0.6
P69905	Hemoglobin subunit alpha OS=HoriMLFSPTTK	0.6	0.4	0.5	0.7	0.5	0.5	0.5	1.1	0.5	0.6	0.5	1.7	2.4	2.9	0.7	0.5	0.5	0.4	1.6	1.3	1.2	1.0	0.8	1.1
P69905	Hemoglobin subunit alpha OS=HoriGAHAGEYGAELER	1.0	0.8	0.8	1.2	0.9	0.9	1.1	1.1	0.9	1.5	1.1	1.3	2.5	2.9	0.8	0.7	0.6	0.7	1.0	1.2	1.0	1.2	0.9	0.9
F8VX15	Aldehyde dehydrogenase, mitochondr eIF9PVMQILK	1.3	0.8	0.9	0.8	1.1	0.5	1.3	1.0	0.9	0.8	1.0	1.2	0.9	0.7	0.8	1.1	1.0	0.9	1.2	0.7	1.1	1.0	0.9	0.9
P00352;F8VX	Retinal dehydrogenase 1 OS=Homo iADLIER	1.0	1.0	0.8	0.4	0.9	0.3	2.1	1.0	1.0	0.3	0.9	0.9	0.8	0.6	2.4	1.1	0.7	0.6	0.9	1.2	1.3	1.0	1.1	1.0
F8VX15	Aldehyde dehydrogenase, mitochondr fVQEDIDYDFEVR	1.2	1.8	0.6	0.6	1.4	0.5	1.5	1.1	2.1	1.6	1.7	3.2	1.2	0.8	1.2	1.5	1.1	1.2	0.9	1.1	1.4	1.2	1.1	0.8
F8VX15	Aldehyde dehydrogenase, mitochondr fVQEDIDYDFEVR	1.1	1.3	0.4	0.6	1.2	0.3	1.0	0.7	0.8	1.2	0.8	1.5	1.2	0.7	0.9	1.4	0.9	0.8	0.9	1.1	1.3	1.2	1.1	0.9
F8VX15	Aldehyde dehydrogenase, mitochondr iVQAAGSSNLK	1.1	1.1	1.6	1.6	1.9	1.4	1.2	0.5	0.6	0.5	0.7	1.0	1.0	0.7	0.9	0.9	0.9	0.6	1.0	1.1	1.0	1.0	1.2	1.0
P00352;P308	Retinal dehydrogenase 1 OS=Homo iVTELEGGK	1.0	1.0	0.9	0.8	1.0	0.7	2.1	0.9	1.0	0.5	1.2	0.9	0.9	0.8	1.9	1.1	1.0	1.0	1.0	0.9	1.2	0.9	1.1	1.0
A8K092	ATP synthase subunit alpha OS=HoriVTFNLAGFEA	1.1	1.2	1.1	1.1	1.2	0.6	1.3	1.1	1.2	1.4	1.5	1.6	1.1	0.9	1.2	1.2	0.9	1.3	1.0	0.9	1.2	0.9	1.0	1.1
A8K092	ATP synthase subunit alpha OS=HoriVTFNLAGFEA	1.3	1.7	1.4	1.2	1.7	1.2	1.6	1.8	1.4	1.8	1.7	1.5	1.2	1.1	1.3	1.3	0.8	1.1	1.4	1.7	2.3	1.5	1.2	1.4
A8K092	ATP synthase subunit alpha OS=HoriVYNAVDAVIAELK	1.2	1.7	1.2	1.4	1.3	1.1	1.5	1.1	1.1	1.2	1.6	1.5	0.9	1.4	0.8	1.1	0.8	0.8	1.0	0.9	1.3	0.9	1.1	1.1
A8K092	ATP synthase subunit alpha OS=HoriVAAFAQFGSLLDAATQQLLSR	0.7	xxx	0.9	1.0	0.8	0.5	1.2	1.2	0.8	1.1	1.1	1.1	1.1	1.4	0.8	0.9	1.0	0.9	1.2	0.9	1.2	0.8	1.3	1.0
A8K092	ATP synthase subunit alpha OS=HoriLIGADTSVLDLEETGR	1.2	1.3	1.1	1.0	1.3	0.6	1.1	1.0	1.0	1.1	1.1	1.2	0.8	0.7	1.0	1.1	0.8	0.9	1.0	1.0	1.1	0.8	1.0	1.0
A8K092	ATP synthase subunit alpha OS=HoriGAIVDVPAGEELLGR	1.1	1.1	1.0	1.0	1.2	0.5	1.3	1.2	1.2	1.5	1.4	1.5	1.2	1.1	1.4	1.2	1.0	1.4	0.9	1.0	0.7	0.8	1.0	1.0
A8K092	ATP synthase subunit alpha OS=HoriWALDGAINDGK	1.0	1.6	0.8	1.0	1.1	0.5	1.9	1.8	1.5	2.0	2.2	1.8	1.1	1.0	1.4	1.2	1.0	1.4	1.0	xxx	0.7	0.6	0.6	0.8
P00352	Retinal dehydrogenase 1 OS=HoriFVEESIDYDFVR	0.9	1.5	1.3	0.5	0.6	0.4	2.6	1.9	1.5	0.3	0.9	0.5	0.6	0.6	3.0	1.1	0.4	0.6	0.7	0.8	2.3	1.2	0.8	1.1
P00352	Retinal dehydrogenase 1 OS=HoriLDLIESGK	1.0	0.8	1.3	0.4	0.7	0.3	3.2	1.6	1.3	0.3	1.4	0.6	0.6	0.7	4.1	1.0	0.6	0.7	1.0	1.0	1.1	0.9	1.0	1.1
P00352	Retinal dehydrogenase 1 OS=HoriYNAVDAVIAELK	1.1	1.5	0.8	1.1	0.9	1.1	1.2	1.6	0.8	1.0	0.9	1.8	4.8	4.1	1.1	0.8	1.0	0.9	1.0	0.9	0.8	0.8	0.9	0.8
P00352	Retinal dehydrogenase 1 OS=HoriWAFGTSTVEGK	1.2	1.1	1.3	1.1	1.1	0.8	2.6	1.5	1.4	0.8	1.4	0.8	0.9	1.0	3.4	1.2	0.9	1.2	0.9	1.0	1.0	0.9	1.1	1.0
E7EUT4	Glyceroldehyde-3-phosphate dehyd gALQNPASTGAAK	1.0	0.9	1.1	1.0	1.2	0.5	1.8	1.6	1.8	2.8	1.7	1.6	1.6	1.1	1.5	1.3	1.0	1.6	0.6	0.7	0.7	0.6	0.7	0.6
E7EUT4	Glyceroldehyde-3-phosphate dehyd gALQNPASTGAAK	1.0	0.9	0.9	1.1	1.1	0.5	1.4	1.2	1.4	1.6	1.3	1.5	1.5	1.1	1.9	1.3	1.1	1.6	1.0	0.9	1.0	0.7	1.1	0.9
E7EUT4	Glyceroldehyde-3-phosphate dehyd iISWYDNEFGYSNR	1.0	0.7	0.7	1.1	0.8	0.5	1.3	1.2	1.5	1.7	1.6	1.3	1.4	0.9	1.5	1.4	1.2	1.4	1.1	1.0	1.5	1.1	1.2	1.1
E7EUT4	Glyceroldehyde-3-phosphate dehyd iVINGNPITFQER	1.2	1.3	0.8	1.0	0.8	1.0	1.7	1.2	1.3	1.9	1.8	1.1	0.9	0.9	1.3	1.6	0.9	0.9	1.1	1.0	1.1	0.9	1.1	1.1
E7EUT4	Glyceroldehyde-3-phosphate dehyd iVINGNPITFQER	0.7	1.3	0.5	1.2	0.7	0.7	2.2	2.0	2.2	2.6	2.0	1.9	1.6	2.2	2.5	1.3	1.6	2.9	1.0	0.9	1.1	0.7	1.2	1.0
E7EUT4	Glyceroldehyde-3-phosphate dehyd vANTLVSVVLDLTCR	0.6	0.7	0.7	0.9	0.5	0.5	1.7	0.8	0.7	0.8	1.8	0.6												

P07900	Heat shock protein HSP 90-alpha CISELLR	1.0	1.0	2.0	1.7	3.2	0.7	1.9	2.7	2.8	2.2	3.6	2.1	2.1	2.3	2.0	1.5	1.9	2.7	1.0	1.1	0.9	0.9	1.2	0.9
B72254	Protein disulfide-isomerase A6 OS= GSTAPVGGGAFFPTIVER	0.9	0.8	1.3	1.2	1.0	0.6	1.5	1.1	1.3	1.7	1.5	1.2	1.0	1.1	1.4	1.3	1.3	1.2	1.0	1.0	1.1	0.8	1.3	1.0
B72254	Protein disulfide-isomerase A6 OS= IAAVDATVnQVLASR	1.0	1.1	1.3	1.4	1.0	0.7	1.6	1.7	1.9	2.4	1.7	2.1	1.6	1.2	1.6	1.7	1.5	2.0	0.9	1.0	1.3	1.1	1.1	0.9
B72254	Protein disulfide-isomerase A6 OS= IAAVDATVnQVLASR	0.9	1.3	1.5	1.2	0.9	0.7	1.0	1.3	1.2	1.6	1.1	1.4	1.4	1.3	1.2	1.2	1.1	1.3	0.9	1.1	1.1	0.9	0.9	1.0
B72254	Protein disulfide-isomerase A6 OS= IGEAIVDAALSALR	1.2	1.0	1.3	1.0	1.3	0.5	1.6	2.0	1.9	2.7	2.0	1.7	1.5	1.6	1.7	1.5	1.7	1.7	1.0	1.3	1.4	1.2	1.3	1.0
P07148	Fatty acid-binding protein, liver OS= aIGLPEELIQ	0.9	0.7	0.3	0.2	0.2	0.2	0.4	0.7	0.4	0.3	0.6	0.3	1.2	0.3	0.2	1.8	0.4	0.5	1.0	1.0	1.0	0.9	1.1	0.9
P07148	Fatty acid-binding protein, liver OS= aIGLPEELIQ	0.7	0.7	0.7	0.5	0.6	0.3	0.4	0.7	0.5	0.4	0.6	0.5	1.3	0.6	0.3	1.6	0.5	0.5	1.1	0.9	0.9	0.6	1.0	1.0
P07148	Fatty acid-binding protein, liver OS= FITTAGSK	0.8	0.8	0.6	0.6	0.4	0.5	0.5	0.7	0.5	0.3	0.8	0.5	1.1	0.5	0.4	1.4	0.6	0.8	1.0	1.6	0.8	1.0	1.0	0.9
P07148	Fatty acid-binding protein, liver OS= sVTELENGDIITMTLGDIVfk	0.7	1.0	0.5	0.5	0.4	xxx	0.5	0.5	0.5	0.3	xxx	0.5	0.1	xxx	0.9	xxx	1.3	xxx	1.0	2.0	0.9	1.5	0.6	0.9
P07148	Fatty acid-binding protein, liver OS= IVVQLEGDNK	1.0	0.9	0.7	0.6	0.5	0.5	0.7	0.9	0.8	0.6	0.8	0.5	1.1	0.2	0.1	1.8	0.4	0.5	1.2	1.0	1.0	0.7	1.1	0.9
P02792	Ferritin light chain OS=Homo sapie IGGPEAGLGEYLFER	4.4	2.2	1.1	0.9	0.7	3.9	1.0	1.6	2.8	1.8	1.1	1.6	1.0	1.4	1.6	1.9	1.1	1.4	0.9	0.6	2.0	1.2	1.0	0.8
P02792	Ferritin light chain OS=Homo sapie INQALLDLHALGFSR	1.8	1.4	1.0	1.2	1.2	1.2	1.1	1.0	2.0	1.3	0.8	1.1	1.4	1.5	1.2	0.9	0.9	0.8	0.7	0.6	1.0	0.7	0.6	0.6
ESRK69	Annexin OS=Homo sapiens GN=AI eLIELATR	1.0	1.4	1.3	0.9	1.1	0.7	1.2	1.1	1.3	0.8	0.9	0.7	1.0	1.4	1.8	1.2	1.0	0.9	1.1	1.1	1.3	1.1	1.0	1.0
ESRK69	Annexin OS=Homo sapiens GN=AI dFAVAIVQSVK	1.0	0.9	1.0	1.0	0.9	0.5	0.6	0.7	0.7	0.5	0.4	0.8	1.2	1.1	1.2	1.0	0.9	0.7	0.9	0.9	1.0	1.0	1.0	1.0
ESRK69	Annexin OS=Homo sapiens GN=AI eALDIITSR	1.1	1.0	1.4	0.9	0.7	0.5	0.9	0.7	1.1	0.5	0.5	0.5	0.7	0.8	1.6	1.1	0.9	0.7	1.0	0.8	1.1	1.0	0.9	0.9
ESRK69	Annexin OS=Homo sapiens GN=AI eEGENLQDAR	1.8	2.3	2.4	1.6	1.3	2.0	1.7	1.5	2.5	1.1	1.1	1.0	1.4	1.3	2.5	1.5	1.4	1.3	1.6	1.3	1.6	1.4	1.6	1.5
ESRK69	Annexin OS=Homo sapiens GN=AI gLGTDEDIITDTHR	1.0	0.9	1.3	0.9	0.9	0.5	1.2	0.8	0.7	0.9	0.7	0.4	0.5	0.7	1.4	1.1	1.0	1.0	0.9	0.7	1.0	1.0	1.0	1.2
P05787	Keratin, type II cytoskeletal 8 OS= sLEAAIAAEQQR	1.0	1.0	1.1	1.3	1.0	0.6	1.5	1.4	2.0	1.1	1.1	2.0	1.1	1.1	0.8	2.0	1.1	1.2	0.9	1.0	1.1	0.8	1.1	0.8
P05787	Keratin, type II cytoskeletal 8 OS= sLEAAIAAEQQR	1.0	1.0	1.1	0.7	0.8	0.9	1.2	1.0	1.5	0.8	0.8	1.5	0.9	0.9	0.7	1.8	1.1	1.2	1.0	1.0	1.1	0.9	1.2	1.0
P05787	Keratin, type II cytoskeletal 8 OS= IEGLTDEINFLR	1.0	1.2	1.2	0.5	0.6	1.3	1.4	1.2	2.2	1.0	1.1	2.6	1.0	1.0	0.2	2.7	1.3	1.3	1.3	1.2	1.1	0.9	1.2	0.9
P05787	Keratin, type II cytoskeletal 8 OS= ISELEAALQR	1.1	1.3	1.3	0.5	0.7	1.1	1.3	1.0	2.2	0.7	0.8	2.3	1.2	1.0	0.3	2.6	1.2	1.2	1.0	1.2	1.1	0.8	1.2	0.9
P05787	Keratin, type II cytoskeletal 8 OS= ISELEAALQR	1.0	1.7	0.8	1.2	1.1	0.9	1.2	1.1	1.9	0.9	0.8	1.8	1.2	1.0	0.5	2.5	1.2	1.4	1.1	1.2	1.3	1.0	1.2	1.1
P35579	Myosin-9 OS=Homo sapiens GN=I MQLDLDLVDLHQR	0.7	xxx	1.1	1.1	1.2	1.4	1.5	2.3	1.5	3.5	1.2	2.1	0.8	1.4	0.8	1.4	0.9	1.6	1.0	1.3	1.0	1.3	1.5	1.1
P35579	Myosin-9 OS=Homo sapiens GN=I VSHLLGINVDFTR	0.8	0.7	1.0	0.6	0.7	xxx	0.8	0.6	0.7	0.7	0.6	0.7	0.6	0.5	0.7	0.6	1.0	0.7	0.8	0.8	1.0	0.9	0.8	0.9
HOYB3	Ribonuclease UK114 (Fragment) O: IEIEAVAIQGPLTASL	1.0	1.5	1.1	1.0	1.0	0.9	0.9	0.8	1.0	0.8	0.7	0.8	0.8	0.8	0.8	1.1	0.8	1.0	xxx	xxx	0.7	0.8	xxx	0.8
HOYB3	Ribonuclease UK114 (Fragment) O: ITVLLADINDFTVNEIYk	1.4	1.3	0.9	0.9	1.2	0.5	1.3	1.2	1.2	1.8	1.4	1.3	1.6	1.1	1.6	1.7	1.1	1.8	0.9	1.0	1.3	0.9	1.1	1.3
Q93088	Betaine-homocysteine S-methyltra GPWPTEAAVHEPEAVR	1.0	1.1	1.1	1.0	1.0	0.9	1.1	1.8	1.2	1.3	1.0	1.1	1.0	1.6	0.9	1.0	1.0	1.2	1.0	0.9	1.1	0.9	1.0	1.3
Q93088	Betaine-homocysteine S-methyltra GQVEVNEAACDIAR	0.3	0.3	0.3	0.2	0.2	0.2	0.7	0.6	0.3	0.2	0.3	0.2	0.1	0.2	0.2	0.2	0.5	0.3	0.9	0.8	0.9	0.9	1.0	1.0
P30086	Phosphatidylethanolamine-binding IYEQLSGK	1.0	1.0	0.7	0.9	0.8	0.6	0.9	0.7	1.1	0.9	0.9	0.6	1.2	1.1	1.6	1.2	1.1	0.9	0.9	0.9	1.1	0.9	1.1	0.7
P30086	Phosphatidylethanolamine-binding IYTLTDPDAPSR	0.9	0.9	0.6	0.5	0.8	0.3	1.0	1.0	1.1	0.9	0.8	0.7	0.9	0.7	1.4	0.8	1.0	0.5	0.7	1.1	0.9	0.8	1.1	0.7
B1ALW1	Thioredoxin OS=Homo sapiens GN: IEATINELV	1.0	1.0	1.4	1.2	1.7	0.6	2.4	2.0	3.0	1.7	2.0	1.3	1.6	1.8	2.4	1.7	1.6	2.3	0.9	0.9	1.8	1.1	1.1	1.3
B1ALW1	Thioredoxin OS=Homo sapiens GN: IAFQEALDAAGDK	0.9	1.0	0.8	0.7	0.9	0.4	1.8	1.8	2.2	1.4	1.6	1.2	1.3	1.4	2.5	1.5	1.8	1.9	0.9	0.9	1.1	1.0	1.0	1.1
B1ALW1	Thioredoxin OS=Homo sapiens GN: IGFVEFSGANK	1.2	1.1	1.6	1.1	1.3	0.6	2.5	2.1	2.7	1.8	2.3	1.5	1.3	1.8	2.9	1.6	2.0	2.2	1.0	1.0	1.2	0.9	1.2	1.0
BOYJC4	Vimentin OS=Homo sapiens GN=V ILLAEQLK	0.9	1.0	0.8	1.1	1.1	1.8	1.0	1.1	1.4	3.5	1.7	1.0	1.1	1.2	0.9	2.8	1.6	3.8	1.5	0.8	1.3	0.9	1.0	1.0
B0Y591	Cytosolic 10-formyltetrahydrofolate aLILFGNDK	1.4	xxx	1.2	1.0	1.3	xxx	1.1	1.1	1.0	1.1	1.0	1.1	1.1	1.1	2.2	1.1	0.8	0.9	1.0	1.1	1.0	1.0	1.0	1.0
O75891	Cytosolic 10-formyltetrahydrofolate dLEAALNEYLK	1.0	1.3	0.7	0.7	0.6	0.3	1.0	0.8	0.8	0.5	0.8	0.8	0.8	0.6	2.8	1.0	0.7	0.7	0.8	1.2	1.3	1.4	0.9	1.0
O75891	Cytosolic 10-formyltetrahydrofolate gAASSVLELTAELVTAEAVR	0.5	0.4	0.6	0.5	0.6	0.4	1.1	1.0	0.7	0.6	0.9	0.7	0.9	1.4	2.9	0.8	0.6	0.5	0.7	0.9	1.2	1.5	0.9	0.8
O75891	Cytosolic 10-formyltetrahydrofolate iNWDQPAEAIHNWIR	1.1	1.3	0.9	0.9	0.9	0.4	1.6	1.5	1.3	1.1	1.4	1.4	1.7	1.8	2.8	0.9	0.5	0.9	1.3	1.1	1.6	1.2	1.0	0.9
O75891	Cytosolic 10-formyltetrahydrofolate iVDESIDHDFVR	xxx	xxx	0.9	xxx	1.0	xxx	1.4	1.4	1.4	0.7	1.0	1.4	1.5	1.3	3.9	1.2	0.9	0.9	0.9	1.1	1.2	1.3	1.1	0.9
P11509	Cytochrome P450 2A6 OS=Homo s: IEQEAAGLIDALR	1.2	2.0	2.9	0.7	1.2	0.4	1.1	0.5	0.5	0.5	0.2	0.9	0.3	0.3	1.5	1.2	1.4	0.4	0.9	1.3	1.5	1.2	0.7	0.8
P09467	Fructose-1,6-bisphosphatase 1 OS: sPVLGSPDDVLEFLK	1.1	0.8	0.5	0.2	0.3	0.2	0.4	0.2	0.2	0.1	0.2	0.1	0.3	0.2	0.3	1.0	0.6	0.3	1.1	0.9	1.4	1.0	1.0	1.3
P09467	Fructose-1,6-bisphosphatase 1 OS: eAVLVDIPIIDHQR	0.9	1.0	0.9	0.9	0.8	0.4	1.2	1.1	1.1	1.3	1.2	1.5	1.1	0.9	1.1	1.1	0.9	1.1	0.9	0.9	1.1	1.2	1.3	1.3
P09467	Fructose-1,6-bisphosphatase 1 OS: gTGLTQLLNSLcTAVK	0.6	0.4	0.4	0.7	0.6	0.5	0.5	0.3	0.2	0.3	0.5	0.4	0.8	0.7	0.3	0.4	0.6	0.2	1.1	0.8	1.0	0.8	1.0	1.2
P00441	Superoxide dismutase [Cu-Zn] OS= gDVAVDIESDVISLGGDHCIIGR	0.8	1.0	0.9	0.9	1.0	0.7	1.6	0.7	0.8	0.7	0.8	0.8	0.9	0.7	2.3	1.2	1.0	0.8	0.8	1.1	1.0	0.9	1.1	0.8
P00441	Superoxide dismutase [Cu-Zn] OS= gDGPVGGINFEQK	1.1	1.1	1.1	0.9	1.1	0.5	1.1	0.9	1.3	1.0	1.3	1.0	0.8	0.8	1.6	1.2	0.9	1.1	0.9	1.0	1.1	0.9	1.1	1.0
P00441	Superoxide dismutase [Cu-Zn] OS= iVGDGLGNVTADK	1.1	1.0	0.9	1.0	1.0	0.6	1.5	1.1	1.1	0.8	0.9	1.1	0.7	0.7	1.4	1.3	1.0	0.9	0.8	1.1	1.0	1.0	1.0	0.9
H7BZJ3	Thioredoxin (Fragment) OS=Homo s: IFEELNADLFER	1.0	1.2	1.6	1.3	1.1	0.7	1.4	1.2	1.6	1.5	1.2	1.1	1.0	1.1	1.6	1.4	1.2	1.4	0.9	1.1	1.0	1.0	1.0	1.0
Q3LXA3	Bifunctional ATP-dependent dihydrdc sPADLQLQLTK	1.1	1.2	1.1	1.1	2.7	0.5	1.7	1.5	1.4	2.4	2.8	2.2	1.6	1.5	1.2	3.1	0.8	0.9	1.4	1.5	1.2	1.6	1.1	1.1
Q3LXA3	Bifunctional ATP-dependent dihydrdc sGTLGLEELHNLDR	0.7	xxx	0.8	xxx	0.8	xxx	xxx	xxx	xxx	xxx	0.9	0.6	0.8	1.1	xxx	xxx	1.3	0.9	1.0	0.8	0.7	0.6	0.7	1.0
P14625	Endoplasmic OS=Homo sapiens G: eLISNASDALDK	1.5	1.4	1.6	0.7	0.9	0.9	1.6	1.7	1.9	1.8	2.3	1.6	1.6	1.4	1.6	1.6	1.7	1.7	0.9	1.0	1.7	1.0	0.6	1.2
P14625	Endoplasmic OS=Homo sapiens G: eVEEEDYK	0.9	1.1	1.7	1.4	1.2	1.1	1.8	1.1	1.8	2.0	2.0	1.4	1.6	1.2	1.7	2.0	1.9	2.3	1.0	1.0	1.0	0.9	1.1	1.0
P14625	Endoplasmic OS=Homo sapiens G: gVDDSDLLPLNVSRR	1.1	1.1	1.7	1.4	1.2	0.9	1.4	1.3	1.5	1.8	1.8	1.4	1.3	1.9	1.5	1.7	1.7	2.0	1.0	1.0	1.0	0.9	1.1	1.0
P14625	Endoplasmic OS=Homo sapiens G: sILFVPTASAPR	1.3	1.2	1.3	0.8	0.9	0.5	1.5	1.3	1.5	1.1	1.6	1.2	1.1	1.1	1.2	1.5	1.3	1.2	1.2	1.2	1.1	0.9	1.4	1.3
P24752	Acetyl-CoA acetyltransferase, mito: IGNEVIPVTVTVK	1.2	0.9	1.1	1.1	1.0	1.0	1.3	0.7	0.7	1.3	1.1	1.1	0.9	0										

