HAL
open science

Le secretome, un modèle adapté à l'étude des protéines sécrétées par les tumeurs.Application à l'analyse du rôle joué par p53 sur la sécrétion de protéines par des cellules du cancer du poumon.

Jérôme Chenau

▶ **To cite this version:**

Jérôme Chenau. Le secretome, un modèle adapté à l'étude des protéines sécrétées par les tumeurs.Application à l'analyse du rôle joué par p53 sur la sécrétion de protéines par des cellules du cancer du poumon.. Sciences du Vivant [q-bio]. Université Joseph-Fourier - Grenoble I, 2009. Français. NNT : . tel-00402098

HAL Id: tel-00402098

https://theses.hal.science/tel-00402098

Submitted on 6 Jul 2009

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITE JOSEPH FOURIER – GRENOBLE 1

U.F.R de Pharmacie

THESE

Pour obtenir le grade de

DOCTEUR DE l'UNIVERSITE JOSEPH FOURIER – GRENOBLE 1

Spécialité : Biotechnologie

Soutenue publiquement par :

Jérôme CHENAU

Le secretome, un modèle adapté à l'étude des protéines sécrétées par les tumeurs

Application à l'analyse du rôle joué par p53 sur la sécrétion de protéines par des cellules du cancer du poumon

Date de soutenance : le 24 juin 2009

Composition du Jury :

Pr. Christian BRAMBILLA, Président
Dr. Claude CARON DE FROMENTEL, Rapporteur
Dr. Catherine GUETTE, Rapporteur
Dr. Sandrine UTTENWEILER, Examinateur
Pr. Mehmet OZTURK, Examinateur
Pr. Michel SEVE, Directeur de thèse

Thèse préparée au sein de la plate-forme de protéomique médicale
Institut Albert Bonniot INSERM U823 Equipe 5

Avant-propos et Remerciements

Me voilà donc arrivé (ou presque !) au terme de cette aventure ô combien enrichissante sur le plan scientifique et personnel. Je tiens à exprimer ici toute ma gratitude aux personnes qui ont contribué à l'aboutissement de ce travail.

Je remercie chaleureusement Michel Seve pour la confiance qu'il m'a témoignée en m'accueillant en thèse dans son équipe. Merci Michel pour ton soutien et pour m'avoir permis de découvrir les joies de la protéomique.

J'adresse mes vifs remerciements au Professeur Christian Brambilla pour avoir accepté de présider ce jury de thèse. J'exprime également toute ma reconnaissance à Mesdames Catherine Guette, Claude Caron de Fromentel, Sandrine Uttenweiler et à Monsieur Mehmet Oztürk pour l'honneur qu'ils me font en acceptant de juger mon travail.

Un grand merci à tous les membres de l'équipe « Protéomique » pour leur aide, leur soutien et tous les moments de bonne humeur que l'on a partagés. Merci Nadège (« P'tite Nadou ») pour ta précieuse aide technique et pour m'avoir supporté durant tout ce temps !! Merci les deux Sylvie (agents B et M) pour votre soutien de tous les instants. Merci Morgane pour tous tes précieux conseils en protéomique et autres. Merci aussi à Valérie (Mme Chromato), à Christian (le « Lance Armstrong » du labo) et aux trois « miss » du secrétariat (Aurélie, Isa et Elizabeth) pour leur aide.

Merci à tous les stagiaires qui ont égayé la vie du labo durant leur passage (Ali, Yacine, Sonia, Sylvain, Noé, Betty, Manel, Anthony, Thomas et Maud). Merci pour toutes les sorties, ainsi que pour le pep's et la bonne humeur que vous nous avez amenés au quotidien. Un merci tout particulier à Jonathan (« Jojo »), stagiaire récidiviste avec qui j'ai partagé mon bureau (bureau n°P53, ça ne s'invente pas !!). Certains de nos « pétages de plombs » de fin de journée resteront de très bons souvenirs pour moi.

Merci également aux autres membres du pavillon B (Madame Marie-Jeanne Richard, Monsieur Patrice Faure, Marie-Pierre, Eliane, Arlette, Nicole, Corine, Valérie,… et tous les autres !!). Merci au Professeur Marie-Christine Favrot pour son aide au début de ma thèse. Merci au Professeur Alain Favier pour les précieux conseils quil m'a distillé lors de nos discussions de fin de journées. Merci à Florence pour son aide précieuse dans plusieurs domaines.

Merci à Jacotte pour m'avoir formé à la culture cellulaire, ainsi que pour toutes nos discussions passionnées autours du GF38 !!

Merci aussi à tous les membres de l'IAB, notamment à ceux de l'équipe 5. Merci Jean-Luc pour les discussions scientifiques très bénéfiques que l'on a pu avoir. Merci Amandine pour tes conseils sur la lignée H358. Merci Corine et Véronique pour votre aide primordiale dans les manips sur les souris. Merci enfin aux différents thésards : Ben, Sandrine, Romuald, et tous les autres.

Merci aux LANistes (et ex-LANistes) pour ne m'avoir jamais oublié après mon stage de Master 2 dans votre labo. Je me souviendrai longtemps de notre célèbre équipe de foot du MiLAN Assez !! En vrac, merci donc à : Séverine, Fabrice, Sylvain, JF, Joce, Sylvie, Jean-Luc, Thierry, Didier, Sandra, Christine, Francette, Stéphane, Peggy, Olivier, Thibaut, Jean, Walid, Pierre-Marie, Zohra et tous les autres.

Merci aussi à tous mes amis, notamment Laurent, François, Mickaël et Olivier, pour avoir été là quand j'avais besoin de me changer les idées.

Merci à mes trois minous, Luna, Vahinée et Casanova pour votre affection quotidienne et avoir su, à votre façon, me remonter le moral quand il le fallait !

Merci à mes grands-parents. Où que vous soyez, je suis sûr que vous êtes fiers de moi.

Mes dernières pensées vont à ma famille et notamment à mes parents et à ma sœur Mélanie. MERCI mille fois pour vos encouragements et votre soutien de tous les instants. Je sais que je pourrai toujours compter sur vous.

A mes parents, à ma sœur

Liste des participations scientifiques

- **Communications écrites (Posters) :**

CHENAU J., MICHELLAND S., de FRAIPONT F., JOSSERAND V. COLL J-L., FAVROT M-C., SEVE M, The cell line secretome, a suitable tool for investigating proteins released in vivo by tumors: Application to the study of p53-modulated proteins secreted in lung cancer cells, *CLARA*, (Archamps 2009).

SIDIBE J., MICHELLAND S., CHENAU J., COUVET M., CUNIN V., SEVE M., A comparison of four high abundant proteins immunoaffinity depletion columns for plasma proteomics and biomarker discovery using iTRAQ labeling, *CLARA*, (Archamps 2009).

CHENAU J., MICHELLAND S., COLL J-L., FAVROT M-C., SEVE M., Analysis of secreted proteins by Non Small Cells Lung Cancer cell line: a differential analysis according to p53 state using iTRAQ, *25e congrès de la SFEAP, (Tours 2009).*

CHENAU J., COUVET M., MICHELLAND S., SIDIBE J., SEVE M., Peptides OFFGEL electrophoresis: a suitable pre-analytical step for complex eukaryotic samples fractionation compatible with quantitative iTRAQ labeling, *25e congrès de la SFEAP,* (Tours 2009).

CHENAU J., MICHELLAND S., COLL J-L., SEVE M, Analyse protéomique différentielle des protéines sécrétées par des cellules du cancer du poumon non à petites cellules en fonction du statut p53, *CLARA*, (Lyon 2008).

SIDIBE J., MICHELLAND S., CHENAU J., COUVET M., SEVE M., Stratégies de pré fractionnement des échantillons plasmatiques pour la recherche de marqueurs en cancérologie, *CLARA*, (Lyon 2008).

- **Communications orales (Conférences) :**

CHENAU J., MICHELLAND S., COLL J-L., FAVROT M-C., SEVE M., Analysis of secreted proteins by Non Small Cells Lung Cancer cell line: a differential analysis according to p53 state using isobaric tags for relative and absolute quantitation (iTRAQ), 8th Siena Meeting: *From Genome to Proteome: Integration and Proteome Completion,* (Siena 2008).

CHENAU J., Analyses différentielles des protéines sécrétées par des cellules du cancer du poumon non à petites cellules en fonction du statut p53, *Journée Protéomique et cancer*; (Angers 2008).

- **Publications :**

CHENAU J., MICHELLAND S., SEVE M., [Secretome: definitions and biomedical interest], *Rev Med Interne.* 2008 Jul;29(7):606-8. Epub 2008 Feb 5.

CHENAU J., MICHELLAND S., SIDIBE J., SEVE M., Peptides OFFGEL electrophoresis: a suitable pre-analytical step for complex eukaryotic samples fractionation compatible with quantitative iTRAQ labeling., *Proteome Science*. 2008 Feb 26;6:9.

CHENAU J., MICHELLAND S., de FRAIPONT F., JOSSERAND V. COLL J-L., FAVROT M-C., SEVE M., The cell line secretome, a suitable tool for investigating proteins released in vivo by tumors: Application to the study of p53-modulated proteins secreted in lung cancer cells, *submitted to publication in Journal of Proteome Research*, 2009.

CHENAU J. et STRANDE V., Analyses différentielles par marquage isotopique, *Lettre 42 de la Société Française d'Electrophorèse et d'analyse Protéomique (bourse SFEAP)*, janvier 2009.

- **Participations scientifiques :**

SIDIBE J., MICHELLAND S., CHENAU J., COUVET M., CUNIN V., SEVE M., A comparison of four high abundant proteins immunoaffinity depletion columns for plasma proteomics and biomarker discovery using iTRAQ labeling, *submitted to publication in BMC Biotechnology*, 2009.

BAILLET A., TROCME C., BERTHIER S., ARLOTTO M., GRANGE L., CHENAU J., QUETANT S., SEVE M., BERGER F., JUVIN R., MOREL F., GAUDIN P. Synovial fluid proteomic fingerprint in arthritis: calgranulins (S100A8, S100A9 and S100A12) discriminate rheumatoid arthritis from other inflammatory joint diseases, *submitted to publication in Arthritis & Rheumatism*, 2009.

SOMMAIRE

PLAN GENERAL DU MANUSCRIT :

Avant-propos et Remerciements 3

Liste des participations scientifiques 7

Liste des Figures et Tableaux : 19

Liste des principales abréviations 23

INTRODUCTION 27

DONNEES BIBLIOGRAPHIQUES 31

CHAPITRE 1 : Les approches et techniques en protéomique 31

CHAPITRE 2 : Les méthodes d'analyse protéomique différentielle 71

CHAPITRE 3 : Le cancer du poumon 91

CHAPITRE 4 : La protéine p53 99

CHAPITRE 5 : La tumorogenèse et la modulation de l'environnement tumoral 109

CHAPITRE 6 : Le secretome 119

OBJECTIFS DU TRAVAIL ET CHOIX METHODOLOGIQUES 125

RESULTATS 131

CHAPITRE 1 : Développement et validation de la technologie OFFGEL pour la séparation de mélanges peptidiques complexes et compatibilité avec le marquage iTRAQ. 131

CHAPITRE 2 : Analyse différentielle des protéines sécrétées modulées par p53 dans des cellules du Cancer du Poumon Non à Petites Cellules. 147

CHAPITRE 3 : Le Secretome : une source de bio-marqueurs ? 227

DISCUSSION GENERALE 237

CONCLUSIONS ET PERSPECTIVES 263

REFERENCES BIBLIOGRAPHIQUES 269

RESUME 289

ABSTRACT 290

PLAN DETAILLE DU MANUSCRIT :

Avant-propos et Remerciements 3

Liste des participations scientifiques 7

Liste des Figures et Tableaux : 19

Liste des principales abréviations 23

INTRODUCTION 27

DONNEES BIBLIOGRAPHIQUES 31

CHAPITRE 1 : Les approches et techniques en protéomique 31

I Introduction 31
1. Définitions de la protéomique 31
2. Origine de la protéomique 31
3. La Protéomique et les autres « Omics » 32

II La Spectrométrie de Masse en Protéomique 35
1. Généralités 35
1.1. Définition 35
1.2. Un intérêt majeur en protéomique 35
1.3. Principe général 35
1.4. Les termes essentiels 36
1.4.1. Masse moléculaire 36
1.4.2. Résolution 37
1.4.3. Précision de mesure 37
1.5. La Spectrométrie de masse en mode tandem 37
1.5.1. Principe 37
1.5.2. Fragmentation des peptides en MS/MS 39
1.6. ESI et MALDI 40
2. La Spectrométrie de Masse à source MALDI 41
2.1. Principe d'ionisation 41
2.1.1. Les matrices 41
2.1.1.1. Les types de matrices 41
2.1.1.2. Les modes de dépôt 42
2.1.1.2.1. Goutte séchée 42
2.1.1.2.2. Couche mince 42
2.1.1.2.3. Sandwich 43
2.1.2. Le laser 43
2.1.3. Les cibles 43
2.1.4. Le mécanisme d'ionisation MALDI 44
2.2. L'analyseur en temps de vol (TOF) 44
2.2.1. Principe général 44
2.2.2. Mode linéaire 46
2.2.3. Mode réflectron 47
2.2.4. Mode MS/MS 48
3. La Spectrométrie de Masse à source ElectroSpray 49

3.1. Principe d'ionisation ... 49
3.2. Les analyseurs compatibles avec une source ESI ... 50
3.2.1. Quadrupôle ... 50
3.2.2. Temps de Vol (TOF) ... 51
3.2.3. Trappe ionique ... 51
3.2.4. La trappe orbitale ... 52
3.2.5. Analyseurs à résonance ionique cyclotronique par transformée de Fourier. 52
3.2.6. L'hybride Q-TOF ... 52
3.2.7. Triple quadrupôle ... 52

III Les méthodes d'analyse protéomique ... 54
1. La préparation de l'échantillon ... 54
1.1. La lyse des tissus et cellules par homogénéisation ... 54
1.2. La solubilisation des protéines ... 55
1.3. La protection contre la protéolyse ... 55
1.4. L'élimination des contaminants ... 55
1.5. La concentration des protéines ... 56
1.6. L'enrichissement de protéines ... 56
1.6.1. Le fractionnement sub-cellulaire ... 56
1.6.2. La purification de complexes protéiques et microdomaines ... 56
1.6.3. Les méthodes d'extractions séquentielles ... 56
1.6.4. Les fluides biologiques ... 57
2. La séparation de l'échantillon ... 57
2.1. Electrophorèse bidimensionnelle ... 57
2.1.1. Le principe ... 57
2.1.2. La coloration ... 58
2.1.3. Avantages et inconvénients ... 59
2.2. Le fractionnement OFFGEL ... 59
2.3. Chromatographie liquide haute pression (CLHP) ... 61
2.3.1. Introduction ... 61
2.3.2. Principe ... 61
2.3.3. Le couplage LC-MS ... 63
2.3.4. La chromatographie multi-dimensions ... 63
3. L'identification des protéines ... 64
3.1. L'empreinte peptidique massique (PMF) ... 65
3.2. Identification des peptides par MS/MS ... 67
3.2.1. Principe ... 67
3.2.2. Le séquençage de novo ... 68
4. Trois grands types d'applications de la protéomique ... 69
4.1. La cartographie ... 69
4.2. L'analyse des modifications post-traductionnelles ... 69
4.3. Les analyses protéomiques différentielles et la quantification ... 70

CHAPITRE 2 : Les méthodes d'analyse protéomique différentielle ... 71
1. Les méthodes « gel-based » ... 71
1.1. Comparaison d'images de gels 2D ... 71
1.2. DIGE ... 72
2. Les méthodes de marquage isotopique ... 73
2.1. Principe général ... 73
2.2. Marquage métabolique (SILAC) ... 75

2.3. Marquage enzymatique ($H_2^{16}O/H_2^{18}O$) ... 78
2.4. Marquage chimique ... 79
2.4.1. ICAT ... 79
2.4.2. iTRAQ ... 82
3. Les méthodes « label-free » ... 86
4. Les méthodes avec standard : vers une quantification absolue ... 87
4.1. Peptide AQUA ... 88
4.2. QCAT ou QconCAT ... 89
4.3. PSAQ ... 90

CHAPITRE 3 : Le cancer du poumon ... 91

1. Introduction ... 91
2. Classification / Histologie ... 92
2.1. Cancers du poumon non à petites cellules ... 92
2.2. Cancers du poumon à petites cellules ... 92
3. Signes et symptômes ... 92
4. Diagnostic ... 93
5. Traitements ... 93
5.1. Cancers du poumon non à petites cellules ... 93
5.2. Cancers du poumon à petites cellules ... 94
6. Les bio-marqueurs : une nécessité ... 94
6.1. Généralités ... 94
6.2. Types de bio-marqueurs tumoraux ... 95
6.3. Applications en diagnostic et thérapeutique ... 95
6.4. Les bio-marqueurs actuels du cancer du poumon ... 96

CHAPITRE 4 : La protéine p53 ... 99

1. Introduction – Historique ... 99
2. Structure de p53 ... 99
3. Rôles de la protéine p53 ... 100
4. Mutations de p53 dans les cancers ... 103
4.1. Généralités ... 103
4.2. Le cas du cancer du poumon ... 105
4.3. Les mutations de TP53 dans l'ADN circulant : 1 source de bio-marqueur ... 106
4.4. Les thérapies anti-p53 : un parcours semé d'embuches ... 106
4.4.1. La restauration par thérapie génique de la fonction de p53 ... 106
4.4.2. Les autres approches ... 107
5. p53 et les protéines sécrétées ... 108

CHAPITRE 5 : La tumorogenèse et la modulation de l'environnement tumoral ... 109

1. Introduction ... 109
2. Invasion et métastase ... 109
2.1. Remaniement de la matrice extracellulaire ... 109
2.1.1. Les cascades protéolytiques ... 110
2.1.1.1. Le système uPA-uPAR ... 110
2.1.1.2. Les métalloprotéases (MMPs) ... 111
2.1.1.3. Les cathepsines ... 112
2.1.2. Les mécanismes d'invasion sans protéolyse ... 113

2.2. La transition épithélio-mésenchymateuse 114
2.3. Les régulateurs de la métastase 115
3. La néo-angiogenèse et la progression tumorale 115
4. La sénescence 117
4.1. La voie p53-p21 117
4.2. La communication SMS des cellules (Senescence-messaging secretome)..... 117

CHAPITRE 6 : Le secretome 119

1. Définitions 119
2. Les différentes voies de sécrétion 119
3. Les méthodes d'analyse du secretome 121
3.1. Analyse indirecte in silico 121
3.2. Analyse directe par protéomique 121
4. Intérêts du secretome 122

OBJECTIFS DU TRAVAIL ET CHOIX METHODOLOGIQUES 125

RESULTATS 131

CHAPITRE 1 : Développement et validation de la technologie OFFGEL pour la séparation de mélanges peptidiques complexes et compatibilité avec le marquage iTRAQ. 131

1.roblématique et objectifs du travail 131

2.Moyens mis en œuvre 132

3.Résultats 133

4.Conclusion 134

5.Publication I 134

6.Résultat complémentaire : Validation de l'approche iTRAQ 143

CHAPITRE 2 : Analyse différentielle des protéines sécrétées modulées par p53 dans des cellules du Cancer du Poumon Non à Petites Cellules. 147

1.Problématique et objectifs du travail 147

2.Moyens mis en œuvre 147

3.Résultats 150

4.Conclusion 151

5.Publication II 151

6.Résultats complémentaires 209
6.1. Dosage des protéines GDF-15, FGF-19 et Kallikrein-6 dans les échantillons de secretome de la lignée H358/TetOn/p53 209
6.2. Autres analyses bioinformatiques des données 210
6.2.1. PLIPS et le croisement des données protéomique 210
6.2.2. Ingenuity systems 211
6.3. Etude de l'influence de p53 sur la modulation d'autres sous-protéomes de la lignée H358/TetOn/p53 212
6.3.1. Sous-protéome cytosolique 215
6.3.2. Sous-protéome des organites et des membranes 219

6.4. Synthèse 222

CHAPITRE 3 : Le Secretome : une source de bio-marqueurs ? 227

1.Problématique et objectifs du travail 227

2.Matériels et méthodes 228

2.1. Echantillons 228

2.2. Dot-blot 229

2.3. ELISA 229

3.Résultats 230

3.1. Analyses par dot blot 230

3.2. Dosage par ELISA 231

4.Conclusion 234

DISCUSSION GENERALE 237

1. Le modèle cellulaire et le procédé d'analyse 237

2. Les protéines modulées par p53 et la tumorogenèse 241

2.1. L'étude du secretome 241

2.1.1. Influence sur les voies classiques de dégradation de la MEC 241

2.1.2. L'ostéopontine et la tumorogenèse 242

2.1.3. Influence sur la transition épithélio-mésenchymateuse (TEM) 243

2.1.4. Influence sur les voies de néo-angiogenèse 245

2.1.5. Influence sur la sénescence 246

2.1.6. La kallikrein-6 et la modulation du micro-environnement tumoral 248

2.1.7. FGF-19 250

2.1.8. FGF-BP1 251

2.1.9. Le transporteur de zinc ZIP10 252

2.1.10. GDF-15 253

2.2. L'étude des protéomes cytosoliques et des organelles&membranes 253

2.2.1. Le complexe MCM 254

2.2.2. Carboxylestérase 2 254

2.2.3. Ferredoxin reductase (NADPH:adrenodoxin oxidoreductase) 254

2.2.4. Ribonucleotide reductase M2B 255

3. Les cibles transcriptionnelles de p53 régulées dans la lignée H358 255

4. Protéomique et oncologie 257

CONCLUSIONS ET PERSPECTIVES 263

REFERENCES BIBLIOGRAPHIQUES 269

RESUME 289

ABSTRACT 290

Liste des Figures et Tableaux :

Figure 1 : Nombre d'articles publiés en protéomique de 1995 à 2008. 32
Figure 2 : Les « Omics » et le dogme central de la biologie. 33
Figure 3 : Un même génome mais différents protéomes. 34
Figure 4 : Principe général d'un spectromètre de masse. 36
Figure 5 : Exemple d'un massif isotopique observé sur un spectre MS. 36
Figure 6 : La résolution d'un pic en spectrométrie de masse. 37
Figure 7 : La spectrométrie de masse en mode tandem. 38
Figure 8 : Exemple d'un spectre MS et d'un spectre MS/MS. 38
Figure 9 : Structure chimique d'un dipeptide. 39
Figure 10 : Types de fragments observés en MS/MS pour un tripeptide. 39
Figure 11 : Structure chimique d'un ion immonium. 40
Figure 12 : Vues au microscope des différents types de cristallisation obtenus selon la matrice. 42
Figure 13 : Cibles MALDI 43
Figure 14 : Le processus d'ionisation MALDI. 44
Figure 15 : L'analyseur en temps de vol (TOF). 45
Figure 16 : Les trois modes d'utilisation du MALDI TOF/TOF 4800. 45
Figure 17 : L'extraction retardée. 46
Figure 18 : Principe du mode réflectron. 47
Figure 19 : La fragmentation peptidique par CID. 48
Figure 20 : L'ionisation de type Electrospray (ESI). 49
Figure 21 : L'analyseur quadrupôle. 50
Figure 22 : L'analyseur trappe ionique. 51
Figure 23 : Aperçu général des deux grandes voies d'analyse protéomique. 54
Figure 24 : Principe de l'électrophorèse bidimensionnelle. 57
Figure 25 : L'isoélectrofocalisation OFFGEL. 60
Figure 26 : La chromatographie liquide haute pression (CLHP). 62
Figure 27 : Les systèmes de chromatographie multi-dimensions de type MudPIT et 2D-LC. 64
Figure 28 : La digestion trypsique des protéines. 65
Figure 29 : L'identification des protéines par Empreinte Peptidique Massique (PMF). 65
Figure 30 : L'identification des protéines par analyse MS/MS. 68
Figure 31 : La comparaison de cartes 2D par le logiciel ImageMaster. 72
Figure 32 : Le principe de la méthode DIGE (Differential In Gel Electrophoresis). 73
Figure 33 : Principe général de la quantification relative par marquage isotopique différentiel. 74
Figure 34 : Représentation de différentes méthodes d'analyse différentielle reposant sur un marquage isotopique. 74
Figure 35 : Différents niveaux de marquage des échantillons selon les approches. 75
Figure 36 : Le principe de la méthode SILAC. 76
Figure 37 : Paires de pics obtenues pour deux peptides sur les spectres MS à l'issue d'une analyse SILAC 77
Figure 38 : Principe de la méthode de marquage enzymatique ($H_2{}^{16}O/H_2{}^{18}O$). 78
Figure 39 : Principe chimique du marquage enzymatique 79