P54868	Hydroxymethylglutaryl-CoA synthase IVS2SDLPk	1.2	1.1	1.0	0.7	0.8	0.5	1.5	0.6	1.1	0.7	1.1	0.6	0.2	0.2	0.3	0.8	1.5	0.3	1.2	0.9	0.9	0.8	1.0	1.0
Q96HR9	Receptor expression-enhancing protein LVE1LGALEAk	0.7	1.4	0.9	0.6	1.1	0.5	0.8	0.5	0.8	1.2	0.7	0.3	0.9	0.5	0.4	1.5	1.3	0.8	0.5	0.6	1.4	0.5	0.9	0.8
HOYNC2	Dehydrogenase/reductase (SDR) for diVLVNNAYAGVQTILNTR	0.9	1.5	1.0	1.3	1.2	0.8	1.1	1.8	2.0	1.3	1.6	1.7	1.0	1.1	0.7	0.9	0.7	0.7	0.9	0.8	1.7	1.1	1.0	1.0
H7C131	3-ketoacyl-CoA thiolase, peroxisomal EELGLPILGVLR	1.0	1.0	0.9	0.5	0.9	0.4	0.6	1.4	2.8	0.5	0.9	2.6	1.2	0.4	0.9	1.0	1.1	0.7	1.3	1.1	1.2	1.2	1.0	1.1
H7C342	D-dopachrome decarboxylase (Frag) FPFLSEWQIGk	0.9	0.6	0.6	0.7	0.6	0.3	0.5	0.4	0.9	0.5	0.5	0.6	1.3	0.6	1.1	0.6	0.8	0.7	0.9	0.7	0.9	1.2	1.0	1.5
P05089	Arginase-1 OS=Homo sapiens GN=eGLYTEEYk	1.1	1.1	1.0	0.7	1.1	0.9	0.9	0.9	0.8	0.6	0.9	0.7	0.5	0.6	0.9	0.8	1.4	0.7	1.0	1.2	1.1	1.1	1.0	1.7
P05089	Arginase-1 OS=Homo sapiens GN=eGLVEEPTVLR	1.1	1.2	1.3	1.4	1.7	0.7	1.6	1.5	2.5	1.4	1.9	1.2	0.8	1.1	1.4	1.1	1.5	1.4	1.0	1.0	1.1	0.9	1.1	1.1
P62158	Calmodulin OS=Homo sapiens GN=eAFSLFDk	1.1	1.0	1.4	1.4	1.6	0.9	1.4	1.4	1.2	2.3	1.2	1.7	1.1	1.1	1.3	1.6	1.0	1.5	0.9	0.9	1.0	0.8	1.1	1.1
Q5VU59	Tropomyosin 3 OS=Homo sapiens GN=IQLEEELDR	1.0	1.1	1.1	1.2	1.2	2.3	1.7	1.7	1.4	2.4	1.6	1.2	1.3	1.6	1.0	1.9	1.3	2.4	1.0	1.2	1.1	1.0	1.1	1.1
B7WPK3	Formimidoyltransferase-cyclodeaminase IAEELDPVVPYLYGGAAR	1.2	1.4	0.7	0.8	0.8	xxx	0.4	0.4	0.3	0.7	0.4	0.2	0.2	0.2	0.4	0.7	0.5	0.4	1.1	1.1	1.5	1.3	1.0	1.2
E7ENL2	Sulfurtransferase OS=Homo sapiens aLVSQAQWVAEALR	1.0	1.0	0.9	1.4	1.2	0.6	0.7	1.3	1.0	0.8	0.8	1.3	2.7	3.0	0.9	0.8	0.8	0.6	0.9	1.1	0.9	1.1	0.9	0.9
E7ENL2	Sulfurtransferase OS=Homo sapiens aRPEDVISEGR	1.0	1.4	1.2	1.2	1.3	0.9	1.1	1.6	1.2	1.6	1.5	1.2	1.0	1.3	1.0	1.1	1.0	1.1	0.9	0.9	1.1	0.8	1.0	0.8
P02671	Fibrinogen alpha chain OS=Homo sapiens sGLIDEVNDQDFNTR	1.0	1.2	1.4	1.5	1.2	0.9	1.3	1.7	1.3	2.3	2.0	2.3	1.9	2.4	0.9	2.7	1.4	1.3	0.7	1.2	1.4	1.5	0.9	0.8
Q06830	Peroxiredoxin-1 OS=Homo sapiens gLFIDDK	0.7	0.5	0.9	0.7	0.9	0.6	0.9	1.0	0.8	0.4	0.7	0.8	0.6	0.5	0.8	0.6	1.0	0.6	0.9	0.9	0.9	1.0	1.0	1.0
Q06830	Peroxiredoxin-1 OS=Homo sapiens lVQAQFQFTDk	xxx	xxx	0.8	0.5	0.8	xxx	1.3	1.3	1.3	0.8	1.1	0.9	1.0	0.5	0.6	1.0	1.5	1.3	1.1	1.0	1.1	0.9	1.0	1.1
Q06830	Peroxiredoxin-1 OS=Homo sapiens lQADYGVk	0.8	1.4	1.0	0.7	1.1	0.8	2.0	2.1	2.1	2.0	1.8	1.9	1.5	1.1	1.4	1.6	1.8	1.8	0.9	1.0	1.2	1.0	1.1	1.1
P55072	Transitional endoplasmic reticulum gGNIGDGGGAADR	1.7	2.2	1.3	1.5	1.5	2.9	2.3	3.2	4.4	3.9	2.8	2.0	2.6	2.3	2.6	2.1	2.9	3.2	1.4	1.4	1.1	1.0	1.1	1.0
P55072	Transitional endoplasmic reticulum nPAIFIDELDAIAPk	0.7	xxx	0.8	1.1	1.0	0.6	1.1	1.5	1.9	1.8	1.3	1.2	1.9	1.4	1.2	1.0	1.1	1.1	1.0	1.1	1.4	1.0	0.7	1.1
G3XAL0	Malate dehydrogenase OS=Homo sapiens iFQVTTLDIVR	1.0	0.9	1.2	0.8	0.9	0.5	1.5	1.5	1.2	1.3	2.2	1.9	1.0	1.2	1.3	1.3	1.1	1.6	1.2	1.0	1.0	0.7	1.0	0.9
G3XAL0	Malate dehydrogenase OS=Homo sapiens vDFPDQALTAALTR	1.1	1.1	1.2	1.2	1.0	0.8	1.4	1.9	1.8	1.4	1.9	2.3	1.5	1.5	1.8	1.5	1.3	1.9	1.0	1.0	1.0	0.9	1.1	0.9
Q5JP53	Tubulin beta chain OS=Homo sapiens gHYTEGAELVDSVLDVVR	xxx	xxx	xxx	1.0	0.6	xxx	0.8	1.7	1.4	1.7	1.2	1.1	2.1	1.6	0.9	1.5	1.4	2.1	1.1	1.1	1.1	1.1	0.9	0.7
Q5JP53	Tubulin beta chain OS=Homo sapiens lAVNMMVFPFR	0.9	1.2	1.2	1.2	1.1	0.6	0.9	1.2	1.2	1.2	1.0	1.2	1.3	1.6	1.1	1.3	1.1	1.0	1.1	0.9	1.1	0.8	1.2	0.9
B4DFK7	Carbonyl reductase 1, isoform CRA fHQLDIDDLQSIIR	1.1	0.8	1.6	1.2	1.4	0.4	1.7	1.7	1.7	1.4	1.2	1.3	1.2	1.2	1.1	1.2	0.9	1.0	1.0	1.1	1.0	0.7	1.2	0.9
B4DFK7	Carbonyl reductase 1, isoform CRA hQLDIDDLQSIIR	1.3	1.6	0.9	0.8	1.0	0.6	1.3	0.8	0.8	1.8	1.1	0.8	0.7	0.6	1.3	0.6	0.4	0.6	1.0	1.1	0.9	0.8	xxx	0.9
E7ERW2	Aspartate aminotransferase OS=Homo sapiens aIAAILNTPDLR	0.9	0.9	1.0	0.9	1.0	0.5	0.8	1.2	1.0	1.1	1.0	1.2	0.9	0.9	1.0	1.1	1.1	0.8	0.7	1.1	0.9	0.8	0.9	1.1
F5H3L8	Stress-70 protein, mitochondrial OS=QVPEGIVTDLIR	1.1	1.3	1.6	2.3	1.6	0.8	1.7	2.1	1.8	2.4	3.2	3.1	1.8	1.7	2.1	1.6	1.6	2.2	1.0	1.0	1.1	0.9	1.0	1.1
F5H3L8	Stress-70 protein, mitochondrial OS=QVQYDQLFQIR	1.0	1.0	0.9	1.2	1.2	0.4	0.4	1.5	1.4	1.4	1.9	2.0	1.4	1.8	1.3	1.3	1.3	1.5	1.1	1.1	1.1	0.8	1.1	0.9
F5H816	Phosphorylase OS=Homo sapiens dQNVNDYIQAVLDR	1.0	1.3	0.7	0.3	1.0	0.3	0.5	0.3	0.4	0.6	0.6	0.3	0.2	0.3	0.4	0.7	0.5	0.5	0.9	0.9	1.6	1.3	0.9	0.9
F5H816	Phosphorylase OS=Homo sapiens lHSLFGDDVFLR	1.1	1.8	0.9	0.8	1.2	0.7	0.7	0.9	0.7	0.8	0.7	0.9	0.9	1.0	0.8	1.0	0.6	0.7	1.3	1.2	1.2	1.1	0.9	1.1
P49411	Elongation factor Tu, mitochondrial lDAVDYTIYVPAPR	1.1	1.2	1.6	1.1	1.6	0.5	1.6	1.3	1.5	1.6	1.7	2.4	0.9	0.9	1.4	1.4	0.9	1.2	0.9	1.1	1.2	1.0	1.2	1.1
P00338	L-lactate dehydrogenase A chain O dLADELALVDVIEKd	0.8	1.9	1.0	1.2	1.6	0.7	1.6	1.4	0.5	1.1	0.9	0.8	1.3	0.6	0.4	1.4	1.0	1.1	0.6	0.6	1.6	1.1	1.0	1.1
Q02252	Methylmalonate-semialdehyde dehydrogenase PPAWADTSVLSR	1.3	1.2	1.0	0.8	0.8	0.5	1.3	0.9	0.9	0.8	0.8	1.4	1.3	1.3	1.1	1.2	1.2	0.7	0.9	1.1	1.1	0.9	1.2	0.9
Q13228	Selenium-binding protein 1 OS=Homo sapiens ePLGPAALHELIR	0.9	1.2	0.8	0.8	0.9	0.6	1.9	1.7	1.5	1.2	1.4	1.8	1.6	1.9	1.6	1.3	0.9	1.1	1.0	0.8	0.9	0.8	1.2	0.8
Q13228	Selenium-binding protein 1 OS=Homo sapiens iYVVVDGSEPR	1.1	1.3	1.6	1.0	1.0	0.7	1.7	1.4	1.4	1.3	1.2	1.5	0.9	1.2	2.0	1.3	1.0	1.2	1.1	1.1	1.1	1.0	1.1	1.3
F5H7H0	Glutathione S-transferase omega-1 gSAPPQVPEGSIR	1.0	1.1	1.4	1.0	1.1	0.8	0.9	1.1	1.0	0.7	0.8	0.8	0.9	0.9	1.4	1.0	1.0	0.9	0.9	1.1	1.0	1.0	1.2	1.1
F5GYQ4	Glutamate dehydrogenase OS=Homo sapiens dDGSWEVIEGYR	1.2	1.0	1.7	0.6	0.7	0.5	0.7	1.1	0.8	1.3	1.4	0.9	0.8	0.5	1.4	0.9	1.0	0.6	0.7	1.0	1.2	1.1	1.0	0.9
F5GYQ4	Glutamate dehydrogenase OS=Homo sapiens hGGTIPIVPTAEFQDR	1.1	1.4	0.9	0.9	1.0	0.6	1.2	1.3	1.3	1.5	1.4	1.4	1.1	1.5	0.9	1.2	1.0	1.1	1.0	1.0	1.0	0.8	1.0	1.1
HOYK49	Electron transfer flavoprotein subunit H1 VAPISDIIAlk	1.1	0.6	0.7	0.7	0.8	0.4	1.6	1.3	1.2	1.0	1.3	1.9	1.1	1.1	1.2	1.4	1.0	1.1	0.8	0.8	0.8	0.7	0.8	0.7
D6PKX4	Alpha actinin 4 short isoform OS=H1GYEELWNEIR	0.9	1.3	1.2	1.0	1.0	1.3	1.6	1.8	1.7	2.3	1.6	1.3	2.6	1.1	1.5	0.7	1.8	0.9	0.7	1.4	1.1	0.9	1.1	1.1
D6PKX4	Alpha actinin 4 short isoform OS=H1ASDLEEWIR	1.0	1.5	1.6	1.4	1.1	2.0	2.0	2.4	1.7	4.1	1.6	2.1	1.4	2.5	1.0	2.6	1.0	2.8	1.0	1.0	1.4	1.1	1.1	1.1
P51570	Galactokinase OS=Homo sapiens eVQLEELEAAR	1.0	1.3	1.5	1.1	1.2	0.8	1.2	1.3	1.5	1.5	1.2	1.0	1.1	1.0	1.0	1.1	1.0	1.4	1.0	0.9	1.1	1.0	1.1	1.1
P68431	Histone H3.1 OS=Homo sapiens GlEaQDFk	1.0	1.0	0.7	0.8	0.7	0.5	0.9	1.2	1.0	1.0	1.1	1.1	1.0	1.2	1.4	1.1	1.0	1.4	0.9	1.4	0.9	0.8	1.2	1.1
P01877	Ig alpha-2 chain C region OS=Homo sapiens sAVQGGPPER	1.4	1.8	2.1	1.9	1.0	3.1	1.2	2.5	1.8	1.8	1.2	1.9	1.2	9.8	1.5	1.1	1.4	3.0	0.9	1.0	1.2	1.4	1.1	0.9
D2DA43	Delta-1-pyrroline-5-carboxylate dehydrogenase lVNEPVLAFQTGSPER	1.0	0.9	1.0	1.3	0.9	0.9	1.0	0.9	1.0	1.5	1.4	2.5	1.4	0.6	0.8	1.3	0.9	0.8	0.9	0.8	1.0	0.8	1.2	1.2
B4DYF2	Trifunctional enzyme subunit alpha, lVLTGPEVLLGALPGAGGTQR	0.7	0.8	0.8	0.8	0.8	0.6	0.8	1.1	0.9	1.1	1.1	1.7	1.7	1.6	1.0	1.2	0.9	0.7	0.9	1.1	1.1	0.8	1.0	0.9
Q61Q15	EEF1A1 protein OS=Homo sapiens GIGGIVPVYGR	1.1	0.9	1.4	1.1	1.2	0.7	1.1	0.9	0.8	1.6	1.1	0.7	1.2	1.3	1.1	1.4	1.4	2.3	1.0	0.9	1.1	1.0	1.2	1.4
Q61Q15	EEF1A1 protein OS=Homo sapiens YYVTIDAPGHR	0.7	0.7	0.8	0.7	0.6	0.4	1.4	1.2	1.0	1.3	1.3	1.1	0.5	0.4	0.5	0.5	0.8	0.4	1.0	1.1	1.1	0.9	1.0	0.9
F2Z3K5	Dolichyl-diphosphooligosaccharide-4-epimerase lVEEIEDLVAR	1.1	1.3	1.3	1.6	1.2	0.9	0.9	1.2	1.0	1.7	1.0	0.9	1.2	0.7	0.9	1.2	1.2	1.1	0.8	0.7	1.7	1.2	0.9	1.4
P16403	Histone H1.2 OS=Homo sapiens GlAaAAGYDVEK	1.0	1.7	1.3	1.5	1.4	0.6	1.9	1.4	1.5	3.9	1.6	1.7	0.9	1.2	1.2	1.3	1.2	2.2	1.1	1.1	1.0	0.9	1.1	1.0
P16403	Histone H1.2 OS=Homo sapiens GlASGPPVLSLIR	0.9	1.1	1.0	1.1	1.1	0.7	1.9	1.4	1.4	4.0	1.7	1.8	0.9	1.1	2.2	1.2	1.0	2.2	0.9	1.0	0.9	1.0	0.9	0.8
P30041	Peroxiredoxin-6 OS=Homo sapiens lSILYPATTGR	1.0	1.2	1.2	1.2	0.9	0.6	1.7	1.2	1.3	1.8	1.4	1.1	0.9	1.4	2.2	1.2	1.0	1.1	1.0	1.0	1.0	0.9	1.1	1.1
P36542	ATP synthase subunit gamma, mitochondrial lEISGAALD	0.9	1.1	1.1	1.4	1.1	1.3	1.0	1.2	1.3	1.8	1.4	1.2	0.8	0.9	1.1	1.7	0.9	1.0	0.9	1.1	1.4	1.6	1.0	1.1
P22626	Heterogeneous nuclear ribonucleoprotein lFIFGGLSFETTESLIR	1.0	1.3	1.3	1.9	2.2	0.8	1.6	1.4	1.4	2.4	1.4	1.3	1.5	1.4	1.1	1.3	1.1	1.5	1.0	1.1	1.3	1.2	1.1	1.0
B8ZL8	10 kDa heat shock protein, mitochondrial lLVEYGGTK	1.0	1.0	0.9	0.7	0.6	0.8	1.6	2.6	1.8	1.8	3.7	3.6	2											

Classification hiérarchique par biopsie et par peptide.



Test non paramétrique *Rank Product*

Protéines surexprimées dans les tumeurs (FRD<0.005)

Swiss Prot	Gene	Name	Peptides	Function
A8K092	ATP5A1	ATP synthase subunit alpha	1	energy
P06576	ATP5B	ATP synthase subunit beta, mitochondrial	1	energy
F8W7G7	FN1	Fibronectin	1	ECM
P16403	HIST1H1C	Histone H1.2	1	DNA binding
Q96KK5	HIST1H2AH	Histone H2A type 1-H	4	DNA binding
O60814	HIST1H2BK	Histone H2B type 1-K	3	DNA binding
P62805	HIST1H4A	Histone H4	2	DNA binding
H7BY16	NCL	Nucleolin (Fragment)	1	DNA binding
Q5TCJ4	LMNA	Lamin A/C	2	DNA/nucleus stability
E7EPB3	RPL14	60S ribosomal protein L14	1	ribosome
P05387	RPLP2	60S acidic ribosomal protein P2	1	ribosome
P62857	RPS28	40S ribosomal protein S28	1	ribosome
B4DUP0	EEF1G	Elongation factor 1-gamma	1	ribosome binding
Q5T6W1	HNRNPK	Heterogeneous nuclear ribonucleoprotein K	1	Pre-mRNA preprocessing
P07900	HSP90AA1	Heat shock protein HSP 90-alpha	4	protein folding
P14625	HSP90B1	Endoplasmic	4	protein folding
F8VZJ4	HSPA1B	Heat shock 70 kDa protein 1A/1B	4	protein folding
P11021	HSPA5	78 kDa glucose-regulated protein	1	protein folding
F5H3L8	HSPA9	Stress-70 protein, mitochondrial	2	protein folding
P10809	HSPD1	60 kDa heat shock protein, mitochondrial	3	protein folding
F8WE65	PPIA	Peptidyl-prolyl cis-trans isomerase	2	protein folding
B7Z254	PDIA6	Protein disulfide-isomerase A6	1	protein folding & transport
P51148	RAB5C	Ras-related protein Rab-5C	1	protein folding & transport
P55072	VCP	Transitional endoplasmic reticulum ATPase	1	protein folding
P22314	UBA1	Ubiquitin-like modifier-activating enzyme 1	1	proteasome
F5H6Q2	UBC	Ubiquitin (Fragment)	2	proteasome
E9PAU8	CES1	Liver carboxylesterase 1	2	detoxification
P07099	EPHX1	Epoxide hydrolase 1	2	detoxification
P06733	ENO1	Alpha-enolase	1	glycolysis
E7EUT4	GAPDH	Glyceraldehyde-3-phosphate dehydrogenase	1	glycolysis
B4E1H9	PGK1	Phosphoglycerate kinase	1	glycolysis
B7Z5D8	TPI1	Triosephosphate isomerase	1	Glycolysis and gluconeogenesis
G3XAL0	MDH2	Malate dehydrogenase	1	Krebs cycle
O75874	IDH1	Isocitrate dehydrogenase [NADP] cytoplasmic	1	Krebs cycle, lipid synthesis
E1B6W5	SCP2	Non-specific lipid-transfer protein	1	lipid transport, steroid metabolism
B4E2S7	LAMP2	Lysosome-associated membrane glycoprotein 2	1	Tumor induced
D6PXK4	ACTN4	Alpha actinin 4 short isoform	1	cytoskeleton
P62158	CALM1	Calmodulin	1	cytoskeleton
P19105	MYL12A	Myosin regulatory light chain 12A	1	cytoskeleton
H0YI43	MYL6	Myosin light polypeptide 6 (Fragment)	1	cytoskeleton
Q5VU59	TPM3	Tropomyosin 3	1	cytoskeleton
Q9BQE3	TUBA1C	Tubulin alpha-1C chain	2	cytoskeleton
Q5JP53	TUBB	Tubulin beta chain	1	cytoskeleton
B0YJC4	VIM	Vimentin	1	cytoskeleton
P07237	P4HB	Protein disulfide-isomerase	1	protein folding
Q06830	PRDX1	Peroxiredoxin-1	1	anti-oxidant
B1ALW1	TXN	Thioredoxin	1	anti-oxidant
P01877	IGHA2	Ig alpha-2 chain C region	1	Antibody
O95831	AIFM1	Apoptosis-inducing factor 1, mitochondrial	1	apoptosis factor

Protéines sous exprimées dans les tumeurs (FRD<0.005)

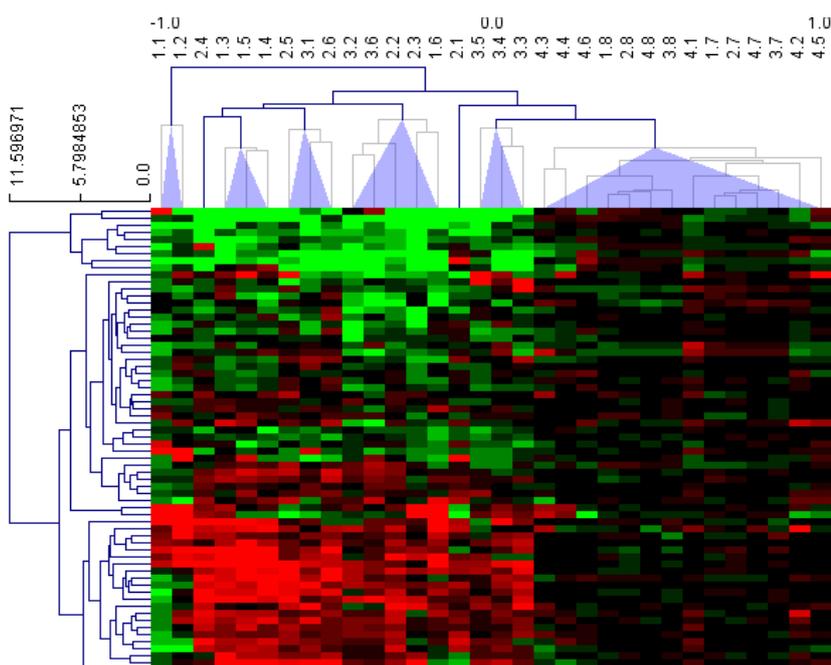
Swiss Prot	Gene	Name	Peptides	Function
P07327;P00326	ADH1A	Alcohol dehydrogenase 1A	7	Détoxification
F5HB16;P00326	ADH1B	Alcohol dehydrogenase 1B	3	Détoxification
P08319	ADH4	Alcohol dehydrogenase 4	5	Détoxification
O95154	AKR7A3	Aflatoxin B1 aldehyde reductase member 3	1	Détoxification
P30837	ALDH1B1	Aldehyde dehydrogenase X, mitochondrial	1	Détoxification
P09210	GSTA2	Glutathione S-transferase A2	1	detoxification, oxidative stress
H7C342	DDT	D-dopachrome decarboxylase (Fragment)	1	Redox regulation
Q06830	PRDX1	Peroxiredoxin-1	1	Redox regulation
P02768	ALB	Serum albumin	2	Blood component
P68871	HBB	Hemoglobin subunit beta	2	Blood component
O75891	ALDH1L1	Cytosolic 10-formyltetrahydrofolate dehydrogenase	2	Metabolism
B7WPK3	FTCD	Formimidoyltransferase-cyclodeaminase	1	Metabolism
P07148	FABP1	Fatty acid-binding protein, liver	4	Lipid transport
D6RER0	ACSL1	Long-chain-fatty-acid--CoA ligase 1	1	Lipid metabolism
P30084	ECHS1	Enoyl-CoA hydratase, mitochondrial	3	lipids catabolism
Q14117	DPYS	Dihydropyrimidinase	1	Metabolism
P05062	ALDOB	Fructose-bisphosphate aldolase B	2	Glycolysis
P09467	FBP1	Fructose-1,6-bisphosphatase 1	2	Glycolysis
E7EUT4	GAPDH	Glyceraldehyde-3-phosphate dehydrogenase	1	Glycolysis
F5H816	PYGL	Phosphorylase	1	Glycogenolysis
D6W5E6	UGP2	UDP-glucose pyrophosphorylase 2, isoform CRA_b	1	Gluconeogenesis
Q16822	PCK2	Phosphoenolpyruvate carboxykinase [GTP], mitochondrial	1	Gluconeogenesis
Q93088	BHMT	Betaine--homocysteine S-methyltransferase 1	1	AA metabolism
E7EW08	SHMT1	Serine hydroxymethyltransferase	1	AA metabolism
P27797	CALR	Calreticulin	1	Protein folding
P00966	ASS1	Argininosuccinate synthase	1	Urea cycle
P31327	CPS1	Carbamoyl-phosphate synthase [ammonia], mitochondrial	4	Urea cycle
Q6IQ15	EEF1A1	EEF1A1 protein	1	Translation factor
P35579	MYH9	Myosin-9	1	Cytoskeleton

2. Fraction purifiée par TiO2.

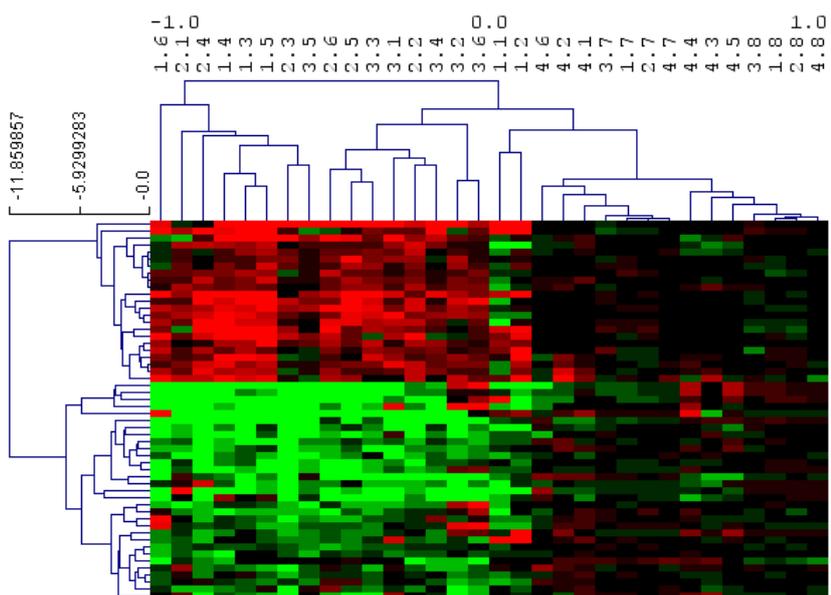
Ratios pour les tumeurs (série 1 à 3) et les non-tumeurs (série 4). (rouge : ratio >2, vert : ratio <0.5)

Acc.	Name	Peptide	1.1	1.2	1.3	1.4	1.5	1.6	2.1	2.2	2.3	2.4	2.5	2.6	3.1	3.2	3.3	3.4	3.5	3.6	4.1	4.2	4.3	4.4	4.5	4.6
B7Z1L3	Membrane-associated	gDQPAAsGDSDDEPPPLPR_1P	0.8	1.0	1.8	1.3	1.8	1.0	0.9	1.1	0.9	1.1	1.0	1.4	0.8	1.0	1.9	2.0	1.3	1.2	1.0	1.1	1.0	1.2	0.9	1.1
Q9H1E3	Nuclear ubiquitinous c	eEDEEPESpPEK_1P_OautresModif	0.7	1.1	1.7	2.1	1.8	1.1	1.1	1.3	1.1	2.2	2.3	2.1	1.6	1.7	1.5	1.6	1.2	1.2	1.0	1.1	0.8	1.1	1.1	1.2
Q9H1E3	Nuclear ubiquitinous c	WDSYQFQESDDAEDYGR_1P_0a	0.7	0.8	1.7	1.8	1.3	1.2	0.8	0.8	0.7	2.5	1.6	1.6	1.2	1.3	1.6	1.3	1.1	0.8	1.2	1.3	0.8	1.2	0.8	1.2
P35579	Myosin-9 OS=Homo	gAGDGsDEEVGdk_1P_OautresModif	0.8	0.5	1.0	1.3	1.2	2.2	1.6	1.0	0.9	1.6	1.8	1.3	0.7	1.2	0.9	1.1	0.9	1.8	1.0	1.3	0.6	1.0	1.3	1.0
F5H4D6	Ras GTPase-activati	sSsPAPADIAQTQVEDLR_1P_0autr	0.8	1.0	1.6	1.9	1.4	1.6	1.6	1.4	1.3	1.9	2.0	1.5	2.1	1.7	1.8	1.5	1.1	1.6	1.1	1.1	1.1	1.0	1.0	0.9
F5H4D6	Ras GTPase-activati	sSsPAPADIAQTQVEDLR_1P_1autr	0.6	1.0	1.1	1.8	1.2	1.7	1.6	1.3	1.4	1.9	2.4	1.7	1.5	1.2	1.3	1.2	0.9	1.2	0.8	0.9	0.9	0.8	1.0	0.8
B4E2T8	Calnexin OS=Homo	aEEDeILNrsPR_1P_OautresModif	0.9	1.0	2.2	2.0	1.8	0.8	0.6	1.6	2.3	1.2	2.1	2.5	2.2	1.3	1.3	1.7	1.5	1.2	1.1	1.1	0.9	0.7	1.2	0.8
B4E2T8	Calnexin OS=Homo	aEEDeILNrsPR_1P_1autresModif	0.8	1.1	1.4	1.3	1.1	0.7	0.7	1.3	1.6	1.0	1.5	1.7	1.4	1.1	1.2	1.2	1.2	1.1	0.9	1.2	0.8	1.0	1.0	1.0
C9J4K3	Fructose-bisphosph	gLAADeSvGTmGNR_1P_1autresM	0.8	0.7	1.0	0.7	0.8	0.8	0.9	0.9	0.7	0.8	0.9	1.0	1.0	0.6	0.9	1.0	1.0	0.7	1.2	1.0	1.1	1.0	0.9	1.1
P62258	14-3-3 protein epsilo	aFDDAIAELDLSEESYK_1P_0autr	1.0	1.3	0.9	0.7	0.8	0.9	1.1	0.7	0.9	1.0	1.5	0.8	0.7	0.6	1.3	0.6	1.1	0.9	1.3	1.3	1.2	1.5	1.1	1.1
Q9UQ35	Serine/arginine repe	rPSPQSPR_1P_OautresModif	0.9	1.5	1.7	1.8	1.7	1.3	1.5	1.8	1.2	1.9	1.5	1.4	1.9	1.5	1.5	1.3	1.2	1.4	1.1	1.3	1.0	1.2	1.0	0.9
Q9UQ35	Serine/arginine repe	rPSPQSPR_2P_0autresModif	1.4	2.1	1.8	2.5	1.9	2.4	1.7	1.7	1.1	1.8	1.6	1.7	1.0	1.7	1.1	1.0	0.8	1.3	1.3	2.6	1.6	0.8	0.8	0.9
Q9UQ35	Serine/arginine repe	sSsPVTEIAsR_1P_OautresModif	0.7	0.9	1.4	1.3	1.2	1.3	1.6	1.6	1.0	1.5	1.6	1.3	1.2	1.5	1.6	1.2	1.0	1.4	0.9	1.2	1.3	1.1	1.1	1.1
O14745	Na(+)/H(+) exchange	aSSTeELNSQDSPPk_1P_0autr	0.4	0.4	0.8	0.6	0.8	0.7	1.2	1.0	1.0	0.9	1.3	1.4	1.4	0.4	0.8	0.9	1.7	0.6	1.9	0.7	0.5	0.8	1.1	1.1
P08238	Heat shock protein H	esDDEAEEEK_1P_OautresModif	0.8	1.0	1.2	1.1	1.3	0.8	1.1	1.4	1.1	0.9	1.7	1.2	1.0	1.3	1.1	0.8	0.7	1.7	0.8	0.9	0.8	1.1	1.1	0.8
B4DU91	Epsin-1 OS=Homo s	tALPISGSSAGELLEAGEVPR_1f	1.0	1.1	1.9	0.9	0.9	0.7	1.0	0.8	0.9	1.3	1.1	1.5	1.7	0.8	1.6	1.7	0.9	0.8	1.4	1.0	1.0	1.2	0.9	0.8
G3V1D1	Ferritin OS=Homo s	tLGDsDNES_1P_OautresModif	3.6	2.4	1.4	0.8	0.8	3.5	0.4	0.9	2.7	1.6	0.5	0.9	0.6	1.0	1.4	1.5	1.4	0.8	0.8	0.6	1.7	1.4	0.7	0.5
G3V1D1	Ferritin OS=Homo s	tLGDsDNES_1P_1autresModif	2.2	2.6	1.6	1.2	1.1	2.5	0.6	1.4	2.4	1.5	0.7	1.2	0.9	1.4	1.3	1.4	1.1	1.0	0.9	1.2	1.5	1.5	0.9	0.8
F8W1K8	60S acidic ribosoma	eSEeSDEDmGFLFD_2P_1autres	0.7	xxx	0.9	1.1	0.8	1.1	1.1	0.9	0.7	1.0	1.1	1.3	1.2	0.6	0.6	1.4	0.9	0.4	1.4	0.7	1.7	1.1	0.8	0.8
P16455	Methylated-DNA-prc	gAGATSGsPPAGR_1P_OautresMoc	0.9	1.5	1.5	1.2	3.0	1.0	0.4	0.6	0.6	1.1	3.8	0.8	0.6	0.7	0.7	1.5	2.4	0.5	1.5	1.3	0.8	1.7	4.2	1.0
B4DN25	UDP-glucose 6-dehy	iPYAPsGEIPk_1P_OautresModif	0.4	0.4	0.4	0.3	0.6	0.3	2.9	0.6	0.3	0.4	0.6	0.3	0.5	0.3	0.5	0.4	0.7	0.5	0.7	0.8	1.0	0.9	0.8	1.5
B4DN25	UDP-glucose 6-dehy	iPYAPsGEIPk_1P_0autresModif	0.8	2.0	1.1	0.7	1.4	0.8	2.7	1.1	0.7	1.0	0.7	0.6	0.4	0.9	0.9	0.7	0.7	1.4	0.7	0.9	1.1	0.9	1.2	0.9
Q8N1C0	CTNNA1 protein OS	iFEYLDSDFEDEFDVR_1P_0autr	0.9	1.0	0.8	0.8	0.8	1.1	1.2	0.8	0.6	1.2	1.1	0.9	0.6	0.5	0.9	1.3	0.7	0.7	1.0	0.8	1.0	0.9	1.0	1.0
E9PK09	Bcl-2-associated tra	dLFDYsPPLHk_1P_OautresModif	0.7	0.7	1.5	1.6	1.5	1.3	1.8	1.6	0.9	3.0	1.6	1.4	1.4	1.2	0.8	1.3	1.3	1.3	1.2	0.9	1.0	1.3	1.0	1.1
P10644	cAMP-dependent pro	eSsPPNPVvK_1P_OautresMoc	1.5	0.9	1.0	0.8	0.7	0.9	1.3	1.1	0.8	1.1	1.1	1.0	0.8	0.9	0.6	0.8	0.9	1.1	0.8	0.9	0.8	1.0	1.2	0.9
Q9Y2W1	Thyroid hormone rec	aSsVSELSPR_1P_OautresModif	0.8	0.9	1.1	1.0	1.2	1.3	1.2	1.4	0.9	1.4	1.0	1.1	1.3	1.0	1.0	1.0	1.1	1.3	1.0	0.9	0.9	1.0	1.0	1.0
Q9Y2W1	Thyroid hormone rec	iDIsPSTFR_1P_OautresModif	0.6	0.6	1.5	1.6	1.1	1.1	1.6	1.8	1.0	1.8	1.2	1.3	1.5	1.1	1.2	1.4	1.2	1.1	1.1	0.9	0.8	1.2	0.8	1.0
B1APG2	Protein kinase, cAM	iWLcGTPeYLAPEILLK_1P_1autres	0.4	xxx	xxx	0.5	xxx	0.7	0.9	0.5	xxx	0.8	1.2	xxx	xxx	0.6	xxx	0.6	1.0	0.5	1.1	1.2	1.4	0.9	xxx	0.9
H3BTK3	Calcium-regulated h	gNVVPSPLTR_1P_OautresModif	0.8	0.8	0.4	0.5	0.7	0.4	1.2	0.5	0.5	0.8	0.5	0.5	0.4	0.4	0.4	0.5	1.1	0.5	1.2	1.3	0.6	0.8	0.6	1.4
H0YEV2	Src substrate contact	iPSPVVEDAASFk_1P_OautresMoc	0.9	0.8	0.4	0.5	0.5	0.8	0.8	0.6	0.5	0.6	0.9	0.7	0.8	0.7	0.7	0.9	0.7	0.6	0.8	1.3	1.0	0.9	1.1	1.2
B8WCJ4	Acetyl-coenzyme A s	sWSPPEVSR_1P_OautresModif	0.5	0.4	0.5	0.7	0.9	0.4	0.6	0.4	0.4	0.4	0.9	0.8	0.5	0.5	0.7	0.9	0.5	0.7	0.8	0.9	1.0	1.0	0.7	1.1
H3BP21	3-hydroxyacyl-CoA d	wLDEsDAEmELR_1P_1autresModi	0.8	0.8	1.3	0.8	0.7	0.8	0.7	0.9	0.9	0.9	1.4	1.2	0.7	0.8	2.2	1.1	0.8	0.7	1.3	1.1	1.2	0.8	0.9	0.8
Q90666	Neuroblast differenti	iPSGGAASPTGSVDIR_1P_0autr	0.7	1.0	1.2	1.5	1.4	1.4	1.0	1.1	0.9	1.0	0.9	1.2	0.9	1.1	1.1	1.2	1.3	1.0	1.1	1.2	1.0	1.0	1.0	1.1
E9PK06	Elongation factor 1-d	tAPQTQVhSPMR_1P_1autresMod	0.8	1.5	1.3	1.2	1.1	1.2	0.9	1.2	1.1	1.3	1.3	1.2	0.9	1.1	1.0	0.9	0.9	1.0	1.2	1.1	0.8	1.0	0.9	1.0
Q92882	Osteoclast-stimulati	nSLAEDVLDEDS_1P_OautresM	0.9	xxx	0.8	xxx	1.0	0.9	1.1	1.0	0.9	0.8	1.0	xxx	0.9	0.4	xxx	0.7	0.7	0.9	0.9	0.9	0.9	0.9	0.9	0.8
Q8IYB3	Serine/arginine repe	rSPsPAPPR_2P_OautresModif	0.9	1.8	1.6	1.4	1.4	1.6	1.3	1.7	0.9	2.0	1.2	1.6	1.3	1.3	1.3	1.4	1.1	1.3	1.2	1.7	1.2	1.1	1.0	0.9
A8K8G0	Hepatoma-derived g	GDLLSDSPK_1P_OautresModif	0.8	0.9	1.0	1.1	0.9	0.9	1.4	0.9	0.7	1.2	1.4	1.5	1.5	1.0	1.7	1.0	1.0	0.8	1.1	0.9	0.9	0.8	0.9	0.9
B3KPE7	Bcl-2-like protein 13	sPATSFLVELDEEEV_1P_0autr	0.9	1.0	0.8	0.8	0.8	1.6	1.0	0.8	0.5	0.6	0.8	0.6	1.6	0.7	xxx	0.9	0.7	0.6	2.0	1.6	1.2	1.2	1.0	1.6
Q13576	Ras GTPase-activati	gSVVDDER_1P_OautresModif	0.9	0.9	0.3	0.4	0.3	0.4	0.5	0.5	0.4	0.4	0.5	1.0	0.3	0.4	0.8	0.4	0.3	0.4	0.7	0.7	1.2	1.1	1.2	1.2
Q15388	Mitochondrial import	iVSAQSLAEDDVE_1P_OautresModi	1.1	1.2	1.3	0.9	1.1	1.3	1.2	1.0	0.8	0.8	1.1	1.0	1.3	0.9	1.2	1.1	0.9	0.9	0.9	1.0	0.9	0.8	1.0	1.0
Q5JW85	Glutathione S-transf	eLQPGsPR_1P_OautresModif	0.7	0.9	1.0	1.4	1.2	0.3	1.0	0.7	0.5	0.2	0.3	0.2	0.2	0.6	0.4	0.5	0.9	0.4	0.6	1.0	0.8	0.6	0.9	1.2
E9PPJ4	Solute carrier organi	vLAVTDSPAR_1P_OautresModif	0.9	1.0	1.0	0.5	0.9	0.8	0.8	0.9	0.9	0.7	0.9	1.0	0.7	1.0	0.9	0.9	0.6	0.8	1.0	1.0	1.0	1.1	1.0	0.9
E9PGR8	Tubulin-tyrosine liga	sSPQTPPEGAALAEFAALHGPA	0.9	1.5	1.9	0.9	1.5	0.9	1.3	1.3	0.9	1.0	1.3	1.7	1.1	0.6	1.1	0.9	0.9	0.5	1.2	1.7	xxx	xxx	1.1	1.5
H7BZK1	Sorbin and SH3 dom	IGEVTSPPPR_1P_OautresModi	0.7	0.9	0.9	1.1	0.9	1.0	1.1	1.2	0.7	1.1	1.3	1.3	1.0	0.8	1.0	1.0	0.8	1.0	1.1	1.1	1.1	1.1	1.0	1.0
B4DS13	Eukaryotic translati	IGeSSQTGSTTSR_1P_0autres	0.6	1.0	0.4	0.8	0.6	1.0	1.1	0.7	0.6	1.9	1.1	0.8	0.6	0.3	0.6	0.7	0.7	0.8	0.9	1.0	1.0	0.7	0.8	1.1
B3KRG5	RNA-binding motif p	iVYLSPR_																								

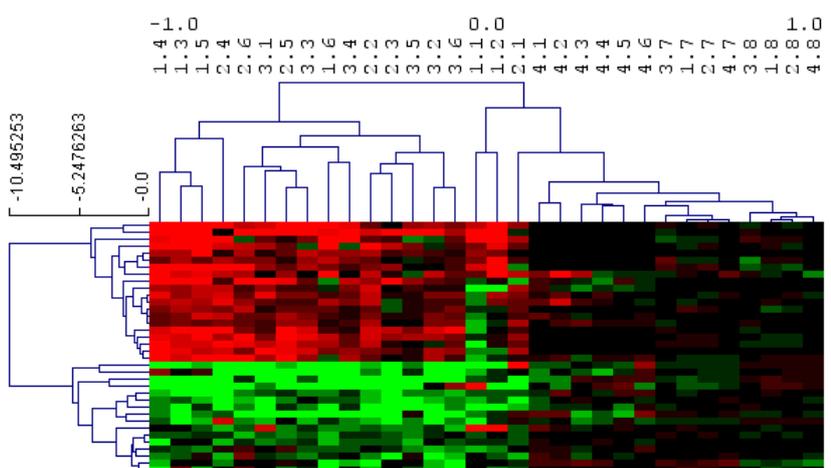
Classification hiérarchique.



Tous les peptides



Peptides significatifs



Phosphopeptides significatifs

Test non paramétrique *Rank Product*

Peptides surexprimés dans les tumeurs (F<0.03)

Swiss-Prot	Phos. sites	Gene	Name	Peptides
B4E2T8	0P	ANXA6	Annexin	eEGGENLDQAR
C9J9W2	0P	HSP90AB1	Heat shock protein HSP 90-beta	eDQTEYLEER
B4E2T8	0P	IGKC	Ig kappa chain C region	vDNALQSGnSQESVTEQDSK
Q9Y2W1	0P	LMNA	Lamin A/C	eDLQELNDR
P08238	1P	AHNAK	Neuroblast differentiation-associated protein AHNAK	sNsFSDER
P01834	1P	CANX	Calnexin	sDAEEDGGTVsQEEEDR
Q5TCJ4	1P	CANX	Calnexin	aEEDEILNRsPR
H3BQZ7	1P	G3BP1	Ras GTPase-activating protein-binding protein 1	sSsPAPADIAQTVQEDLR
E5RIS7	1P	hCG_204479	HCG2044799	sGDETPGSEVPGDK
Q9UQ35	1P	LASP1	LIM and SH3 domain protein 1 (Fragment)	rDsQDGSYSR
Q9H1E3	1P	LMNA	Lamin A/C	sGAQASSTPLsPTR
Q9H1E3	1P	NUCKS1	Nuclear ubiquitous casein and cyclin-dependent kinase substrate 1	tSTsPPPEK
Q5TCJ4	1P	NUCKS1	Nuclear ubiquitous casein and cyclin-dependent kinase substrate 1	aTVTPsPVK
F5H4D6	1P	PAK4	Serine/threonine-protein kinase PAK 4	rDsPPPPAR
Q9UQ35	1P	SRRM2	Serine/arginine repetitive matrix protein 2	sSsPVTELASR
O96013	1P	SRRM2	Serine/arginine repetitive matrix protein 2	rPsPQPSPR
C9J5V9	1P	TCEA1	Transcription elongation factor A protein 1	ePAITSQNsPEAR
Q09666	1P	THRAP3	Thyroid hormone receptor-associated protein 3	iDisPSTFR
Q5VUA4	1P	YBX1	Nuclease-sensitive element-binding protein 1 (Fragment)	nEGsESAPEGQAQQR
E7EMC6	1P	ZNF318	Zinc finger protein 318	rSsPPPPSGSSSR
Q9UQ35	2P	SRRM1	Serine/arginine repetitive matrix protein 1	rsPsPAPPPR
Q8IYB3	2P	SRRM2	Serine/arginine repetitive matrix protein 2	sStPPRQsPSR
Q9UQ35	2P	SRRM2	Serine/arginine repetitive matrix protein 2	rPsPQPSPR

Peptides sous-exprimés dans les tumeurs (FDR<0.03)

Swiss-Prot	Phos. sites	Gene	Name	Peptides
Q5T8M8	0P	ACTA1	Actin, alpha 1, skeletal muscle	sYELPDGQVITIGNER
Q5T8M8	0P	ACTA1	Actin, alpha 1, skeletal muscle	dSYVGDQEAQSk
F8VXI5	0P	ALDH2	Aldehyde dehydrogenase, mitochondrial	tFVQEDIYDEFVER
C9J4K3	0P	ALDOB	Fructose-bisphosphate aldolase	eLSEIAQSIVAnGk
C9J4K3	0P	ALDOB	Fructose-bisphosphate aldolase	eTTIQGLDGLSER
B4DDF9	0P	ANXA4	Annexin	vLLVLcGGDD
E7ENY8	0P	COL3A1	Collagen alpha-1(III) chain	dGSpGEPGAnGLpGAAGER
P31327	0P	CPS1	Carbamoyl-phosphate synthase [ammonia], mitochondrial	iAPsFAVESIEDALK
P31327	0P	CPS1	Carbamoyl-phosphate synthase [ammonia], mitochondrial	iEFEGQPVDFVDPNK
A8MW49	0P	FABP1	Fatty acid-binding protein, liver	aIGLPEELIQk
A8MQC5	0P	H2AFV	Histone H2A	aGLQFPVGR
P69905	0P	HBA1	Hemoglobin subunit alpha	vGAHAGEYGAEALER
F8W6P5	0P	HBB	LVV-hemorphin-7 (Fragment)	vNVDEVGGEALGR
F5GZQ4	0P	LDHA	L-lactate dehydrogenase A chain (Fragment)	dLADELALVDVIEDk
P05387	0P	RPLP2	60S acidic ribosomal protein P2	iLDSVIGIEADDDR
H0YL80	0P	TPM1	Tropomyosin alpha-1 chain (Fragment)	iQLVEEELDR
F8WCJ4	1P	ACSS2	Acetyl-coenzyme A synthetase, cytoplasmic	sWSPPPEVSR
F5GXU9	1P	BCKDHA	2-oxoisovalerate dehydrogenase subunit alpha, mitochondrial (Fragment)	sVDEVNYWDk
H3BTK3	1P	CARHSP1	Calcium-regulated heat stable protein 1 (Fragment)	gNVVPSPLPTR
H0YEV2	1P	CTTN	Src substrate cortactin (Fragment)	IPsSPVYEDAASFk
B4DS13	1P	EIF4B	Eukaryotic translation initiation factor 4B	tGsESSQTGTSTSSR
Q5JW85	1P	GSTA3	Glutathione S-transferase	flQPGsPR
Q13576	1P	IQGAP2	Ras GTPase-activating-like protein IQGAP2	yGsiVDDER
C9JSU1	1P	LRRFIP2	Leucine-rich repeat flightless-interacting protein 2 (Fragment)	rGsGDTSSLIDPDTLSSELR
H0YJS7	1P	NDRG2	Protein NDRG2 (Fragment)	tAsLTSASVDGNNR
B1APG2	1P	PRKACB	Protein kinase, cAMP-dependent, catalytic, beta (Fragment)	tWTLGTPPEYLAPeIILSk
E9PPJ4	1P	SLCO2B1	Solute carrier organic anion transporter family member 2B1	vLAVTDsPAR
B3KMK8	1P	STBD1	Starch-binding domain-containing protein 1	hsSWGDVGVGGSLK
B4DN25	1P	UGDH	UDP-glucose 6-dehydrogenase	iPYAPsGEIPk
O95218	1P	ZRANB2	Zinc finger Ran-binding domain-containing protein 2	eEsDGGEYDEFGR
Q9COC2	2P	TNKS1BP1	182 kDa tankyrase-1-binding protein	vPssDEEVVEEQSR

3. Communications par poster

Hepatocellular Carcinoma in Non Fibrotic Liver: a Phosphoproteomic Approach

Luc Négroni¹, Stéphane Claverol², Eric Chevet³, Jean Rosenbaum³ and Jean-Marie Schmitter^{1*}

¹ University of Bordeaux, UMR5248, Centre de Génomique Fonctionnelle, Bordeaux, France

² University of Bordeaux, Pôle Protéomique, Centre de Génomique Fonctionnelle, Bordeaux, France

³ University of Bordeaux, GREF, Bordeaux, France

INTRODUCTION

Hepatocellular carcinoma (HCC) most frequently develops on a pre-existing cirrhosis. 10 to 40% of HCC occur in the absence of cirrhosis and about 5% in non-fibrotic liver (nHCC). In 20 to 40% of nHCC, no classical risk factor of HCC can be identified. Therefore, nHCC offers a unique opportunity to investigate factors that are directly carcinogenic in the absence of cirrhosis. In particular, the identification of specific phosphorylation sites associated with pathological states may be indicative of the activation of signaling axes, and these latter ones may represent interesting therapeutic targets.