Figure 40 : Le tag ICAT de première génération. 80
Figure 41 : Le principe de ma méthode de marquage ICAT. 81
Figure 42 : Structure schématique du tag iTRAQ 4-plex. 82
Figure 43 : Structure chimique des différents tags iTRAQ 114, 115, 116, et 117. 83
Figure 44 : La méthode iTRAQ 4-plex. 84
Figure 45 : Exemple de spectre MS/MS obtenu avec la méthode iTRAQ 8-plex. 85
Figure 46 : Structure chimique des tags iTRAQ 8-plex. 85
Figure 47 : La quantification absolue en protéomique. 88
Figure 48 : Principe de l'approche AQUA (Sigma). 88
Figure 49 : Principe de l'approche QconCAT. 89
Figure 50 : Diagramme représentant le nombre estimé de nouveaux cas de cancer (incidence) et de décès (mortalité) dans le monde en 2002 91
Figure 51 : Les domaines structuraux de la protéine p53. 99
Figure 52 : Schéma simplifié de la régulation de p53. 101
Figure 53 : Schéma synthétique de la voie de signalisation p53. 103
Figure 54 : La voie uPA-uPAR. 111
Figure 55 : Les MMPs favorisent le développement des cellules cancéreuses. 112
Figure 56 : La cascade protéolytique de la cathepsine B. 113
Figure 57 : La transition épithélio-mésenchymateuse (TEM) durant la tumorogenèse. 114
Figure 58 : Les acteurs de l'environnement tumoral impliqués dans la néo-angiogenèse. 116
Figure 59 : Les facteurs du SMS (Senescence-Messaging Secretome) 117
Figure 60 : Les différentes voies de sécrétion de protéines par la cellule. 120
Figure 61 : Validation de la quantification relative par marquage iTRAQ 2-plex. 143
Figure 62 : Validation de la quantification relative par marquage iTRAQ 8-plex. 144
Figure 63 : Représentation du procédé d'analyse protéomique utilisé pour l'analyse différentielle des protéines p53-modulées relarguées par la lignée H358. 149
Figure 66 : Niveaux d'expression des protéines GDF-15, Kallikrein-6 et FGF-19 dans le secretome des cellules H358/TetOn/p53. 210
Figure 67 : Relation entre p53 et les protéines différentielles du secretome des cellules H358 établi par le logiciel Ingenuity. 211
Figure 68 : Comparaison de trois sous-protéomes de la lignée cellulaire H358/TetOn/p53. 214
Figure 69 : Protéines modulées par p53 dans le secretome *versus* protéome cytosolique. 217
Figure 70 : Protéines modulées par p53 dans le secretome *versus* protéome des organelles et des membranes 222
Figure 71 : Spécificité des protéines modulées par p53 dans le Secretome. 223
Figure 72 : Criblage des protéines par la technique dot-blot. 230
Figure 73 : Exemple de dosages ELISA de la protéine FGF-19 dans les plasmas. 233
Figure 74 : Influence de p53 sur les cascades protéolytiques dégradant la MEC 242
Figure 75 : Influence de la voie médiée par l'ostéopontine sur la tumorogenèse. 243
Figure 76 : Influence sur le phénotype TEM des cellules. 244
Figure 77 : Le double rôle de la β-caténine. 245
Figure 78 : Influence sur l'angiogenèse 246
Figure 79 : Les différentes voies de transduction du signal de senescence 247
Figure 80 : Modèle proposé pour expliquer l'expression différentielle de KLK6 249

Figure 81 : Modèle de l'action de FGF-19 sur le développement tumoral. 250
Figure 82 : Modèle du mécanisme d'action de la protéine FGF-BP1 252
Figure 83 : Régulation de l'expression de GDF-15 253
Figure 84 : Les cibles transcriptionnelles de p53 régulées dans la lignée H358 256

Tableau 1 : Matrices classiques pour l'analyse protéomique par MALDI-TOF 41
Tableau 2 : Caractéristiques générales de différents spectromètres de masse utilisés en protéomique 50
Tableau 3 : Les pré-requis pour une coloration optimale de gels SDS-PAGE. 59
Tableau 4 : Les bio-marqueurs protéiques actuellement utilisés pour le cancer du poumon 97
Tableau 5: Les bio-marqueurs protéiques potentiels du cancer du poumon 97
Tableau 6 : Spectre de mutation des 6 codons « hotspots » dans les cancers 105
Tableau 7 : Les effecteurs indirects de l'angiogenèse. 116
Tableau 8 : Dosages ELISA des protéines GDF-15, FGF-19 et Kallikrein-6 dans le secretome après ou non ré-expression de p53 par les cellules H358/TetOn/p53. 209
Tableau 9 : Les protéines modulées par p53 dans la fraction cytosolique des cellules H358/TetOn/p53. 216
Tableau 10 : Représentation des protéines modulées par p53 dans le secretome de la lignée H358/TetOn/p53 avec leur ratio iTRAQ éventuel obtenu dans les échantillons « cytosol » et « organelles & membranes ». 219
Tableau 11 : Les protéines modulées par p53 dans la fraction « organelles & membranes » des cellules H358/TetOn/p53. 221
Tableau 12 : Liste des patients atteints d'un NSCLC de type ADK inclus dans l'étude ELISA. 229
Tableau 13 : Dosages ELISA des protéines GDF-15, FGF-19 et Kallikrein-6. 231
Tableau 14 : Dosage plasmatique par ELISA des protéines KLK6, FGF19 et GDF15 232

Liste des principales abréviations

A : Absorbance
ACN : Acétonitrile
AQUA : Absolute quantification
CHCA : α-cyano-4-hydroxycinnamic acid
CID : Collision-Induced Dissociation
CLHP : Chromatographie Liquide Haute Performance (Pression)
CPNPC : Cancer du Poumon Non à Petites Cellules
Da : Dalton (masse moléculaire)
Dox : Doxycycline
2-DE : Two-dimensional gel electrophoresis
CM : Conditioned medium
DIGE : Differencial in gel electrophoresis
DTT : Dithiothréitol
ECL : Enhanced ChemiLuminescence
ELISA : Enzyme-Linked ImmunoSorbent Assay
ESI : Electrospray Ionization
FT-ICR : Fourier transform-ion cyclotron resonance
FWHM : Full-Width Half-Maximum
HPLC : High-performance liquid chromatography
HUPO : Human Proteome Organization
IARC : International Agency for Research on Cancer
ICAT : Isotope-coded affinity tag
id : Inside diameter
IEF : Isoelectrofocalisation
IPG : Immoblized pH gradient
iTRAQ : Isobaric tag for relative and absolute quantitation
LC : Liquid Chromatography
MALDI : Matrix-Assited Laser Desorption/Ionization
MEC : Matrice extra-cellulaire
MMTS : Methyl methanethiosulfonate
MRM : Multiple Reaction Monitoring
MS : Mass spectrometry
MS/MS : Tandem mass spectrometry
m/z : Ratio masse / charge
NSCLC : Non-Small Cell Lung Cancers
pI : Point isolectrique
PLIPS : Protein Lists Identified in Proteomics Studies
PMF : Peptide Mass Fingerprint
ppm : partie par million (part per million)
PSAQ : Protein Standard Absolute Quantification
PTM : Post traductional modification
QconCAT : Quantitative Concatemere
RP : Reverse phase

S/N : Signal to Noise ratio
SCX : Stong Cation eXchange
SELDI : Surface Enhanced Laser Desorption Ionisation
SILAC : Stable Isotope Labeling with Amino acid in Cell culture
SVF : Sérum de veau fœtal
TAP-tag : tandem-affinity purification tag
TCA : TriChloroacetic Acid
TCEP : Tris-(2-carboxyethyl) phosphine
TEAB : Triethylammonium bicarbonate
TEM : Transition épithélio-mésenchymateuse
TetOn : Promoteur inductible par la tetracycline
TFA : Trifluoroacetic acid
TIS : Time Ion Selector
TMT : Tandem Mass Tags
TOF : Time Of Flight
Vh : Volt x hours
WB : Western blot
wt : Wild type

INTRODUCTION

INTRODUCTION

Le cancer est un véritable problème de santé publique. Selon le rapport annuel de l'Organisation Mondiale de la Santé (OMS), publié en décembre 2008, cette maladie sera la première cause de mortalité dans le monde d'ici 2010, détrônant les maladies cardio-vasculaires, et le nombre de cas devrait doubler d'ici 2030. La lutte contre le cancer représente donc plus que jamais un enjeu primordial. La recherche porte aujourd'hui sur la compréhension des bases moléculaires et génétiques impliquées dans les mécanismes de l'oncogenèse, la recherche de cibles diagnostique, pronostique et thérapeutique, et le développement d'un arsenal thérapeutique approprié. Parmi tous ces mécanismes, un évènement majeur de la tumorogenèse est lié à la modulation de l'interaction des cellules tumorales entre elles et avec leur micro-environnement. La mutation ou la variation d'expression de facteurs pro- ou anti-oncogénique joue un rôle important dans ce phénomène.

Lors de ce travail de thèse, nous avons voulu apporter des éléments de réponse à la problématique suivante : l'étude des protéines sécrétées par la tumeur apporte t'elle des éléments contribuant à la compréhension des mécanismes de la tumorogenèse et peut elle permettre de mettre en valeur de nouvelles cibles d'intérêt clinique ? Nous avons ciblé notre étude sur le cancer du poumon, le plus fréquent au niveau mondial tant en terme d'incidence que de mortalité. Ce cancer présente le plus fort taux de mutation du gène p53, un anti-oncogène clé de la cellule. Cette perte de fonction induite par ces mutations module la sécrétion d'un grand nombre de facteurs protéiques dont la compréhension des conséquences physiopathologiques est primordiale, bien que paradoxalement, peu étudiée. Notre travail s'est ainsi focalisé sur l'étude de l'influence de p53 sur la modulation de la sécrétion de protéines par les tumeurs du poumon.

Ce manuscrit de thèse s'articule autour de plusieurs axes. Dans une première partie, nous décrirons les données bibliographiques en rapport avec le contexte scientifique de cette étude. Différents chapitres seront ainsi abordés : la protéomique et les approches d'analyses différentielles, le cancer du poumon, la tumorogenèse, la protéine p53 et le secretome. Dans un second temps, nous présenterons les résultats obtenus. Ils ont nécessité la mise en place d'un modèle d'étude cellulaire et d'un procédé d'analyse protéomique adéquats. La transposition vers un modèle *in vivo* permettra de d'investiguer la présence et la régulation des protéines d'intérêt au niveau de la tumeur et des fluides biologiques. Ces résultats seront discutés et les nombreuses perspectives qui en découlent détaillées. Ces travaux participeront à une meilleure compréhension du rôle primordial joué par l'altération de p53 dans la modulation de l'environnement tumoral. De plus, ils permettront d'évaluer l'utilisation du secretome comme source potentielle pour la découverte de marqueurs tumoraux.

DONNEES BIBLIOGRAPHIQUES

DONNEES BIBLIOGRAPHIQUES

CHAPITRE 1 : Les approches et techniques en protéomique

I Introduction

1. Définitions de la protéomique

C'est en 1994 que le terme « protéome », défini alors comme l'ensemble des PROTEines exprimées par un génOME, a été pour la première fois employé par Marc Wilkins, puis repris dans ses publications [1, 2]. La science nommée « protéomique » était née. Aujourd'hui, le terme « protéomique » est utilisé de façon plus spécifique pour décrire l'étude de l'ensemble des protéines exprimées par un compartiment cellulaire (ex : les protéines nucléaires, les protéines mitochondriales, les protéines membranaires…), une cellule, un tissu ou un organisme à un temps donné et selon un ensemble de conditions connues [3]. Son but n'est pas seulement d'identifier toutes les protéines mais aussi de créer un réseau multidimensionnel des protéines d'une cellule intégrant leurs interactions et localisations. Le protéome d'une cellule est dynamique car directement lié à l'environnement immédiat dans lequel il est étudié. En réponse à un signal interne ou externe, les protéines peuvent subir des modifications post-traductionnelles, changer de compartiment cellulaire (sécrétion,…), être synthétisées ou dégradées. En conséquence, l'étude du protéome doit s'effectuer dans un cadre bien défini.

2. Origine de la protéomique

Le véritable point de départ de la protéomique se situe en 1975, quand simultanément O'Farrell [4], Klose [5] et Scheel [6] publièrent les méthodes donnant naissance à l'électrophorèse bidimensionnelle. Leurs études ont permis, pour la première fois, la séparation d'un mélange complexe de protéines respectivement chez *Escherichia coli*, la souris et le cochon de Guinée. Ainsi, 1100 composés protéiques ont, par exemple, pu être séparés par O'Farrell chez *Escherichia coli* [4]. Néanmoins, les limites technologiques de l'époque ne permettaient pas l'identification de ces protéines et faisaient donc de l'électrophorèse bidimensionnelle une technique purement descriptive. Il faudra attendre l'avènement de la spectrométrie de masse des peptides dans les années 1990 pour que la protéomique prenne « son envol ». Le développement simultané de différentes techniques de séparation des protéines (électrophorèse bidimensionnelle, chromatographie liquide) de bio-informatique (bases de données, moteurs de recherche,…) ou de spectrométrie de masse (amélioration de sensibilité, résolution, précision de masse,…) contribua

grandement à cet essor de la protéomique [7]. Depuis, le nombre de publications en protéomique croît de façon exponentielle (Figure 1). La première publication protéomique date de 1995 et l'on pouvait en dénombrer plus de 25 000 au premier trimestre 2008.

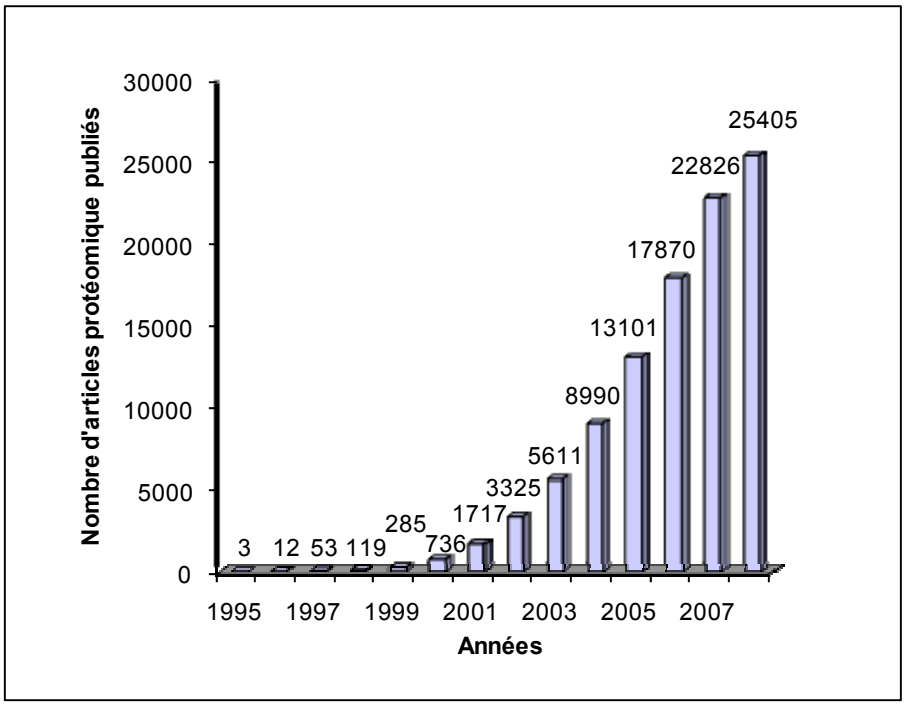


Figure 1 : Nombre d'articles publiés en protéomique de 1995 à 2008.
La recherche est réalisée dans PubMed (www.ncbi.nlm.nih.gov/sites/entrez) en utilisant les mots clés "proteomics", "proteome" ou "proteomic".

3. La Protéomique et les autres « Omics »

Le premier séquençage complet d'un génome, celui de la bactérie *Haemophilia influenzae,* a été publié en1995 [8]. Depuis, de nombreux génomes ont été obtenus dont le très attendu génome humain en 2003 [9]. Néanmoins, le séquençage de génomes ne permet pas de connaître les protéines exprimées par une cellule à un instant donné de sa vie. En effet, de nombreux processus augmentent la diversité chimique des protéines comme l'épissage alternatif, l'édition des ARN, les modifications post-traductionnelles (plus de 100 connues [10]), la protéolyse, les interactions des protéines entre elles,… (Figure 2).

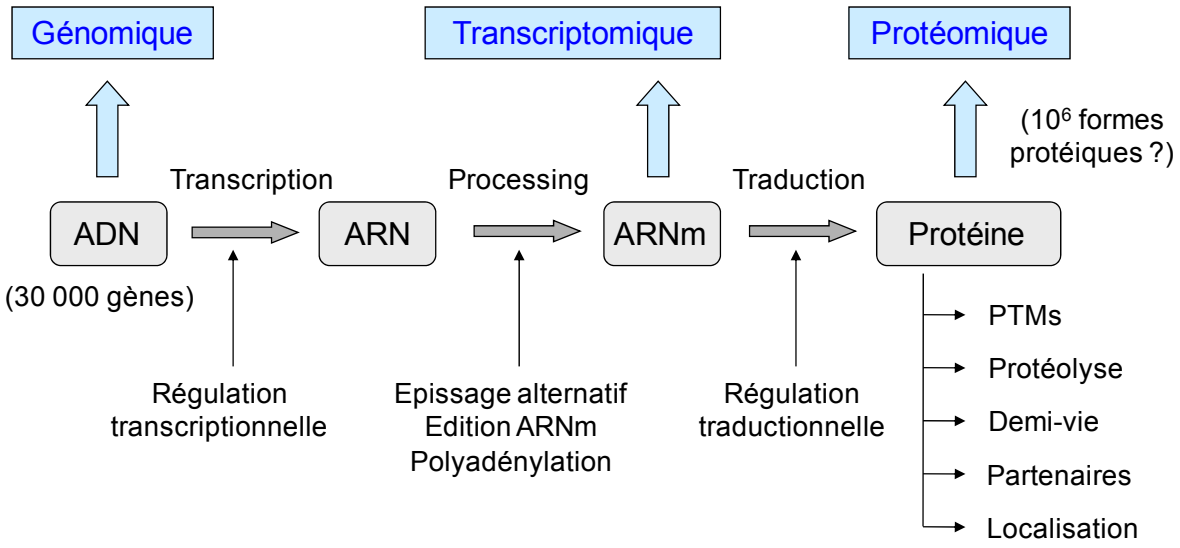


Figure 2 : Les « Omics » et le dogme central de la biologie.

La complexité du protéome dépasse ainsi de très loin la séquence du génome. Le nombre de gènes eucaryotes relativement faible (20488 gènes identifiés dans le génome humain [11]), ainsi que leur stabilité durant toute la vie d'une cellule, met l'accent sur le rôle primordial joué par la modulation du profil protéique dans la complexité des différents organismes. De plus, il est clairement défini que ce sont les protéines et non les gènes qui accomplissent les différentes fonctions cellulaires. L'exemple le plus connu est celui exposé par Friedrich Lottspeich en 1999 dans sa revue intitulée « One genome - different proteomes » [12]. Il prend comme exemple la chenille qui se transforme en papillon (Figure 3). Avec les mêmes gènes, cet insecte présente une morphologie clairement différente selon son stade de développement. Ces différences sont liées à une modulation du protéome et non du génome.

La transcriptomique étudie les ARNm, c'est-à-dire le niveau d'expression des gènes et permet donc d'appréhender les protéines exprimées par la cellule. Néanmoins, plusieurs études, comme celle de Gygi *et al.*, démontrent l'absence de relation linéaire entre transcriptome et protéome [13]. Les auteurs concluent que les niveaux d'ARNm exprimés ne corrèlent pas toujours avec les niveaux protéiques et que l'analyse des transcrits n'est donc pas suffisante.. La protéomique est un domaine à part entière, différent mais complémentaire de la génomique ou la transcriptomique [14] (Figure 2).

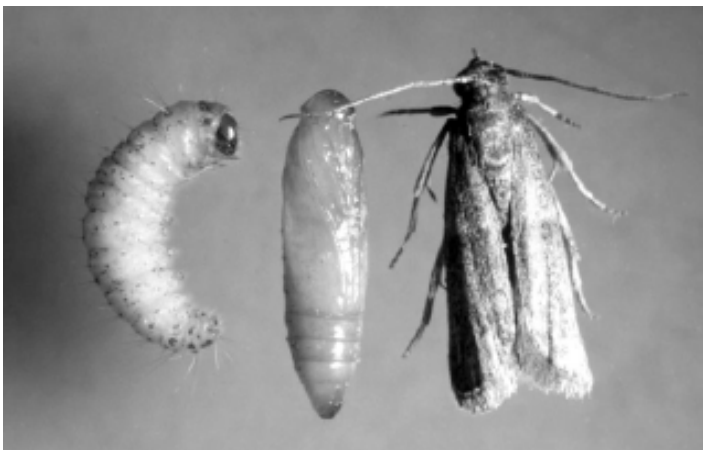


Figure 3 : Un même génome mais différents protéomes.
Exemple des différents stades de développement du papillon (d'après Lottspeich F [12])

La protéomique n'est pas pour autant une finalité en soi. Connaître les protéines exprimées est important mais pas suffisant car celles-ci ne sont pas isolées dans la cellule mais en interaction avec de nombreux autres facteurs : on parle alors d'interactome.

La métabolomique étudie les métabolites, ce qui englobe notamment les réseaux fonctionnels de protéines impliquées dans les voies métaboliques.

Les « omics » s'intéressent donc à un domaine d'étude très vaste et ambitieux, où de nombreuses découvertes sont encore à faire.

II La Spectrométrie de Masse en Protéomique

1. Généralités

1.1. Définition

La spectrométrie de masse est une technique d'analyse chimique permettant de détecter et d'identifier des molécules d'intérêt par mesure de leur masse monoisotopique. Elle est aussi utile pour caractériser la structure chimique des molécules en les fragmentant. Son principe réside dans la séparation en phase gazeuse de molécules chargées (ions) en fonction de leur rapport masse-sur-charge (*m/z*). La spectrométrie de masse est utilisée dans pratiquement tous les domaines scientifiques : physique, astrophysique, chimie en phase gazeuse, chimie organique, dosages, biologie, médecine...

1.2. Un intérêt majeur en protéomique

La spectrométrie de masse est devenue une approche quasi incontournable pour l'identification des protéines ainsi que leur caractérisation (modifications post-traductionnelles, chimiques, mutations d'acides aminés, quantification…). Les techniques d'ionisation MALDI et ESI, couplées à des analyseurs de plus en plus résolutifs et sensibles, rendent possible l'analyse de quantité très faible d'échantillon (quelques femtomoles à sub-femtomoles) [15].

1.3. Principe général

Un spectromètre de masse permet de mesurer le rapport masse-sur-charge (*m/z*) des ions formés à partir de l'échantillon analysé. Il est toujours constitué des parties suivantes (Figure 4) :

- ***Une source d'ionisation :*** elle permet l'ionisation des molécules à analyser, ainsi que leur émission en phase gazeuse (vaporisation, sublimation ou désorption).
- ***Un analyseur :*** il permet de séparer les ions en fonction de leur rapport *m/z*.
- ***Un détecteur :*** il permet de compter les ions en leur associant un rapport *m/z*.
- ***Un système d'enregistrement (informatique) :*** Il permet de traiter le signal enregistré et de visualiser les résultats sous forme de spectres.

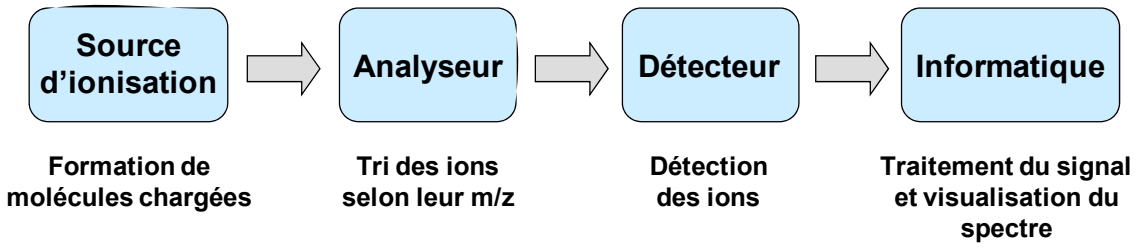


Figure 4 : Principe général d'un spectromètre de masse.

1.4. Les termes essentiels

1.4.1. Masse moléculaire

La masse moléculaire d'un ion est déduite de la valeur *m/z* du pic moléculaire observé sur le spectre de masse grâce à la formule $m/z = (M+zH^+)/z$ où M est le masse moléculaire, z l'état de charge et H^+ le nombre de protons. Ce pic correspond à un ion qui contient tous les atomes de la molécule étudiée, sans qu'il y ait de rupture de liaison. La masse moléculaire correspond donc à la composition élémentaire (formule brute) de l'ion moléculaire.