With this in view, we have undertaken the characterization of the phosphoproteome from several subsets of nHCC, using ITRAQ 8-plex labeling for quantitative analysis. However, this methodology allows to label a max. of 100µg peptide sample per ITRAQ unit. Thus, we tried to develop a robust protocol aimed at phosphopeptide enrichment prior to ITRAQ labeling.

Large scale purification of phosphopeptides from liver tissue by a tandem enrichment method using SCX followed by IMAC has been presented [1]. We did not select this approach, because it would have required labeling a great number of SCX fractions, and we have chosen phosphopeptide enrichment as first step. However, both IMAC and TiO₂ enrichment methods suffer from co-elution of acidic peptides. Efforts to reduce the contamination led to several protocols. For instance, TiO₂ was shown to be most effective when 5%TFA, and excluders such as glycolic acid or lactic acid were used [2], whereas IMAC (Phos-Select) with 0.25M acetic acid, 30% ACN loading buffer was presented as the most effective method [3]. These controversial results suggest that some parameters were still uncontrolled. Therefore, we started tests in a realistic situation, i.e. liver tissue rather than a mixture of purified proteins, to select suitable conditions for our phosphopeptide enrichment step.

METHODS

Sample preparation

100 mg of liver tissue was homogenized in 0.5 ml of buffer containing 0.1M TEAB pH 9.3, 0.15 M NaCl, 0.10% SDS, 5mM NaF, 2mM orthovanadate and 10 mM DTT, then alkylated with 50 mM IAA and digested with 100 µg trypsin (Sigma). A Sep-Pack cartridge was used for subsequent sample clean-up.

Phosphopeptide enrichment

POROS 20 MC (Applied Biosystems), Phos-Select (Sigma) and TiO₂ were packed in 2 mm x 5 cm columns. The chromatographic system (Beckman) had a 500 µl sample loop, with a column oven set to 40°C. Washing buffers were 50 mM Acetic Acid for POROS, 0.25M Acetic Acid for Phos-Select, 0.25M Acetic Acid or 5%TFA for TiO₂.

IMACs were activated with 100 mM FeCl₃. Three loading buffers were tested with POROS 20 MC. Buffer 1: 0.1M Acetic Acid 20% ACN pH 2.9; buffer 2: 0.1M Formic Acid 20% ACN pH 2.3; buffer 3: 0.1M TFA, 20% ACN pH 1.3. Phos-Select was used with loading buffer 4: 0.25M Acetic Acid 30% ACN. TiO₂ was used with buffer 4 and buffer 5: 5% TFA, 30% ACN. Peptides were always eluted with 0.4 M NH₄OH.

LC-MS/MS and data processing

Samples were analyzed with a LTQ orbitrap (Thermo Scientific) coupled to a nanoLC system (Dionex). Run time was 2h, with a "TOP 10" acquisition method. Data processing was performed with Proteome Discoverer (Thermo Scientific), or Xlndem. The presented results were obtained with FDR<1% and sequence redundancy was removed.

1- REPEATABILITY OF LC-MS/MS

An IMAC eluate was analyzed twice by LC-MS/MS. Only minor variations of retention time appear between the 2 runs (Fig. 1A). Numbers of identified peptides, phosphopeptides and phospho-Serine, Threonine, Tyrosine ratio are almost identical for the 2 runs (1B). However, there is an important variation in the identified sequences. Ven's diagram shows that ~30% of phosphorylated sequences are specific for either run 1 or run 2 (1C). With our conditions, a maximum of ~2500 peptides/run may be identified. This indicates that the scan time of the Orbitrap is not fast enough for the analysis of all peptides and points out (i) the importance of the development of a highly selective enrichment step for phosphopeptides, (ii) the probable positive impact of tandem purification by SCX and (iii) the interest of repeated LC-MS/MS runs when applicable.

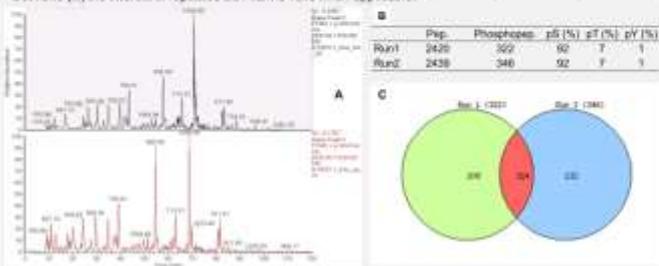


Fig.1: A: LC-MS/MS base peak of IMAC purified fraction (POROS 20 MC loaded with buffer 1). B: the number of total peptides, phosphopeptides and % of pS, pT and pY in the 2 runs. C: Ven's diagram of phosphopeptides identified in runs 1 and 2.

2 - TOP 10 IDENTIFICATIONS

A peptide counting approach was used to classify the first 10 proteins for the different conditions. Contamination with unphosphorylated peptides clearly decreases with TFA loading buffer, for both POROS and TiO₂.

	POROS/AA	POROS/FA	POROS/TFA	Phos-Select	TiO ₂ /AA	TiO ₂ /TFA
1	MUP 2	CPS	HSP 90- α	FABP	MUP 11 and 8	HSP 90- β
2	MUP 11 and 8	GST P 1	CPS	GST P 1	Myosin-9	CPS
3	FABP	HSP 90- β	SIA RPMP 1	MUP 11 and 8	FABP	NDRG2
4	GST P 1	HSP 90- α	HSP 90- α	Myosin-9	SAB	SIA RPMP 1
5	78 kDa GRP	Endoplasmic	SAB	Serum albumin	Catecholamin	HSP 90- α
6	CPS	Catecholamin	NDRG2	Endoplasmic	FD	EP 3 B
7	GD 1	MAPRC 1	SIA RPMP 2	Regucalcin	Catalse	SAB
8	Myosin-9	Myosin-9	Myosin-9	Catecholamin	GST P 1	3-K-CoA T
9	Catecholamin	MUP11 and 8	WD RCP C2	CA 3	CA 3	60S ARP P2
10	60 kDa HSP	GST My 1	RGAP 12	CPS	Utrich oxidase	HGF

Table1: Top ten identified proteins for different enrichment conditions. Bold corresponds to proteins identified with phosphopeptides. 3-K CoA T: 3-ketolacyl-CoA thioester; ARP: acidic ribosomal protein; FABP: Fatty acid-binding protein; CPS: Carbamoyl-phosphate synthase; HGF: Hepatoma-derived growth factor; MAPRC: Membrane-associated progesterone receptor component; MUP: Major urinary proteins; RGAP: Rho GTPase-activating protein; WD RCP: WD repeat-containing protein.

CONCLUSION

We found that POROS, Phos-Select and TiO₂ enrichment protocols led to phosphopeptide fractions highly contaminated with non-phosphorylated peptides (essentially acidic ones) when acetic acid was used as loading/washing buffer. This lack of selectivity can be reduced with a lower pH, but this solution cannot be applied to IMAC columns due to an important Fe³⁺ leaching and a decrease of their phosphopeptide binding capacity. The best results were obtained with TiO₂ and a high concentration of TFA. With 5%TFA and 30% ACN, 312 non-redundant phosphopeptides were identified from liver tissue, corresponding to 55% of the identified peptides. This ratio can probably be increased with the use of excluders such as glycolic acid or lactic acid, and thus allowing efficient labeling with ITRAQ reagent in the rest of our process for quantitative analysis of liver biopsies.

3 - PHOSHOPEPTIDE ENRICHMENT AND CONTAMINATION

POROS 20MC enrichment efficiency is dramatically affected by loading conditions. TFA gives the highest proportion of phosphopeptides (80%), but only 81 phosphopeptides are identified with this condition. Phos-select used with its recommended buffer presents an important contamination. TiO₂ produced either the worst results (with acetic acid) or the best results (312 phosphopeptides identified with TFA). TFA and the corresponding low pH seems to be a key factor for a successful purification.

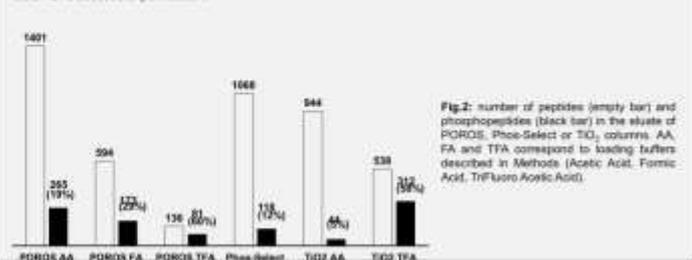


Fig.2: number of peptides (empty bar) and phosphopeptides (black bar) in the eluate of POROS, Phos-Select or TiO₂ columns. AA, FA and TFA correspond to loading buffers described in Methods (Acetic Acid, Formic Acid, Trifluoro Acetic Acid).

4 - IRON LEACHING

Decreasing the pH is not recommended with IMAC supports due to a Fe³⁺ leaching at low pH. Absorbance monitoring at 280nm for a POROS 20MC column shows that a major peak corresponding to iron appears with Formic Acid or TFA, explaining the decrease of the number of phosphopeptides identified with these conditions.

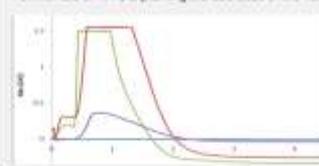


Fig.3: UV monitoring (280 nm) of a POROS 20MC eluate after injection of loading buffer. Blue trace: TFA without FeCl₃ activation; Green trace: TFA with FeCl₃ activation; Red trace: Formic Acid with FeCl₃ activation; Purple trace: Acetic Acid with FeCl₃ activation.

5 - PHOSPHORYLATION LEVEL

Chromatographic media and buffers affect slightly the level of phosphorylation found in identified peptides: higher proportions of 2p and 3p phosphopeptides are identified when the pH of loading buffer is decreased.

Table2: proportion of 1, 2 and 3 phosphate groups per phosphopeptide for the different conditions of chromatography.

1. Villa, I., et al., Large-scale phosphorylation analysis of mouse liver. *PNAS USA*, 2007, **104**, 1488-1493. 2. Jensen, S.S. and M.R. Larsen, Evaluation of the impact of some experimental procedures on different phosphopeptide enrichment techniques. *RCMS*, 2007, **12**, 365-367. 3. Bednarek, R., et al., Reproducible isolation of distinct, overlapping segments of the phosphoproteome. *Nat Methods*, 2007, **4**, 211-217.

Comparison between IMAC and TiO₂ for on-line purification of phosphopeptides: application for quantitative analysis of hepatic biopsies

Luc Negroni¹, Sophie Geiger¹, Stéphane Claverol², Marc Bonneau², Eric Chevet³, Jean Rosenbaum³, Jean-Marie Schmitter¹

¹ University of Bordeaux, UMR5248-CNRS, Centre de Génomique Fonctionnelle, Bordeaux, France

² University of Bordeaux, Pôle Protéomique, Centre de Génomique Fonctionnelle, Bordeaux, France

³ University of Bordeaux, GREF-INSERM, Bordeaux, France

INTRODUCTION

On one hand, research for biomarkers needs quantitative methods applied to a large number of samples in order to allow the statistical analysis required to validate biomarkers. On the other hand phosphoproteomics require an enrichment of phosphopeptides that is usually done by affinity chromatography with disposable tips or in-vial techniques, i.e. processes that are not suitable in terms of reproducibility for large series of samples. As we project to identify biomarkers specific to the phosphoproteome, we developed an enrichment protocol relying on the use of an affinity support packed in a HPLC column. Test experiments showed unambiguously the superior performance of TiO₂ in comparison to IMAC-Fe³⁺ (POROS 20 MC). Further, the protocol was optimized using a simple mixture (tryptic digest of casein-albumin) and confirmed the results of Jensen et al [1]: high concentrations of trifluoroacetic acid (TFA) decrease the non specific binding of mono-acidic peptides. We successfully applied the method to complex samples; in particular, a protocol with the implementation of ITRAQ labeling and ion exchange (SCX) fractionation has been used for the analysis of biopsies from hepatic carcinoma.

METHODS

Samples

The mixture of Casein/albumin tryptic peptides (Cas/BSA) was prepared with a 1:10 molar ratio. Tryptic peptides corresponding to complex samples were obtained from 50 mg of liver (mouse or biopsies) and 20 mg of vine roots (Vitis Vifera) respectively.

Phosphopeptide enrichment

POROS 20 MC (Applied Biosystems) and TiO₂ (GL Science) were packed in 2 mm x 5 cm columns and mounted in a HPLC system with UV (280 nm) detector (Beckman). All the chromatographic steps (sample loading, washing and elution) were carried out via the 500 µl sample loop. POROS was activated with 100 mM FeCl₃ and regenerated with 0.2 M EDTA. Peptides were eluted with 0.5M NH₄OH.

LC-MS/MS and data processing

Cas/BSA was analyzed with a CapLC - Qtof Premier (Waters) LC/MS/MS system. Complex samples were analyzed with a LTQ orbitrap (Thermo Scientific) coupled to a nanoLC system (Donex). Peptides identification was carried out with Proteome Discover™ (Thermo Scientific).

1- LIMITATION OF IMAC FOR ON-LINE PURIFICATION: Fe³⁺ LEACHING



Fig.1: Fe³⁺ leaching and precipitation

Left: UV trace obtained with IMAC-Fe³⁺ column and injection of different loading buffers. The significant peak is obtained when Fe³⁺ is not chelated on resin. Upper: color of the resin after the 1st FeCl₃ activation (A), then EDTA regeneration (B) and after multiple EDTA regenerations (C).

IMAC-Fe³⁺ is not a stable affinity medium after FeCl₃ loading and washing. Fe³⁺ is still released. Fe³⁺ leaching increases with the strength of the acid (acetic acid < formic acid < TFA). Furthermore, with time, the support keeps its brown color even when EDTA regeneration is carried out. NH₄OH used for phosphopeptides elution converts Fe³⁺ in Fe(OH)₃, which spontaneously forms a brown precipitate of Fe₂O₃·3H₂O (i.e. rust).

3 - APPLICATION TO COMPLEX SAMPLES

Complex mixtures of peptides (500 µg from mouse liver or 100 µg from root extract of Vitis Vifera) was purified by TiO₂ column chromatography using 5% TFA, 1M Glycolic, 20% ACN as loading buffer. Phosphopeptides represents up to 60% of the identified peptides, corresponding to 450 and 201 non redundant phosphopeptides of liver and root proteins respectively.

Table1: Number of peptides classified according their phosphorylation level

Phosphorylation	Mouse liver	Vitis
0	165	63
1	259	128
2	24	9
3	2	1
>3	0	0
Total	450	201
Phos./Total (%)	63 %	69 %

* The S/T/Y ratio observed with both samples is roughly 100/10/1

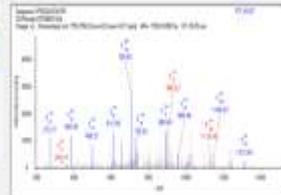


Fig.3: CID spectra of phosphorylated peptides.

Phosphopeptides are identified with their mass (MS) and their fragments (MS/MS). Here S, T, Y, unambiguously sign the C-terminal part of the peptide. On the contrary, S/T is characterized by low intensity ions (y₁ and b₁). In most cases, MS/MS spectra contain no ions characteristic of the phosphorylation site when S and T are phosphorylated.

2- EFFECT OF ACID AND NaCl ON TiO₂ SPECIFICITY

TiO₂ specificity depends on the acid used in the loading buffer. A high concentration of TFA dramatically decreases the contamination of mono-acidic peptides (D/E rich). NaCl does not modify the elution profile, confirming that TiO₂ chromatography is not affected by salts and does not involves ionic interaction.

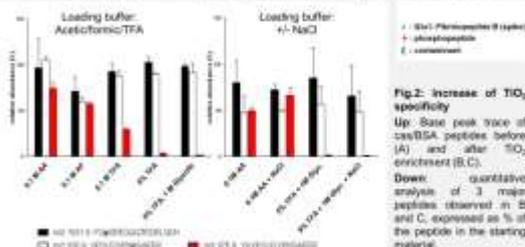
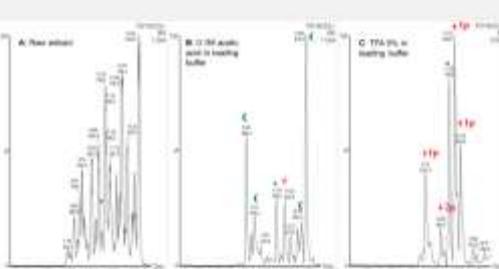


Fig.2: Increase of TiO₂ specificity

Up: Base peak trace of cas/BSA peptide before (A) and after TiO₂ enrichment (B,C).
Down: quantitative analysis of 3 major peptides observed in B and C, expressed as % of the peptide in the starting material.

4- QUANTITATIVE ANALYSIS OF PHOSPHOPEPTIDES

The following protocol is proposed for quantitative analysis. Preliminary results obtained with 3 liver biopsies showed that fraction 3 from SCX contains essentially labeled phosphopeptides (see below). Fraction 1 contains the unretained material, corresponding mainly to multi-phosphorylated and unlabeled peptides. Fraction 4 corresponds to acidic (D/E rich) but unphosphorylated peptides.

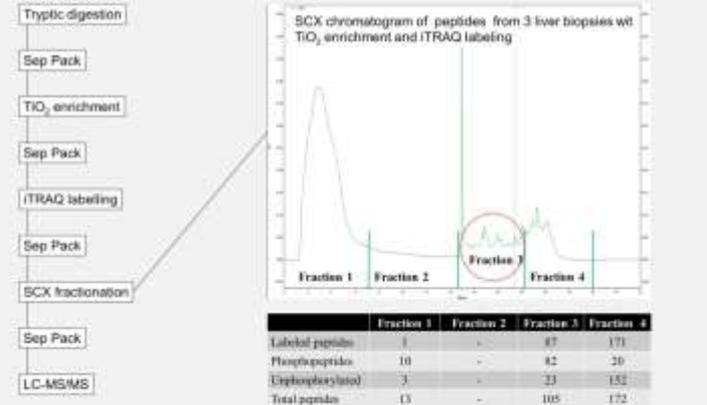


Fig.4: Process for quantitative analysis of phosphopeptides

The process involves multiple purification steps which are carried out by HPLC (including Sep Pack). The SCX chromatogram is obtained after TiO₂ enrichment and ITRAQ labeling of 3 liver biopsies. Fraction 1 and 2 corresponds to flow through and gradient delay respectively. The table presents the number of non-redundant peptides identified in the four SCX fractions.

CONCLUSION

Fe³⁺ leaching renders POROS-Fe³⁺ inappropriate for on-line chromatography of phosphopeptides. On the contrary, TiO₂ is a stable and reusable affinity medium. It has been used for the purification of plant and animal extracts, hundreds of phosphopeptides were identified, and the proportion of contaminating mono-acidic peptides was limited by the use of a highly acidic loading buffer. The TiO₂ protocol has been integrated in a larger process in order to propose a method for the quantification of phosphopeptides with large series of samples. Results obtained with liver biopsies show that SCX fractionation after ITRAQ labeling concentrates the phosphorylated and labeled peptides in the first part of the salt gradient. The corresponding fraction contains nearly 80% phosphopeptides. Thus, the dual TiO₂-SCX column purification of ITRAQ phosphopeptides appears well adapted to large sample sets.

Jensen SS, Larsen MR. Evaluation of the impact of some experimental procedures on different phosphopeptide enrichment techniques. *Appl Commun Mass Spectrom* 2007; 21:3635-3645.

Dosage d'extraits peptidiques avant analyse quantitative par la méthodologie iTRAQ

Luc Negroni, Robin Rullaud, Jean-Marie Schmitter

UMR 5248, Centre de Génomique Fonctionnelle, Université de Bordeaux – Mèl : l.negroni@cbmn.u-bordeaux.fr

INTRODUCTION

L'analyse protéomique utilise souvent un marquage isotopique pour permettre l'analyse quantitative simultanée par spectrométrie de masse de peptides dans des échantillons différents (de 2 à 8). Le réactif de marquage est toujours utilisé en excès pour assurer un marquage total des groupements réactifs; une quantification du substrat (peptides ou protéines) est donc nécessaire pour respecter cette condition. En général, la quantification est réalisée sur le mélange protéique avant digestion. Par contre, un dosage des peptides avant marquage devient nécessaire quand une étape de fractionnement ou enrichissement sélectif est réalisée.

Nous avons été confrontés à ce problème dans le cadre d'une analyse de phosphoprotéome utilisant le marquage iTRAQ. Partant de 1 mg de peptides, nous avons dû déterminer quelle quantité de phosphopeptides était purifiée en sortie de chromatographie d'affinité, afin de respecter le protocole iTRAQ (100 µg maximum de protéines/peptides par unité de réactif iTRAQ).

Pour quantifier les fractions enrichies en phosphopeptides, nous avons comparé différents dosages colorimétriques (dosages de Bradford, Lowry et BCA) avec un dosage par LC-MS d'acides aminés.

MATERIELS et METHODES

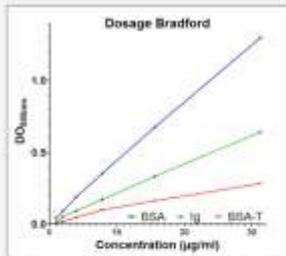
- Pour les dosages colorimétriques, les courbes dose-réponse ont été réalisées avec l'albumine (bsa), une solution d'immunoglobulines (Ig) et une solution de peptides typiques de l'albumine (bsa-T).
- Pour la composition en acides aminés, l'hydrolyse acide a été réalisée avec le system Protog (WATERS). L'analyse LC-MS/MS avec un system capLC-Qtof Premier (WATERS) sur colonne C18 (250 µm x 15 cm), gradient de 20% à 100% B en 30 min. (A= 0.1% ac. formique 1% ACN, B= 0.1% ac. formique, 99% ACN).

Dosage Bradford*

*Bio-Rad Protein Assay

Principe: le colorant G250 se lie aux acides aminés aromatiques et résidus hydrophobes des acides aminés.

Protocole: 400 µl échantillon + 100 µl réactif, vortex, repos de 15 min, lecture à 595 nm



Gamme 1-25 µg/ml, Sensibilité 0.5 µg

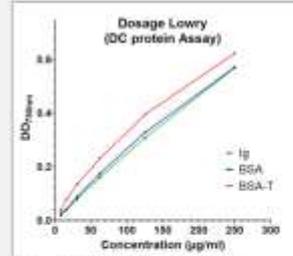
- Les peptides typiques ne lient pas/peu le colorant.
- Forte variation entre protéines (pertes différentes).
- ⇒ ce colorant ne permet pas le dosage de protéines ou peptides en mélange.

Dosage Lowry**

**DC Protein Assay de Biorad, contient sol. A (sol. alcaline de Cu²⁺) et sol. B (sol. de Folin)

Principe: en milieu alcalin, le cuivre Cu²⁺ est complexé et réduit en Cu⁺ par les liaisons peptidiques, le Cu⁺ formé réduit alors le réactif de Folin pour développer une coloration bleue.

Protocole: 100 µl échantillon + 50 µl A + 400 µl B; vortex, repos 15 min, lecture à 750 nm.



Gamme 5-250 µg/ml, sensibilité 1 µg

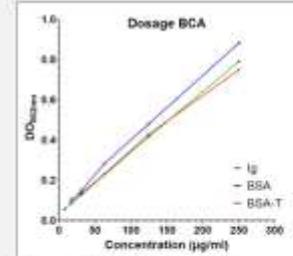
- Faible variation entre protéines.
- Légère augmentation pour les peptides.
- ⇒ Dosage utilisable pour les peptides et protéines.

Dosage BCA***

*** BCA Protein assay de Thermo Fisher, contient une sol. A (ac. bicinchoninique) et sol. B(sol. alcaline de Cu²⁺).

Principe: comme pour le dosage Lowry, le cuivre Cu²⁺ est complexé et réduit en Cu⁺; dans un deuxième temps, chaque Cu⁺ est complexé par 2 molécules d'acide bicinchoninique pour développer alors une coloration violette.

Protocole: 62.5 µl échantillon + 500 µl réactif (50 part A+1 partB); vortex, repos 30 min à 37°C, lecture à 562 nm.



Gamme 5-250 µg/ml, sensibilité 1 µg

- Faible variation entre protéines.
- Légère diminution pour les peptides.
- ⇒ Dosage utilisable pour les peptides et protéines.

Dosage des acides aminés par LC-MS

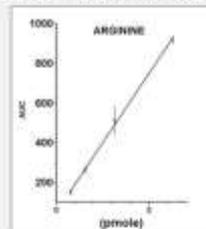
Principe: Après hydrolyse acide des protéines, les acides aminés libérés sont quantifiés par LC-MS. Afin d'obtenir une rétention suffisante sur colonne C18, ils sont dérivés par le Fluorénylméthyl(oxy)carbonyl chloride (FMOC)

Réactifs: standard d'acides aminés (Sigma), FluorénylMéthylOxyCarbonyl chloride (FMOC) 10 mM (2.59 mg/ml dans l'acétonitrile)

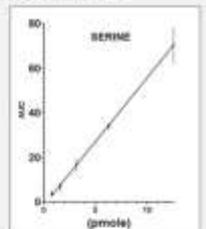
Protocole: (i) Hydrolyse acide : HCl 6N, 115 °C, 5h sous vide (ii) Dérivation : 50 µl borate 0.5 M pH 8.8 + 100 µl FMOC, 5 min à 25°C.



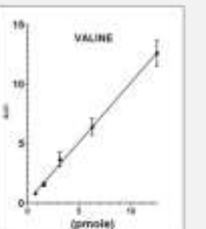
Base Peak Intensity (BPI) et colorant ionique pour 6 acides aminés choisis (25 pmoles injectés).



Gamme: 0.25-10 pmoles



Gamme: 1-20 pmoles



Gamme: 1-20 pmoles

- Quantification possible à partir de plusieurs acides aminés sauf W, N, Q, C, M (l'hydrolyse acide détruit W, transforme les amides en acides [N → D et Q → E], C et M peuvent être oxydés).
- La quantité d'acides aminés permet d'estimer la quantité de protéine correspondante par la formule suivante: $\text{protéine}_{\text{µg}} = \text{acide aminé}_{\text{pmoles}} \times \text{PM}_{\text{acide aminé}} \times \text{fp}$ (fp = proportion de l'acide aminé considéré dans une séquence protéique type, soit 0.05, 0.088, 0.066 pour respectivement R, S, V).
Exemple: 1 pmole R correspond à $1 \times 156 \times 1/0.05 = 3.12 \mu\text{g}$ protéine/peptide.

Choix du dosage

- Nous avons utilisé le dosage Lowry (DC protein assay) pour quantifier nos échantillons. Dans nos mains il est apparu aussi sensible que le dosage BCA tout en étant plus facile à mettre en œuvre (incubation plus courte, pas de chauffage à 37°C).
- La composition en acides aminés par LC-MS n'a pas été choisie comme technique de quantification car elle est plus difficile à mettre en œuvre. Dans le cadre de notre marquage iTRAQ, elle ne se justifiait pas. C'est cependant la seule méthode testée qui permet un étalonnage précis.

CONCLUSION

- le dosage Bradford est déconseillé pour le dosage des protéines, car sa gamme d'étalonnage varie fortement suivant le type de protéines. Par ailleurs, sa sensibilité vis-à-vis des peptides est quasi nulle.
- Nous proposons l'emploi d'un dosage de Lowry pour déterminer la quantité de peptides ou de protéines avant marquage iTRAQ. Ce dosage est rapide et permet de doser indifféremment les peptides et les protéines. L'information fournie permet alors de limiter les quantités d'iTRAQ à utiliser pour le marquage.
- Application: le dosage Lowry nous a permis de déterminer que 10 µg de phosphopeptides étaient purifiés par une colonne d'affinité TiO₂ à partir de 1 mg de peptides typiques provenant de biopsies hépatiques. Grâce à cette mesure, nous n'avons utilisé que 1/3 d'unité iTRAQ, soit une économie de plus de 60% du kit de marquage commercial.

Références bibliographiques

- Abian, J., Oosterkamp, A. J., and Gelpí, E. (1999). Comparison of conventional, narrow-bore and capillary liquid chromatography/mass spectrometry for electrospray ionization mass spectrometry: practical considerations. *J Mass Spectrom* **34**, 244-254.
- Ahn, Y. H., Park, E. J., Cho, K., Kim, J. Y., Ha, S. H., Ryu, S. H., and Yoo, J. S. (2004). Dynamic identification of phosphopeptides using immobilized metal ion affinity chromatography enrichment, subsequent partial beta-elimination/chemical tagging and matrix-assisted laser desorption/ionization mass spectrometric analysis. *Rapid Commun Mass Spectrom* **18**, 2495-501.
- Alais, C., and Linden, G. (1987). "Biochimie alimentaire," Masson.
- Alkofer, B., Lepennec, V., and Chiche, L. (2011). Hepatocellular cancer in the non-cirrhotic liver. *J Visc Surg*. **148**, 3-11. doi: 10.1016/j.jviscsurg.2010.12.012.
- Alpert, A. J. (2008). Electrostatic repulsion hydrophilic interaction chromatography for isocratic separation of charged solutes and selective isolation of phosphopeptides. *Anal Chem* **80**, 62-76.
- Antberg L Fau - Cifani, P., Cifani P Fau - Sandin, M., Sandin M Fau - Levander, F., Levander F Fau - James, P., and James, P. (2012). Critical comparison of multidimensional separation methods for increasing protein expression coverage. *J Proteome Res*.
- Armstrong, M., Jonscher, K., and Reisdorph, N. A. (2007). Analysis of 25 underivatized amino acids in human plasma using ion-pairing reversed-phase liquid chromatography/time-of-flight mass spectrometry. *Rapid Commun Mass Spectrom* **21**, 2717-26.
- Aryal, U. K., Olson, D. J., and Ross, A. R. (2008). Optimization of immobilized gallium (III) ion affinity chromatography for selective binding and recovery of phosphopeptides from protein digests. *J Biomol Tech*. **19**, 296-310.
- Aryal, U. K., and Ross, A. R. (2009). Enrichment and analysis of phosphopeptides under different experimental conditions using titanium dioxide affinity chromatography and mass spectrometry. *Rapid Commun Mass Spectrom* **24**, 219-31.
- Attwood, P. V. (2013). P-N bond protein phosphatases. *Biochim Biophys Acta* **1834**, 470-8.
- Balbuena, T. S., He, R., Salvato, F., Gang, D. R., and Thelen, J. J. (2012). Large-scale proteome comparative analysis of developing rhizomes of the ancient vascular plant *equisetum hyemale*. *Front Plant Sci* **3:131**, 10.3389/fpls.2012.00131. Epub 2012 Jun 26.
- Bank, R. A., Jansen, E. J., Beekman, B., and te Koppele, J. M. (1996). Amino acid analysis by reverse-phase high-performance liquid chromatography: improved derivatization and detection conditions with 9-fluorenylmethyl chloroformate. *Anal Biochem* **240**, 167-76.
- Bardou, M., Barkun, A. N., and Martel, M. (2013). Obesity and colorectal cancer. *Gut*. **62**, 933-47. doi: 10.1136/gutjnl-2013-304701. Epub 2013 Mar 12.
- Belghazi, M., Bathany, K., Hountondji, C., Grandier-Vazeille, X., Manon, S., and Schmitter, J. M. (2001). Analysis of protein sequences and protein complexes by matrix-assisted laser desorption/ionization mass spectrometry. *Proteomics* **1**, 946-54.
- Belot, A., Grosclaude, P., Bossard, N., Jouglu, E., Benhamou, E., Delafosse, P., Guizard, A., Molinié, F., Danzon, A., Bara, S., Bouvier, A., Trétarre, B., Binder-Foucard, F., Colonna, M., Daubisse, L., Hédelin, G., Launoy, G., Le Stang, N., Maynadié, M., Monnereau, A., Troussard, X., Faivre,

- J., Collignon, A., Janoray, I., Arveux, P., Buemi, A., Raverdy, N., Schwartz, C., Bovet, M., Chérié-Challine, L., Estève, J., Remontet, L., and Velte, n. M. (2008). Cancer incidence and mortality in France over the period 1980-2005. *Rev Epidemiol Sante Publique* **56**, 159-75.
- Berry, W. L., and Janknecht, R. (2013). KDM4/JMJD2 histone demethylases: epigenetic regulators in cancer cells. *Cancer Res.* **73**, 2936-42. doi: 10.1158/0008-5472.CAN-12-4300. Epub 2013 May 3.
- Besant, P. G., and Attwood, P. V. (2012). Histone H4 histidine phosphorylation: kinases, phosphatases, liver regeneration and cancer. *Biochem Soc Trans* **40**, 290-3.
- Besant, P. G., Tan, E., and Attwood, P. V. (2003). Mammalian protein histidine kinases. *Int J Biochem Cell Biol* **35**, 297-309.
- Biemann, K. (1990). Appendix 5. Nomenclature for peptide fragment ions (positive ions). *Methods Enzymol* **193**, 886-7.
- Binder-Foucard, F., Belot, A., Delafosse, P., Remontet, L., Woronoff, A. S., and Bossard, N. (2013). Estimation nationale de l'incidence et de la mortalité par cancer en France entre 1980 et 2012. Étude à partir des registres des cancers du réseau Francim - Partie 1 – Tumeurs solides. *InVS, INCa*.
- Blackwell, J. A., and Carr, P. W. (1992). The role of Lewis acid-base processes in ligand-exchange chromatography of benzoic acid derivatives on zirconium oxide. *Anal Chem* **64**, 853-62.
- Blanc, J. F., Lalanne, C., Plomion, C., Schmitter, J. M., Bathany, K., Gion, J. M., Bioulac-Sage, P., Balabaud, C., Bonneu, M., and Rosenbaum, J. (2005). Proteomic analysis of differentially expressed proteins in hepatocellular carcinoma developed in patients with chronic viral hepatitis C. *Proteomics* **5**, 3778-89.
- Bleiholder, C., Suhai, S., Harrison, A. G., and Paizs, B. (2011). Towards understanding the tandem mass spectra of protonated oligopeptides. 2: The proline effect in collision-induced dissociation of protonated Ala-Ala-Xxx-Pro-Ala (Xxx = Ala, Ser, Leu, Val, Phe, and Trp). *J Am Soc Mass Spectrom.* **22**, 1032-9. doi: 10.1007/s13361-011-0092-1. Epub 2011 Apr 12.
- Bodenmiller, B., Mueller, L. N., Mueller, M., Domon, B., and Aebersold, R. (2007). Reproducible isolation of distinct, overlapping segments of the phosphoproteome. *Nat Methods* **4**, 231-7.
- Boersema, P. J., Geiger, T., Wisniewski, J. R., and Mann, M. (2013). Quantification of the N-glycosylated secretome by super-SILAC during breast cancer progression and in human blood samples. *Mol Cell Proteomics.* **12**, 158-71. doi: 10.1074/mcp.M112.023614. Epub 2012 Oct 22.
- Boja, E. S., Phillips, D., French, S. A., Harris, R. A., and Balaban, R. S. (2009). Quantitative mitochondrial phosphoproteomics using iTRAQ on an LTQ-Orbitrap with high energy collision dissociation. *J Proteome Res.* **8**, 4665-75. doi: 10.1021/pr900387b.
- Bonhomme, L., Valot, B., Tardieu, F., and Zivy, M. (2012). Phosphoproteome dynamics upon changes in plant water status reveal early events associated with rapid growth adjustment in maize leaves. *Mol Cell Proteomics.* **11**, 957-72. Epub 2012 Jul 10.
- Borie, F., Bouvier, A. M., Herrero, A., Faivre, J., Launoy, G., Delafosse, P., Velten, M., Buemi, A., Peng, J., Grosclaude, P., and Tretarre, B. (2008). Treatment and prognosis of hepatocellular carcinoma: a population based study in France. *J Surg Oncol.* **98**, 505-9. doi: 10.1002/jso.21159.
- Byford, M. F. (1991). Rapid and selective modification of phosphoserine residues catalysed by Ba²⁺ ions for their detection during peptide microsequencing. *Biochem J* **280 (Pt 1)**, 261-5.
- Bythell, B. J., Maitre, P., and Paizs, B. (2010). Cyclization and rearrangement reactions of a(n) fragment ions of protonated peptides. *J Am Chem Soc.* **132**, 14766-79. doi: 10.1021/ja101556g.

- Bythell, B. J., Somogyi, A., and Paizs, B. (2009). What is the structure of b(2) ions generated from doubly protonated tryptic peptides? *J Am Soc Mass Spectrom.* **20**, 618-24. doi: 10.1016/j.jasms.2008.11.021. Epub 2008 Dec 6.
- Calderaro, J., Rebouissou, S., de Koning, L., Masmoudi, A., Heralut, A., Dubois, T., Maille, P., Soyeux, P., Sibony, M., de la Taille, A., Vordos, D., Lebret, T., Radvanyi, F., and Allory, Y. (2013). PI3K/AKT pathway activation in bladder carcinogenesis. *Int J Cancer.*
- Capony, J. P., and Demaille, J. G. (1983). A rapid microdetermination of phosphoserine, phosphothreonine, and phosphotyrosine in proteins by automatic cation exchange on a conventional amino acid analyzer. *Anal Biochem* **128**, 206-12.
- Cappello, P., Tomaino, B., Chiarle, R., Ceruti, P., Novarino, A., Castagnoli, C., Migliorini, P., Perconti, G., Giallongo, A., Milella, M., Monsurro, V., Barbi, S., Scarpa, A., Nistico, P., Giovarelli, M., and Novelli, F. (2009). An integrated humoral and cellular response is elicited in pancreatic cancer by alpha-enolase, a novel pancreatic ductal adenocarcinoma-associated antigen. *Int J Cancer.* **125**, 639-48. doi: 10.1002/ijc.24355.
- Cargile, B. J., Bundy, J. L., Freeman, T. W., and Stephenson, J. L., Jr. (2004). Gel based isoelectric focusing of peptides and the utility of isoelectric point in protein identification. *J Proteome Res.* **3**, 112-9.
- Carlomagno, L., Huebner, V. D., and Matthews, H. R. (1985). Rapid separation of phosphoamino acids including the phosphohistidines by isocratic high-performance liquid chromatography of the orthophthalaldehyde derivatives. *Anal Biochem* **149**, 344-8.
- Catherman, A. D., Durbin, K. R., Ahlf, D. R., Early, B. P., Fellers, R. T., Tran, J. C., Thomas, P. M., and Kelleher, N. L. (2013). Large-scale top down proteomics of the human proteome: membrane proteins, mitochondria, and senescence. *Mol Cell Proteomics.*
- Chaerkady, R., Harsha, H. C., Nalli, A., Gucek, M., Vivekanandan, P., Akhtar, J., Cole, R. N., Simmers, J., Schulick, R. D., Singh, S., Torbenson, M., Pandey, A., and Thuluvath, P. J. (2008). A quantitative proteomic approach for identification of potential biomarkers in hepatocellular carcinoma. *J Proteome Res* **7**, 4289-98.
- Chelius, D., and Bondarenko, P. V. (2002). Quantitative profiling of proteins in complex mixtures using liquid chromatography and mass spectrometry. *J Proteome Res.* **1**, 317-23.
- Chen, M., Su, X., Yang, J., Jenkins, C. M., Cedars, A. M., and Gross, R. W. (2010). Facile identification and quantitation of protein phosphorylation via beta-elimination and Michael addition with natural abundance and stable isotope labeled thiocholine. *Anal Chem* **82**, 163-71.
- Cheng, G., Luo, R., Hu, C., Lin, J., Bai, Z., Zhang, B., and Wang, H. (2013). TiO₂-based phosphoproteomic analysis of schistosomes: characterization of phosphorylated proteins in the different stages and sex of *Schistosoma japonicum*. *J Proteome Res.* **12**, 729-42. doi: 10.1021/pr3007864. Epub 2013 Jan 18.
- Chernushevich, I. V., Loboda, A. V., and Thomson, B. A. (2001). An introduction to quadrupole-time-of-flight mass spectrometry. *J Mass Spectrom* **36**, 849-65.
- Chervet, J. P., Ursem, M., and Salzman, J. P. (1996). Instrumental requirements for nanoscale liquid chromatography. *Anal Chem.* **68**, 1507-12. doi: 10.1021/ac9508964.
- Chevet, E., Smirle, J., Cameron, P. H., Thomas, D. Y., and Bergeron, J. J. (2010). Calnexin phosphorylation: linking cytoplasmic signalling to endoplasmic reticulum luminal functions. *Semin Cell Dev Biol.* **21**, 486-90. doi: 10.1016/j.semcd.2009.12.005. Epub 2010 Jan 6.
- Chien, K. Y., Liu, H. C., and Goshe, M. B. (2011). Development and application of a phosphoproteomic method using electrostatic repulsion-hydrophilic interaction chromatography (ERLIC), IMAC, and LC-MS/MS analysis to study Marek's Disease Virus infection. *J Proteome Res* **10**, 4041-53.

- Chignard, N., Shang, S., Wang, H., Marrero, J., Brechot, C., Hanash, S., and Beretta, L. (2006). Cleavage of endoplasmic reticulum proteins in hepatocellular carcinoma: Detection of generated fragments in patient sera. *Gastroenterology* **130**, 2010-22.
- Cihova, M., Altanerova, V., and Altaner, C. (2011). Stem cell based cancer gene therapy. *Mol Pharm.* **8**, 1480-7. doi: 10.1021/mp200151a. Epub 2011 Jul 25.
- Cleary, S. P., Jeck, W. R., Zhao, X., Kuichen, Selitsky, S. R., Savich, G. L., Tan, T. X., Wu, M. C., Getz, G., Lawrence, M. S., Parker, J. S., Li, J., Powers, S., Kim, H., Fischer, S., Giundi, M., Ghanekar, A., and Chiang, D. Y. (2013). Identification of driver genes in hepatocellular carcinoma by exome sequencing. *Hepatology*.
- Cohen, S. A., and Michaud, D. P. (1993). Synthesis of a fluorescent derivatizing reagent, 6-aminoquinolyl-N-hydroxysuccinimidyl carbamate, and its application for the analysis of hydrolysate amino acids via high-performance liquid chromatography. *Anal Biochem* **211**, 279-87.
- Coller, J. K., Fritz, P., Zanger, U. M., Siegle, I., Eichelbaum, M., Kroemer, H. K., and Mürdter, T. E. (2001). Distribution of microsomal epoxide hydrolase in humans: an immunohistochemical study in normal tissues, and benign and malignant tumours. *Histochem J.* **33**, 329-36.
- Cui, L., and Reid, G. E. (2013). Examining factors that influence erroneous phosphorylation site localization via competing fragmentation and rearrangement reactions during ion trap CID-MS/MS and -MS(3.). *Proteomics* **13**, 964-73.
- Cydzik, M., Rudowska, M., Stefanowicz, P., and Szewczuk, Z. (2011). The competition of charge remote and charge directed fragmentation mechanisms in quaternary ammonium salt derivatized peptides--an isotopic exchange study. *J Am Soc Mass Spectrom.* **22**, 2103-7. doi: 10.1007/s13361-011-0245-2. Epub 2011 Oct 5.
- Dai, J., Jin, W. H., Sheng, Q. H., Shieh, C. H., Wu, J. R., and Zeng, R. (2007). Protein phosphorylation and expression profiling by Yin-yang multidimensional liquid chromatography (Yin-yang MDLC) mass spectrometry. *J Proteome Res* **6**, 250-62.
- Dai, J., Wang, L. S., Wu, Y. B., Sheng, Q. H., Wu, J. R., Shieh, C. H., and Zeng, R. (2009). Fully automatic separation and identification of phosphopeptides by continuous pH-gradient anion exchange online coupled with reversed-phase liquid chromatography mass spectrometry. *J Proteome Res* **8**, 133-41.
- DeGnove, J. P., and Qin, J. (1998). Fragmentation of phosphopeptides in an ion trap mass spectrometer. *J Am Soc Mass Spectrom* **9**, 1175-88.
- Deng, Y. H., Li, R. J., Zhang, H. S., Du, X. L., and Wang, H. (2007). Liquid chromatographic analysis of phosphoamino acids at femtomole level using chemical derivatization with N-hydroxysuccinimidyl fluorescein-O-acetate. *Anal Chim Acta* **601**, 118-24.
- Ding, C., Jiang, J., Wei, J., Liu, W., Zhang, W., Liu, M., Fu, T., Lu, T., Song, L., Ying, W., Chang, C., Zhang, Y., Ma, J., Wei, L., Malovannaya, A., Jia, L., Zhen, B., Wang, Y., He, F., Qian, X., and Qin, J. (2013). A Fast Workflow for Identification and Quantification of Proteomes. *Mol Cell Proteomics* **13**, 13.
- Dole, M., Mack, L. L., Hines, R. L., Mobley, R. C., Ferguson, L. D., and Alice, M. B. (1968). Molecular Beams of Macroions. *J Chem Phys* **49**, 2240-2249.
- Donato, P., Cacciola, F., Sommella, E., Fanali, C., Dugo, L., Dachà, M., Campiglia, P., Novellino, E., Dugo, P., and Mondello, L. (2011). Online Comprehensive RPLC × RPLC with Mass Spectrometry Detection for the Analysis of Proteome Samples. *Analytical Chemistry* **83**, 2485-2491.

- Dongre, A. R., Somogyi, A., and Wysocki, V. H. (1996). Surface-induced dissociation: an effective tool to probe structure, energetics and fragmentation mechanisms of protonated peptides. *J Mass Spectrom* **31**, 339-50.
- Emadali, A., and Gallagher-Gambarelli, M. (2009). [Quantitative proteomics by SILAC: practicalities and perspectives for an evolving approach]. *Med Sci (Paris)*. **25**, 835-42. doi: 10.1051/medsci/20092510835.
- Eng, J. K., McCormack, A. L., and Yates Iii, J. R. (1994). An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. *Journal of the American Society for Mass Spectrometry* **5**, 976-989.
- Engholm-Keller, K., Birck, P., Stirling, J., Pociot, F., Mandrup-Poulsen, T., and Larsen, M. R. (2012). TiSH--a robust and sensitive global phosphoproteomics strategy employing a combination of TiO₂, SIMAC, and HILIC. *J Proteomics* **75**, 5749-61.
- Evans, C., Noirel, J., Ow, S. Y., Salim, M., Pereira-Medrano, A. G., Couto, N., Pandhal, J., Smith, D., Pham, T. K., Karunakaran, E., Zou, X., Biggs, C. A., and Wright, P. C. (2012). An insight into iTRAQ: where do we stand now? *Anal Bioanal Chem*. **404**, 1011-27. doi: 10.1007/s00216-012-5918-6. Epub 2012 Mar 27.
- Fan, Y., Guo, Y., Hamblin, M., Chang, L., Zhang, J., and Chen, Y. E. (2011). Inhibition of gluconeogenic genes by calcium-regulated heat-stable protein 1 via repression of peroxisome proliferator-activated receptor alpha. *J Biol Chem*. **286**, 40584-94. doi: 10.1074/jbc.M111.232918. Epub 2011 Oct 11.
- Fearon, E. R., and Vogelstein, B. (1990). A genetic model for colorectal tumorigenesis. *Cell*. **61**, 759-67.
- Fedirko, V., Tramacere, I., Bagnardi, V., Rota, M., Scotti, L., Islami, F., Negri, E., Straif, K., Romieu, I., La Vecchia, C., Boffetta, P., and Jenab, M. (2011). Alcohol drinking and colorectal cancer risk: an overall and dose-response meta-analysis of published studies. *Ann Oncol*. **22**, 1958-72. doi: 10.1093/annonc/mdq653. Epub 2011 Feb 9.
- Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F., and Whitehouse, C. M. (1989). Electrospray ionization for mass spectrometry of large biomolecules. *Science* **246**, 64-71.
- Fenselau, C., and Yao, X. (2009). 18O₂-labeling in quantitative proteomic strategies: a status report. *J Proteome Res*. **8**, 2140-3. doi: 10.1021/pr8009879.
- Ficarro, S. B., McClelland, M. L., Stukenberg, P. T., Burke, D. J., Ross, M. M., Shabanowitz, J., Hunt, D. F., and White, F. M. (2002). Phosphoproteome analysis by mass spectrometry and its application to *Saccharomyces cerevisiae*. *Nat Biotechnol* **20**, 301-5.
- Ficarro, S. B., Salomon, A. R., Brill, L. M., Mason, D. E., Stettler-Gill, M., Brock, A., and Peters, E. C. (2005). Automated immobilized metal affinity chromatography/nano-liquid chromatography/electrospray ionization mass spectrometry platform for profiling protein phosphorylation sites. *Rapid Commun Mass Spectrom* **19**, 57-71.
- Fiechter, G., Sivec, G., and Mayer, H. K. (2012). Application of UHPLC for the simultaneous analysis of free amino acids and biogenic amines in ripened acid-curd cheeses. *J Chromatogr B Analyt Technol Biomed Life Sci*.
- Fonslow, B. R., Niessen, S. M., Singh, M., Wong, C. C., Xu, T., Carvalho, P. C., Choi, J., Park, S. K., and Yates, J. R., 3rd (2012). Single-step inline hydroxyapatite enrichment facilitates identification and quantitation of phosphopeptides from mass-limited proteomes with MudPIT. *J Proteome Res*. **11**, 2697-709. doi: 10.1021/pr300200x. Epub 2012 Apr 17.
- Fountoulakis, M., and Lahm, H. W. (1998). Hydrolysis and amino acid composition of proteins. *J Chromatogr A* **826**, 109-34.