L'existence d'isotopes se traduit par la présence de plusieurs pics moléculaires. Une molécule observée n'est donc pas visible sous la forme d'un pic moléculaire unitaire mais d'un groupe de pics, appelé massif isotopique (Figure 5).

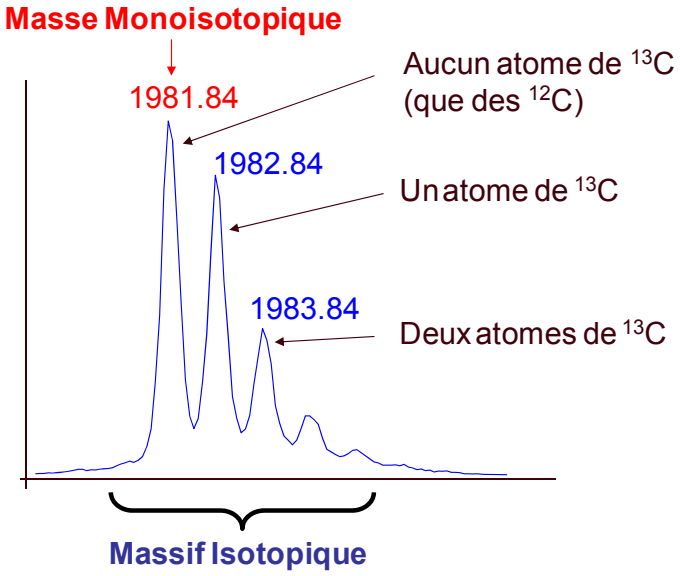


Figure 5 : Exemple d'un massif isotopique observé sur un spectre MS.

La masse monoisotopique correspond à la masse du premier pic du massif isotopique, c'est-à-dire celle qui ne prend en compte que les masses des isotopes les plus stables (^{12}C, ^{1}H, ^{16}O, ^{14}N,…).

Les masses moléculaires sont mesurées en Dalton (Da). Un Dalton correspond à 1/12 de la masse d'un atome de ^{12}C (c'est-à-dire 1,66. 10^{-27} Kg).

1.4.2. Résolution

La résolution d'un analyseur en spectrométrie de masse correspond au rapport M/FWHM où M est la masse du pic en Da et FWHM (Full-Width Half-Maximum) la largeur du pic à mi-hauteur. Plus la résolution est élevée, plus le pic sera fin et plus la mesure de masse sera précise (Figure 6).

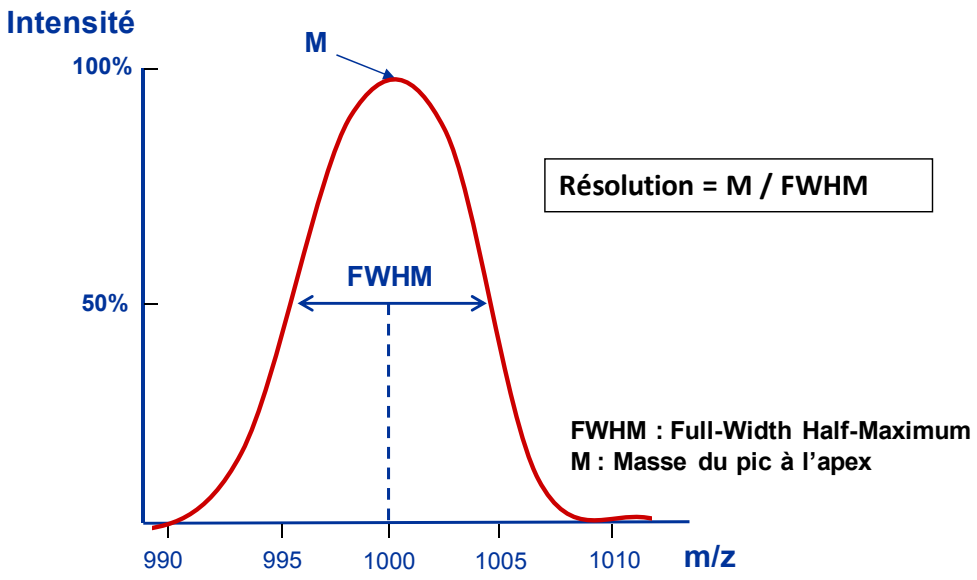


Figure 6 : La résolution d'un pic en spectrométrie de masse.

Voici quelques ordres de grandeur selon l'analyseur utilisé : quadrupôle ou trappe à ions (résolution de 100 à 5000), TOF (résolution de 100 à 20 000), FT-ICR (résolution de 1000 à 1 000 000).

1.4.3. Précision de mesure

La précision de mesure traduit la différence de masse entre la masse exacte de la molécule et sa masse observée en spectrométrie de masse. Elle est généralement exprimée en ppm (parts per million) par la formule suivante : 10^6 x (Mo-Mt)/Mt où Mo est la masse observée en Da et Mt la masse théorique.

1.5. La Spectrométrie de masse en mode tandem

1.5.1. Principe

La spectrométrie de masse en mode tandem (ou MS/MS) combine deux analyseurs, afin d'étudier les caractéristiques de fragmentation des échantillons. Elle permet de compléter la mesure de masse par des informations structurales et notamment réaliser le séquençage d'un peptide d'intérêt.

Le principe général d'analyse est énoncé Figure 7. Les ions formés dans la source pénètrent dans le premier analyseur MS1. Parmi ceux-ci, un ion précurseur est sélectionné et spécifiquement fragmenté dans une cellule de collision contenant des molécules sous forme gazeuse. Les ions fragments issus de la décomposition de l'ion précurseur sont alors étudiés dans un deuxième analyseur MS2.

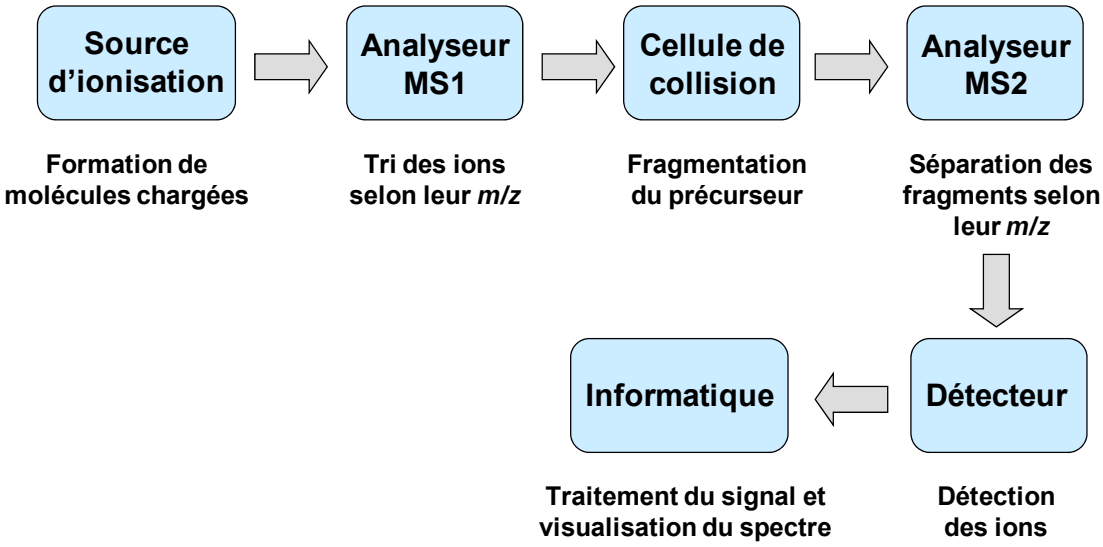


Figure 7 : La spectrométrie de masse en mode tandem.

La formation des ions fragments suit des règles précises. Les groupements chimiques ont des caractéristiques propres de fragmentation, et le profil d'un spectre MS/MS donne en général des informations complètes et fiables sur la structure du composé précurseur (Figure 8). En biologie, ce type d'expérience est notamment utilisé pour déterminer des informations de séquence des peptides, ou rechercher et caractériser des modifications post-traductionnelles.

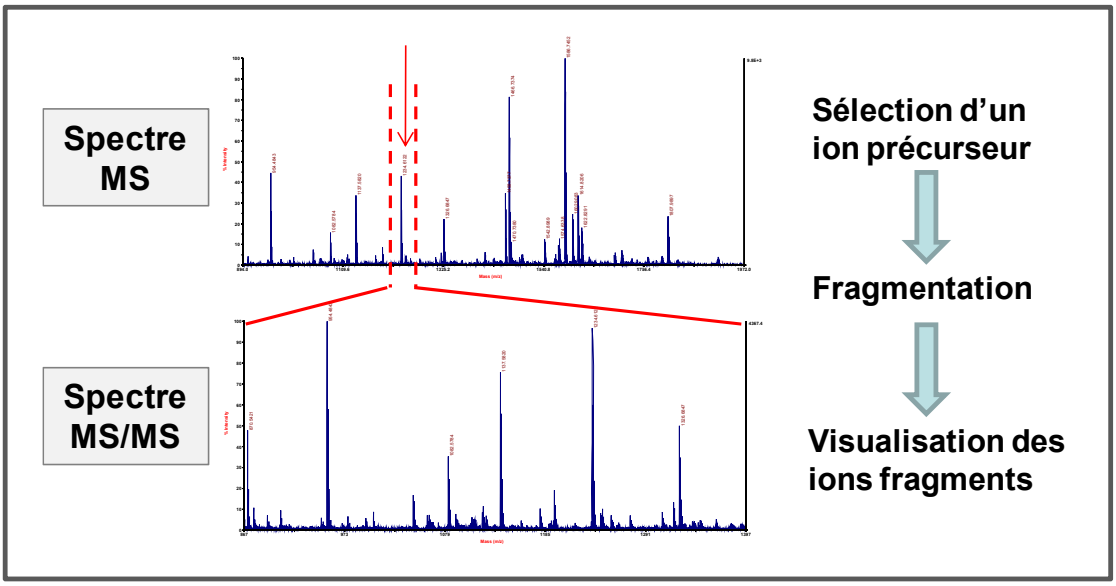


Figure 8 : Exemple d'un spectre MS et d'un spectre MS/MS obtenu après fragmentation d'un ion précurseur d'intérêt.

1.5.2. Fragmentation des peptides en MS/MS

Un peptide est un copolymère d'acides aminés reliés par des liaisons peptidiques -CO-NH- (Figure 9). Il possède une extrémité N-terminale (groupement $–NH_2$) et C-terminale (groupement –COOH).

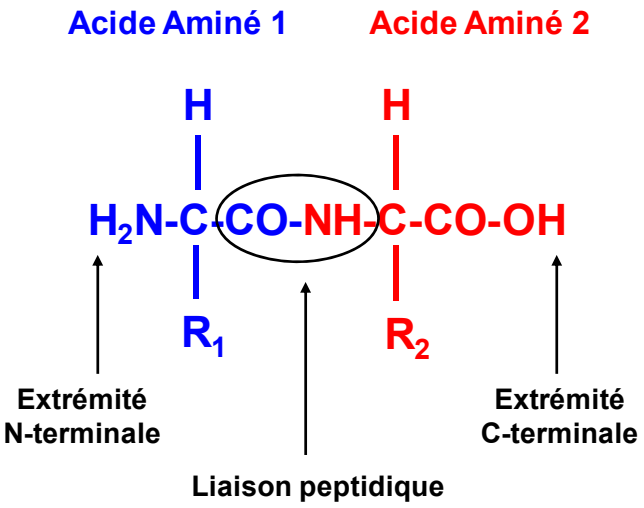


Figure 9 : Structure chimique d'un dipeptide.

En spectrométrie de masse de type MS/MS, les peptides se fragmentent majoritairement au niveau des liaisons peptidiques. Cela aboutit à la formation d'ions b si la charge est conservée par le fragment N-terminal et d'ions y si elle est conservée par le fragment C-terminal. La Figure 10 montre les types de fragments générés lors de la fragmentation d'un peptide en accord avec la nomenclature de Roepstorff-Fohlmann-Biemann [16, 17].

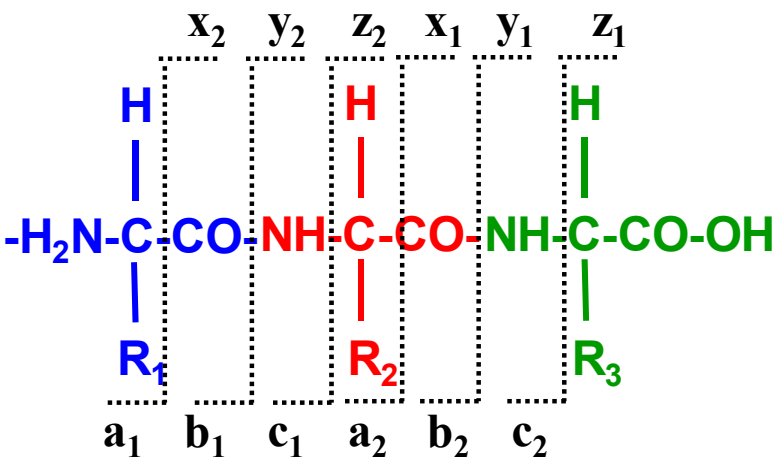


Figure 10 : Types de fragments observés en MS/MS pour un tripeptide.

On observe trois types de fragments :

- ***Les fragments N-terminaux*** : ils sont obtenus lorsque la charge de l'ion est conservée par la partie N-terminale du peptide. Les fragments obtenus sont nommés a, b ou c selon le lieu de fragmentation par rapport à la liaison peptidique (Figure 10).
- ***Les fragments C-terminaux*** : ils sont obtenus lorsque la charge de l'ion est conservée par la partie C-terminale du peptide. Les fragments obtenus sont nommés x, y ou z selon le lieu de fragmentation (Figure 10).

Ces fragments sont notés avec un indice qui précise le nombre de résidus portés par l'ion fragment.

- ***Les fragments internes*** : ils sont formés d'une combinaison d'un clivage de type a et b, et appelés ions immoniums (Figure 11). On les retrouve dans la région de faible *m/z* du spectre MS/MS.

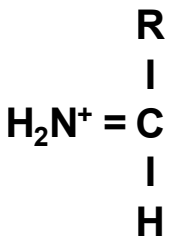


Figure 11 : Structure chimique d'un ion immonium.

1.6. ESI et MALDI

En 1988, deux nouvelles méthodes d'ionisation des molécules ont été décrites : l'électrospray (ESI pour ElectroSpray Ionization) et le MALDI (Matrix-Assisted Laser Desorption/Ionization). Ces deux techniques ont grandement contribué à l'utilisation de la spectrométrie de masse pour l'analyse de biopolymères dont les protéines et peptides en permettant l'ionisation douce de ces molécules.

Ces approches sont présentées dans les deux paragraphes suivants.

2. La Spectrométrie de Masse à source MALDI

La spectrométrie de masse par Désorption et Ionisation Laser Assistée par Matrice (MALDI) a été développée en 1988 par Karas et Hillenkamp [18], ainsi que par Tanaka [19]. Cette technique consiste à co-cristalliser les analytes avec une matrice organique. Leur excitation avec un laser va génèrer des molécules chargées généralement analysées par un analyseur en temps de vol (TOF).***Principe d'ionisation***

2.1.1. Les matrices

2.1.1.1. Les types de matrices

La première étape d'une analyse par MALDI-MS consiste à co-cristalliser une faible quantité de molécules à analyser avec une grande quantité de molécules de matrice.

Cette matrice doit :

- Etre une petite molécule capable de co-cristalliser avec l'échantillon.
- Protéger l'échantillon en absorbant l'essentiel de l'énergie du laser.
- Etre donneur ou accepteur de protons pour permettre l'ionisation de l'analyte.

Pour l'analyse protéomique, trois principales matrices acides sont classiquement utilisées (Tableau 1 et Figure 12) :

- La CHCA (α-cyano-4-hydroxycinnamic acid),
- L'acide sinapinic (3,5-dimethoxy-4-hydroxy cinnamic acid),
- La DHB (2,5-dihydroxybenzoic acid).

Abréviation	Nom	Structure	MW (Da)	Applications
CHCA	alpha-cyano-4-hydroxycinnamic acid		189,04	peptides, protéines, lipides, oligonucléotides
SA	3,5-dimethoxy-4-hydroxy cinnamic acid (acide sinapinique)		224,07	peptides, protéines, glycoprotéines
DHB	2,5-dihydroxybenzoic acid		154,03	peptides, protéines, glycoprotéines, polymères synthétiques, lipides, carbohydrates

Tableau 1 : Matrices classiques pour l'analyse protéomique par MALDI-TOF

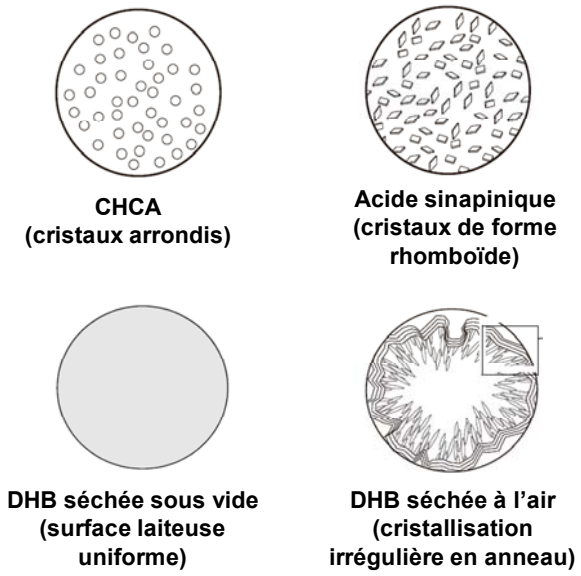


Figure 12 : Vues au microscope des différents types de cristallisation obtenus selon la matrice utilisée.

Le choix de la matrice est un élément primordial de l'analyse de type MALDI-TOF [20]. La CHCA est la matrice la plus couramment employée pour l'analyse de peptides. L'analyse des protéines de haut poids moléculaire se fait préférentiellement avec l'acide sinapinique. La DHB est quant à elle plutôt utilisée pour étudier des modifications post-traductionnelles telles que les phosphorylations ou glycosylations.

2.1.1.2. Les modes de dépôt

Trois principaux modes de dépôt ont été décrits [21] : les dépôts en goutte séchée, en couche mince ou en sandwich.

2.1.1.2.1. Goutte séchée

C'est la méthode la plus simple à mettre en œuvre. Une solution acide de matrice est mélangée à l'analyte et une goutte de cette préparation est déposée sur la surface de la cible du spectromètre de masse. Le mélange est soit réalisé au préalable (pré-mix), soit effectué sur la cible elle-même. La cristallisation obtenue par séchage de la goutte à température ambiante peut être accélérée par chauffage, ventilation ou aspiration sous vide. Ce mode de dépôt est adapté à l'analyse des petites molécules et des peptides. Par contre, la répartition des cristaux est plutôt hétérogène.

2.1.1.2.2. Couche mince

La matrice est ici mise en solution dans un solvant très volatil, tel que l'acétone. Une goutte de solution de matrice est déposée sur la cible et l'évaporation rapide du solvant produit une couche mince et homogène de cristaux. L'échantillon est alors déposé sur cette couche mince et évaporé.

Ce type de dépôt est plus sensible car les peptides sont répartis de manière uniforme. Par contre, le dépôt est plus fragile, et le nombre de tirs laser est donc limité.

2.1.1.2.3. Sandwich

Ce mode consiste à déposer l'échantillon en sandwich entre deux couches de matrices. Une couche mince de matrice est d'abord appliquée, puis l'analyte est déposé par dessus et recouvert d'une deuxième couche de matrice. Cette méthode génère des cristaux de petite taille très homogènes et permet d'augmenter notablement la résolution et la reproductibilité des différentes analyses. Elle est adaptée pour l'analyse des macromolécules et protéines.

La technique MALDI est relativement tolérante aux sels (par rapport à l'ESI). Cette tolérance est notamment due à la nature solide du dépôt, ainsi qu'à la faible solubilisation de la plupart des matrices dans l'eau. On peut ainsi facilement laver un dépôt MALDI par l'ajout d'une goutte d'eau.

2.1.2. Le laser

Le laser permet l'ionisation de l'échantillon en phase gazeuse. Son intensité doit être réglée au minimum nécessaire afin de diminuer la dégradation de l'échantillon et le bruit de fond. L'ionisation des peptides se produit dans la gamme de longueur d'onde UV 266-420 nm en présence de CHCA [22]. Or, les peptides absorbent la lumière jusqu'à 280 nm à cause des cycles aromatiques des acides aminés tyrosine et tryptophane. En conséquence, les longueurs d'onde utilisées sont souvent supérieures à 280 nm. Le laser le plus commun est le Nd:YAG (355 nm), utilisé par les spectromètres de masse MALDI-TOF/TOF de dernière génération.

2.1.3. Les cibles

Les supports sur lesquels sont réalisés les dépôts sont appelés cibles (Figure 13). Elles sont généralement en acier inoxydable, conductrices et traitées en surface de manière à être chimiquement inertes par rapport à l'échantillon.

Figure 13 : Cibles MALDI 384 spots (à droite) et LC-MALDI 1600 spots (à gauche) pour MALDI-TOF/TOF 4800 (Applied Biosystems).

2.1.4. Le mécanisme d'ionisation MALDI

L'ionisation MALDI s'effectue sous vide (≈ 10^{-7} mbar). Les ions sont produits en phase gazeuse sous l'effet d'une impulsion laser qui irradie la surface cristalline de la matrice entourant l'analyte (Figure 14). Le faisceau laser excite et ionise la matrice par l'absorption d'un photon. La matrice se dissocie alors et se sublime en entraînant l'échantillon. L'ionisation s'effectue en deux phases. La première consiste en l'ionisation de la matrice résultant de l'absorption d'un photon émis par le laser. L'ionisation de l'échantillon se réalise dans un second temps grâce au transfert en phase gazeuse d'un proton de la matrice à l'analyte [23].

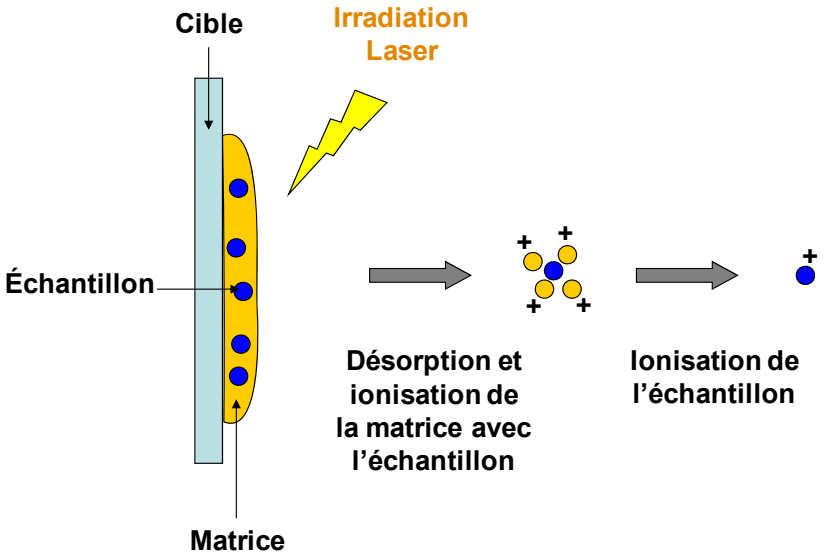


Figure 14 : Le processus d'ionisation MALDI.

L'ionisation MALDI génère généralement des ions monochargés (contrairement à l'ionisation ESI). Les peptides ou protéines ionisés sont ensuite accélérés par l'application d'une différence de potentiel entre la cible et la lentille d'extraction en vue d'être séparés par l'analyseur en temps de vol.

2.2. L'analyseur en temps de vol (TOF)

2.2.1. Principe général

Les spectromètres de masse à source MALDI sont généralement couplés à un analyseur en temps de vol (TOF pour « Time Of Flight »). Ce dernier est adapté à la nature pulsée de la désorption MALDI. Il s'agit d'un cylindre vide dans lequel règne une basse pression avec à une extrémité la source MALDI et à l'autre un détecteur (Figure 15). Son principe est le suivant : si des ions sont accélérés, par un même potentiel, d'une position initiale donnée et à un temps initial donné et

qu'on les fait voler dans un tube sous vide, alors ces ions seront séparés en fonction de leur rapport *m/z*.