- Freund, D. M., and Prenni, J. E. (2013). Improved Detection of Quantitative Differences Using a Combination of Spectral Counting and MS/MS Total Ion Current. *J Proteome Res* **12**, 12.
- Friedman, S. L. (2008). Hepatic fibrosis -- overview. *Toxicology* **254**, 120-9.
- Fu, N. N., Wang, H., and Zhang, H. S. (2011). Rapid analysis of phosphoamino acids from phosphotyrosine by near-infrared cyanine 1-(epsilon-succinimidyloxyhexanoate)-1'-methyl-3,3',3'-tetramethylindocarbocyanine-5,5'-disulfonate potassium derivatization and polyacrylamide-coated CE with LIF detection. *Electrophoresis* **32**, 712-9.
- Gan, C. S., Guo, T., Zhang, H., Lim, S. K., and Sze, S. K. (2008). A comparative study of electrostatic repulsion-hydrophilic interaction chromatography (ERLIC) versus SCX-IMAC-based methods for phosphopeptide isolation/enrichment. *J Proteome Res* **7**, 4869-77.
- Ganapathy-Kanniappan, S., Kunjithapatham, R., and Geschwind, J. F. (2012). Glyceraldehyde-3-phosphate dehydrogenase: a promising target for molecular therapy in hepatocellular carcinoma. *Oncotarget*. **3**, 940-53.
- Gao, L., Uttamchandani, M., and Yao, S. Q. (2012). Comparative proteomic profiling of mammalian cell lysates using phosphopeptide microarrays. *Chem Commun (Camb)* **48**, 2240-2.
- Garcia, M. C. (2005). The effect of the mobile phase additives on sensitivity in the analysis of peptides and proteins by high-performance liquid chromatography-electrospray mass spectrometry. *J Chromatogr B Analyt Technol Biomed Life Sci*.
- Gazzana, G., and Borlak, J. (2008). Mapping of the serum proteome of hepatocellular carcinoma induced by targeted overexpression of epidermal growth factor to liver cells of transgenic mice. *J Proteome Res* **7**, 928-37.
- Geiger, T., Wisniewski, J. R., Cox, J., Zanivan, S., Kruger, M., Ishihama, Y., and Mann, M. (2011). Use of stable isotope labeling by amino acids in cell culture as a spike-in standard in quantitative proteomics. *Nat Protoc*. **6**, 147-57. doi: 10.1038/nprot.2010.192.
- Gillet, L. C., Navarro, P., Tate, S., Rost, H., Selevsek, N., Reiter, L., Bonner, R., and Aebersold, R. (2012). Targeted data extraction of the MS/MS spectra generated by data-independent acquisition: a new concept for consistent and accurate proteome analysis. *Mol Cell Proteomics* **11**, O111 016717.
- Giorgio Righetti, P. (1984). Isoelectric focusing in immobilized pH gradients. *Journal of Chromatography A* **300**, 165-224.
- Godugu, B., Neta, P., Simon-Manso, Y., and Stein, S. E. (2010). Effect of N-terminal glutamic acid and glutamine on fragmentation of peptide ions. *J Am Soc Mass Spectrom* **21**, 1169-76.
- Gonzalez-Gonzalez, M., Garcia, J. G., Montero, J. A., Fernandez, L. M., Bengoechea, O., Munez, O. B., Orfao, A., Sayagues, J. M., and Fuentes, M. (2013). Genomics and proteomics approaches for biomarker discovery in sporadic colorectal cancer with metastasis. *Cancer Genomics Proteomics*. **10**, 19-25.
- Goshe, M. B., Veenstra, T. D., Panisko, E. A., Conrads, T. P., Angell, N. H., and Smith, R. D. (2002). Phosphoprotein isotope-coded affinity tags: application to the enrichment and identification of low-abundance phosphoproteins. *Anal Chem* **74**, 607-16.
- Gronborg, M., Kristiansen, T. Z., Stensballe, A., Andersen, J. S., Ohara, O., Mann, M., Jensen, O. N., and Pandey, A. (2002). A mass spectrometry-based proteomic approach for identification of serine/threonine-phosphorylated proteins by enrichment with phospho-specific antibodies: identification of a novel protein, Frigg, as a protein kinase A substrate. *Mol Cell Proteomics* **1**, 517-27.
- Grumbach, I. M., and Veh, R. W. (1991). Sulpho-N-hydroxysuccinimide activated long chain biotin. A new microtitre plate assay for the determination of its stability at different pH values and its reaction rate with protein bound amino groups. *J Immunol Methods*. **140**, 205-10.

- Gu, C., Tsapraillis, G., Brechi, L., and Wysocki, V. H. (2000). Selective gas-phase cleavage at the peptide bond C-terminal to aspartic acid in fixed-charge derivatives of Asp-containing peptides. *Anal Chem.* **72**, 5804-13.
- Guichard, C., Amaddeo, G., Imbeaud, S., Ladeiro, Y., Pelletier, L., Maad, I. B., Calderaro, J., Bioulac-Sage, P., Letexier, M., Degos, F., Clement, B., Balabaud, C., Chevet, E., Laurent, A., Couchy, G., Letouze, E., Calvo, F., and Zucman-Rossi, J. (2012). Integrated analysis of somatic mutations and focal copy-number changes identifies key genes and pathways in hepatocellular carcinoma. *Nat Genet* **44**, 694-8.
- Guo, Y., Zhao, J., Bi, J., Wu, Q., Wang, X., and Lai, Q. (2012). Heterogeneous nuclear ribonucleoprotein K (hnRNP K) is a tissue biomarker for detection of early hepatocellular carcinoma in patients with cirrhosis. *J Hematol Oncol* **5**, 37.
- Guzel, Y., Rainer, M., Mirza, M. R., Messner, C. B., and Bonn, G. K. (2013). Highly selective recovery of phosphopeptides using trypsin-assisted digestion of precipitated lanthanide-phosphoprotein complexes. *Analyst.* **138**, 2897-905. doi: 10.1039/c3an00066d.
- Haaf, E., and Schlosser, A. (2012). Peptide and protein quantitation by acid-catalyzed 18O-labeling of carboxyl groups. *Anal Chem.* **84**, 304-11. doi: 10.1021/ac202561m. Epub 2011 Dec 2.
- Han, D., Moon, S., Kim, Y., Ho, W. K., Kim, K., Kang, Y., Jun, H., and Kim, Y. (2012). Comprehensive phosphoproteome analysis of INS-1 pancreatic beta-cells using various digestion strategies coupled with liquid chromatography-tandem mass spectrometry. *J Proteome Res.* **11**, 2206-23. doi: 10.1021/pr200990b. Epub 2012 Mar 14.
- Han, G., Ye, M., Liu, H., Song, C., Sun, D., Wu, Y., Jiang, X., Chen, R., Wang, C., Wang, L., and Zou, H. (2010). Phosphoproteome analysis of human liver tissue by long-gradient nanoflow LC coupled with multiple stage MS analysis. *Electrophoresis.* **31**, 1080-9. doi: 10.1002/elps.200900493.
- Han, G., Ye, M., Zhou, H., Jiang, X., Feng, S., Jiang, X., Tian, R., Wan, D., Zou, H., and Gu, J. (2008). Large-scale phosphoproteome analysis of human liver tissue by enrichment and fractionation of phosphopeptides with strong anion exchange chromatography. *Proteomics* **8**, 1346-61.
- Harrison, A. G. (1997). The gas-phase basicities and proton affinities of amino acids and peptides. *Mass Spectrom Rev* **16**, 201-217.
- Harrison, A. G. (1999). Linear free energy correlations in mass spectrometry. *J Mass Spectrom* **34**, 577-589.
- Harrison, A. G. (2009a). Charge-separation reactions of doubly-protonated peptides: effect of peptide chain length. *J Am Soc Mass Spectrom.* **20**, 1890-5.
- Harrison, A. G. (2009b). To b or not to b: the ongoing saga of peptide b ions. *Mass Spectrom Rev* **28**, 640-54.
- Harrison, A. G., Csizmadia, I. G., Tang, T. H., and Tu, Y. P. (2000). Reaction competition in the fragmentation of protonated dipeptides. *J Mass Spectrom* **35**, 683-8.
- Heggie, G. D., Sommadossi, J. P., Cross, D. S., Huster, W. J., and Diasio, R. B. (1987). Clinical pharmacokinetics of 5-fluorouracil and its metabolites in plasma, urine, and bile. *Cancer Res.* **47**, 2203-6.
- Heller, M., Mattou, H., Menzel, C., and Yao, X. (2003). Trypsin catalyzed 16O-to-18O exchange for comparative proteomics: tandem mass spectrometry comparison using MALDI-TOF, ESI-QTOF, and ESI-ion trap mass spectrometers. *J Am Soc Mass Spectrom.* **14**, 704-18.
- Henrich, M. L., Marino, F., Groenewold, V., Kops, G. J., Mohammed, S., and Heck, A. J. (2013). Universal Quantitative Kinase Assay Based on Diagonal SCX Chromatography and Stable

- Isotope Dimethyl Labeling Provides High-definition Kinase Consensus Motifs for PKA and Human Mps1. *J Proteome Res.* **12**, 2214-24. doi: 10.1021/pr400074f. Epub 2013 Apr 3.
- Herencia, C., Martinez-Moreno, J. M., Herrera, C., Corrales, F., Santiago-Mora, R., Espejo, I., Barco, M., Almaden, Y., de la Mata, M., Rodriguez-Ariza, A., and Munoz-Castaneda, J. R. (2012). Nuclear translocation of beta-catenin during mesenchymal stem cells differentiation into hepatocytes is associated with a tumoral phenotype. *PLoS One* **7**, e34656.
- Ho, H. Y., Cheng, M. L., Shiao, M. S., and Chiu, D. T. (2013). Characterization of global metabolic responses of glucose-6-phosphate dehydrogenase-deficient hepatoma cells to diamide-induced oxidative stress. *Free Radic Biol Med.* **54:71-84.**, 10.1016/j.freeradbiomed.2012.10.557. Epub 2012 Nov 6.
- Horie K Fau - Sato, Y., Sato Y Fau - Kimura, T., Kimura T Fau - Nakamura, T., Nakamura T Fau - Ishihama, Y., Ishihama Y Fau - Oda, Y., Oda Y Fau - Ikegami, T., Ikegami T Fau - Tanaka, N., and Tanaka, N. (2012). Estimation and optimization of the peak capacity of one-dimensional gradient high performance liquid chromatography using a long monolithic silica capillary column. *J Chromatogr A.*
- Hornbeck, P. V., Kornhauser, J. M., Tkachev, S., Zhang, B., Skrzypek, E., Murray, B., Latham, V., and Sullivan, M. (2012). PhosphoSitePlus: a comprehensive resource for investigating the structure and function of experimentally determined post-translational modifications in man and mouse. *Nucleic Acids Res* **40**, D261-70.
- Hunter, T., and Sefton, B. M. (1980). Transforming gene product of Rous sarcoma virus phosphorylates tyrosine. *Proc Natl Acad Sci U S A* **77**, 1311-5.
- Huttlin, E. L., Jedrychowski, M. P., Elias, J. E., Goswami, T., Rad, R., Beausoleil, S. A., Villen, J., Haas, W., Sowa, M. E., and Gygi, S. P. (2010). A tissue-specific atlas of mouse protein phosphorylation and expression. *Cell* **143**, 1174-89.
- IARC (2008). Cancer Mortality Database, cancer map, <http://globocan.iarc.fr/>.
- InVS, and INCa (2013). Estimation nationale de l'incidence et de la mortalité par cancer en France entre 1980 et 2012. Partie1 - tumeurs solides. Synthèse.
- Iribarne, J., and Thomson, B. A. (1976). On the evaporation of small ions from charged droplets. *J Chem Phys* **64**, 2287-2294.
- Jaffe, H., Veeranna, and Pant, H. C. (1998). Characterization of serine and threonine phosphorylation sites in beta-elimination/ethanethiol addition-modified proteins by electrospray tandem mass spectrometry and database searching. *Biochemistry* **37**, 16211-24.
- Jensen, S. S., and Larsen, M. R. (2007). Evaluation of the impact of some experimental procedures on different phosphopeptide enrichment techniques. *Rapid Commun Mass Spectrom* **21**, 3635-45.
- Kanamori, H., Kawakami, T., Effendi, K., Yamazaki, K., Mori, T., Ebinuma, H., Masugi, Y., Du, W., Nagasaka, K., Ogiwara, A., Kyono, Y., Tanabe, M., Saito, H., Hibi, T., and Sakamoto, M. (2011). Identification by differential tissue proteome analysis of talin-1 as a novel molecular marker of progression of hepatocellular carcinoma. *Oncology* **80**, 406-15.
- Karas, M., and Hillenkamp, F. (1988). Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Anal Chem* **60**, 2299-2301.
- Key, T. J. (2011). Fruit and vegetables and cancer risk. *Br J Cancer.* **104**, 6-11. doi: 10.1038/sj.bjc.6606032. Epub 2010 Nov 30.
- Khalil, A. A., Kabapy, N. F., Deraz, S. F., and Smith, C. (2011). Heat shock proteins in oncology: diagnostic biomarkers or therapeutic targets? *Biochim Biophys Acta.* **1816**, 89-104. doi: 10.1016/j.bbcan.2011.05.001. Epub 2011 May 14.

- Kim, J. S., Kim, J., Oh, J. M., and Kim, H. J. (2011). Tandem mass spectrometric method for definitive localization of phosphorylation sites using bromine signature. *Anal Biochem* **414**, 294-6.
- Kocher T Fau - Swart, R., Swart R Fau - Mechtler, K., and Mechtler, K. (2011). Ultra-high-pressure RPLC hyphenated to an LTQ-Orbitrap Velos reveals a linear relation between peak capacity and number of identified peptides. *Anal Chem*.
- Konzer, A., Ruhs, A., Braun, T., and Kruger, M. (2013). Global protein quantification of mouse heart tissue based on the SILAC mouse. *Methods Mol Biol* **1005:39-52.**, 10.1007/978-1-62703-386-2_4.
- Kornblau, S. M., and Coombes, K. R. (2011). Use of reverse phase protein microarrays to study protein expression in leukemia: technical and methodological lessons learned. *Methods Mol Biol* **785**, 141-55.
- Kovanich, D., Cappadona, S., Raijmakers, R., Mohammed, S., Scholten, A., and Heck, A. J. (2012). Applications of stable isotope dimethyl labeling in quantitative proteomics. *Anal Bioanal Chem*. **404**, 991-1009. doi: 10.1007/s00216-012-6070-z. Epub 2012 May 27.
- Krishnamoorthy, S., Liu, Z., Hong, A., Zhu, R., Chen, H., Li, T., Zhou, X., and Gao, X. (2013). A Novel Phosphopeptide Microarray Based Interactome Map in Breast Cancer Cells Reveals Phosphoprotein-GRB2 Cell Signaling Networks. *PLoS One* **8**, e67634.
- Kucerova, L., Matuskova, M., Pastorakova, A., Tyciakova, S., Jakubikova, J., Bohovic, R., Altanerova, V., and Altaner, C. (2008). Cytosine deaminase expressing human mesenchymal stem cells mediated tumour regression in melanoma bearing mice. *J Gene Med*. **10**, 1071-82. doi: 10.1002/jgm.1239.
- Kumar, C., and Mann, M. (2009). Bioinformatics analysis of mass spectrometry-based proteomics data sets. *FEBS Lett*. **583**, 1703-12. doi: 10.1016/j.febslet.2009.03.035. Epub 2009 Mar 21.
- Kuramitsu, Y., and Nakamura, K. (2006). Proteomic analysis of cancer tissues: shedding light on carcinogenesis and possible biomarkers. *Proteomics* **6**, 5650-61.
- Kuroda, I., Shintani, Y., Motokawa, M., Abe, S., and Furuno, M. (2004). Phosphopeptide-selective column-switching RP-HPLC with a titania precolumn. *Anal Sci*. **20**, 1313-9.
- Kurosky, A., Miller, B. T., and Knock, S. L. (1993). Kinetic analysis of biotinylation of specific residues of peptides by high-performance liquid chromatography. *J Chromatogr*. **631**, 281-7.
- Kuyama, H., Toda, C., Watanabe, M., Tanaka, K., and Nishimura, O. (2003). An efficient chemical method for dephosphorylation of phosphopeptides. *Rapid Commun Mass Spectrom* **17**, 1493-6.
- Kuzyk, M. A., Ohlund, L. B., Elliott, M. H., Smith, D., Qian, H., Delaney, A., Hunter, C. L., and Borchers, C. H. (2009). A comparison of MS/MS-based, stable-isotope-labeled, quantitation performance on ESI-quadrupole TOF and MALDI-TOF/TOF mass spectrometers. *Proteomics* **9**, 3328-3340.
- Kyono, Y., Sugiyama, N., Imami, K., Tomita, M., and Ishihama, Y. (2008). Successive and selective release of phosphorylated peptides captured by hydroxy acid-modified metal oxide chromatography. *J Proteome Res* **7**, 4585-93.
- Lai, K. K., Kolippakkam, D., and Beretta, L. (2008). Comprehensive and quantitative proteome profiling of the mouse liver and plasma. *Hepatology* **47**, 1043-51.
- Lan, K., and Jorgenson, J. W. (1999). Automated measurement of peak widths for the determination of peak capacity in complex chromatograms. *Anal Chem*. **71**, 709-14.
- Larsen, M. R., Thingholm, T. E., Jensen, O. N., Roepstorff, P., and Jorgensen, T. J. (2005). Highly selective enrichment of phosphorylated peptides from peptide mixtures using titanium dioxide microcolumns. *Mol Cell Proteomics* **4**, 873-86.

- Laurent-Puig, P., Lievre, A., Ducreux, M., and Lorient, M. A. (2008). [The biological point of view on pharmacogenetics of anticancer agents in colorectal cancer]. *Bull Cancer*. **95**, 935-42. doi: 10.1684/bdc.2008.0728.
- Le Naour, F., Brichory, F., Misek, D. E., Brechot, C., Hanash, S. M., and Beretta, L. (2002). A distinct repertoire of autoantibodies in hepatocellular carcinoma identified by proteomic analysis. *Mol Cell Proteomics* **1**, 197-203.
- Lecomte, T., Laurent-Puig, P., and Lorient, M. A. (2006). Pharmacogénétique en hépatogastroentérologie. *Hépatogastro* **13**.
- Lee, H. J., Na, K., Kwon, M. S., Kim, H., Kim, K. S., and Paik, Y. K. (2009). Quantitative analysis of phosphopeptides in search of the disease biomarker from the hepatocellular carcinoma specimen. *Proteomics* **9**, 3395-408.
- Lee, J., Xu, Y., Chen, Y., Sprung, R., Kim, S. C., Xie, S., and Zhao, Y. (2007a). Mitochondrial phosphoproteome revealed by an improved IMAC method and MS/MS/MS. *Mol Cell Proteomics*. **6**, 669-76. Epub 2007 Jan 5.
- Lee, J., Xu, Y., Chen, Y., Sprung, R., Kim, S. C., Xie, S., and Zhao, Y. (2007b). Mitochondrial phosphoproteome revealed by an improved IMAC method and MS/MS/MS. *Mol Cell Proteomics* **6**, 669-76.
- Lee, K. A., Farnsworth, C., Yu, W., and Bonilla, L. E. (2011). 24-hour lock mass protection. *J Proteome Res*. **10**, 880-5. doi: 10.1021/pr100780b. Epub 2010 Dec 27.
- Li, L., Chen, S. H., Yu, C. H., Li, Y. M., and Wang, S. Q. (2008). Identification of hepatocellular carcinoma-associated antigens and autoantibodies by serological proteome analysis combined with protein microarray. *J Proteome Res* **7**, 611-20.
- Li, S., and Mao, M. (2012). Next generation sequencing reveals genetic landscape of hepatocellular carcinomas. *Cancer Lett* **12**, 00591-5.
- Li, W., Backlund, P. S., Boykins, R. A., Wang, G., and Chen, H. C. (2003). Susceptibility of the hydroxyl groups in serine and threonine to beta-elimination/Michael addition under commonly used moderately high-temperature conditions. *Anal Biochem* **323**, 94-102.
- Li, W., Boykins, R. A., Backlund, P. S., Wang, G., and Chen, H. C. (2002). Identification of phosphoserine and phosphothreonine as cysteic acid and beta-methylcysteic acid residues in peptides by tandem mass spectrometric sequencing. *Anal Chem* **74**, 5701-10.
- Li, Z., Adams, R. M., Chourey, K., Hurst, G. B., Hettich, R. L., and Pan, C. (2012). Systematic comparison of label-free, metabolic labeling, and isobaric chemical labeling for quantitative proteomics on LTQ Orbitrap Velos. *J Proteome Res*. **11**, 1582-90. doi: 10.1021/pr200748h. Epub 2012 Feb 16.
- Liberski, A. R., Al-Noubi, M. N., Rahman, Z. H., Halabi, N. M., Dib, S. S., Al-Mismar, R., Billing, A. M., Krishnankutty, R., Ahmad, F. S., Raynaud, C., Tabrizi, J. A., Engholm-Keller, K., and Graumann, J. (2013). Adaptation of a commonly used, chemically defined medium for human embryonic stem cells to stable isotope labeling with amino acids in cell culture (SILAC). *J Proteome Res* **5**, 5.
- Lim, H., Eng, J., Yates, J. R., 3rd, Tollaksen, S. L., Giometti, C. S., Holden, J. F., Adams, M. W., Reich, C. I., Olsen, G. J., and Hays, L. G. (2003). Identification of 2D-gel proteins: a comparison of MALDI/TOF peptide mass mapping to mu LC-ESI tandem mass spectrometry. *J Am Soc Mass Spectrom*. **14**, 957-70.
- Lim, S. O., Park, S. J., Kim, W., Park, S. G., Kim, H. J., Kim, Y. I., Sohn, T. S., Noh, J. H., and Jung, G. (2002). Proteome analysis of hepatocellular carcinoma. *Biochem Biophys Res Commun*. **291**, 1031-7.

- Link, A. J., Eng, J., Schieltz, D. M., Carmack, E., Mize, G. J., Morris, D. R., Garvik, B. M., and Yates, J. R., 3rd (1999). Direct analysis of protein complexes using mass spectrometry. *Nat Biotechnol* **17**, 676-82.
- Liu, H., Finch, J. W., Lavalley, M. J., Collamati, R. A., Benevides, C. C., and Gebler, J. C. (2007). Effects of column length, particle size, gradient length and flow rate on peak capacity of nano-scale liquid chromatography for peptide separations. *J Chromatogr A* **1147**, 30-6. Epub 2007 Feb 11.
- Liu, Z., Ma, Y., Yang, J., and Qin, H. (2011). Upregulated and downregulated proteins in hepatocellular carcinoma: a systematic review of proteomic profiling studies. *OMICS*. **15**, 61-71. doi: 10.1089/omi.2010.0061. Epub 2010 Aug 20.
- Longley, D. B., Harkin, D. P., and Johnston, P. G. (2003). 5-fluorouracil: mechanisms of action and clinical strategies. *Nat Rev Cancer*. **3**, 330-8.
- Louwagie, M., Kieffer-Jaquinod, S., Dupierris, V., Coute, Y., Bruley, C., Garin, J., Dupuis, A., Jaquinod, M., and Brun, V. (2012). Introducing AAA-MS, a rapid and sensitive method for amino acid analysis using isotope dilution and high-resolution mass spectrometry. *J Proteome Res* **11**, 3929-36.
- Lundby, A., Secher, A., Lage, K., Nordsborg, N. B., Dmytriiev, A., Lundby, C., and Olsen, J. V. (2012). Quantitative maps of protein phosphorylation sites across 14 different rat organs and tissues. *Nat Commun*. **3:876**, 10.1038/ncomms1871.
- Macek, B., Mann, M., and Olsen, J. V. (2009). Global and site-specific quantitative phosphoproteomics: principles and applications. *Annu Rev Pharmacol Toxicol* **49**, 199-221.
- Makarov, A. (2000). Electrostatic axially harmonic orbital trapping: a high-performance technique of mass analysis. *Anal Chem*. **72**, 1156-62.
- Makarov, A., Denisov, E., Kholomeev, A., Balschun, W., Lange, O., Strupat, K., and Horning, S. (2006). Performance evaluation of a hybrid linear ion trap/orbitrap mass spectrometer. *Anal Chem*. **78**, 2113-20.
- Malencik, D. A., Zhao, Z. Z., and Anderson, S. R. (1990). Determination of dityrosine, phosphotyrosine, phosphothreonine, and phosphoserine by high-performance liquid chromatography. *Anal Biochem* **184**, 353-9.
- Malinowsky, K., Wolff, C., Ergin, B., Berg, D., and Becker, K. F. (2010). Deciphering signaling pathways in clinical tissues for personalized medicine using protein microarrays. *J Cell Physiol* **225**, 364-70.
- Malmer, M. F., and Schroeder, L. A. (1990). Amino acid analysis by high-performance liquid chromatography with methanesulfonic acid hydrolysis and 9-fluorenylmethylchloroformate derivatization. *J Chromatogr* **514**, 227-39.
- Mann, M., Ong, S. E., Gronborg, M., Steen, H., Jensen, O. N., and Pandey, A. (2002). Analysis of protein phosphorylation using mass spectrometry: deciphering the phosphoproteome. *Trends Biotechnol* **20**, 261-8.
- Manning, G., Whyte, D. B., Martinez, R., Hunter, T., and Sudarsanam, S. (2002). The protein kinase complement of the human genome. *Science* **298**, 1912-34.
- Martin, J., and Dufour, J. F. (2008). Tumor suppressor and hepatocellular carcinoma. *World J Gastroenterol* **14**, 1720-33.
- Mazanek, M., Roitinger, E., Hudecz, O., Hutchins, J. R., Hegemann, B., Mitulovic, G., Taus, T., Stingl, C., Peters, J. M., and Mechtler, K. (2010). A new acid mix enhances phosphopeptide enrichment on titanium- and zirconium dioxide for mapping of phosphorylation sites on protein complexes. *J Chromatogr B Analyt Technol Biomed Life Sci* **878**, 515-24.

- McAlister, G. C., Huttlin, E. L., Haas, W., Ting, L., Jedrychowski, M. P., Rogers, J. C., Kuhn, K., Pike, I., Grothe, R. A., Blethrow, J. D., and Gygi, S. P. (2012). Increasing the multiplexing capacity of TMTs using reporter ion isotopologues with isobaric masses. *Anal Chem.* **84**, 7469-78. doi: 10.1021/ac301572t. Epub 2012 Aug 20.
- McIlwain, S., Mathews, M., Bereman, M. S., Rubel, E. W., MacCoss, M. J., and Noble, W. S. (2012). Estimating relative abundances of proteins from shotgun proteomics data. *BMC Bioinformatics.* **13:308.**, 10.1186/1471-2105-13-308.
- McLachlin, D. T., and Chait, B. T. (2003). Improved beta-elimination-based affinity purification strategy for enrichment of phosphopeptides. *Anal Chem* **75**, 6826-36.
- Medina, J. C., Wu, N., and Lee, M. L. (2001). Comparison of empirical peak capacities for high-efficiency capillary chromatographic techniques. *Anal Chem.* **73**, 1301-6.
- Medzihradzky, K. F., and Trinidad, J. C. (2012). Unusual fragmentation of Pro-Ser/Thr-containing peptides detected in collision-induced dissociation spectra. *J Am Soc Mass Spectrom* **23**, 602-7.
- Megger, D. A., Pott, L. L., Ahrens, M., Padden, J., Bracht, T., Kuhlmann, K., Eisenacher, M., Meyer, H. E., and Sitek, B. (2013). Comparison of label-free and label-based strategies for proteome analysis of hepatoma cell lines. *Biochim Biophys Acta* **14**, 00289-6.
- Mertins, P., Udeshi, N. D., Clauser, K. R., Mani, D. R., Patel, J., Ong, S. E., Jaffe, J. D., and Carr, S. A. (2012). iTRAQ labeling is superior to mTRAQ for quantitative global proteomics and phosphoproteomics. *Mol Cell Proteomics.* **11**, M111.014423. doi: 10.1074/mcp.M111.014423. Epub 2011 Dec 30.
- Meyer, H. E., Hoffmann-Posorske, E., Korte, H., and Heilmeyer, L. M., Jr. (1986). Sequence analysis of phosphoserine-containing peptides. Modification for picomolar sensitivity. *FEBS Lett* **204**, 61-6.
- Minagawa, H., Honda, M., Miyazaki, K., Tabuse, Y., Teramoto, R., Yamashita, T., Nishino, R., Takatori, H., Ueda, T., Kamijo, K., and Kaneko, S. (2008). Comparative proteomic and transcriptomic profiling of the human hepatocellular carcinoma. *Biochem Biophys Res Commun* **366**, 186-92.
- Molle, D., Jardin, J., Piot, M., Pasco, M., Leonil, J., and Gagnaire, V. (2009). Comparison of electrospray and matrix-assisted laser desorption ionization on the same hybrid quadrupole time-of-flight tandem mass spectrometer: application to bidimensional liquid chromatography of proteins from bovine milk fraction. *J Chromatogr A* **1216**, 2424-32. doi: 10.1016/j.chroma.2009.01.017. Epub 2009 Jan 15.
- Molnar-Perl, I. (2001). Derivatization and chromatographic behavior of the o-phthalaldehyde amino acid derivatives obtained with various SH-group-containing additives. *J Chromatogr A* **913**, 283-302.
- Molnar-Perl, I. (2011). Advancement in the derivatizations of the amino groups with the o-phthalaldehyde-thiol and with the 9-fluorenylmethyloxycarbonyl chloride reagents. *J Chromatogr B Analyt Technol Biomed Life Sci* **879**, 1241-69.
- Moore, S., and Stein, W. H. (1954). A modified ninhydrin reagent for the photometric determination of amino acids and related compounds. *J Biol Chem* **211**, 907-13.
- Morrice, N., and Aitken, A. (1985). A simple and rapid method of quantitative analysis of phosphoamino acids by high-performance liquid chromatography. *Anal Biochem* **148**, 207-12.
- Moser, K., and White, F. M. (2006). Phosphoproteomic analysis of rat liver by high capacity IMAC and LC-MS/MS. *J Proteome Res* **5**, 98-104.

- Murthy, L. R., and Iqbal, K. (1991). Measurement of picomoles of phosphoamino acids by high-performance liquid chromatography. *Anal Biochem* **193**, 299-305.
- Mustafa, G. M., Petersen, J. R., Ju, H., Cicalese, L., Snyder, N., Haidacher, S. J., Denner, L., and Elferink, C. (2013). Biomarker Discovery for Early Detection of Hepatocellular Carcinoma (HCC) in Hepatitis C (HCV) Infected Patients. *Mol Cell Proteomics*.
- Na, K., Jeong, S. K., Lee, M. J., Cho, S. Y., Kim, S. A., Lee, M. J., Song, S. Y., Kim, H., Kim, K. S., Lee, H. W., and Paik, Y. K. (2013). Human liver carboxylesterase 1 outperforms alpha-fetoprotein as biomarker to discriminate hepatocellular carcinoma from other liver diseases in Korean patients. *Int J Cancer*. **133**, 408-15. doi: 10.1002/ijc.28020. Epub 2013 Feb 13.
- Nagaraj N Fau - Kulak, N. A., Kulak Na Fau - Cox, J., Cox J Fau - Neuhauser, N., Neuhauser N Fau - Mayr, K., Mayr K Fau - Hoerning, O., Hoerning O Fau - Vorm, O., Vorm O Fau - Mann, M., and Mann, M. (2012). System-wide perturbation analysis with nearly complete coverage of the yeast proteome by single-shot ultra HPLC runs on a bench top Orbitrap. *Mol Cell Proteomics*.
- Nahnsen, S., Bielow, C., Reinert, K., and Kohlbacher, O. (2013). Tools for label-free peptide quantification. *Mol Cell Proteomics*. **12**, 549-56. doi: 10.1074/mcp.R112.025163. Epub 2012 Dec 17.
- Nawrocki, J., Dunlap, C., McCormick, A., and Carr, P. W. (2004). Part I. Chromatography using ultra-stable metal oxide-based stationary phases for HPLC. *J Chromatogr A* **1028**, 1-30.
- Nawrocki, J., Rigney, M. P., McCormick, A., and Carr, P. W. (1993). Chemistry of zirconia and its use in chromatography. *J Chromatogr A*. **657**, 229-82.
- Ndassa, Y. M., Orsi, C., Marto, J. A., Chen, S., and Ross, M. M. (2006). Improved immobilized metal affinity chromatography for large-scale phosphoproteomics applications. *J Proteome Res* **5**, 2789-99.
- Neilson, K. A., Ali, N. A., Muralidharan, S., Mirzaei, M., Mariani, M., Assadourian, G., Lee, A., van Sluyter, S. C., and Haynes, P. A. (2011). Less label, more free: approaches in label-free quantitative mass spectrometry. *Proteomics*. **11**, 535-53. doi: 10.1002/pmic.201000553. Epub 2011 Jan 17.
- Nguyen, S., and Fenn, J. B. (2007). Gas-phase ions of solute species from charged droplets of solutions. *Proc Natl Acad Sci U S A* **104**, 1111-7.
- Nie, S., Dai, J., Ning, Z. B., Cao, X. J., Sheng, Q. H., and Zeng, R. (2012). Comprehensive profiling of phosphopeptides based on anion exchange followed by flow-through enrichment with titanium dioxide (AFET). *J Proteome Res* **9**, 4585-94.
- Niedbalski, J. S., and Ringer, D. P. (1986). Separation and quantitative analysis of O-linked phosphoamino acids by isocratic high-performance liquid chromatography of the 9-fluorenylmethyl chloroformate derivatives. *Anal Biochem* **158**, 138-45.
- Nordenstedt, H., White, D. L., and El-Serag, H. B. (2010). The changing pattern of epidemiology in hepatocellular carcinoma. *Dig Liver Dis* **42 Suppl 3**, S206-14.
- Novakova, L., Vaast, A., Stassen, C., Broeckhoven, K., De Pra, M., Swart, R., Desmet, G., and Eeltink, S. (2013). High-resolution peptide separations using nano-LC at ultra-high pressure. *J Sep Sci* **36**, 1192-9. doi: 10.1002/jssc.201201087. Epub 2013 Mar 4.
- Nuhse, T. S., Stensballe, A., Jensen, O. N., and Peck, S. C. (2003). Large-scale analysis of in vivo phosphorylated membrane proteins by immobilized metal ion affinity chromatography and mass spectrometry. *Mol Cell Proteomics*. **2**, 1234-43. Epub 2003 Sep 22.
- O'Farrell, P. H. (1975). High resolution two-dimensional electrophoresis of proteins. *J Biol Chem*. **250**, 4007-21.
- Oda, Y., Nagasu, T., and Chait, B. T. (2001). Enrichment analysis of phosphorylated proteins as a tool for probing the phosphoproteome. *Nat Biotechnol* **19**, 379-82.

- Olsen, J. V., Blagoev, B., Gnad, F., Macek, B., Kumar, C., Mortensen, P., and Mann, M. (2006). Global, in vivo, and site-specific phosphorylation dynamics in signaling networks. *Cell*. **127**, 635-48.
- Oosterkamp, A. J., Gelpi, E., and Abian, J. (1998). Quantitative peptide bioanalysis using column-switching nano liquid chromatography/mass spectrometry. *J Mass Spectrom* **33**, 976-83.
- Ow, S. Y., Salim, M., Noirel, J., Evans, C., Rehman, I., and Wright, P. C. (2009). iTRAQ underestimation in simple and complex mixtures: "the good, the bad and the ugly". *J Proteome Res*. **8**, 5347-55. doi: 10.1021/pr900634c.
- Paizs, B., and Suhai, S. (2004). Towards understanding the tandem mass spectra of protonated oligopeptides. 1: mechanism of amide bond cleavage. *J Am Soc Mass Spectrom* **15**, 103-13.
- Paizs, B., and Suhai, S. (2005). Fragmentation pathways of protonated peptides. *Mass Spectrom Rev* **24**, 508-48.
- Palumbo, A. M., and Reid, G. E. (2008). Evaluation of gas-phase rearrangement and competing fragmentation reactions on protein phosphorylation site assignment using collision induced dissociation-MS/MS and MS3. *Anal Chem* **80**, 9735-47.
- Pan, C., Gnad, F., Olsen, J. V., and Mann, M. (2008). Quantitative phosphoproteome analysis of a mouse liver cell line reveals specificity of phosphatase inhibitors. *Proteomics* **8**, 4534-46.
- Pappin, D. J., Hojrup, P., and Bleasby, A. J. (1993). Rapid identification of proteins by peptide-mass fingerprinting. *Curr Biol*. **3**, 327-32.
- Parsons, D. W., Jones, S., Zhang, X., Lin, J. C., Leary, R. J., Angenendt, P., Mankoo, P., Carter, H., Siu, I. M., Gallia, G. L., Olivi, A., McLendon, R., Rasheed, B. A., Keir, S., Nikolskaya, T., Nikolsky, Y., Busam, D. A., Tekleab, H., Diaz, L. A., Jr., Hartigan, J., Smith, D. R., Strausberg, R. L., Marie, S. K., Shinjo, S. M., Yan, H., Riggins, G. J., Bigner, D. D., Karchin, R., Papadopoulos, N., Parmigiani, G., Vogelstein, B., Velculescu, V. E., and Kinzler, K. W. (2008). An integrated genomic analysis of human glioblastoma multiforme. *Science*. **321**, 1807-12. doi: 10.1126/science.1164382. Epub 2008 Sep 4.
- Patel, V. J., Thalassinou, K., Slade, S. E., Connolly, J. B., Crombie, A., Murrell, J. C., and Scrivens, J. H. (2009). A comparison of labeling and label-free mass spectrometry-based proteomics approaches. *J Proteome Res*. **8**, 3752-9. doi: 10.1021/pr900080y.
- Paul, W. (1990). Electromagnetic traps for charged and neutral particles. *Reviews of Modern Physics* **62**, 531-540.
- Pflieger, D., Junger, M. A., Muller, M., Rinner, O., Lee, H., Gehrig, P. M., Gstaiger, M., and Aebersold, R. (2008). Quantitative proteomic analysis of protein complexes: concurrent identification of interactors and their state of phosphorylation. *Mol Cell Proteomics* **7**, 326-46.
- Phanstiel, D., Unwin, R., McAlister, G. C., and Coon, J. J. (2009). Peptide quantification using 8-plex isobaric tags and electron transfer dissociation tandem mass spectrometry. *Anal Chem*. **81**, 1693-8. doi: 10.1021/ac8019202.
- Phanstiel, D., Zhang, Y., Marto, J. A., and Coon, J. J. (2008). Peptide and protein quantification using iTRAQ with electron transfer dissociation. *J Am Soc Mass Spectrom*. **19**, 1255-62. doi: 10.1016/j.jasms.2008.05.023. Epub 2008 Jun 17.
- Pichler, P., Kocher, T., Holzmann, J., Mazanek, M., Taus, T., Ammerer, G., and Mechtler, K. (2010). Peptide labeling with isobaric tags yields higher identification rates using iTRAQ 4-plex compared to TMT 6-plex and iTRAQ 8-plex on LTQ Orbitrap. *Anal Chem* **82**, 6549-58.
- Pierce, A., Unwin, R. D., Evans, C. A., Griffiths, S., Carney, L., Zhang, L., Jaworska, E., Lee, C. F., Blinco, D., Okoniewski, M. J., Miller, C. J., Bitton, D. A., Spooncer, E., and Whetton, A. D. (2008). Eight-channel iTRAQ enables comparison of the activity of six leukemogenic tyrosine kinases. *Mol Cell Proteomics*. **7**, 853-63. Epub 2007 Oct 21.

- Piersma, S. R., Warmoes, M. O., de Wit, M., de Reus, I., Knol, J. C., and Jimenez, C. R. (2013). Whole gel processing procedure for GeLC-MS/MS based proteomics. *Proteome Sci.* **11**, 17. doi: 10.1186/1477-5956-11-17.
- Pinkse, M. W., Uitto, P. M., Hilhorst, M. J., Ooms, B., and Heck, A. J. (2004). Selective isolation at the femtomole level of phosphopeptides from proteolytic digests using 2D-NanoLC-ESI-MS/MS and titanium oxide precolumns. *Anal Chem* **76**, 3935-43.
- Polesel, J., Montella, M., Dal Maso, L., Crispo, A., Serraino, D., and Talamini, R. (2012). Re: hepatocellular carcinoma risk factors and disease burden in a european cohort: a nested case-control study. *J Natl Cancer Inst* **104**, 1681-3; author reply 1683-4.
- Poppe, H. (1997). Some reflections on speed and efficiency of modern chromatographic methods. *Journal of Chromatography A* **778**, 3-21.
- Posewitz, M. C., and Tempst, P. (1999). Immobilized gallium(III) affinity chromatography of phosphopeptides. *Anal Chem.* **71**, 2883-92.
- Pote, N., Alexandrov, T., Le Faouder, J., Laouirem, S., Leger, T., Mebarki, M., Belghiti, J., Camadro, J. M., Bedossa, P., and Paradis, V. (2013). Imaging mass spectrometry reveals modified forms of histone H4 as new biomarkers of microvascular invasion in hepatocellular carcinomas. *Hepatology* **58**, 983-94.
- Premstaller A Fau - Oberacher, H., Oberacher H Fau - Walcher, W., Walcher W Fau - Timperio, A. M., Timperio Am Fau - Zolla, L., Zolla L Fau - Chervet, J. P., Chervet Jp Fau - Cavusoglu, N., Cavusoglu N Fau - van Dorsselaer, A., van Dorsselaer A Fau - Huber, C. G., and Huber, C. G. (2001). High-performance liquid chromatography-electrospray ionization mass spectrometry using monolithic capillary columns for proteomic studies. *Anal Chem.*
- Pritchard, C. C., and Grady, W. M. (2011). Colorectal cancer molecular biology moves into clinical practice. *Gut.* **60**, 116-29. doi: 10.1136/gut.2009.206250. Epub 2010 Oct 4.
- Putz, S. M., Boehm, A. M., Stiewe, T., and Sickmann, A. (2012). iTRAQ analysis of a cell culture model for malignant transformation, including comparison with 2D-PAGE and SILAC. *J Proteome Res.* **11**, 2140-53. doi: 10.1021/pr200881c. Epub 2012 Mar 23.
- Qian, W. J., Goshe, M. B., Camp, D. G., 2nd, Yu, L. R., Tang, K., and Smith, R. D. (2003). Phosphoprotein isotope-coded solid-phase tag approach for enrichment and quantitative analysis of phosphopeptides from complex mixtures. *Anal Chem* **75**, 5441-50.
- Qu, J., Wang, Y., Luo, G., Wu, Z., and Yang, C. (2002). Validated quantitation of underivatized amino acids in human blood samples by volatile ion-pair reversed-phase liquid chromatography coupled to isotope dilution tandem mass spectrometry. *Anal Chem* **74**, 2034-40.
- Quidville, V., Alsafadi, S., Goubar, A., Commo, F., Scott, V., Pioche-Durieu, C., Girault, I., Baconnais, S., Le Cam, E., Lazar, V., Delalogue, S., Saghatchian, M., Pautier, P., Morice, P., Dessen, P., Vagner, S., and Andre, F. (2013). Targeting the deregulated spliceosome core machinery in cancer cells triggers mTOR blockade and autophagy. *Cancer Res.* **73**, 2247-58. doi: 10.1158/0008-5472.CAN-12-2501. Epub 2013 Jan 28.
- Rabilloud, T. (2009). Membrane proteins and proteomics: love is possible, but so difficult. *Electrophoresis.* **30**, S174-80. doi: 10.1002/elps.200900050.
- Reid, G. E., Simpson, R. J., and O'Hair, R. A. (2000). Leaving group and gas phase neighboring group effects in the side chain losses from protonated serine and its derivatives. *J Am Soc Mass Spectrom* **11**, 1047-60.
- Ringer, D. P. (1991). Separation of phosphotyrosine, phosphoserine, and phosphothreonine by high-performance liquid chromatography. *Methods Enzymol* **201**, 3-10.

- Ritorto, M. S., Cook, K., Tyagi, K., Pedrioli, P. G., and Trost, M. (2013). Hydrophilic Strong Anion Exchange (hSAX) Chromatography for Highly Orthogonal Peptide Separation of Complex Proteomes. *J Proteome Res* **7**, 7.
- Rodeck, U., Herlyn, M., Herlyn, D., Molthoff, C., Atkinson, B., Varello, M., Steplewski, Z., and Koprowski, H. (1987). Tumor growth modulation by a monoclonal antibody to the epidermal growth factor receptor: immunologically mediated and effector cell-independent effects. *Cancer Res.* **47**, 3692-6.
- Rodriguez-Suarez, E., and Whetton, A. D. (2013). The application of quantification techniques in proteomics for biomedical research. *Mass Spectrom Rev.* **32**, 1-26. doi: 10.1002/mas.21347. Epub 2012 Jul 30.
- Roepstorff, P., and Fohlman, J. (1984). Proposal for a common nomenclature for sequence ions in mass spectra of peptides. *Biomed Mass Spectrom* **11**, 601.
- Rogowska-Wrzesinska, A., Le Bihan, M. C., Thaysen-Andersen, M., and Roepstorff, P. (2013). 2D gels still have a niche in proteomics. *J Proteomics* **24**, 00038-9.
- Roskoski, R., Jr. (2012). ERK1/2 MAP kinases: structure, function, and regulation. *Pharmacol Res.* **66**, 105-43. doi: 10.1016/j.phrs.2012.04.005. Epub 2012 Apr 27.
- Ross, P. L., Huang, Y. N., Marchese, J. N., 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., and Pappin, D. J. (2004). Multiplexed protein quantitation in *Saccharomyces cerevisiae* using amine-reactive isobaric tagging reagents. *Mol Cell Proteomics.* **3**, 1154-69. Epub 2004 Sep 22.
- Rozman, M. (2011). Modelling of the gas-phase phosphate group loss and rearrangement in phosphorylated peptides. *J Mass Spectrom* **46**, 949-55.
- Rusnak, F., Zhou, J., and Hathaway, G. M. (2004). Reaction of phosphorylated and O-glycosylated peptides by chemically targeted identification at ambient temperature. *J Biomol Tech* **15**, 296-304.
- Sacco, F., Perfetto, L., Castagnoli, L., and Cesareni, G. (2012). The human phosphatase interactome: An intricate family portrait. *FEBS Lett* **586**, 2732-9.
- Saibil, H. (2013). Chaperone machines for protein folding, unfolding and disaggregation. *Nat Rev Cancer* **13**, 630-42.
- Sandra, K., Moshir, M., D'Hondt, F., Verleysen, K., Kas, K., and Sandra, P. (2008). Highly efficient peptide separations in proteomics Part 1. Unidimensional high performance liquid chromatography. *J Chromatogr B Analyt Technol Biomed Life Sci* **866**, 48-63. Epub 2007 Oct 30.
- Sano, A., and Nakamura, H. (2004). Chemo-affinity of titania for the column-switching HPLC analysis of phosphopeptides. *Anal Sci* **20**, 565-6.
- Sano, A., and Nakamura, H. (2007). Evaluation of titanium and titanium oxides as chemo-affinity sorbents for the selective enrichment of organic phosphates. *Anal Sci* **23**, 1285-9.
- Santoni, V., Molloy, M., and Rabilloud, T. (2000). Membrane proteins and proteomics: un amour impossible? *Electrophoresis.* **21**, 1054-70.
- Savitski, M. M., Falth, M., Fung, Y. M., Adams, C. M., and Zubarev, R. A. (2008). Bifurcating fragmentation behavior of gas-phase tryptic peptide dications in collisional activation. *J Am Soc Mass Spectrom.* **19**, 1755-63. doi: 10.1016/j.jasms.2008.08.003. Epub 2008 Aug 9.
- Savitski, M. M., Kjeldsen, F., Nielsen, M. L., and Zubarev, R. A. (2007). Relative specificities of water and ammonia losses from backbone fragments in collision-activated dissociation. *J Proteome Res* **6**, 2669-73.