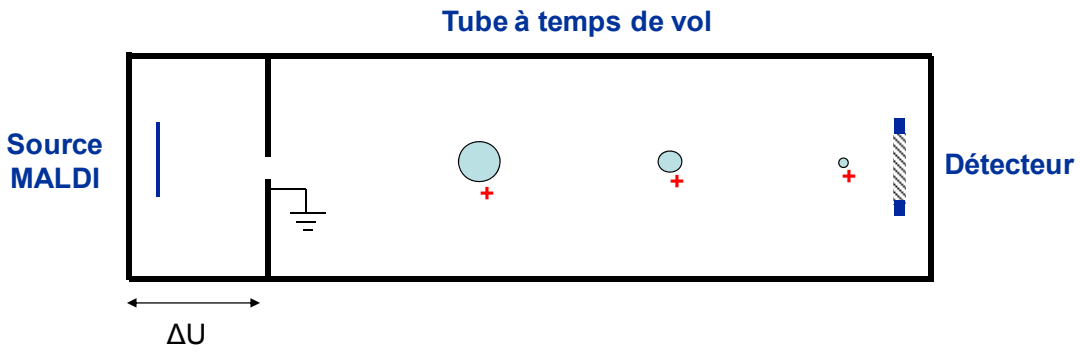


Figure 15 : L'analyseur en temps de vol (TOF).

Le ratio *m/z* d'un ion est directement proportionnel au carré de son temps de vol dans l'analyseur selon la formule : $m/z = 2t^2K/L^2$ où m est la masse de l'analyte, z sa charge, t son temps de vol, K son énergie cinétique et L la longueur du tube.

La Figure 16 schématise le principe du MALDI-TOF/TOF 4800 de la société Applied Biosystems dont est équipée la plate-forme protéomique médicale de Grenoble. Cet appareil est utilisable selon trois modes différents : mode linéaire, mode réflectron et mode MS/MS.

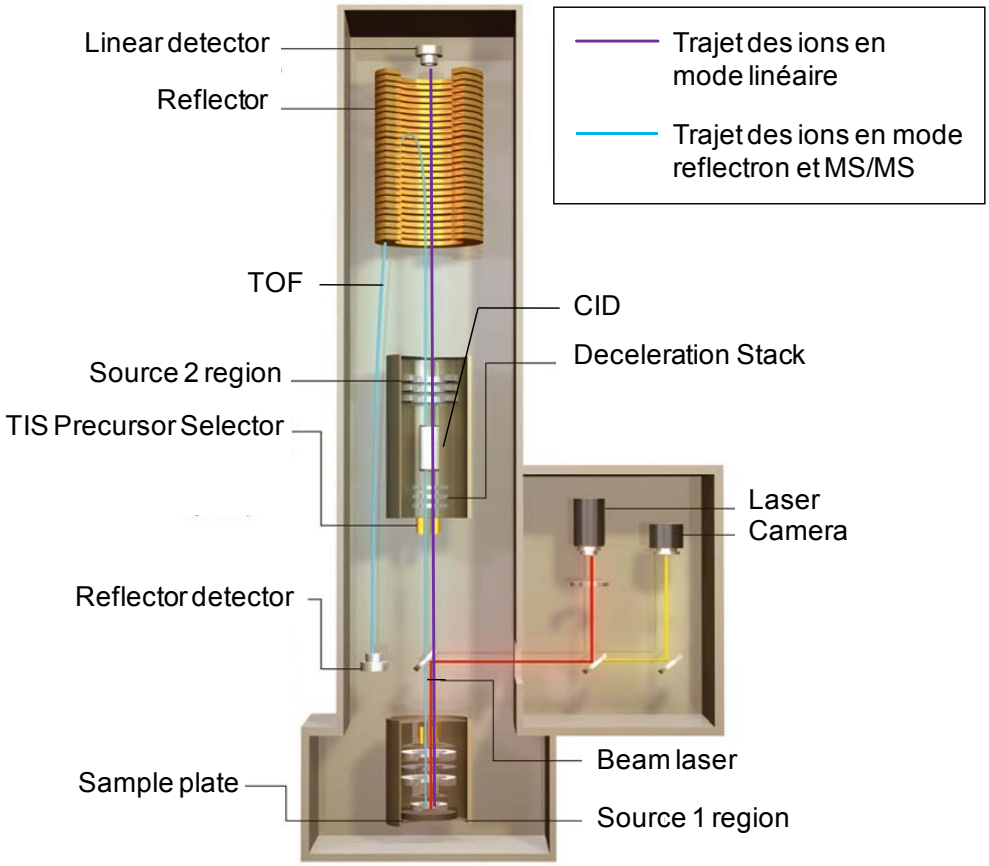


Figure 16 : Les trois modes d'utilisation du MALDI TOF/TOF 4800 (Applied Biosystems).

2.2.2. Mode linéaire

Le mode linéaire correspond au mode d'utilisation le plus simple. Les ions sont générés de façon pulsée en fonction des impulsions du laser utilisé pour l'ionisation. Cet ensemble d'ions est accéléré et expulsé de la source par un gradient de potentiel décroissant à une vitesse dépendante de leur *m/z* (Figure 15).

Ce mode est adapté à l'étude des molécules de haut poids moléculaire (protéines) et possède une sensibilité correcte. Elle a, par contre, une faible résolution (de l'ordre de 1000) à cause d'une :

- distribution d'énergie cinétique différente pour des ions de même rapport *m/z*.
- distribution spatiale des ions différente à cause de la désorption laser et de l'épaisseur du dépôt.

Ces problèmes sont partiellement résolus par la mise en place d'une extraction retardée qui diminue au maximum la dispersion des ions générés par le processus MALDI (dispersion en énergie cinétique initiale et spatiale). Celle-ci repose sur l'utilisation d'une différence de potentiels entre la cible et une lentille située avant le TOF (Figure 17). Elle permet de réaligner les ions de vélocité différente pour qu'ils pénètrent simultanément dans le tube de temps de vol et augmente ainsi sensiblement la résolution des pics [24, 25].

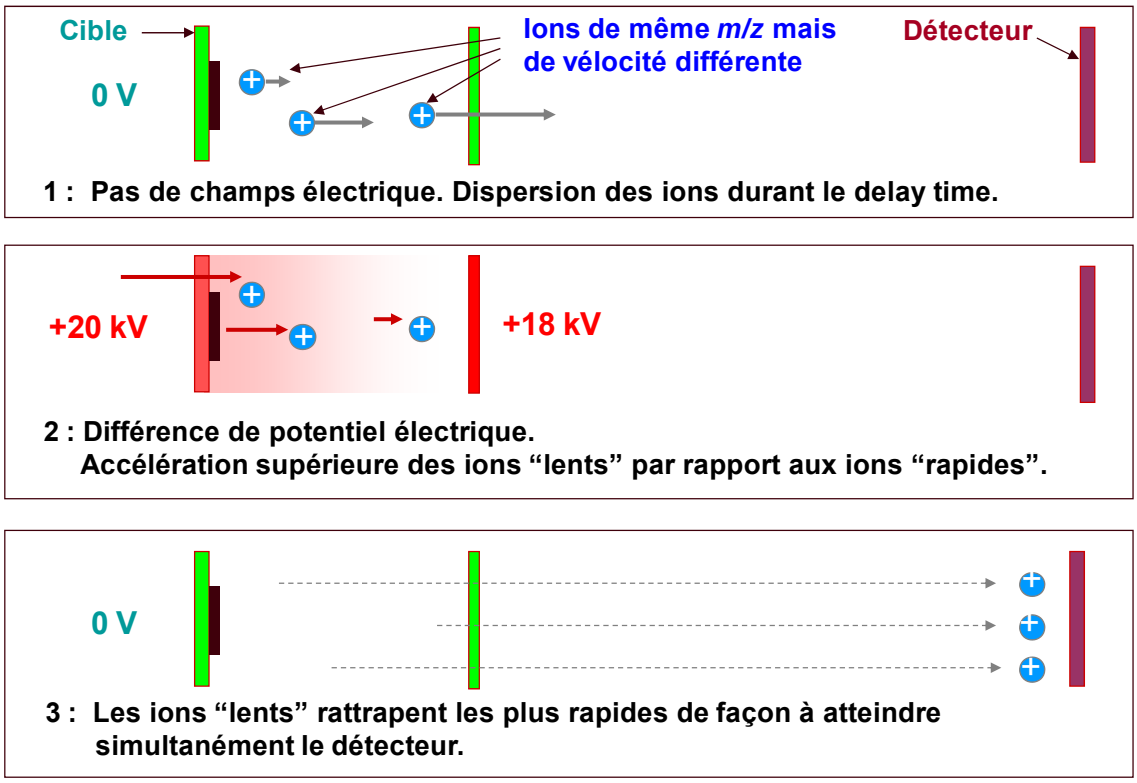


Figure 17 : L'extraction retardée.

2.2.3. Mode réflectron

Le mode réflectron permet de remédier en partie au problème de basse résolution observé en mode linéaire [26]. Il repose sur l'utilisation d'un miroir électrostatique, appelé réflectron, qui impose un champ électrique s'opposant à la progression des ions (Figure 18). Ces derniers vont pénétrer plus ou moins profondément dans le champ électrique du réflectron en fonction de leur énergie cinétique initiale (vélocité). Lorsque les ions atteignent leur zéro d'énergie, ils font demi-tour et sont ré-accélérés en sens inverse. Cette stratégie permet de réaligner en sortie de réflectron des ions de même rapport *m/z* mais de vélocité différente. Ceux-ci vont alors frapper en phase un second détecteur, aboutissant ainsi à une meilleure résolution des pics donnant accès aux masses isotopiques (résolution > 20000 avec le MALDI-TOF/TOF 4800). La mesure de masse est aussi plus précise (< 10 ppm). En contrepartie, ce mode n'est utilisable que pour des petites masses (< 5000 Da). Il est donc adapté à l'analyse de peptides mais ne permet pas d'étudier des protéines entières. Le réflectron corrige donc la dispersion en énergie cinétique et allonge la distance de vol sans augmenter la taille de l'instrument.

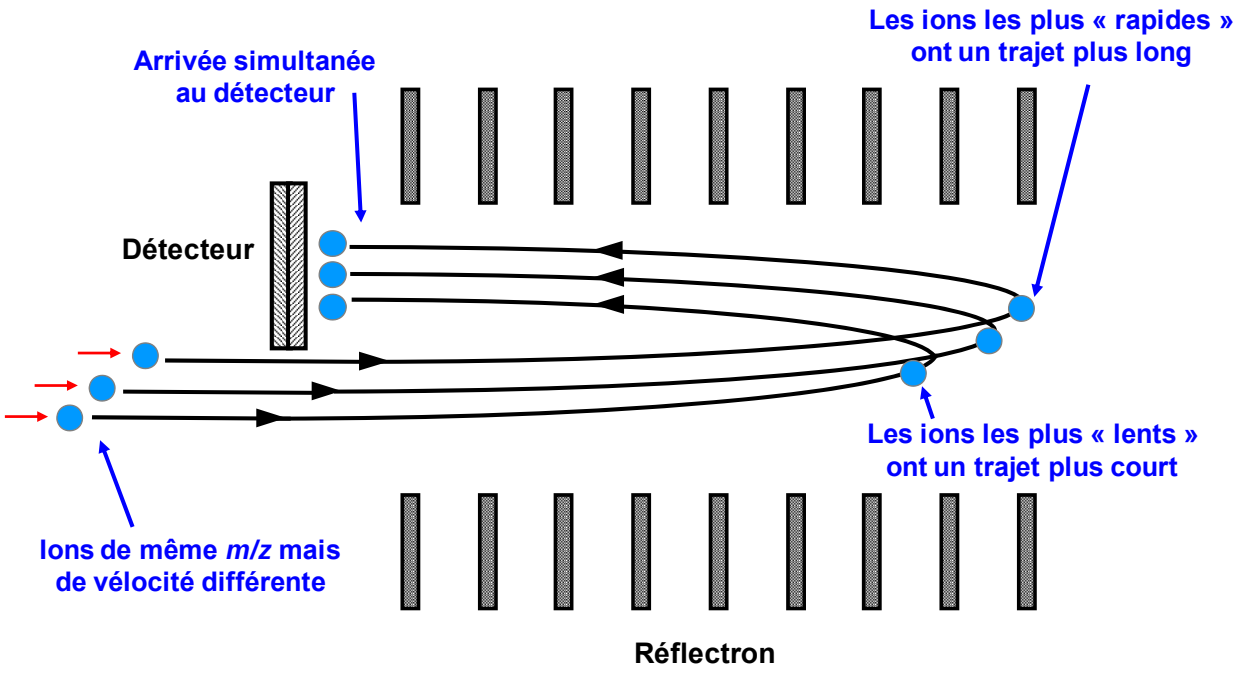


Figure 18 : Principe du mode réflectron.

2.2.4. Mode MS/MS

Le mode MS/MS permet de fragmenter les peptides d'intérêt pour obtenir des informations de séquence. Le tube de temps de vol est alors séparé en deux par une chambre de collision (CID : Collision Induced Dissociation), contenant un gaz inerte (Ar, He, N_2, Air) (Figure 16). Le premier TOF sépare les différents peptides de l'échantillon. A l'entrée de la chambre de collision, le TIS (Time Ion Selector) sélectionne uniquement l'ion précurseur d'intérêt qui va alors entrer en collision avec les molécules de gaz du CID (Figure 19). Un transfert d'énergie se produit durant cette collision induisant la fragmentation de l'ion précurseur et la génération de fragments chargés. Ceux-ci sont ré-accélérés pour être séparés dans le « deuxième » TOF.

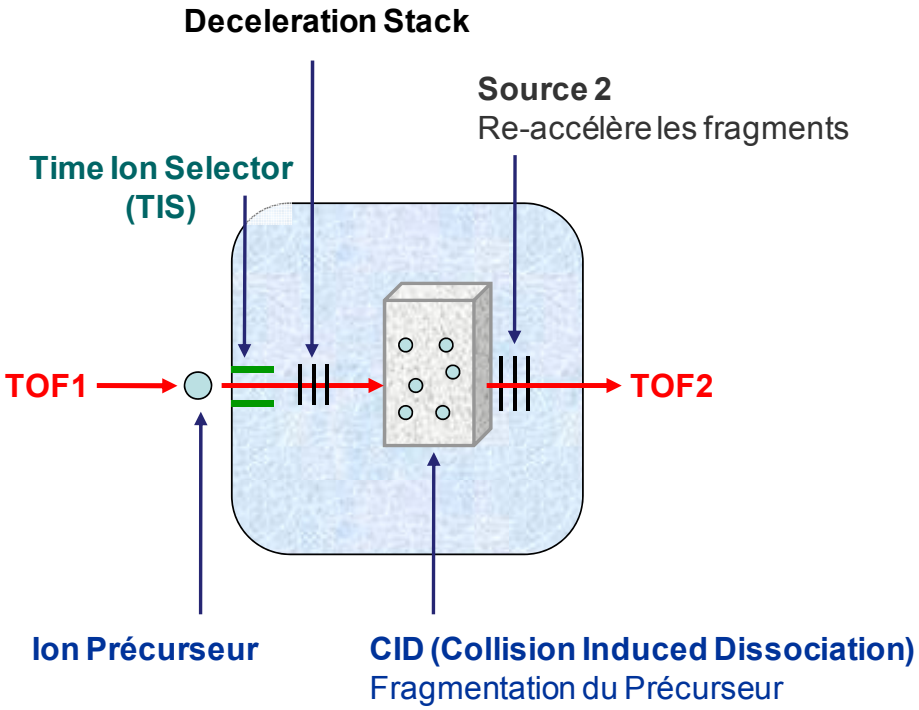


Figure 19 : La fragmentation peptidique par CID (Collision Induced Dissociation) du spectromètre de masse MALDI TOF/TOF 4800.

Ce mode a grandement élargi le champ d'application des spectromètres de masse à source MALDI. Ceux-ci peuvent désormais être couplés en « off-line » avec un système chromatographique dans le but de réaliser des analyses LC-MS/MS jusque là uniquement accessibles aux spectromètres de masse à source electrospray.

3. La Spectrométrie de Masse à source ElectroSpray

Le phénomène d'ionisation par électrospray (ESI) a été décrit par les travaux de Dole en 1968 [27]. Au milieu des années 80, ce mode d'ionisation est couplé à un analyseur quadrupolaire par l'équipe de Fenn, et vaudra à ce dernier un prix Nobel en 2002 [28]. Il démontre lors de ses expériences que des molécules non volatiles peuvent être ionisées en phase gazeuse sans fragmentation ou destruction. La source ESI génère des ions en phase gazeuse à partir d'analytes en solution et à pression atmosphérique. Cette technique peut donc être couplée directement (en « on-line ») à une chromatographie liquide (LC-MS) [29].

3.1. Principe d'ionisation

L'échantillon est introduit dans un capillaire porté à haut potentiel électrique. Ce champ électrique intense provoque, en sortie du capillaire, la formation d'un nuage de gouttelettes chargées (spray) qui traversent simultanément un gradient de champ électrique et un gradient de pression en direction de l'analyseur. Pendant ce trajet, la taille des gouttelettes diminue par évaporation du solvant et explosions "coulombiennes" successives (divisions spontanées de la gouttelette chargée en gouttelettes plus petites, dûes à une charge en surface très élevée) (Figure 20). Les spectres ESI obtenus en mode positif présentent en général un ensemble de pics correspondant aux ions multichargés de type $[M+nH]^{n+}$ où M correspond à la masse moléculaire de l'analyte, et n au nombre de charges portées par la molécule ionisée.

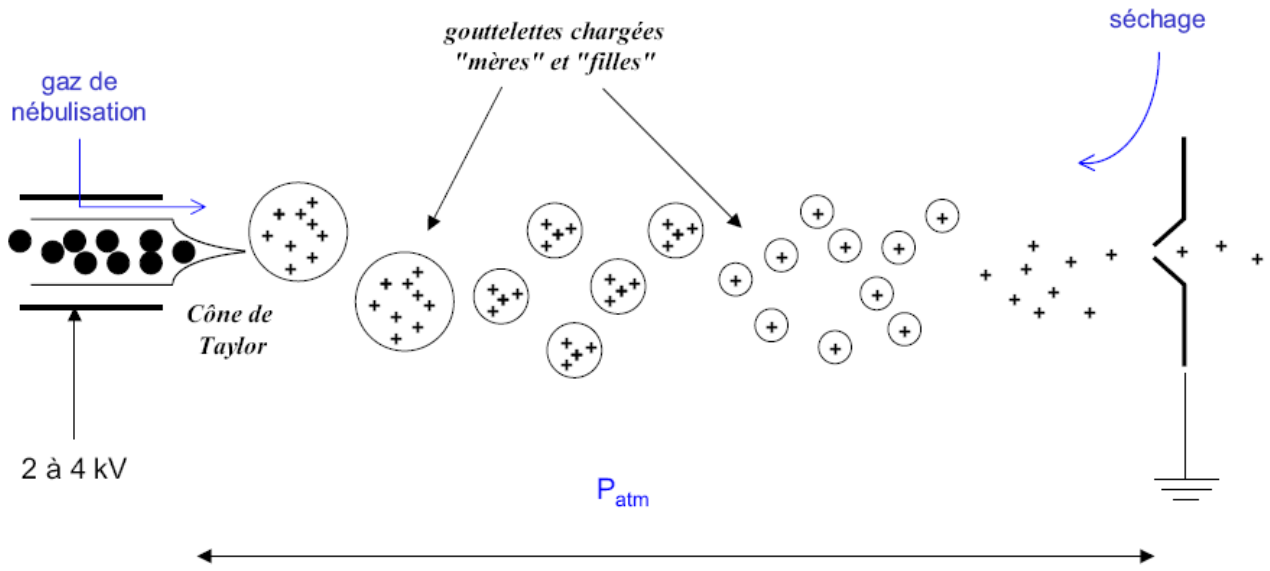


Figure 20 : L'ionisation de type Electrospray (ESI).

Puisque l'électrospray fournit des ions multichargés, l'analyse de protéines de hautes masses moléculaires (supérieures à 100 kDa) peut être réalisée avec des analyseurs de gamme de masse limitée. La source électrospray est souvent couplée à un analyseur quadrupolaire Il a aussi été couplé à d'autres analyseurs tels que la trappe ionique, l'analyseur à résonance ionique cyclotronique par transformée de Fourier, le temps de vol, la trappe orbitale ou l'analyseur à secteur magnétique (Tableau 2).

Instruments	Résolution	Gamme de masse (Da)
MALDI TOF/TOF	20 000	500 000
ESI-Q-TOF	10 000	40 000
ESI-trappe ionique	5 000	6 000
ESI-FT-ICR	>100 000	5 000
ESI-Orbitrap	100 000	5 000
ESI-Q-TRAP	5 000	3 000

Tableau 2 : Caractéristiques générales de différents spectromètres de masse utilisés en protéomique.

3.2. Les analyseurs compatibles avec une source ESI

3.2.1. Quadrupôle

Le quadrupôle, développé par Paul [30], est l'analyseur le plus couramment utilisé en couplage avec les sources électrospray. Il est constitué de quatre barreaux métalliques parallèles disposés par paire et portés à un potentiel opposé (Figure 21). Cet analyseur est à balayage et utilise la stabilité de la trajectoire des ions entre les barreaux pour les séparer selon leur rapport *m/z*.

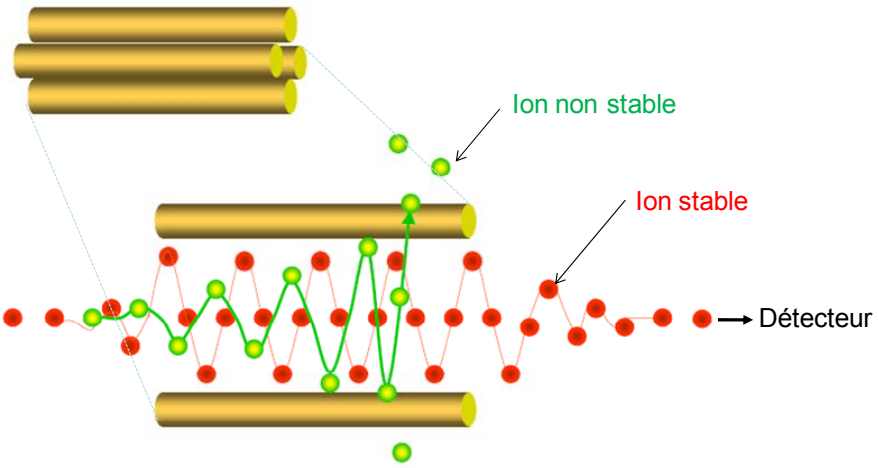


Figure 21 : L'analyseur quadrupôle.

Ses avantages principaux sont sa souplesse d'utilisation, sa vitesse de balayage satisfaisante, ainsi que son adaptabilité à différentes interfaces permettant le couplage avec la chromatographie gazeuse ou liquide. La gamme de masse analysée est toutefois faible (généralement comprise entre 10 et 4000 Da) et sa résolution unitaire (environ 1000).

3.2.2. Temps de Vol (TOF)

Un analyseur en temps de vol peut être aussi couplé à une source électrospray (voir paragraphe II.2.2 pour le principe du TOF). Le couplage entre une source électrospray et un analyseur TOF nécessite une injection orthogonale pulsée des ions.

3.2.3. Trappe ionique

C'est un piège à ions où l'analyse et la détection des ions s'effectuent dans un même espace, suivant des séquences temporelles successives. Il est constitué de trois électrodes à section hyperbolique : une électrode annulaire encadrée par deux électrodes (d'entrée et de sortie) qui forment les calottes supérieure et inférieure du dispositif [31] (Figure 22). En mode MS, le fonctionnement de la trappe à ions se décompose en deux étapes : le piégeage des ions dans la trappe et leur éjection vers le détecteur. En mode MS/MS, il est composé de quatre étapes : le piégeage des ions, l'isolement des ions parents à fragmenter, leur fragmentation et l'éjection des ions fragments vers le détecteur.

Cet analyseur a l'avantage d'être très sensible et robuste. Il présente, par contre, une résolution relativement faible (environ 2000). On observe aussi, en présence d'une grande quantité d'ions dans la cellule, un « effet charge » faussant les mesures.

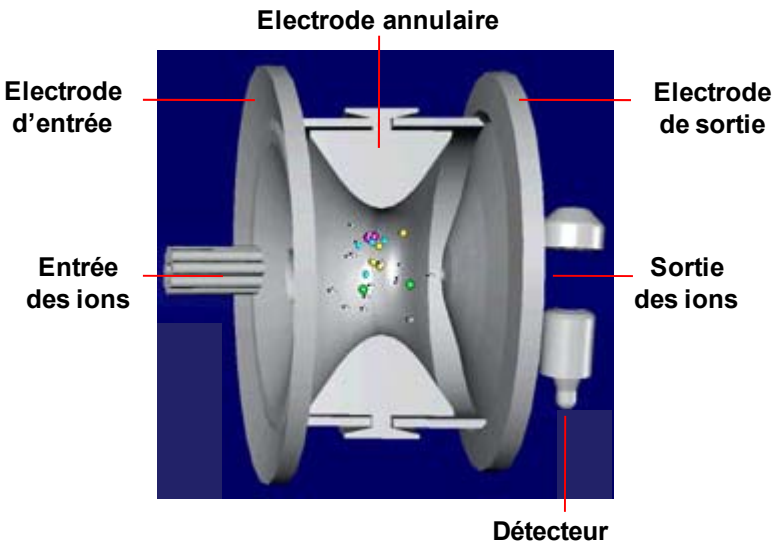


Figure 22 : L'analyseur trappe ionique.

3.2.4. La trappe orbitale

La trappe orbitale est composée d'une électrode creuse, à l'intérieur de laquelle est placée co-axialement une deuxième électrode en forme de fuseau. La forme particulière de ces deux électrodes permet d'appliquer un champ électrostatique piégeant les ions. Les ions d'un *m/z* donné gravitent sur la même trajectoire circulaire. Le courant induit par ces oscillations permet, grâce à la transformée de Fourier, d'accéder aux mesures de *m/z*. Cet appareil est actuellement commercialisé par la société Thermo sous le nom « Orbitrap ».

La précision des mesures est particulièrement bonne (1-2 ppm) et la résolution (jusque 100 000) rivalise avec celle du FT-ICR. La trappe orbitale est principalement utilisée en spectrométrie de masse en tandem, associée à un piège linéaire.

3.2.5. Analyseurs à résonance ionique cyclotronique par transformée de Fourier (FT-ICR)

L'analyseur à résonance ionique cyclotronique est composé d'une cellule ICR (Ion Cyclotron Resonance) constituée de six plaques sous tension, isolées les unes des autres. Un ion se déplace dans le champ magnétique selon une trajectoire circulaire à une fréquence dépendant de sa masse. Ce mouvement cyclotron peut être excité (résonance), détecté et analysé par transformée de Fourier.

Cet analyseur a l'une des meilleures résolutions qui soient (>100 000), ainsi qu'une très bonne précision de mesure de masse (~ 1 ppm). C'est par contre un équipement très lourd en termes d'utilisation et d'entretien. On observe aussi, en présence d'une grande quantité d'ions dans la cellule, un « effet charge » faussant les mesures.

3.2.6. L'hybride Q-TOF

Le Q-TOF est un appareil hybride constitué de deux analyseurs, un quadrupôle (Q) et un temps de vol (TOF) séparés par une cellule de collision. Les ions sont sélectionnés par le quadrupôle, puis accélérés et fragmentés dans la cellule de collision. Les fragments sont refocalisés grâce à un hexapôle et analysés en TOF.

Son principal avantage réside en l'augmentation de la sensibilité et de la résolution des pics.

3.2.7. Triple quadrupôle

Un triple quadrupôle découle de l'association de deux analyseurs quadrupolaires en série(Q1 et Q3), séparés par une cellule de collision constituée d'un quadrupôle plus court (Q2). Cette combinaison permet de travailler en modes MS ou MS/MS. Lors d'une acquisition en mode MS/MS, la cellule de collision est remplie d'un gaz inerte (argon par exemple) à pression élevée

(10^{-2} torr). Différents modes d'analyse peuvent être entrepris selon l'information recherchée. Le mode descendant permet d'obtenir des informations structurales. Les modes ascendant et perte de neutre sont utiles pour rechercher des ions possédant des caractéristiques communes (séquence, modifications post-traductionnelles,…). Le quatrième mode (Multiple Reaction Monitoring ou MRM), dérivé du mode descendant, est dédié à la quantification.

- en *mode descendant*, l'ion précurseur est sélectionné en Q1, fragmenté en Q2 et ses ions fragments analysés en Q3.
- en *mode ascendant*, le Q1 balaie une gamme de masse tandis que Q2 est focalisé sur une valeur de *m/z* particulière. Tous les ions précurseurs capables de donner un ion fragment de *m/z* similaire sont détectés en Q3.
- en *mode perte de neutre*, Q1 et Q3 balaient simultanément une gamme de masse avec un décalage de masse constant. Le spectre établi présente alors tous les ions précurseurs capables de se fragmenter en générant une masse neutre égale au décalage imposé.
- en *mode MRM*, l'ion précurseur d'intérêt est sélectionné en Q1 et fragmenté en Q2, comme pour le mode descendant. En revanche, Q3 est focalisé sur l'ion fragment d'intérêt. Ce mode présente une double sélectivité : au niveau de l'ion précurseur et de l'ion fragment. En outre, Q1 et Q3 sont à tensions constantes, ce qui augmentee la sensibilité de détection par rapport aux autres modes de balayage. En conséquence, le mode MRM est très utile pour la quantification.

III Les méthodes d'analyse protéomique

Une analyse protéomique est constituée généralement de trois étapes principales : la préparation, la séparation, et l'identification/caractérisation des protéines de l'échantillon (Figure 23).

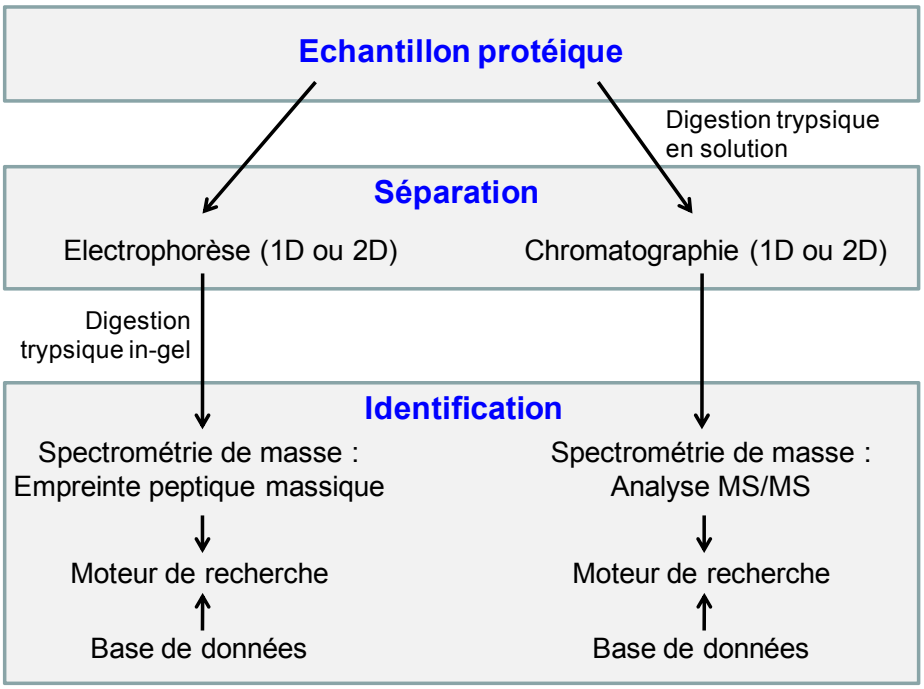


Figure 23 : Aperçu général des deux grandes voies d'analyse protéomique.

1. La préparation de l'échantillon

Une préparation adéquate de l'échantillon est essentielle à la réussite ultérieure de toute analyse protéomique. Différentes techniques permettent d'extraire, enrichir, concentrer les protéines, ou d'éliminer les contaminants tels que les sels,…[32]. Idéalement, ces traitements doivent être aussi réduits que possible pour minimiser les pertes ou dégradations des protéines de l'échantillon.
La revue de Ahmed dresse une liste assez exhaustive des différentes approches développées [33]. Quelques étapes clés sont décrites ci-dessous.

1.1. La lyse des tissus et cellules par homogénéisation

L'homogénéisation vise à détruire les cellules ou tissus pour récupérer l'extrait protéique. En protéomique, différentes méthodes sont définies :

- ***La lyse mécanique*** est réalisée grâce à des appareils de broyage tels que les rotor-stator, les potters (ex : Warring blender, Polytron) ou par PCT (pressure cycling technology) tel que le French press [34, 35].

- ***La lyse par ultrasons*** (sonication) repose sur un effet piézoélectrique générant des ondes ultrasoniques de haute énergie détruisant les cellules [36].
- ***La lyse par cycles de congélation/décongélation*** utilise l'effet destructeur de la formation de cristaux de glace durant la congélation. Elle alterne généralement des cycles azote liquide / bain marie [36].
- ***La lyse par choc osmotique*** utilise des détergents dont les plus communs sont le Triton X-100, le Tween 80 et le NP 40 [36].

1.2. La solubilisation des protéines

Chaque protéine possède une composition chimique, une structure et des propriétés qui lui sont propres : caractère hydrophobe des protéines membranaires, structure quaternaire multimérique,… Par conséquent, une protéine extraite de son environnement cellulaire n'est pas forcément soluble et risque de précipiter. De par l'hétérogénéité biochimique des protéines, il n'existe pas de technique idéale de solubilisation. Les tampons utilisés sont généralement une combinaison d'agents chaotropiques (urée, thiourée,…), de détergents ioniques, non-ioniques ou zwitterion (SDS, CHAPS, Triton X-100,…), d'agents réducteurs de ponts disulfures (DTT, TEAB,…) et d'inhibiteurs de protéases [37]. La solubilisation peut aussi être améliorée par agitation ou sonication.

1.3. La protection contre la protéolyse

On estime à près de 7000 le nombre de protéases codées par le génome humain. Si celles-ci ne sont pas inhibées, leur libération durant la lyse cellulaire peut entrainer une dégradation des protéines de l'échantillon. Cette dégradation peut être minimisée en réalisant une extraction rapide. De plus, il est souvent recommandé d'ajouter des inhibiteurs de protéases spécifiques tels que le PMSF, l'AEBSF, l'EDTA ou des coktails d'anti-protéases [36].

1.4. L'élimination des contaminants

Les sels, tampons, détergents, acides nucléiques, polysaccharides, lipides et particules fréquemment présentes dans les échantillons ont tendance à interférer avec la séparation des protéines, inhiber leur digestion ou entrer en conflit avec l'analyse MS. Ces contaminants doivent donc, dans la mesure du possible, être éliminés. Il faut néanmoins être attentif à ce que le traitement nécessaire ne soit pas plus perturbant pour l'échantillon que son absence !! [38].

- Les sels peuvent être éliminés par dialyse, ultrafiltration, gel filtration, précipitation au TCA ou à l'acétone, ou par des kits commerciaux (clean-up) [36].
- Les détergents sont classiquement éliminés par dialyse, chromatographie phase inverse ou précipitation des protéines.
- Les lipides sont présents en grande quantité dans les fluides biologiques, tels que le sang, et se lient aux protéines. Cette interaction diminue la stabilité des protéines et peut affecter leur pI ou PM. Les tampons à base de CHAPS utilisés en électrophorèse bidimensionnelle ou la précipitation à l'acétone éliminent efficacement les lipides [36].
- La présence de polysaccharides, surtout chargés, doit être évitée car ils peuvent s'agréger, obstruer les pores de gels SDS-PAGE ou faire précipiter les protéines [39]. La précipitation des protéines au TCA, à l'acétone, au sulfate d'ammonium suivi par une ultracentrifugation est efficace pour les éliminer.

1.5. La concentration des protéines

Certains échantillons protéiques très dilués nécessitent d'être concentrés avant de pouvoir être analysés. Il existe différentes méthodes : la précipitation au TCA ou à l'acétone, la chromatographie phase inverse (C18, Zip-tip), l'ultrafiltration.

1.6. L'enrichissement de protéines

1.6.1. Le fractionnement sub-cellulaire

Le fractionnement sub-cellulaire permet de cibler et purifier les protéines d'un sous-protéome particulier (cytosol, organites, noyaux,…). Il est généralement réalisé par centrifugation sur gradient de densité. Néanmoins, obtenir des fractions complètement pures est difficile.

1.6.2. La purification de complexes protéiques et microdomaines

Les protéines fonctionnent rarement seules mais sont plutôt organisées en unités fonctionnelles qui sont différentes en taille, en nombre de partenaires interagissant et en stabilité. Ces complexes multi-protéiques peuvent être isolés et purifiés par des techniques d'affinité (co-IP, TAP-tag..), de pull-downs, chromatographiques, d'électrophorèse en condition native (blue native PAGE),…

1.6.3. Les méthodes d'extractions séquentielles

Différents kits commerciaux (ProteoExtract Subcellular Proteome Extraction Kit de Calbiochem) offrent la possibilité d'extraire les protéines de façon séquentielle afin d'obtenir 4 sous-protéomes

successifs enrichis : les protéines cytosoliques, les protéines des membranes et organites, les protéines nucléaires et les protéines du cytosquelette.

1.6.4. Les fluides biologiques

Beaucoup des marqueurs diagnostiques, pronostiques, ou de suivi de thérapie utilisés en clinique sont retrouvés dans les fluides biologiques. Les deux fluides les plus communs sont le sang et l'urine, deux sources non invasives pour le patient. Ces fluides contiennent un mélange protéique très complexe avec une gamme de polarité, d'hydrophobicité et de taille de plusieurs ordres de magnitude. Dans l'idéal, l'analyse protéomique devrait être réalisée sur l'échantillon natif afin d'éviter toute perte artificielle ou biais expérimentaux. Néanmoins, ces fluides ont une très grande complexité gênant leur analyse directe en protéomique. Des colonnes de déplétion sont, par exemple, souvent utilisées pour éliminer les protéines majoritaires du plasma (albumine,…). Néanmoins, ces approches sont limitées quant à leur efficacité et reproductibilité.

2. La séparation de l'échantillon

Deux grandes approches de séparation sont utilisées en protéomique : les techniques électrophorétiques et chromatographiques.

2.1. Electrophorèse bidimensionnelle

2.1.1. Le principe

L'électrophorèse bidimensionnelle a vu le jour au début des années 1970 [40]. Elle repose sur une séparation des protéines en deux étapes (Figure 24).

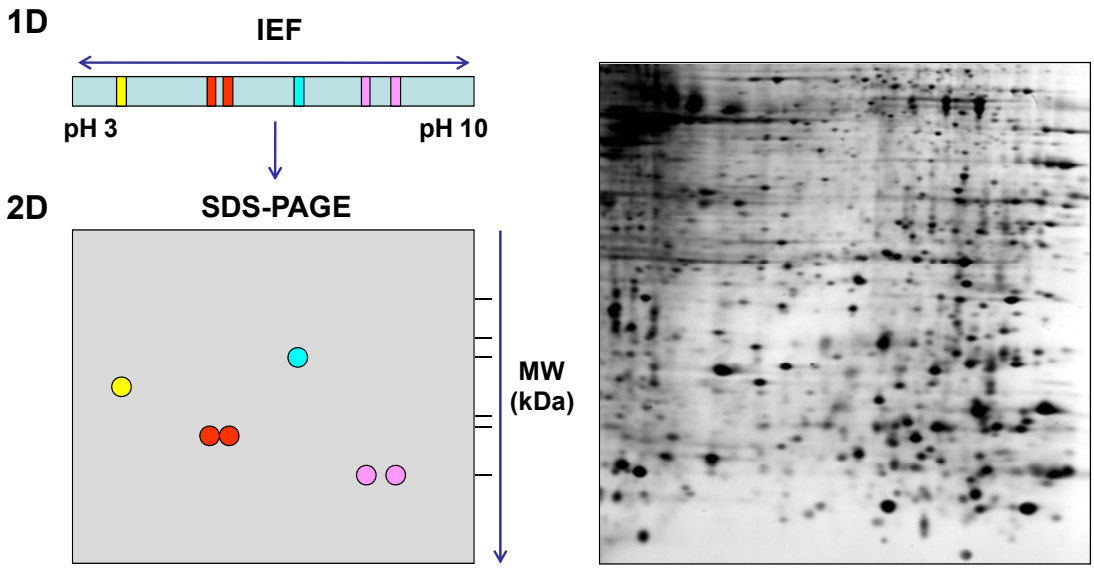


Figure 24 : Principe de l'électrophorèse bidimensionnelle.

La première dimension consiste en une isoélectrofocalisation (IEF). Les protéines sont séparées en fonction de leur point isoélectrique (pI) sur une strip de gel contenant un gradient immobilisé de pH (IPG). Sous l'effet d'un champ électrique, les protéines migrent dans le gel en fonction de leur charge nette. Quand elles arrivent à la valeur de pH qui équivaut à leur pI, leur charge globale devient nulle et elles s'immobilisent. On peut ainsi focaliser les protéines sur une gamme allant de 3 à 10 unités de pH ou des gammes plus ciblées. La seconde dimension est une électrophorèse SDS-PAGE qui sépare les protéines selon leur masse moléculaire. La combinaison de ces deux séparations permet d'obtenir une « carte 2D » où chaque spot représente une forme protéique de l'échantillon.

2.1.2. La coloration

L'étape de coloration permet de révéler les protéines séparées lors de l'électrophorèse 2D [41]. De nombreuses colorations sont possibles dont les trois plus communes sont celles au bleu de Coomassie, au nitrate d'argent et par fluorescence. Chaque colorant a des caractéristiques propres et le choix dépend donc de l'application. Les paramètres à prendre en compte sont la sensibilité de détection, l'homogénéité de la coloration (selon le type de protéines), sa linéarité (proportionnalité avec la quantité de protéines) et sa compatibilité avec l'analyse ultérieure en spectrométrie de masse (Tableau 3).

- ***Le bleu de Coomassie*** : bien que peu sensible (détection de l'ordre du µg à une centaine de ng avec le bleu colloïdal), il présente les avantages d'être un colorant linéaire et homogène et donc d'être parfaitement adapté aux études comparatives. Cette coloration est par ailleurs compatible avec l'analyse en spectrométrie de masse.
- ***Le nitrate d'argent*** : c'est le colorant le plus sensible (détection de l'ordre du ng de protéines) mais il présente les inconvénients d'être très peu linéaire (coloration dépendant des propriétés intrinsèques de chaque protéine) et peu compatible avec l'analyse en spectrométrie de masse. Néanmoins, des protocoles adaptés ont permis de surmonter en partie ce problème [42].
- ***La fluorescence*** (exemple : SYPRO ruby, orange,…) : les fluorochromes présentent actuellement le meilleur compromis. Ils permettent une coloration linéaire, homogène, compatible avec la spectrométrie de masse et d'une sensibilité quasiment aussi bonne que la coloration au nitrate d'argent. Elle nécessite néanmoins un appareillage UV adéquat pour sa visualisation.

	Nitrate d'argent	Bleu de Coomassie colloïdal	Fluorescence (SYPRO)
Sensibilité	+++	+	++
Homogénéité	++	+++	++
Linéarité	+	++	+++
Compatibilité avec la MS	+	++	++

Tableau 3 : Les pré-requis pour une coloration optimale de gels SDS-PAGE.

2.1.3. Avantages et inconvénients

L'électrophorèse bidimensionnelle est une technique très utile pour l'analyse protéomique. Elle permet de séparer des mélanges très complexes de protéines avec un bon pouvoir de résolution et de distinguer les différents isoformes d'une même protéine (modifications post-traductionnelles, clivages…).

Elle présente néanmoins certains inconvénients. C'est une appproche longue à mettre en œuvre et pas toujours reproductible. Elle est peu adaptée à l'analyse de protéines membranaires. Enfin, elle ne permet d'analyser qu'une « fenêtre » du protéome, les protéines de poids moléculaires extrêmes (< 10 kDa ou > 100 kDa) et de pI extrêmes (pI < 4 ou > 9) n'étant que très difficilement observables.

2.2. Le fractionnement OFFGEL

L'OFFGEL est une approche électrophorétique récemment développée par Rossier et son équipe [43] et actuellement commercialisée sous le nom d'OFFGEL Fractionator par la société Agilent Technologies. Elle sépare les protéines ou peptides en fonction de leur point isoélectrique, mais à la différence d'une isoélectrofocalisation classique, cette séparation se fait en milieu liquide.

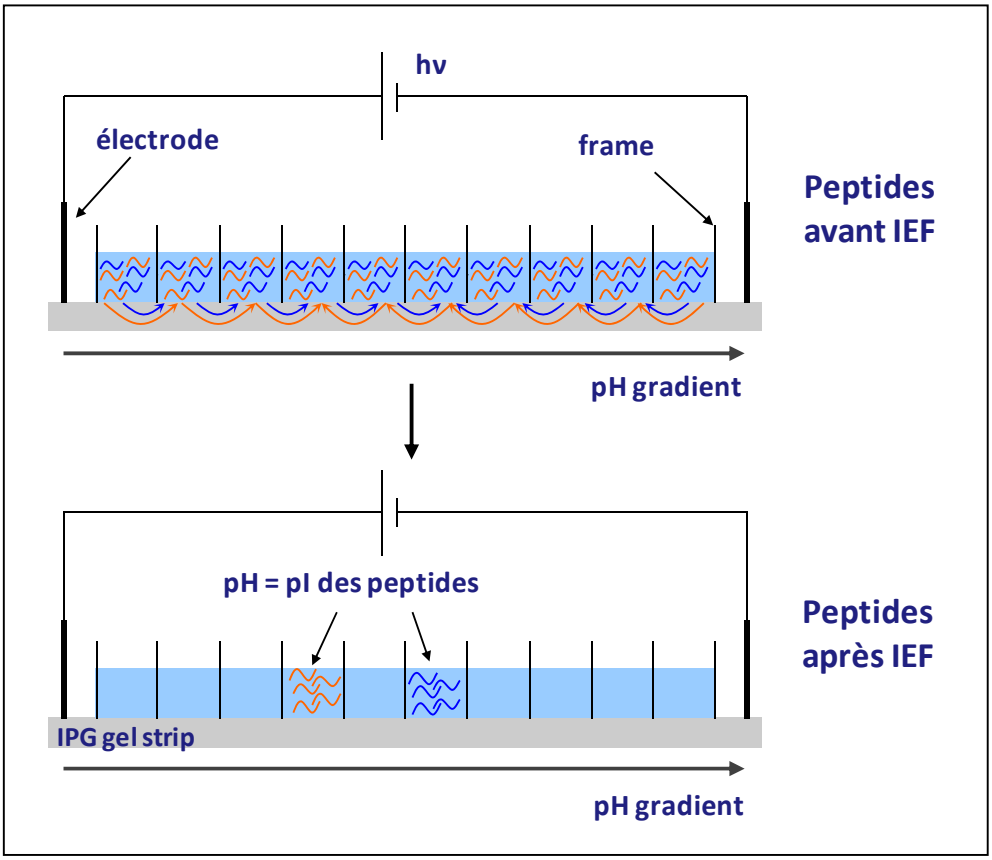


Figure 25 : L'isoélectrofocalisation OFFGEL.

Son principe est énoncé Figure 25. Une bandelette de gel (strip), contenant un gradient de pH immobilisé (le même type que celles utilisées en IEF classique), est positionnée dans une des pistes du plateau. Un support (« frame ») vient se clipper par-dessus de façon à s'accoler au gel et délimiter jusqu'à 24 fractions étanches tout au long de la strip. L'échantillon protéique ou peptidique, dilué dans un tampon contenant des ampholytes, est alors réparti dans chacune des fractions. L'application d'un courant induit la migration des peptides (ou protéines), la strip servant de « pont » reliant les différentes fractions les unes aux autres. A la fin de l'IEF, chaque fraction liquide contient l'ensemble des peptides (ou protéines) ayant un pI correspondant à la gamme de pH de la fraction. La résolution est de 0.1 à 0.6 unités de pH par fraction.

L'avantage majeur de cette approche est de pouvoir récupérer l'échantillon fractionné sous forme liquide. Cela la rend compatible avec des méthodes de type chromatographie en phase inverse, pour réaliser une séparation multidimensionnelle. L'OFFGEL tolère la charge d'une grande quantité d'échantillon (jusqu'à 5 mg), ce qui permet d'avoir, après la focalisation, suffisamment de matériel biologique pour les analyses ultérieures. Enfin, elle est tolérante à des quantités raisonnables de sels [44].

2.3. Chromatographie liquide haute pression (CLHP)

L'abréviation CLHP signifie indifféremment chromatographie liquide à haute performance ou à haute pression. C'est une technique de choix pour la séparation de mélanges peptidiques (ou protéiques) [45].

2.3.1. Introduction

La chromatographie liquide haute pression (CLHP ou HPLC pour High Performance Liquid Chromatography) a été développée dans le milieu des années 1970 pour diminuer les temps de séparation en augmentant la pression. A la fin des années 1970, le développement de nouvelles chimies des colonnes, comme la chromatographie liquide en phase inverse, a permis d'améliorer la qualité de séparation des analytes. Dans les années 1980, l'informatisation et l'automatisation des systèmes ont permis d'augmenter la fiabilité et la facilité d'utilisation des CLHP. Les colonnes utilisées ont été grandement améliorées lors de la dernière décennie. Des micro-colonnes ou colonnes capillaires ont ainsi été développées avec un diamètre interne de 3 à 200 µm. Toutes ces améliorations (pression stable, faible diamètre, faible débit,…) font de la chromatographie une technique de choix pour la séparation d'analytes.