- Scheri, R. C., Lee, J., Curtis, L. R., and Barofsky, D. F. (2008). A comparison of relative quantification with isobaric tags on a subset of the murine hepatic proteome using electrospray ionization quadrupole time-of-flight and matrix-assisted laser desorption/ionization tandem time-of-flight. *Rapid Commun Mass Spectrom.* **22**, 3137-46. doi: 10.1002/rcm.3719.
- Schmidt, A., Karas, M., and Dulcks, T. (2003). Effect of different solution flow rates on analyte ion signals in nano-ESI MS, or: when does ESI turn into nano-ESI? *J Am Soc Mass Spectrom* **14**, 492-500.
- Schuster Sa Fau - Boyes, B. E., Boyes Be Fau - Wagner, B. M., Wagner Bm Fau - Kirkland, J. J., and Kirkland, J. J. (2012). Fast high performance liquid chromatography separations for proteomic applications using Fused-Core(R) silica particles. *J Chromatogr A*.
- Scigelova, M., Hornshaw, M., Giannakopoulos, A., and Makarov, A. (2011). Fourier transform mass spectrometry. *Mol Cell Proteomics.* **10**, M111.009431. doi: 10.1074/mcp.M111.009431.
- Sheridan, A., Lenahan, M., Condell, O., Bonilla-Santiago, R., Sergeant, K., Renaut, J., Duffy, G., Fanning, S., Nally, J. E., and Burgess, C. M. (2013). Proteomic and phenotypic analysis of triclosan tolerant verocytotoxigenic Escherichia coli O157:H19. *J Proteomics* **80C:78-90.**, 10.1016/j.jprot.2012.12.025.
- Shibue M Fau - Mant, C. T., Mant Ct Fau - Hodges, R. S., and Hodges, R. S. (2005). Effect of anionic ion-pairing reagent concentration (1-60 mM) on reversed-phase liquid chromatography elution behaviour of peptides. *J Chromatogr A*.
- Shiraha, H., Yamamoto, K., and Namba, M. (2013). Human hepatocyte carcinogenesis (review). *Int J Oncol* **42**, 1133-8.
- Siena, S., Sartore-Bianchi, A., Di Nicolantonio, F., Balfour, J., and Bardelli, A. (2009). Biomarkers predicting clinical outcome of epidermal growth factor receptor-targeted therapy in metastatic colorectal cancer. *J Natl Cancer Inst.* **101**, 1308-24. doi: 10.1093/jnci/djp280. Epub 2009 Sep 8.
- Simic, T., Savic-Radojevic, A., Pljesa-Ercegovac, M., Matic, M., and Mimic-Oka, J. (2009). Glutathione S-transferases in kidney and urinary bladder tumors. *Nat Rev Urol.* **6**, 281-9. doi: 10.1038/nrurol.2009.49.
- Sjodin, M. O., Wetterhall, M., Kultima, K., and Artemenko, K. (2013). Comparative study of label and label-free techniques using shotgun proteomics for relative protein quantification. *J Chromatogr B Analyt Technol Biomed Life Sci.* **928:83-92.**, 10.1016/j.jchromb.2013.03.027. Epub 2013 Mar 31.
- Song, C., Wang, F., Ye, M., Cheng, K., Chen, R., Zhu, J., Tan, Y., Wang, H., Figeys, D., and Zou, H. (2011). Improvement of the quantification accuracy and throughput for phosphoproteome analysis by a pseudo triplex stable isotope dimethyl labeling approach. *Anal Chem.* **83**, 7755-62. doi: 10.1021/ac201299j. Epub 2011 Sep 21.
- Stephanowitz, H., Lange, S., Lang, D., Freund, C., and Krause, E. (2011). Improved Two-Dimensional Reversed Phase-Reversed Phase LC-MS/MS Approach for Identification of Peptide-Protein Interactions. *Journal of Proteome Research* **11**, 1175-1183.
- Sugiyama, N., Masuda, T., Shinoda, K., Nakamura, A., Tomita, M., and Ishihama, Y. (2007). Phosphopeptide enrichment by aliphatic hydroxy acid-modified metal oxide chromatography for nano-LC-MS/MS in proteomics applications. *Mol Cell Proteomics* **6**, 1103-9.
- Sun, S., Xu, M. Z., Poon, R. T., Day, P. J., and Luk, J. M. (2010a). Circulating Lamin B1 (LMNB1) biomarker detects early stages of liver cancer in patients. *J Proteome Res* **9**, 70-8.

- Sun, S., Yi, X., Poon, R. T., Yeung, C., Day, P. J., and Luk, J. M. (2009). A protein-based set of reference markers for liver tissues and hepatocellular carcinoma. *BMC Cancer*. **9:309**, 10.1186/1471-2407-9-309.
- Sun, Y., Zang, Z., Xu, X., Zhang, Z., Zhong, L., Zan, W., Zhao, Y., and Sun, L. (2010b). Differential proteomics identification of HSP90 as potential serum biomarker in hepatocellular carcinoma by two-dimensional electrophoresis and mass spectrometry. *Int J Mol Sci*. **11**, 1423-33. doi: 10.3390/ijms11041423.
- Tabb, D. L., Smith, L. L., Brechi, L. A., Wysocki, V. H., Lin, D., and Yates, J. R., 3rd (2003). Statistical characterization of ion trap tandem mass spectra from doubly charged tryptic peptides. *Anal Chem* **75**, 1155-63.
- Tamai, T., Uto, H., Takami, Y., Oda, K., Saishoji, A., Hashiguchi, M., Kumagai, K., Kure, T., Mawatari, S., Moriuchi, A., Oketani, M., Ido, A., and Tsubouchi, H. (2011). Serum manganese superoxide dismutase and thioredoxin are potential prognostic markers for hepatitis C virus-related hepatocellular carcinoma. *World J Gastroenterol*. **17**, 4890-8. doi: 10.3748/wjg.v17.i44.4890.
- Tanaka, K., Waki, H., Ido, Y., Akita, S., Yoshida, Y., and Yoshida, T. (1988). Protein and polymer analysis up to m/z 100.000 by laser ionisation time-of-flight mass spectrometry. *Rapid Commun Mass Spectrom* **2**, 151-153.
- Thingholm, T. E., Jensen, O. N., Robinson, P. J., and Larsen, M. R. (2008). SIMAC (sequential elution from IMAC), a phosphoproteomics strategy for the rapid separation of monophosphorylated from multiply phosphorylated peptides. *Mol Cell Proteomics* **7**, 661-71.
- Thingholm, T. E., Jorgensen, T. J., Jensen, O. N., and Larsen, M. R. (2006). Highly selective enrichment of phosphorylated peptides using titanium dioxide. *Nat Protoc* **1**, 1929-35.
- Thingholm, T. E., Palmisano, G., Kjeldsen, F., and Larsen, M. R. (2010). Undesirable charge-enhancement of isobaric tagged phosphopeptides leads to reduced identification efficiency. *J Proteome Res* **9**, 4045-52.
- Thompson, A. J., Hart, S. R., Franz, C., Barnouin, K., Ridley, A., and Cramer, R. (2003). Characterization of protein phosphorylation by mass spectrometry using immobilized metal ion affinity chromatography with on-resin beta-elimination and Michael addition. *Anal Chem* **75**, 3232-43.
- Tong, S. W., Yang, Y. X., Hu, H. D., An, X., Ye, F., Hu, P., Ren, H., Li, S. L., and Zhang, D. Z. (2012). Proteomic investigation of 5-fluorouracil resistance in a human hepatocellular carcinoma cell line. *J Cell Biochem* **113**, 1671-80.
- Trichopoulos, D., Bamia, C., Lagiou, P., Fedirko, V., Trepo, E., Jenab, M., Pischon, T., Nothlings, U., Overved, K., Tjonneland, A., Outzen, M., Clavel-Chapelon, F., Kaaks, R., Lukanova, A., Boeing, H., Aleksandrova, K., Benetou, V., Zylis, D., Palli, D., Pala, V., Panico, S., Tumino, R., Sacerdote, C., Bueno-De-Mesquita, H. B., Van Kranen, H. J., Peeters, P. H., Lund, E., Quiros, J. R., Gonzalez, C. A., Sanchez Perez, M. J., Navarro, C., Dorronsoro, M., Barricarte, A., Lindkvist, B., Regner, S., Werner, M., Hallmans, G., Khaw, K. T., Wareham, N., Key, T., Romieu, I., Chuang, S. C., Murphy, N., Boffetta, P., Trichopoulou, A., and Riboli, E. (2011). Hepatocellular carcinoma risk factors and disease burden in a European cohort: a nested case-control study. *J Natl Cancer Inst* **103**, 1686-95.
- Trinh, H. V., Grossmann, J., Gehrig, P., Roschitzki, B., Schlapbach, R., Greber, U. F., and Hemmi, S. (2013). iTRAQ-Based and Label-Free Proteomics Approaches for Studies of Human Adenovirus Infections. *Int J Proteomics* **2013:581862**, 10.1155/2013/581862. Epub 2013 Mar 11.

- Tsai, C. F., Wang, Y. T., Chen, Y. R., Lai, C. Y., Lin, P. Y., Pan, K. T., Chen, J. Y., Khoo, K. H., and Chen, Y. J. (2008). Immobilized metal affinity chromatography revisited: pH/acid control toward high selectivity in phosphoproteomics. *J Proteome Res* **7**, 4058-69.
- Urbaniak, M. D., Martin, D. M., and Ferguson, M. A. (2013). Global quantitative SILAC phosphoproteomics reveals differential phosphorylation is widespread between the procyclic and bloodstream form lifecycle stages of *Trypanosoma brucei*. *J Proteome Res.* **12**, 2233-44. doi: 10.1021/pr400086y. Epub 2013 Mar 29.
- van der Veken, P., Dirksen, E. H., Ruijter, E., Elgersma, R. C., Heck, A. J., Rijkers, D. T., Slijper, M., and Liskamp, R. M. (2005). Development of a novel chemical probe for the selective enrichment of phosphorylated serine- and threonine-containing peptides. *Chembiochem* **6**, 2271-80.
- Villen, J., Beausoleil, S. A., Gerber, S. A., and Gygi, S. P. (2007). Large-scale phosphorylation analysis of mouse liver. *Proc Natl Acad Sci U S A* **104**, 1488-93.
- Villen, J., and Gygi, S. P. (2008). The SCX/IMAC enrichment approach for global phosphorylation analysis by mass spectrometry. *Nat Protoc* **3**, 1630-8.
- Wang, C., Guo, K., Gao, D., Kang, X., Jiang, K., Li, Y., Sun, L., Zhang, S., Sun, C., Liu, X., Wu, W., Yang, P., and Liu, Y. (2011a). Identification of transaldolase as a novel serum biomarker for hepatocellular carcinoma metastasis using xenografted mouse model and clinic samples. *Cancer Lett.* **313**, 154-66. doi: 10.1016/j.canlet.2011.08.031. Epub 2011 Oct 8.
- Wang, F., Han, G., Yu, Z., Jiang, X., Sun, S., Chen, R., Ye, M., and Zou, H. (2010). Fractionation of phosphopeptides on strong anion-exchange capillary trap column for large-scale phosphoproteome analysis of microgram samples. *J Sep Sci.* **33**, 1879-87. doi: 10.1002/jssc.200900718.
- Wang, H., Alvarez, S., and Hicks, L. M. (2012a). 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.* **11**, 487-501. doi: 10.1021/pr2008225. Epub 2011 Dec 1.
- Wang, J., Zhang, Y., Jiang, H., Cai, Y., and Qian, X. (2006). Phosphopeptide detection using automated online IMAC-capillary LC-ESI-MS/MS. *Proteomics* **6**, 404-11.
- Wang, W. A., Groenendyk, J., and Michalak, M. (2012b). Calreticulin signaling in health and disease. *Int J Biochem Cell Biol.* **44**, 842-6. doi: 10.1016/j.biocel.2012.02.009. Epub 2012 Feb 21.
- Wang, Y., Kuramitsu, Y., Takashima, M., Yokoyama, Y., Iizuka, N., Tamesa, T., Sakaida, I., Oka, M., and Nakamura, K. (2011b). Identification of four isoforms of aldolase B down-regulated in hepatocellular carcinoma tissues by means of two-dimensional Western blotting. *In Vivo.* **25**, 881-6.
- Warburg, O. (1956). On the origin of cancer cells. *Science.* **123**, 309-14.
- Washburn Mp Fau - Wolters, D., Wolters D Fau - Yates, J. R., 3rd, and Yates, J. R., 3rd (2001). Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat Biotechnol.*
- Webb, K. J., Xu, T., Park, S. K., and Yates, J. R., 3rd (2013). Modified MuDPIT Separation Identified 4488 Proteins in a System-wide Analysis of Quiescence in Yeast. *J Proteome Res.* **12**, 2177-84. doi: 10.1021/pr400027m. Epub 2013 Apr 9.
- Wei, R. R., Zhang, M. Y., Rao, H. L., Pu, H. Y., Zhang, H. Z., and Wang, H. Y. (2012). Identification of ADH4 as a novel and potential prognostic marker in hepatocellular carcinoma. *Med Oncol.* **29**, 2737-43. doi: 10.1007/s12032-011-0126-3. Epub 2011 Dec 7.
- Werner, T., Becher, I., Sweetman, G., Doce, C., Savitski, M. M., and Bantscheff, M. (2012). High-resolution enabled TMT 8-plexing. *Anal Chem.* **84**, 7188-94. doi: 10.1021/ac301553x. Epub 2012 Aug 8.

- Wiktorowicz, J. E., English, R. D., Wu, Z., and Kurosky, A. (2012). Model studies on iTRAQ modification of peptides: sequence-dependent reaction specificity. *J Proteome Res.* **11**, 1512-20. doi: 10.1021/pr2003165. Epub 2012 Feb 15.
- Wilkins, M. R., Ou, K., Appel, R. D., Sanchez, J. C., Yan, J. X., Golaz, O., Farnsworth, V., Cartier, P., Hochstrasser, D. F., Williams, K. L., and Gooley, A. A. (1996a). Rapid protein identification using N-terminal "sequence tag" and amino acid analysis. *Biochem Biophys Res Commun.* **221**, 609-13.
- Wilkins, M. R., Pasquali, C., Appel, R. D., Ou, K., Golaz, O., Sanchez, J. C., Yan, J. X., Gooley, A. A., Hughes, G., Humphery-Smith, I., Williams, K. L., and Hochstrasser, D. F. (1996b). From proteins to proteomes: large scale protein identification by two-dimensional electrophoresis and amino acid analysis. *Biotechnology (N Y).* **14**, 61-5.
- Wilm, M., and Mann, M. (1996). Analytical properties of the nanoelectrospray ion source. *Anal Chem* **68**, 1-8.
- Wu, J., Shakey, Q., Liu, W., Schuller, A., and Follettie, M. T. (2007). Global profiling of phosphopeptides by titania affinity enrichment. *J Proteome Res* **6**, 4684-9.
- Wysocki, V. H., Tsaprailis, G., Smith, L. L., and Breci, L. A. (2000). Mobile and localized protons: a framework for understanding peptide dissociation. *J Mass Spectrom* **35**, 1399-406.
- Xia, Q., Cheng, D., Duong, D. M., Gearing, M., Lah, J. J., Levey, A. I., and Peng, J. (2008). Phosphoproteomic analysis of human brain by calcium phosphate precipitation and mass spectrometry. *J Proteome Res.* **7**, 2845-51. doi: 10.1021/pr8000496. Epub 2008 May 30.
- Xie F Fau - Smith, R. D., Smith Rd Fau - Shen, Y., and Shen, Y. (2012). Advanced proteomic liquid chromatography. *J Chromatogr A.*
- Xu, C. F., Lu, Y., Ma, J., Mohammadi, M., and Neubert, T. A. (2005). Identification of phosphopeptides by MALDI Q-TOF MS in positive and negative ion modes after methyl esterification. *Mol Cell Proteomics* **4**, 809-18.
- Xue, Y., Wei, J., Han, H., Zhao, L., Cao, D., Wang, J., Yang, X., Zhang, Y., and Qian, X. (2009). Application of open tubular capillary columns coated with zirconium phosphonate for enrichment of phosphopeptides. *J Chromatogr B Analyt Technol Biomed Life Sci.* **877**, 757-64. doi: 10.1016/j.jchromb.2009.02.020. Epub 2009 Feb 11.
- Yan, J. X., Packer, N. H., Gooley, A. A., and Williams, K. L. (1998). Protein phosphorylation: technologies for the identification of phosphoamino acids. *J Chromatogr A* **808**, 23-41.
- Yan, J. X., Packer, N. H., Tonella, L., Ou, K., Wilkins, M. R., Sanchez, J. C., Gooley, A. A., Hochstrasser, D. F., and Williams, K. L. (1997). High sample throughput phosphoamino acid analysis of proteins separated by one- and two-dimensional gel electrophoresis. *J Chromatogr A* **764**, 201-210.
- Ye, J., Zhang, X., Young, C., Zhao, X., Hao, Q., Cheng, L., and Jensen, O. N. (2010). Optimized IMAC-IMAC protocol for phosphopeptide recovery from complex biological samples. *J Proteome Res* **9**, 3561-73.
- Yost, R. A., and Enke, C. G. (1979). Triple quadrupole mass spectrometry for direct mixture analysis and structure elucidation. *Anal Chem* **51**, 1251-64.
- Yu, Z., Han, G., Sun, S., Jiang, X., Chen, R., Wang, F., Wu, R., Ye, M., and Zou, H. (2009). Preparation of monodisperse immobilized Ti(4+) affinity chromatography microspheres for specific enrichment of phosphopeptides. *Anal Chim Acta* **636**, 34-41.
- Zarei, M., Sprenger, A., Gretzmeier, C., and Dengjel, J. (2012). Combinatorial use of electrostatic repulsion-hydrophilic interaction chromatography (ERLIC) and strong cation exchange (SCX) chromatography for in-depth phosphoproteome analysis. *J Proteome Res* **11**, 4269-76.

- Zarei, M., Sprenger, A., Metzger, F., Gretzmeier, C., and Dengjel, J. (2011). Comparison of ERLIC-TiO₂, HILIC-TiO₂, and SCX-TiO₂ for global phosphoproteomics approaches. *J Proteome Res* **10**, 3474-83.
- Zeindl-Eberhart, E., Haraida, S., Liebmann, S., Jungblut, P. R., Lamer, S., Mayer, D., Jager, G., Chung, S., and Rabes, H. M. (2004). Detection and identification of tumor-associated protein variants in human hepatocellular carcinomas. *Hepatology*. **39**, 540-9.
- Zhang, X., Ye, J., Jensen, O. N., and Roepstorff, P. (2007). Highly Efficient Phosphopeptide Enrichment by Calcium Phosphate Precipitation Combined with Subsequent IMAC Enrichment. *Mol Cell Proteomics*. **6**, 2032-42. Epub 2007 Aug 4.
- Zhang, Y., Ficarro, S. B., Li, S., and Marto, J. A. (2009). Optimized Orbitrap HCD for quantitative analysis of phosphopeptides. *J Am Soc Mass Spectrom*. **20**, 1425-34. doi: 10.1016/j.jasms.2009.03.019. Epub 2009 Mar 28.
- Zhang, Y., Wen, Z., Washburn, M. P., and Florens, L. (2011). Improving proteomics mass accuracy by dynamic offline lock mass. *Anal Chem*. **83**, 9344-51. doi: 10.1021/ac201867h. Epub 2011 Nov 16.
- Zhang, Z. (2004). Prediction of low-energy collision-induced dissociation spectra of peptides. *Anal Chem* **76**, 3908-22.
- Zhang, Z. (2011). Prediction of collision-induced-dissociation spectra of peptides with post-translational or process-induced modifications. *Anal Chem* **83**, 8642-51.
- Zhao, Y., Kong, R. P., Li, G., Lam, M. P., Law, C. H., Lee, S. M., Lam, H. C., and Chu, I. K. (2012). Fully automatable two-dimensional hydrophilic interaction liquid chromatography-reversed phase liquid chromatography with online tandem mass spectrometry for shotgun proteomics. *J Sep Sci*. **35**, 1755-63. doi: 10.1002/jssc.201200054.
- Zheng, X., Gai, X., Ding, F., Lu, Z., Tu, K., Yao, Y., and Liu, Q. (2013). Histone acetyltransferase PCAF Up-regulated cell apoptosis in hepatocellular carcinoma via acetylating histone H4 and inactivating AKT signaling. *Mol Cancer*. **12**, 96.
- Zhou, H., Ye, M., Dong, J., Han, G., Jiang, X., Wu, R., and Zou, H. (2008). Specific phosphopeptide enrichment with immobilized titanium ion affinity chromatography adsorbent for phosphoproteome analysis. *J Proteome Res* **7**, 3957-67.
- Zhou, W., Liotta, L. A., and Petricoin, E. F. (2012). The spectra count label-free quantitation in cancer proteomics. *Cancer Genomics Proteomics*. **9**, 135-42.
- Zhu, W., Smith, J. W., and Huang, C. M. (2010). Mass spectrometry-based label-free quantitative proteomics. *J Biomed Biotechnol* **2010:840518**, 10.1155/2010/840518. Epub 2009 Nov 10.
- Ziolkowski, P., Gamian, E., Osiecka, B., Zougman, A., and Wisniewski, J. R. (2009). Immunohistochemical and proteomic evaluation of nuclear ubiquitous casein and cyclin-dependent kinases substrate in invasive ductal carcinoma of the breast. *J Biomed Biotechnol* **2009:919645**, 10.1155/2009/919645. Epub 2009 Dec 24.

Analyse phosphoprotéomique pour la recherche de biomarqueurs: développements et applications

L'analyse protéomique est une méthode de choix pour la recherche de biomarqueurs. Sans à priori, elle permet d'établir un catalogue de protéines qui sont exprimées dans une cellule, un tissu ou un organisme entier. Cette approche a été mise en œuvre pour une analyse protéomique de cellules coliques cancéreuses et une analyse phosphoprotéomique de biopsies hépatiques.

L'étude de cellules coliques cancéreuses transformées par le gène cytosine désaminase et traitées par la prodrogue 5-fluorocytosine a été réalisée par électrophorèse bidimensionnelle des extraits protéiques. L'analyse d'image a permis la quantification de 353 protéines et isoformes dont 14 sont surexprimées et 4 sous exprimées lors d'un traitement par la prodrogue. Parmi les protéines dont l'expression est affectée par le traitement, l'HSP90 présente un niveau d'expression constant mais est identifiée sous deux formes qui diffèrent par leur pI. L'analyse par spectrométrie de masse a identifié une phosphorylation de ser 254 qui pourrait contribuer à la régression tumorale.

Après avoir développé une méthode HPLC-TiO₂ pour la purification de phosphopeptides, une analyse protéomique de 24 biopsies humaines provenant de carcinomes hépatocellulaires sur foie non fibreux (nf-CHC) et de tissus sains a été réalisée avec la technologie iTRAQ. Les peptides surexprimés dans les tumeurs correspondent à des protéines de choc thermique, des protéines liées à l'ADN/ARN (histones, protéines du spliceosome), des protéines de la phase 1 de la détoxification (carboxyestérase, époxide hydrolase), les protéines du cytosquelette (actinine, tubuline), des protéines ou enzymes anti-oxydantes (superoxyde dismutase, thioredoxine). Les peptides sous-exprimés correspondent à des protéines du cycle de l'urée, de la détoxification (alcool déhydrogénase) du métabolisme des sucres, des lipides et des acides aminés. Dans la fraction TiO₂, 19 phosphopeptides sont significativement surexprimés et 15 phosphopeptides sont significativement sous exprimés, mettant en évidence une surreprésentation du motif -(S/T)P- parmi les phosphopeptides surexprimés dans les tumeurs. Une activation des *proline directed kinases* ou une inhibition des phosphatases correspondantes est donc probablement un événement caractéristique des nf-CHC. Ces peptides/protéines dérégulées sont autant de biomarqueurs potentiels pour le carcinome hépatocellulaire.

Phosphopeptides - TiO₂ – Cancer

Phosphoproteomics in cancer biomarkers discovery : chromatographic development and applications

Proteomic is a method of choice for biomarker research, without a priori, it establishes a directory of proteins that are expressed in a cell, tissue or an entire organism. This approach has been implemented for proteomic analysis of colon cancer cells and phosphoproteomic analysis of liver biopsies.

The study of colon cancer cells transformed by the cytosine deaminase and treated with the prodrug 5-fluorocytosine was performed by two-dimensional electrophoresis. Image analysis allowed the quantification of 353 proteins, 14 isoforms are overexpressed and 4 under expressed during treatment with the prodrug. Among the proteins whose expression is affected by the treatment, HSP90 has a constant level of expression, but is identified in two isoforms that differ in their pI. Mass spectrometry identified phosphorylation of ser 254 which could contribute to tumour regression.

After developed an HPLC- TiO₂ method for the purification of phosphopeptides, a proteomic analysis of 24 biopsies from human hepatocellular carcinoma on non-fibrous liver (nfCHC) and normal tissue was performed with the iTRAQ technology. Peptides overexpressed in tumours correspond to HSPs, DNA / RNA binding proteins (histones, spliceosome), proteins form Phase 1 detoxification (carboxyesterase, epoxide hydrolase), the cytoskeletal proteins (actinin, tubulin), antioxidant proteins or enzymes (superoxyde dismutase, thioredoxin). The under-expressed peptides belong to proteins of the urea cycle, the detoxification (alcohol dehydrogenase) metabolism of sugars, lipids and amino acids. In the TiO₂ fraction, 19 phosphopeptides are significantly overexpressed and 15 phosphopeptides were significantly under-expressed. This phosphoproteome has demonstrated an overrepresentation of the (S/T)P pattern among overexpressed phosphopeptides, indicating activation of proline-directed kinases in nfCHC or inhibition of the corresponding phosphatases. These deregulated peptides/proteins are as potential biomarkers for hepatocellular carcinoma.

Phosphopeptide - TiO₂ – Cancer