2.3.2. Principe

La chromatographie liquide sépare les différents constituants d'un mélange selon le procédé suivant (Figure 26).

- Les molécules à séparer (analytes) sont portées par un liquide appelé phase mobile.
- Les analytes interagissent ou non avec un support fixe (contenu dans une colonne) appelé phase stationnaire. Il y a donc une distribution ou partition des analytes entre les phases mobile et stationnaire.
- Le flux du fluide vecteur étant continu, c'est la rétention plus ou moins longue des différents analytes sur la colonne qui entraine leur séparation.
- Un détecteur placé en sortie de colonne permet de suivre l'élution et d'obtenir soit des spectres (pics de différentes longueurs d'onde à un temps donné), soit des chromatogrammes (évolution des pics à une longueur d'onde donnée en fonction du temps).

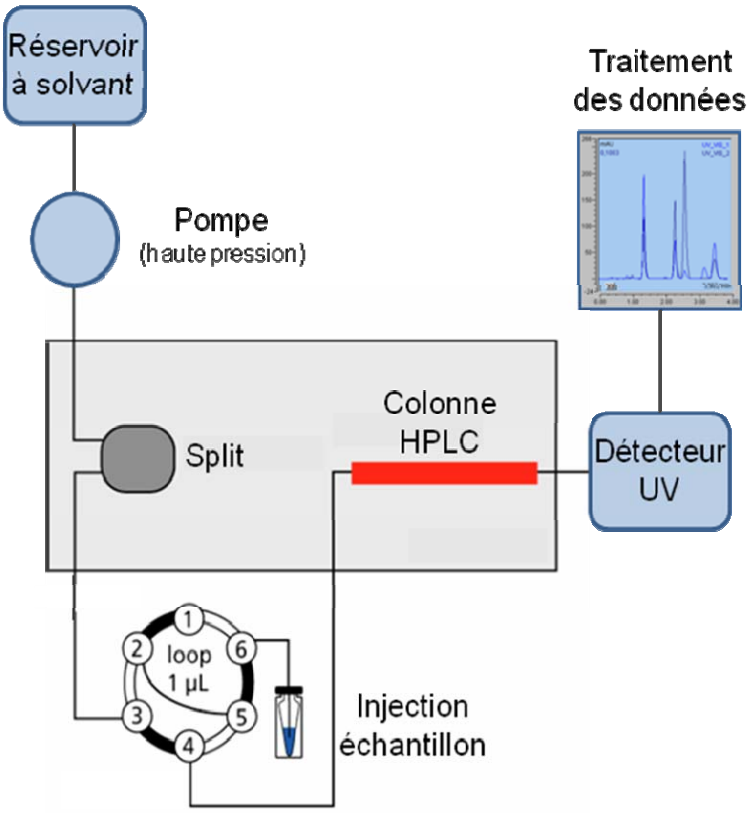


Figure 26 : La chromatographie liquide haute pression (CLHP).

La principale différence avec les chromatographies classiques réside dans la durée d'élution qui est de 100 à 1000 fois plus courte en CLHP. Une telle vitesse est obtenue par l'application d'une pression élevée (de l'ordre de 200 bars) grâce à une pompe qui maintient constant le débit du liquide. L'efficacité des colonnes a aussi été améliorée en uniformisant la granulométrie et en diminuant la grosseur des grains. Ces progrès ont entraîné la réduction de l'espace vide entre les grains et donc augmenté les interactions du soluté avec la phase stationnaire. Actuellement les colonnes possèdent des granulés de l'ordre du micromètre sur lesquels la phase stationnaire est fixée chimiquement; de telles phases sont appelées phases greffées. Ces colonnes sont généralement précédées par une pré-colonne pour les protéger d'un colmatage éventuel.

En protéomique, les analytes sont classiquement les peptides issus de la digestion à la trypsine d'un mélange protéique plus ou moins complexe. On utilise communément des colonnes de type « phase inverse » (RP pour reverse phase), composées de silice greffée à des chaînes linéaires de 8 ou 18 atomes de carbones (C_8 et C_{18}). Elles séparent les peptides selon leur hydrophobicité grâce à un gradient de solvant organique (généralement de l'acétonitrile). Les peptides hydrophiles, peu retenus par la colonne, sont élués rapidement alors que les hydrophobes nécessitent une plus forte concentration de solvant. Les solvants utilisés en RP-LC (acétonitrile + TFA ou acide formique) sont adaptés à l'analyse MS.

2.3.3. Le couplage LC-MS

La miniaturisation des colonnes (75 µm de diamètre interne) a permis une diminution des débits (de l'ordre de la centaine de nanolitres par minute) et donc facilité le couplage avec la spectrométrie de masse.

On distingue deux types de couplage de la chromatographie liquide (LC) à la spectrométrie de masse (MS) :

- ***Le couplage on-line*** : la CLHP est directement couplée à un spectromètre de masse de type ESI. Au fur et à mesure de leur élution, les différents peptides arrivent à la source ESI où ils sont ionisés, puis analysés par le spectromètre de masse. La miniaturisation des colonnes CLHP et la diminution des débits a favorisé le couplage direct LC-MS. Ce couplage a l'avantage d'automatiser le procédé d'analyse.
- ***Le couplage off-line*** : la CLHP n'est pas couplée directement au spectromètre de masse et les différentes fractions sont collectées au fur et à mesure de leur élution en sortie de colonne. Ce mode « off-line » rend compatible la séparation chromatographique et l'analyse par spectrométrie de masse à source MALDI (LC-MALDI-TOF/TOF). Les CLHP dernière génération ont une collecte automatisée des fractions éluées. Par exemple, le système HPLC U3000 (Dionex) est couplé à un automate, le Probot. Celui-ci peut notamment mélanger, en sortie de colonne, les éluats et la matrice, et déposer l'ensemble sur une cible MALDI selon un schéma automatisé.

2.3.4. La chromatographie multi-dimensions

La complexité des échantillons analysés en protéomique nécessite souvent leur séparation en multi-dimensions. Une deuxième séparation chromatographique de type échangeuse de cations (SCX pour « Strong Cation Exchange ») peut être ajoutée en amont de la RP. Les peptides (chargés positivement à pH acide) sont élués de la SCX grâce à un gradient de sels ou par « steps » (palliés croissant de sels). Les peptides de chaque fraction sont alors séparés en RP, puis analysés en spectrométrie de masse.

Deux grandes approches ont été décrites (Figure 27). Les deux colonnes peuvent être couplées physiquement, on parle alors de MudPIT pour Multidimensional Protein Identification Technology [46, 47]. Inversement, les deux colonnes peuvent être découplées, on parle alors de 2D-LC [48].

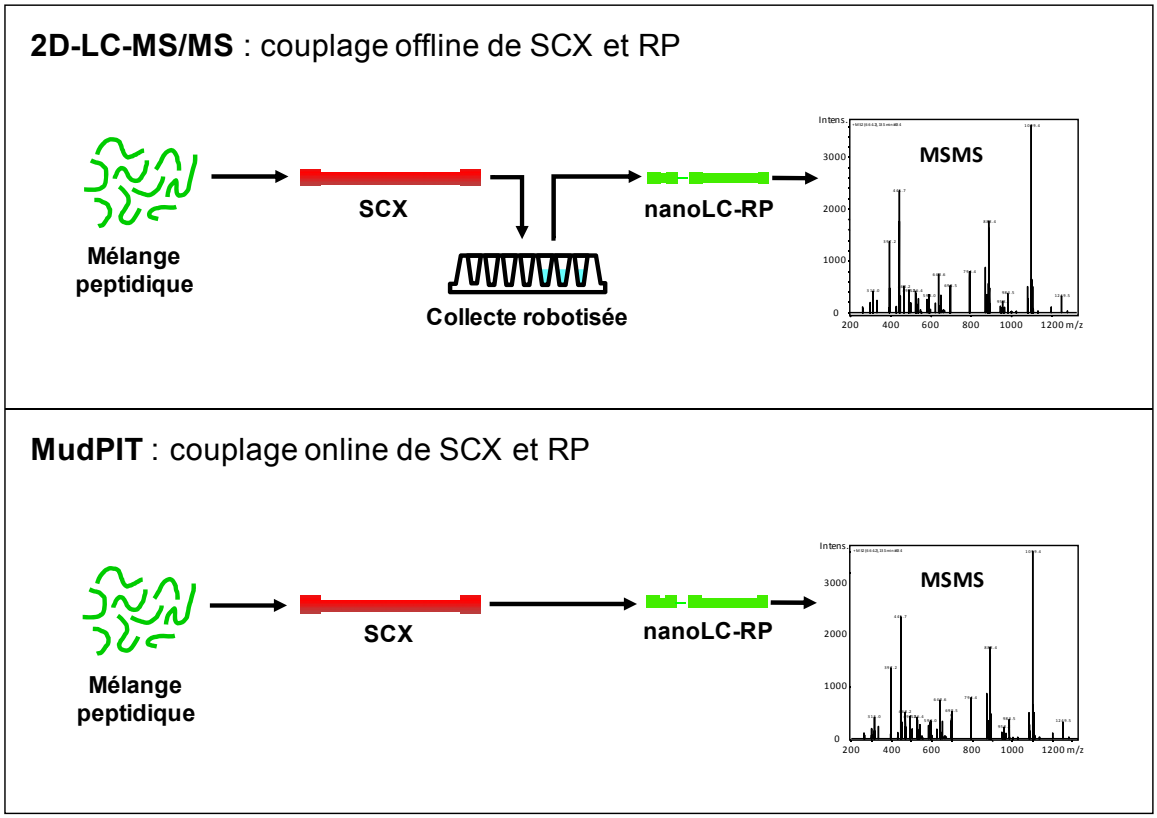


Figure 27 : Les systèmes de chromatographie multi-dimensions de type MudPIT et 2D-LC.

La RP-LC peut aussi être combinée avec une approche électrophorétique. L'électrophorèse SDS-PAGE est ainsi communément utilisée comme première dimension de séparation. Plus récemment, l'isoélectrofocalisation en solution de type OFFGEL a montré son efficacité.

3. L'identification des protéines

L'identification des protéines en protéomique repose sur l'utilisation de la spectrométrie de masse. Cela nécessite une mesure précise de la masse des analytes. Or, la limite maximale pour avoir des masses précises ne dépasse généralement pas les 4000 *m/z*. En conséquence, les protéines ne sont généralement pas analysées telles quelles mais sont au préalable découpées en peptides de façon contrôlée. Des endoprotéases sont ainsi utilisées pour protéolyser de façon spécifique les protéines en peptides. L'endoprotéase la plus couramment utilisée est la trypsine. Elle hydrolyse les liaisons peptidiques dans lesquelles un acide aminé basique (lysine ou arginine) engage sa fonction acide (COOH) (Figure 28). Ces deux acides aminés ont une abondance et une répartition assez homogènes dans les différentes protéines. La trypsine génère donc des peptides de taille adéquate pour l'analyse en spectrométrie de masse.

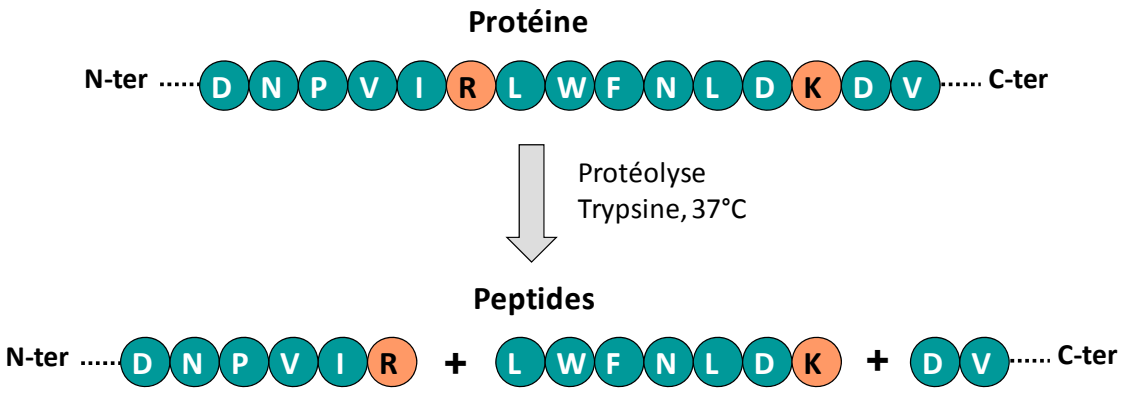


Figure 28 : La digestion trypsique des protéines.

Il existe deux grandes approches pour l'identification des protéines : l'empreinte peptidique massique (PMF) et les informations de séquence issues de la fragmentation spécifique de peptides. Les deux reposent sur un principe général similaire (Figure 23) : les données expérimentales obtenues en spectrométrie de masse sont comparées à des données théoriques contenues dans des bases de données protéiques grâce à l'utilisation d'un moteur de recherche. Si les deux types de données coïncident, la protéine est identifiée avec un score de confiance.

3.1. L'empreinte peptidique massique (PMF)

Cette approche, décrite en 1993 [49-52], est généralement utilisée lorsque la méthode de séparation des protéines est l'électrophorèse bidimensionnelle.

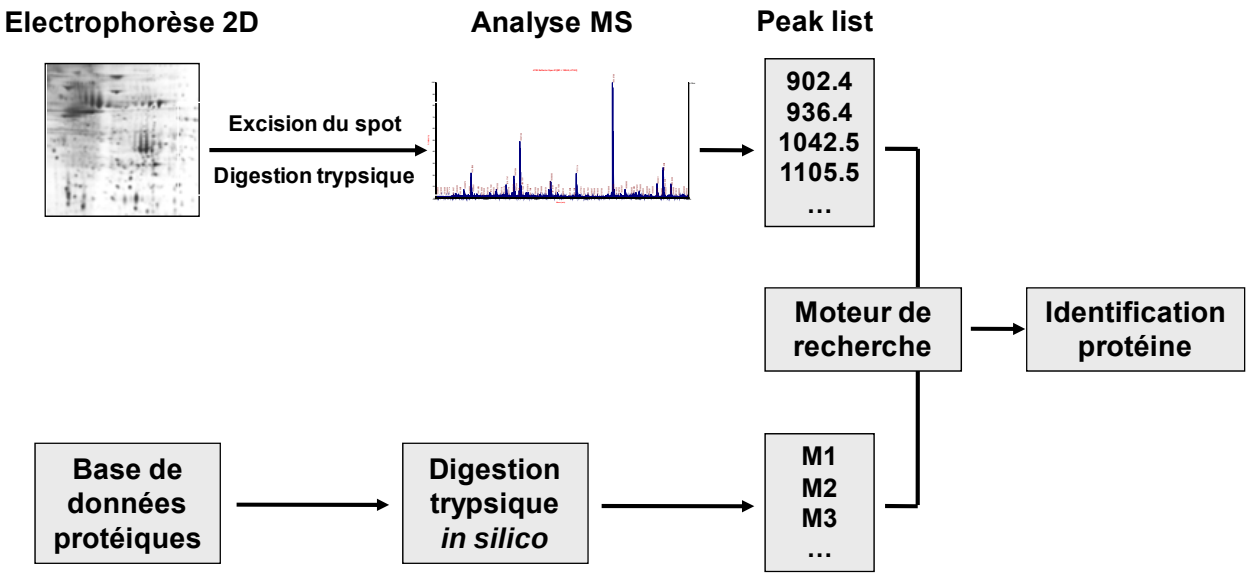


Figure 29 : L'identification des protéines par Empreinte Peptidique Massique (PMF).

Son principe est défini Figure 29. Si la résolution de l'électrophorèse bidimensionnelle est suffisamment bonne, alors chaque spot protéique observé correspond à une seule forme protéique. Le spot protéique d'intérêt peut être excisé soit manuellement, soit grâce à l'utilisation d'un robot piqueur. La protéine contenue dans le morceau de gel est alors digérée par la trypsine et les différents peptides extraits du gel. L'analyse en mode MS génère une liste de masses expérimentales correspondant aux peptides trypsiques observés pour la protéine analysée. Ces peptides représentent une signature de la protéine : on parle d'empreinte peptidique massique ou PMF (Peptide Mass Fingerprint). Cette liste de masses est comparée aux masses théoriques issues de la digestion *in silico* de protéines répertoriées dans des bases de données. Les bases de données classiquement utilisées pour l'analyse de protéines humaines sont UniProt (www.uniprot.org), NCBI (www.ncbi.nlm.nih.gov) ou IPI (www.ebi.ac.uk/IPI). La confrontation entre les masses expérimentales et théoriques est assurée par l'utilisation d'un logiciel algorithmique appelé moteur de recherche. Citons, par exemple, Mascot (www.matrixscience.com), ou Aldente (www.genebio.com/aldente). Le résultat de la recherche conduit ou non à l'identification de la protéine par l'attribution d'un score de confiance. Ce score prend notamment en compte le nombre de masses peptidiques correspondant à la protéine, le pourcentage de recouvrement de la séquence protéique et la précision de mesure des différentes masses peptidiques.

Cette approche, facile à mettre en œuvre, est compatible avec les spectromètres de masse à source MALDI et permet d'obtenir rapidement des résultats.
Elle possède néanmoins certaines limites :

- Un risque d'ambiguïté si le nombre de peptides retrouvés n'est pas suffisant.
- Une nécessité d'avoir une bonne séparation en électrophorèse. Si le spot contient plus d'une protéine, l'analyse en PMF est très compliquée car les peptides ne se rapporteront pas tous à la même protéine.
- Des spectres bien résolus peuvent ne pas donner lieu à une identification si la protéine n'est pas présente dans la base de données.

Si cette stratégie PMF ne permet pas d'identifier les protéines, le recours à une approche MS/MS peut être envisagé afin d'obtenir des informations de séquence.

3.2. Identification des peptides par MS/MS

3.2.1. Principe

L'échantillon protéique à analyser est généralement digéré en solution par la trypsine afin de générer un mélange (complexe) de peptides. Ces peptides sont séparés (généralement par chromatographie multi-dimensionnelle) et analysés en spectrométrie de masse en mode tandem (MS/MS). Dans la mesure où chaque spectre MS obtenu représente un ensemble de peptides ayant des paramètres de rétention similaires en chromatographie mais appartenant à des protéines différentes, l'approche PMF ne peut être appliquée.

Le principe repose donc sur la fragmentation spécifique de chaque peptide précurseur et l'analyse en spectrométrie de masse des différents fragments générés (Figure 30.A). On obtient alors un spectre MS/MS par peptide fragmenté. L'identification du peptide et donc de la protéine se fait par l'intermédiaire d'un moteur de recherche [53]. Celui-ci compare chaque *m/z* du spectre MS/MS expérimental aux profils de fragmentations théoriques des peptides obtenus après digestion *in silico* des protéines de la base de données. Le profil de fragmentation d'un peptide joue ainsi le rôle de « code barre » caractéristique de ce peptide. Des logiciels tels que Mascot (www.matrixscience.com), Sequest (fields.srcipps.edu/sequest), Phenyx (phenyx.vital-it.ch/pwi) ou ProteinPilot (Applied Biosystems) aident à cette analyse et à l'identification des peptides. La digestion des protéines se faisant avant les étapes de séparation, une même protéine a ses différents peptides trypsiques répartis dans plusieurs fractions d'élution. Les logiciels intègrent ce paramètre et recoupent donc l'ensemble des données d'identification. Une identification sera communément jugée fiable si au moins deux peptides uniques, associés à une même protéine, ont été identifiés avec une confiance suffisante.

Néanmoins, cette approche reste dépendante de l'utilisation des bases de données qui sont plus ou moins complètes selon l'espèce étudiée. Ainsi, des spectres MS/MS de qualité ne garantissent pas l'identification du peptide. Il faut alors envisager une approche de séquençage *de novo*.

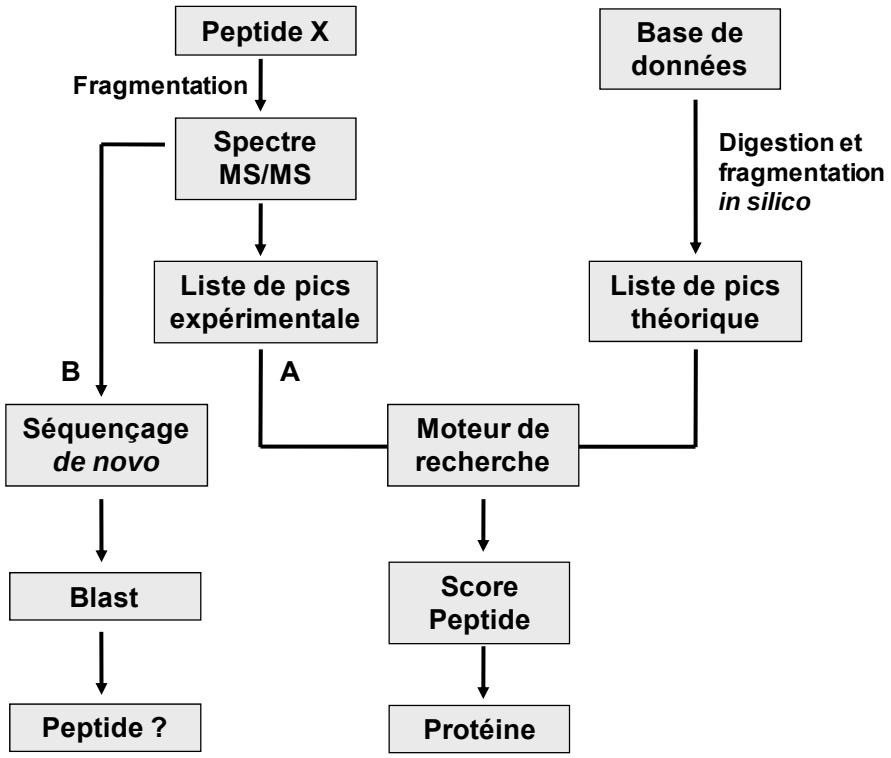


Figure 30 : L'identification des protéines par analyse MS/MS.

A. Empreinte massique des fragments.

B. Séquençage *de novo*.

3.2.2. Le séquençage de novo

Les règles de fragmentation d'un peptide sont clairement définies et la majorité des fragments obtenus sont de types b et y (voir le chapitre 1 de la partie bibliographique). La différence de masse entre deux fragments consécutifs de même type donne une masse correspondant à l'un des 20 acides aminés. Il est ainsi possible de se servir de ces informations pour en déduire la séquence (ou une partie de la séquence) des peptides d'intérêts. Cette approche est appelée séquençage *de novo* (Figure 30.B). Initialement, ce séquençage était réalisé manuellement. Des logiciels, tels que Peaks (www.bioinformaticssolutions.com) ou De novo Explorer (Applied Biosystems) assistent maintenant l'utilisateur dans cette analyse.

Le séquençage *de novo* a l'avantage de permettre l'exploitation de spectres de fragmentation pour lesquels la séquence exacte du peptide n'est pas présente dans la base de données, tels que des peptides ayant des polymorphismes de séquence ou des peptides modifiés. Il est en conséquence principalement utilisé pour l'analyse des protéines d'espèces dont les informations de séquence du génome sont peu ou pas renseignées ou pour identifier des peptides ayant des modifications post-traductionnelles particulières. Cependant, l'analyse *de novo* est longue et fastidieuse et nécessite des spectres MS/MS de grande qualité.

Le séquençage *de novo* permet d'accéder à la séquence du peptide mais pas de connaître à quelle protéine celui-ci appartient. Dans ce but, les séquences identifiées sont comparées (« matchées ») aux séquences de bases de données généralistes par l'utilisation de logiciels de BLAST. Cela peut permettre, par homologie de séquence, d'identifier ou caractériser avec une fiabilité relative une protéine non référencée.

4. Trois grands types d'applications de la protéomique

4.1. La cartographie

La cartographie cellulaire consiste à générer une image globale d'un système biologique via (1) l'identification et la caractérisation exhaustive des protéines qui le constituent, et (2) les interactions mises en jeu entre ces protéines.

Des travaux d'une ampleur considérable ont visé à établir le protéome cellulaire [54] ; cependant, ce genre d'études se trouve vite limité par la grande gamme dynamique des protéines dans la cellule. Certaines protéines sont en effet très majoritaires ; ainsi, l'actine est présente avec près de 10^8 copies par cellule alors que d'autres protéines minoritaires sont en quelques copies par cellule. En conséquence, l'étude de ces protéines nécessite l'analyse de sous-protéomes (protéines cytosoliques, protéines nucléaires, protéines sécrétées, protéines membranaires,…) grâce à des techniques préalables de fractionnement.

Une autre voie consiste à analyser les réseaux d'interactions des protéines les unes avec les autres. Ce concept porte le nom « d'interactome » [55]. La plupart des protéines exercent leur fonction en interagissant avec d'autres protéines et enzymes de façon régulée en des localisations bien définies de la cellule. Ainsi, lorsqu'une nouvelle protéine est identifiée, la recherche de sa fonction peut être abordée par la caractérisation de ses partenaires.

4.2. L'analyse des modifications post-traductionnelles

Les modifications post-traductionnelles (PTMs) représentent un mécanisme de régulation fondamental des protéines et influencent directement leur fonction. La protéomique est un outil de choix pour l'analyse de ces PTMs puisque celles-ci induisent une différence de masse pouvant être théoriquement détectée [56]. Prenons l'exemple de la phosphorylation des protéines. Celles-ci peuvent être analysées par électrophorèse bidimensionnelle [57], une différence de phosphorylation d'une protéine induisant un changement dans son état de charge et donc un shift horizontal. Il existe aussi un colorant particulier, le pro-Q Diamond, capable de colorer sur gel uniquement les protéines phosphorylées [58]. L'étude des protéines phosphorylées repose

également sur leur enrichissement. Les protéines phosphorylées sur leur résidu tyrosine peuvent, par exemple, être enrichies par immunoprécipitation avec des anticorps anti-phosphotyrosine [59]. Enfin, des phosphopeptides peuvent être identifiés directement en spectrométrie de masse en recherchant la présence d'un ion rapporteur correspondant au groupe PO_3- de *m/z* = 79 [60].

4.3. Les analyses protéomiques différentielles et la quantification

L'analyse différentielle des profils d'expression protéiques dans différentes conditions constitue une application importante de la protéomique. De nombreuses approches sont possibles. Certaines reposent sur la comparaison d'images de gels 2D, d'autres sur le marquage isotopique différentiel des échantillons et leur analyse en spectrométrie de masse ou encore sur des techniques « label free » étudiant les signaux MS.

On peut différencier deux types d'analyses différentielles selon qu'elles permettent une quantification relative ou absolue. La quantification relative est exprimée sous forme d'un ratio définissant la quantité d'une protéine par rapport à la mesure de cette même protéine dans une autre condition. La quantification absolue vise à déterminer la quantité « exacte » d'une protéine dans un échantillon (par exemple, en ng/mL pour un bio-marqueur ou en nombre de copies d'une protéine par cellule). Toutes les méthodes protéomiques actuellement disponibles pour la quantification sans *a-priori* de protéines inconnues sont des méthodes de quantification relative. La quantification absolue est possible dans un second temps, une fois que la protéine à quantifier est connue.

Les applications de ces analyses différentielles sont nombreuses. Elles peuvent permettre de mettre en évidence et éventuellement comprendre certains mécanismes de régulation ou de dysfonctionnement de la machinerie cellulaire ou d'étudier la cinétique d'action d'une drogue sur un protéome. Dans la recherche biomédicale, la majorité des études protéomiques vise à comparer le profil d'expression des protéines entre différents tissus ou lignées cellulaires. L'analyse différentielle d'échantillons sains *vs* pathologiques, induits *vs* non-induits, soumis à un stress ou non (etc…) est indispensable à la découverte de nouvelles cibles diagnostiques, pronostiques ou thérapeutiques.

Les différentes approches permettant une analyse protéomique différentielle sont décrites et discutées de façon détaillée dans le chapitre suivant.

CHAPITRE 2 : Les méthodes d'analyse protéomique différentielle

Un grand nombre de méthodes a été développé pour l'analyse quantitative différentielle des protéomes. Celles-ci comparent les profils d'expression des protéines de deux ou plusieurs conditions distinctes [61].

Ces méthodes peuvent être réparties en différentes catégories :

- ***Les méthodes « gel-based »***, reposant sur la quantification des protéines à partir de différentes images de gels.
- ***Les méthodes de marquage isotopique***, reposant sur le marquage différentiel des différents protéomes à comparer grâce à l'incorporation de formes isotopiques.
- ***Les méthodes « label-free »***, permettant une analyse semi-quantitative sans marquage préalable des différents protéomes.
- ***Les méthodes avec standard***, permettant la quantification absolue de protéines ciblées.

1. Les méthodes « gel-based »

1.1. Comparaison d'images de gels 2D

La méthode historique pour comparer les profils d'expressions protéiques issus de différentes conditions est l'analyse comparative d'images de gels 2D. Typiquement, les protéines de chaque condition sont séparées par électrophorèse bidimensionnelle et détectées par coloration du gel. Les cartes 2D ainsi obtenues sont numérisées, et les images alignées, normalisées et comparées grâce à l'utilisation de logiciels dédiés tels que ImageMaster (GE healthcare) ou PDQuest (Biorad). Ces logiciels comparent l'intensité de coloration (ou le volume) des spots protéiques pour chaque condition (Figure 31). Les différences de coloration mesurées sont corrélées aux quantités de protéines observées sur les gels comparés. On peut ainsi établir pour chaque spot un ratio d'expression de la protéine dans une condition par rapport à une autre. Afin d'être statistiquement fiable, cette approche nécessite la réalisation de réplicas de chaque condition (au minimum 3 gels par condition).

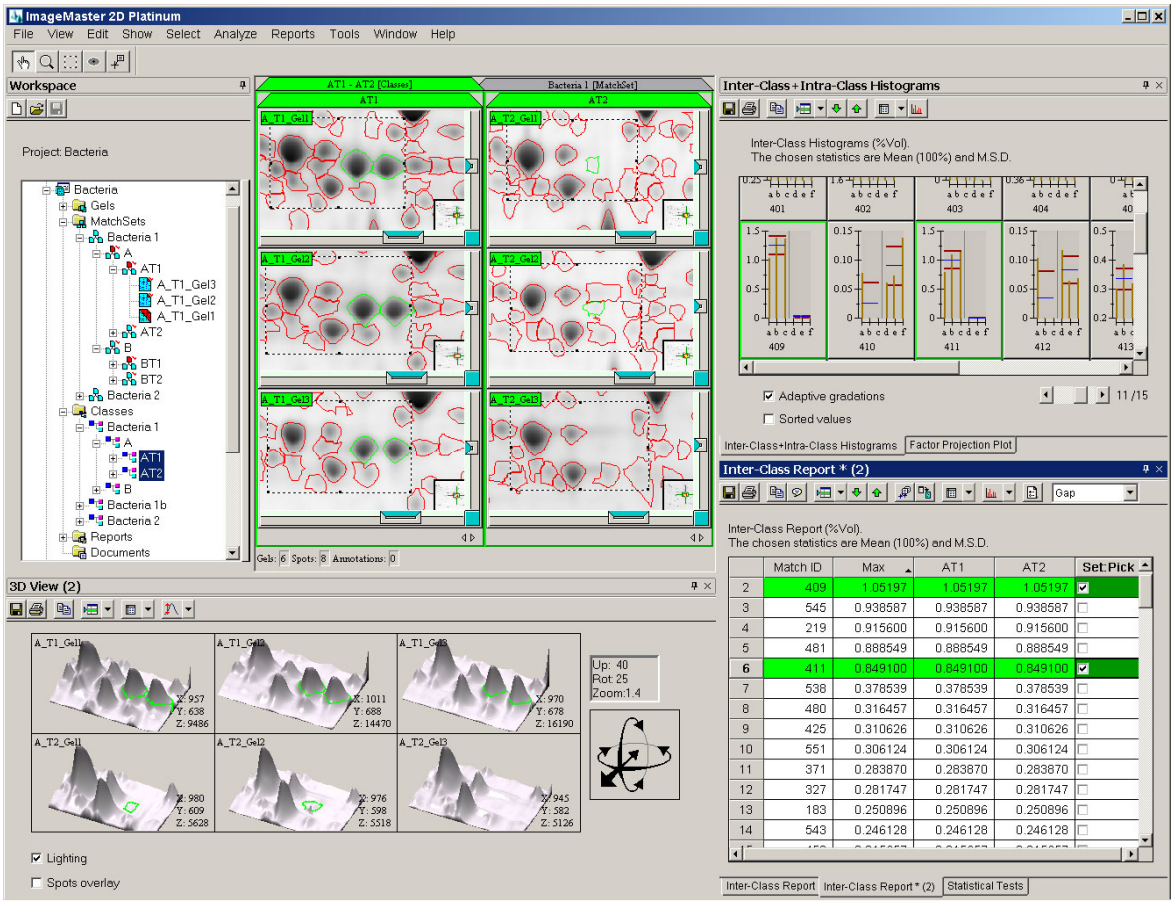


Figure 31 : La comparaison de cartes 2D par le logiciel ImageMaster.

Dans cet exemple, 2 conditions composées chacune d'un panel de 3 gels sont comparées. Le logiciel détecte, associe et quantifie les spots de chaque image de gel. Les spots entourés en vert sont exprimés de façon différentielle entre les deux conditions. Une visualisation en 3D des spots (intensité x surface) est aussi possible (en bas à gauche).

Cette technique a l'avantage de permettre la visualisation des différents états d'une même protéine (PTMs). Cependant, elle se heurte aux limites inhérentes à l'électrophorèse 2D (problème de reproductibilité, de sensibilité, de linéarité de la réponse,…). L'approche DIGE peut palier à certaines de ces limites.

1.2. DIGE

L'approche DIGE ou « Differential In Gel Electrophoresis » a été développée en 1997 par l'équipe de Unlü *et al.* [62]. Son principe repose sur le marquage covalent des protéines de chaque condition par des cyanines fluorescentes spécifiques (Cy2, Cy3 et Cy5) (Figure 32). Les protéines des deux échantillons à comparer sont respectivement marquées par les fluorochromes Cy3 et Cy5. Un mélange de chaque échantillon selon un ratio 1:1 est marqué par le fluorochrome Cy2 afin de servir de standard interne (normalisation des résultats). Les échantillons marqués sont alors mélangés en quantité égale et les protéines séparées sur le même gel 2D. Les images de gels

sont acquises aux longueurs d'onde d'excitation de chaque cyanine. Des logiciels, tel que Decyder (GE Healthcare), analysent ces images. Ils superposent les différents profils et quantifient les différences d'intensité des spots protéiques à chaque longueur d'onde pour en déduire des ratios d'expression. Comme les protéines de chaque échantillon sont séparées sur un seul et même gel, la technique DIGE permet d'éliminer les erreurs dues aux variations inter-gels. De plus, c'est une technique sensible capable de détecter jusqu'à 5 femtomol de protéines et des différences d'expression de l'ordre de 15% [63]. Cependant la réalisation et l'analyse de gels reste une expérience longue et délicate se heurtant aux limitations inhérentes à l'électrophorèse 2D.

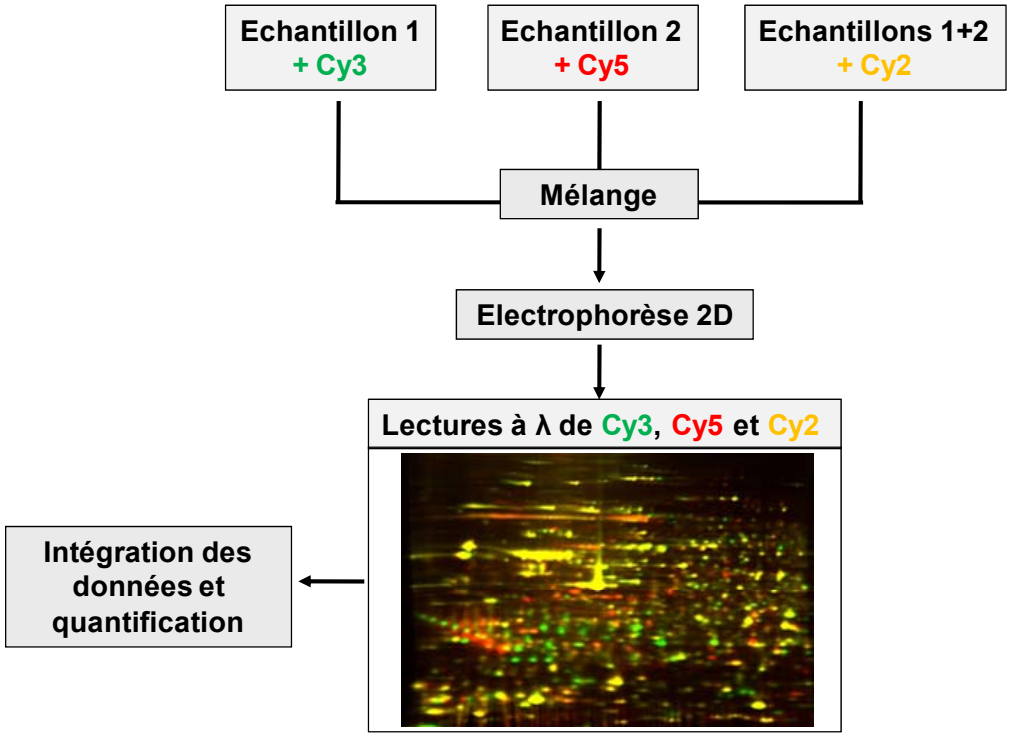


Figure 32 : Le principe de la méthode DIGE (Differential In Gel Electrophoresis).

2. Les méthodes de marquage isotopique

2.1. Principe général

La bonne résolution des dernières générations de spectromètres de masse offre la possibilité de distinguer les différentes formes isotopiques d'une même molécule. Les peptides marqués isotopiquement ont, pour chaque condition, les mêmes propriétés physico-chimiques exceptées en ce qui concerne leur masse. Partant de ce principe les échantillons sont marqués spécifiquement de façon isotopique (proportion d'isotopes lourds ^{2}H, ^{13}C, ^{15}N,…), puis mélangés selon un ratio 1:1 et analysés de façon simultanée (Figure 33). La quantification relative s'effectue alors en

comparant les signaux des tags rapporteurs obtenus sur les spectres de masses pour chaque forme isotopique.

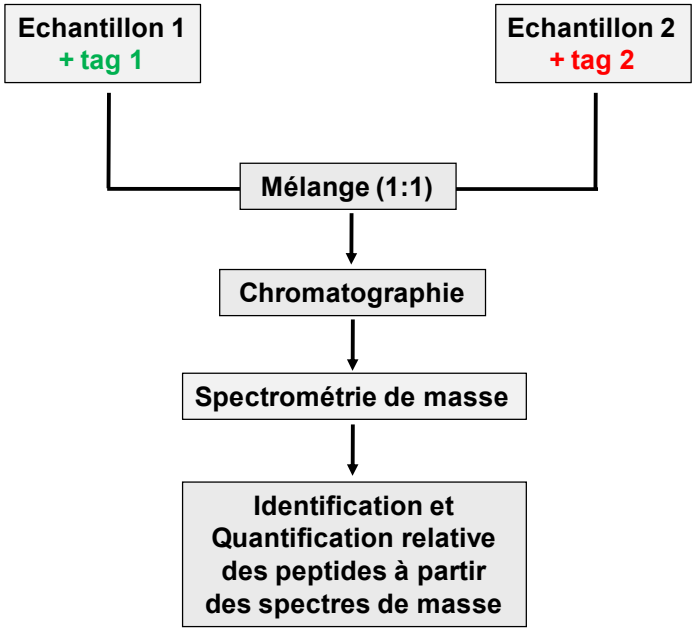


Figure 33 : Principe général de la quantification relative par marquage isotopique différentiel.

Il existe différentes stratégies de marquage isotopique (Figure 34).

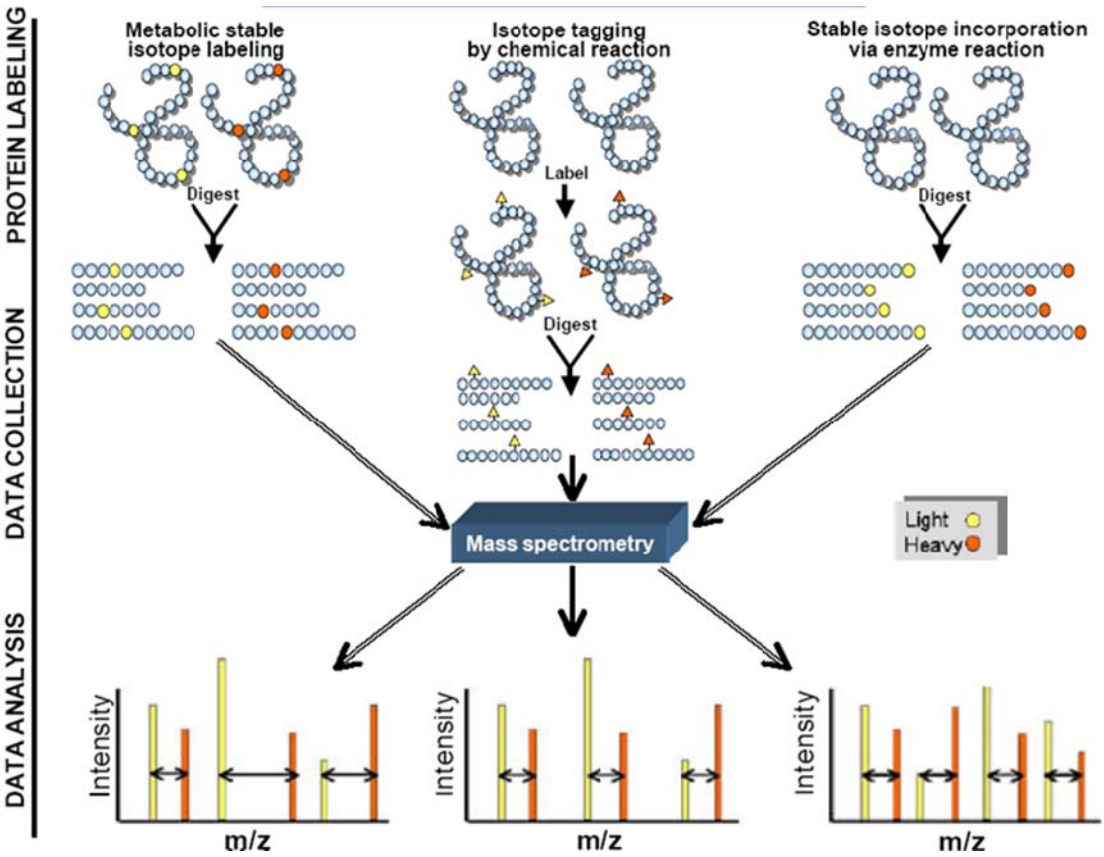


Figure 34 : Représentation de différentes méthodes d'analyse différentielle reposant sur un marquage isotopique (D'après Aebersold et al., 2003) [15].

De gauche à droite : marquages métabolique (SILAC), chimique (ICAT) et enzymatique ($H_2{}^{18}O$).

Ces marquages isotopiques peuvent être réalisés par des approches chimiques (ICPL : *Isotope Coded Protein Labelling* [64], ICAT : *Isotope Coded Affinity Tag* [65], iTRAQ : *isobaric Tags for Relative and Absolute Quantitation*) [66],…), enzymatiques ($H_2^{16}O/H_2^{18}O$) [67], ou au niveau de cellules vivantes *via* une incorporation métabolique (SILAC : *Stable Isotope Labelling by Amino acids in Cell culture*) [68]. Selon les méthodes utilisées, le marquage se fait plus ou moins précocement dans le processus d'analyse protéomique [15, 69, 70] (Figure 35).

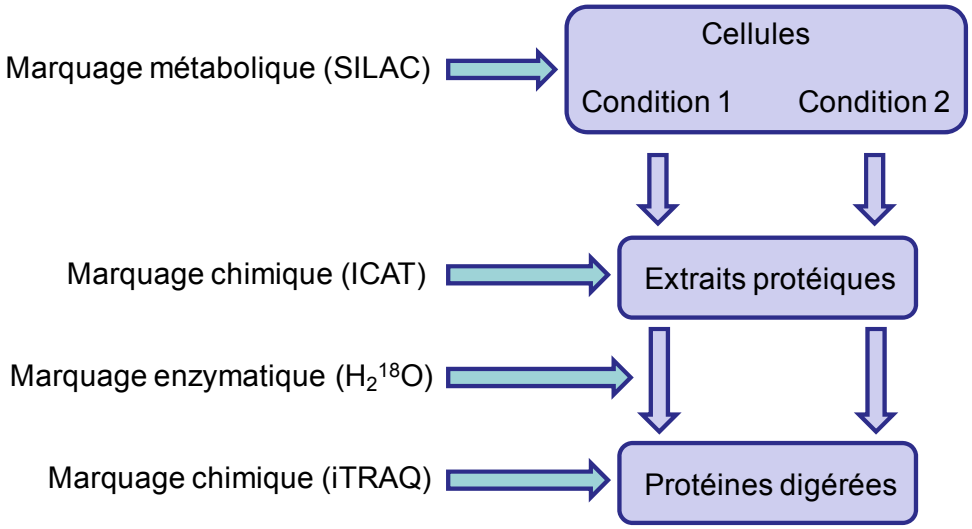


Figure 35 : Différents niveaux de marquage des échantillons selon les approches.

Les approches les plus classiques sont décrites ci-dessous.

2.2. Marquage métabolique (SILAC)

Cette approche est utilisée pour comparer les protéomes de deux types cellulaires cultivables. (Figure 36). Initialement introduite pour des études protéomiques bactériennes par Oda *et al.* [71], cette technique a par la suite été transférée à l'étude de cellules eucaryotes [72], puis finalement optimisée et nommée SILAC (Stable Isotope Labelling by Amino acids in Cell culture) par Ong *et al.* [68].

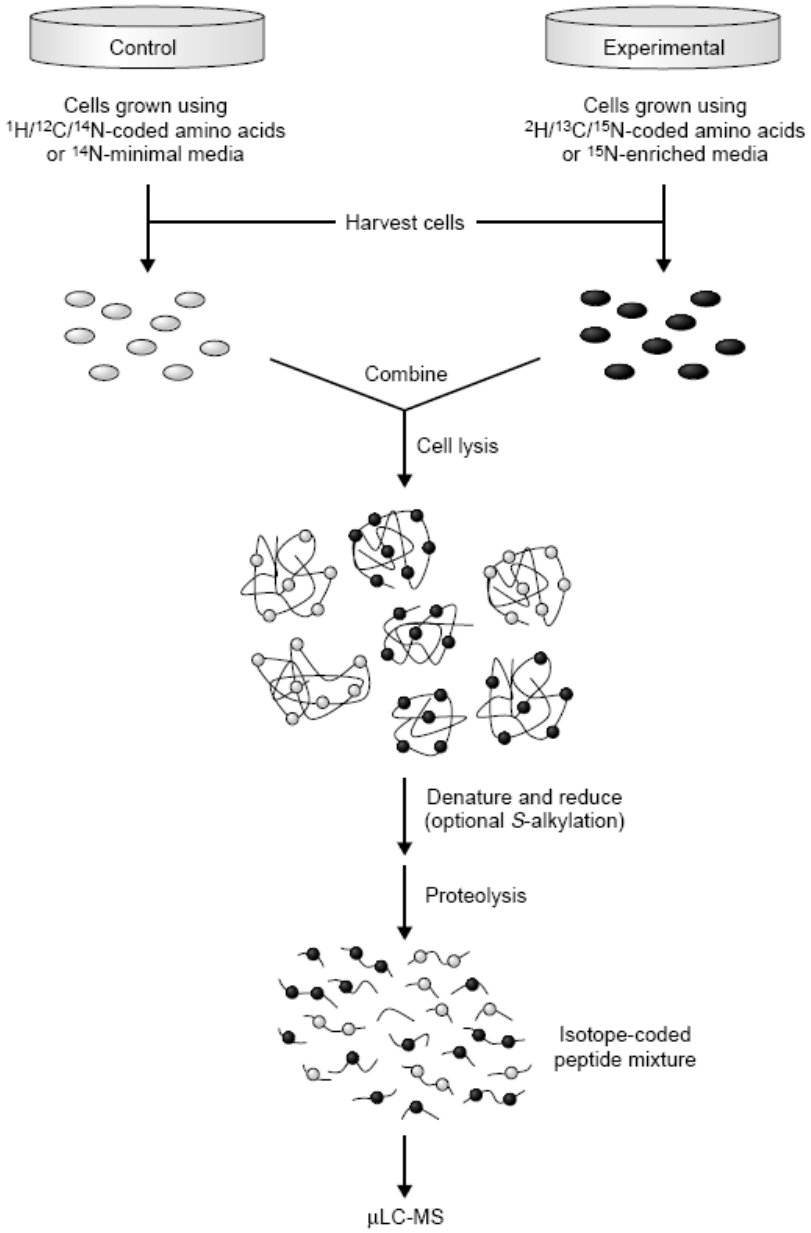


Figure 36 : Le principe de la méthode SILAC.

Les deux cultures de cellules sont réalisées parallèlement. La méthode SILAC consiste à remplacer un acide aminé, dans le milieu de culture de l'un des deux types cellulaires, par son équivalent isotopiquement alourdi. A travers la croissance cellulaire et le *turnover* des protéines, ces acides aminés marqués sont incorporés par les cellules afin de renouveler leur stock protéique. Après un nombre de doublements cellulaires suffisant (afin d'assurer un renouvellement complet du stock protéique), une quantité égale des deux populations de cellules est mélangée. Les protéines sont alors extraites puis digérées à la trypsine. Les peptides résultant sont analysés et quantifiés en spectrométrie de masse. Sur les spectres MS, chaque peptide apparaît alors sous la forme d'une paire de massifs isotopiques qui est utilisée pour calculer la proportion relative du peptide dans chaque condition (Figure 37). L'intégration des données

issues de l'analyse des différents peptides d'une même protéine permet d'avoir des ratios d'expression relative fiables.

Les acides aminés marqués isotopiquement les plus fréquemment utilisés sont la $^{13}C_6$-arginine et la $^{13}C_6$-lysine. En effet, si la digestion trypsique est totale, chaque peptide aura ainsi l'incorporation d'un seul acide aminé alourdi (en C-terminal de la séquence). Cela facilite la quantification car l'écart de masse entre le peptide « alourdi » et « normal » est fixe (écart de masse de 6 *m/z* pour un peptide mono-chargé). C'est un avantage du SILAC par rapport aux premières approches de marquage métabolique au ^{15}N modifiant l'extrémité N-terminale des liaisons peptidiques ainsi que les chaînes latérales des acides aminés contenant de l'azote (lysine, arginine, asparagine, glutamine et histidine). En effet, l'écart de masse entre le peptide alourdi et normal n'est alors pas fixe puisqu'il dépend de la séquence en acides aminés de chaque peptide et la quantification est donc plus ardue.

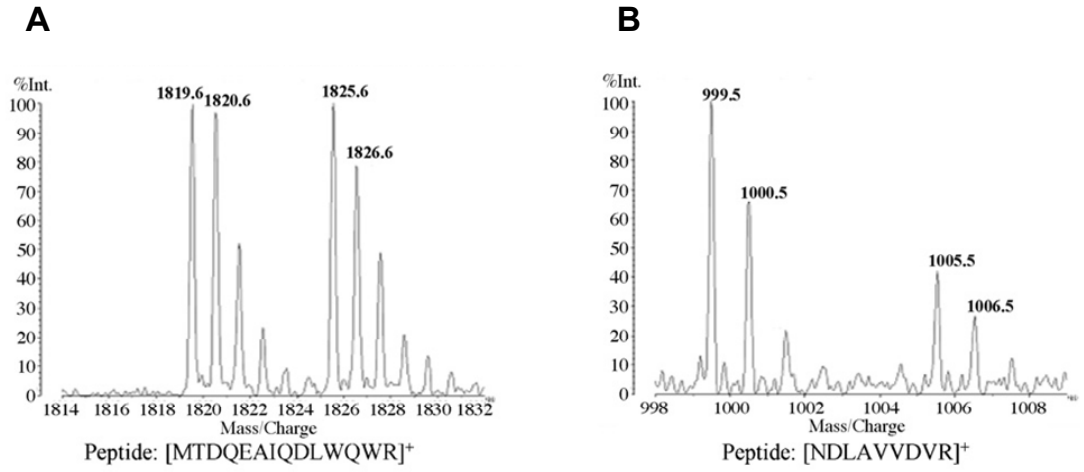


Figure 37 : Paires de pics obtenues pour deux peptides sur les spectres MS à l'issue d'une analyse SILAC (d'après Fenselau, 2007) [70].

A : Exemple d'un peptide exprimé en quantité égale dans les deux conditions.

B : Exemple d'un peptide surexprimé dans la condition native.

L'avantage de la méthode SILAC réside dans l'incorporation précoce de la forme isotopique sur des cellules vivantes en culture. Les échantillons sont mélangés au stade cellulaire, limitant ainsi les biais dus à des pertes inégales de protéines durant la lyse des cellules, l'extraction des protéines, la digestion ou les étapes de séparation. Par contre, elle reste limitée à l'étude de cellules amplifiables en culture et ne peut donc pas être appliquée aux tissus, ou échantillons cliniques (plasma,…). De plus, certains types cellulaires ne sont pas cultivables en présence de sérum dialysé, sûrement dû à la perte de facteurs de croissance pendant le processus de dialyse. Enfin, il faut normaliser les ratios pour limiter tout biais dû à un renouvellement partiel du stock protéique.

2.3. Marquage enzymatique ($H_2^{16}O/H_2^{18}O$)

La technique de marquage enzymatique repose sur l'incorporation d'atomes d'oxygène 18 (^{18}O) à l'extrémité C-terminale des peptides générée au cours de la digestion des protéines par une protéase [67, 73] (Figure 38). Un des échantillons est soumis à l'action d'une protéase dans de l'eau lourde ($H_2^{18}O$) alors que le second est digéré dans de l'eau normale ($H_2^{16}O$). La protéase utilisée pour la digestion est généralement la trypsine. Celle-ci incorpore deux atomes d'oxygène ^{18}O ou ^{16}O au niveau du groupement carboxyl terminal (COOH) généré pour chaque peptide (Figure 39). Comme pour le marquage métabolique, l'analyse en spectrométrie de masse révèle pour chaque peptide des paires de pics, séparés ici d'un *m/z* de 4 si z=1 (incorporation de 2 ^{18}O). Ces paires sont utilisées pour le calcul de la proportion du peptide dans chacun des échantillons.

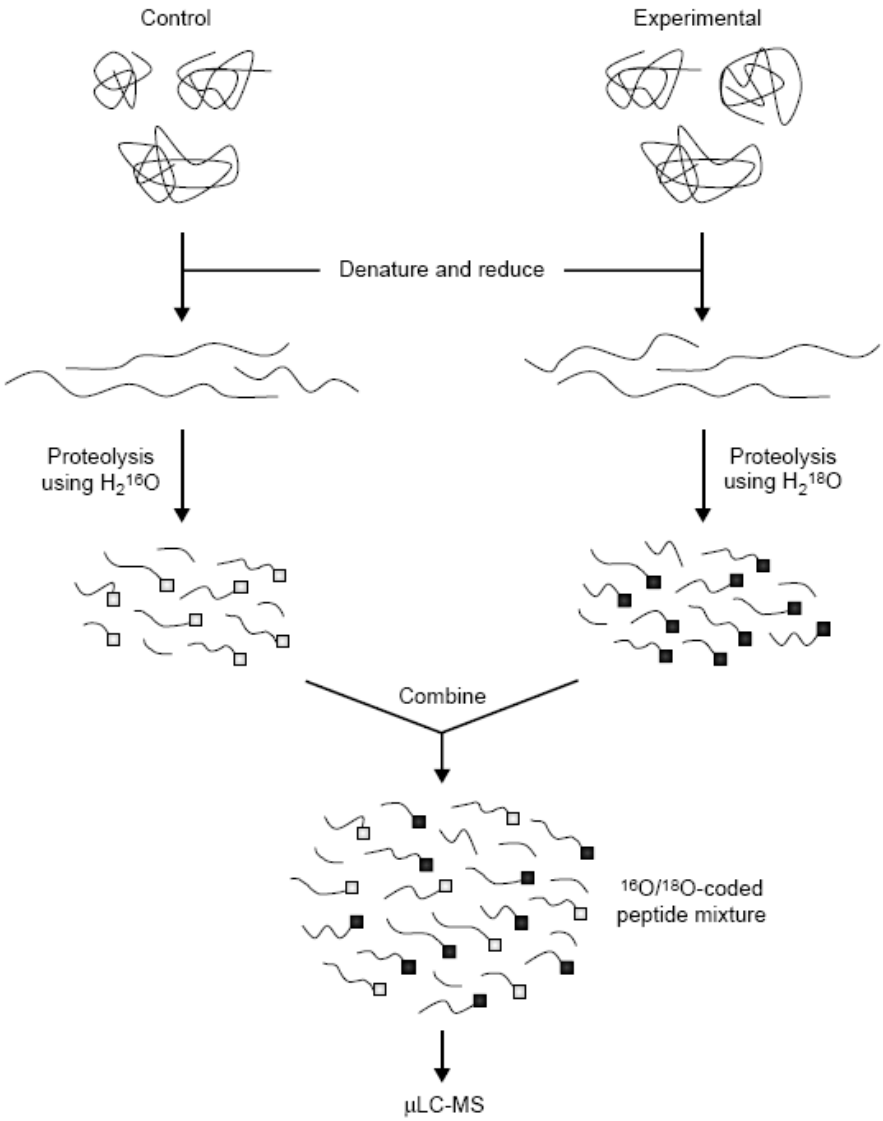


Figure 38 : Principe de la méthode de marquage enzymatique ($H_2^{16}O/H_2^{18}O$).

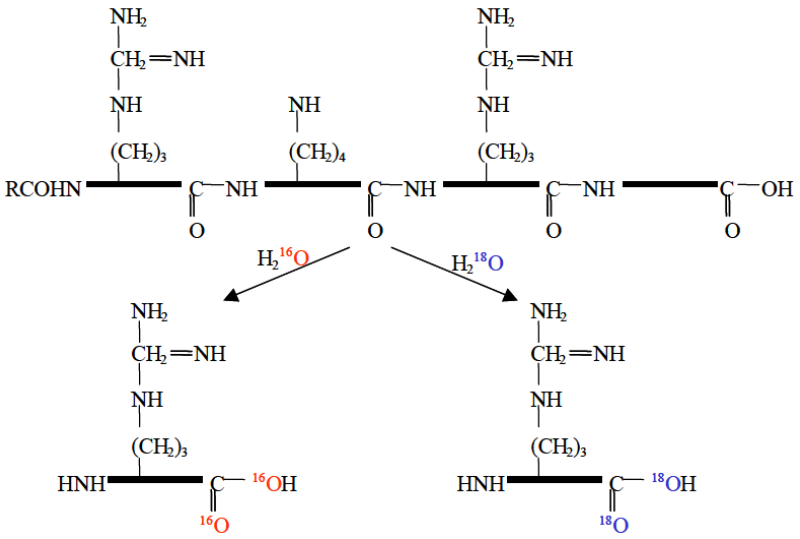


Figure 39 : Principe chimique du marquage enzymatique (digestion à la trypsine dans $H_2{}^{16}O$ ou $H_2{}^{18}O$).

Les avantages de cette approche sont sa facilité de mise en œuvre, ainsi que sa compatibilité avec la plupart des échantillons. Par contre, le mélange tardif des deux échantillons (après la digestion) est susceptible d'entraîner des biais expérimentaux plus importants. De plus, l'écart de 4 *m/z* entre les deux pics est relativement faible et peut entraîner un risque de chevauchement des massifs isotopiques faussant la quantification.

2.4. Marquage chimique

L'analyse différentielle par marquage isotopique est réalisée par l'ajout d'un tag chimique particulier soit au niveau protéique, soit au niveau peptidique. Différentes approches ont été développées et nous présenteront ici les deux plus classiques : l'ICAT et l'iTRAQ. Pour plus d'informations sur les autres approches, se référer à la table 1 de la revue de Panchaud *et al.* [61].

2.4.1. ICAT

La technologie ICAT (Isotope Coded Affinity Tag) a été développée en 1999 par l'équipe de Gygi [65] et est actuellement commercialisée par la société Applied Biosystems. Elle repose sur l'utilisation de deux tags isotopiquement distincts pour marquer spécifiquement les cystéines des protéines par liaison covalente.

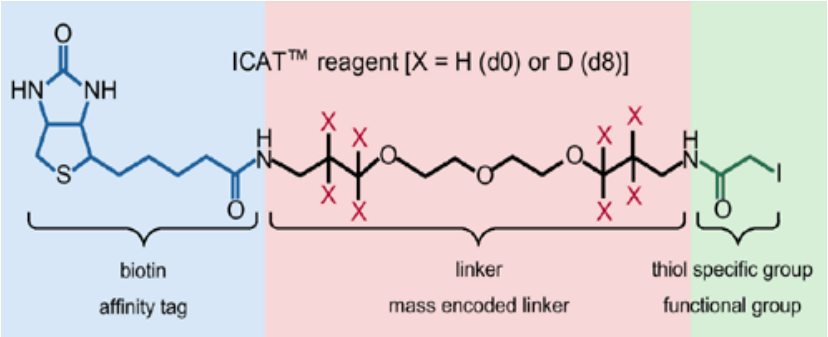


Figure 40 : Le tag ICAT de première génération.

Le tag comprend trois parties (Figure 40) :

- ***Un groupe réactif*** (atome d'iode) qui va réagir de façon covalente (alkylation) avec les groupements thiol des cystéines.
- ***Un bras polyether*** de structure isotopique caractéristique. C'est lui qui induit la différence de masse entre les deux formes du tag selon qu'il possède 8 atomes d'hydrogènes (D_0) ou 8 atomes de deutériums (D_8), introduisant une différence de masse de 8 *m/z* entre les deux tags (si z=1).
- ***Une étiquette biotine*** qui permet la purification des peptides marqués sur colonne d'affinité (colonne d'avidine). Cette biotine permet l'enrichissement des peptides marqués et contribue à une décomplexification de l'échantillon puisque seuls les peptides dont le résidu cystéine a été modifié par le réactif ICAT seront analysés par spectrométrie de masse. L'identification et la quantification de protéines minoritaires peuvent donc s'en trouver facilitées.

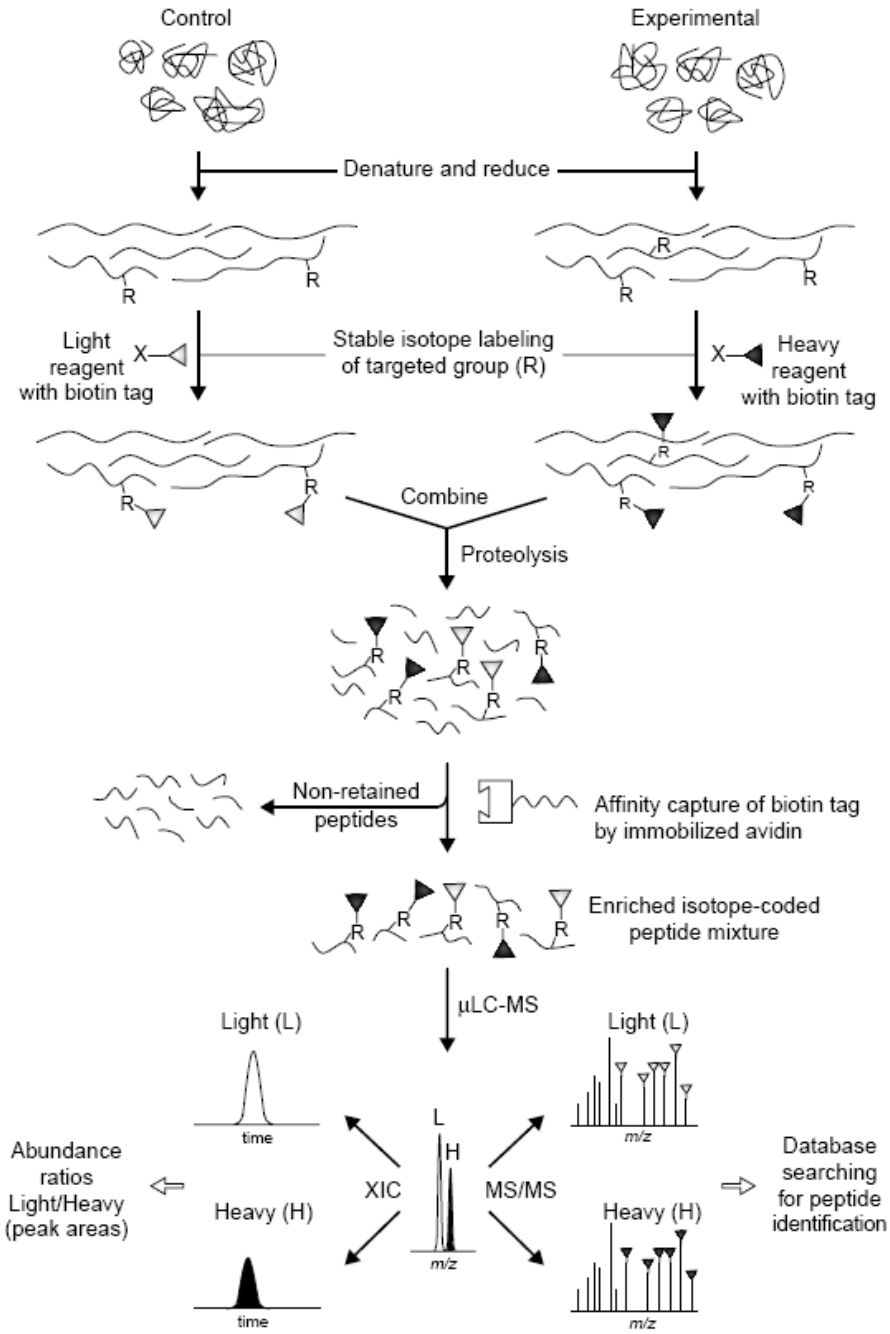


Figure 41 : Le principe de ma méthode de marquage ICAT.

Le principe de la méthode ICAT est présenté Figure 41. Les protéines de chaque échantillon sont marquées par un tag ICAT spécifique. Les protéines sont alors mélangées et digérées par la trypsine. Comme défini ci-dessus, le tag permet un enrichissement des peptides marqués. La quantification se réalise une fois encore sur le spectre MS par l'analyse des paires de peptides et le calcul d'un ratio d'expression.

Deux générations de réactifs ICAT ont été développées. La première contient un tag isotopique constitué d'atomes d'hydrogène D_0 ou D_8 et dont la masse du réactif sous forme légère (D_0) est de 570,5 Da (masse moléculaire moyenne). Le problème est qu'il se crée parfois un décalage

entre le temps d'élution du peptide deutéré (D_8) par rapport à celui non deutéré (D_0). Celui-ci a été résolu par l'apparition d'une deuxième génération de tag ICAT, contenant des atomes de carbone ^{12}C ou ^{13}C, co-élués en chromatographie. De plus, cette deuxième génération possède une liaison acido-labile permettant de cliver la biotine après concentration des peptides lors de l'étape d'affinité, ce qui facilite leur analyse par MS/MS (masse moyenne du réactif léger après clivage = 227 Da).

Le gros désavantage de cette technique est qu'elle ne marque que les cystéines, or la répartition de cet acide aminé est hétérogène selon les protéines, d'où le risque de perte d'information. Notons qu'il existe néanmoins des tags ICAT se fixant sur les lysines (KICAT).

2.4.2. iTRAQ

Les différentes approches de marquage isotopique développées précédemment reposent toutes sur un principe similaire visant à obtenir des paires de peptides sur les spectres MS afin d'accéder à leur quantification relative. Cela pose le problème de complexifier le spectre (nombre de pics deux fois plus importants) et d'avoir un risque de chevauchement des profils faussant la quantification. De plus ces techniques ne comparent que deux conditions différentes.

Dans l'optique de surmonter ces limitations, un nouveau type de marquage chimique a été mis au point : l'iTRAQ (isobaric Tags for Relative and Absolute Quantitation). Ce marquage a été développé en 2004 par Ross *et al.* [66] et est actuellement commercialisé par la société Applied Biosystems. Il permet de comparer jusqu'à 8 conditions différentes simultanément.

Comme pour l'ICAT, ce marquage repose sur la fixation d'un tag chimique particulier, par contre celui-ci est dit isobarique, la masse ajoutée au peptide étant la même quelque soit le tag iTRAQ utilisé.

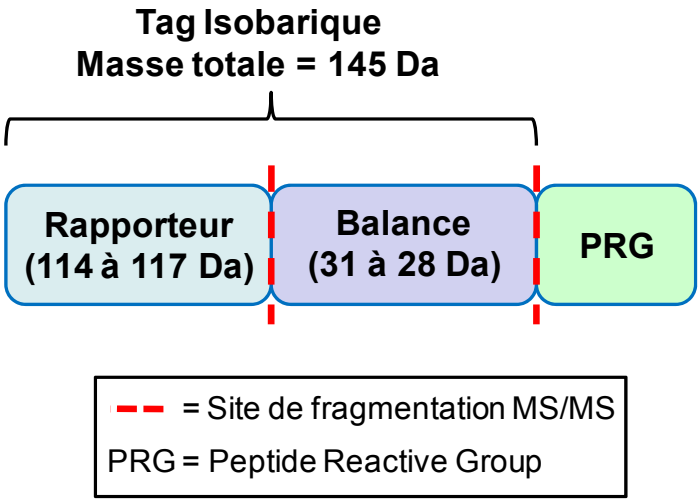


Figure 42 : Structure schématique du tag iTRAQ 4-plex.

La structure du tag iTRAQ est schématisée Figure 42. Il est composé de trois parties :

- ***Un groupement réactif spécifique des fonctions amines*** (groupement N-hydroxy-succinimide ester). Il permet la fixation covalente du tag aux fonctions amines N-terminales des peptides ainsi qu'aux chaînes latérales des lysines.
- ***Un groupement « rapporteur » chargé*** de *m/z* 114, 115, 116 ou 117 selon le tag.
- ***Un groupement « perte de neutre »*** non chargé de 31 à 28 Da qui permet de contrebalancer les masses du groupement rapporteur de façon à avoir quelque soit le tag une masse finale de *m/z* 145 (exemple : perte de neutre de 31 Da pour le groupement rapporteur 114, de 30 Da pour le 115,…). Un site de fragmentation préférentielle en MS/MS est présent de chaque côté de ce groupement.

La Figure 43 décrit la structure chimique des différents tags avec la composition isotopique les différenciant.

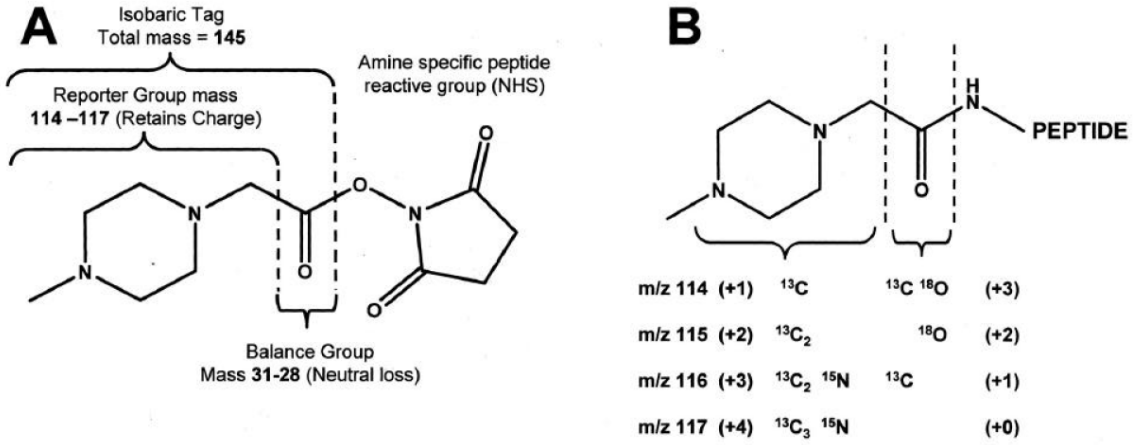


Figure 43 : Structure chimique des différents tags iTRAQ 114, 115, 116, et 117.
A : Structure chimique générale du tag avant fixation au peptide.
B : Structure chimique du tag après fixation au peptide avec les différentes formes isotopiques.

Le principe d'une analyse iTRAQ 4-plex classique est présenté Figure 44. Les différents échantillons protéiques sont digérés à la trypsine et les peptides résultants marqués avec un tag iTRAQ particulier. Après leur marquage spécifique, les peptides sont mélangés, séparés par chromatographie et finalement analysés en spectrométrie de masse. Les tags étant isobariques, les différents peptides marqués apparaissent lors du scan MS sous la forme d'un seul et même pic (correspondant à la masse du peptide + les 145 Da ajoutés par le tag). La quantification se réalise lors de l'analyse MS/MS. En effet, durant la fragmentation du peptide, il y a une coupure préférentielle de part et d'autre du groupement perte de neutre. La libération des groupements rapporteurs de *m/z* 114 à 117 est donc proportionnelle à la quantité relative du peptide dans

chacune des conditions auxquelles ils se rapportent. La fenêtre de *m/z* 114 à 117 a été choisie car aucun fragment ou ion immonium pouvant fausser la quantification ne se situe dans cette zone. Le fragment perte de neutre n'étant pas chargé, il n'est pas analysé en spectrométrie de masse. Le reste du spectre MS/MS est classique et permet de visualiser les différents fragments du peptide et donc d'identifier ce peptide. Les étapes d'identification et de quantification sont donc réalisées simultanément durant l'analyse MS/MS.

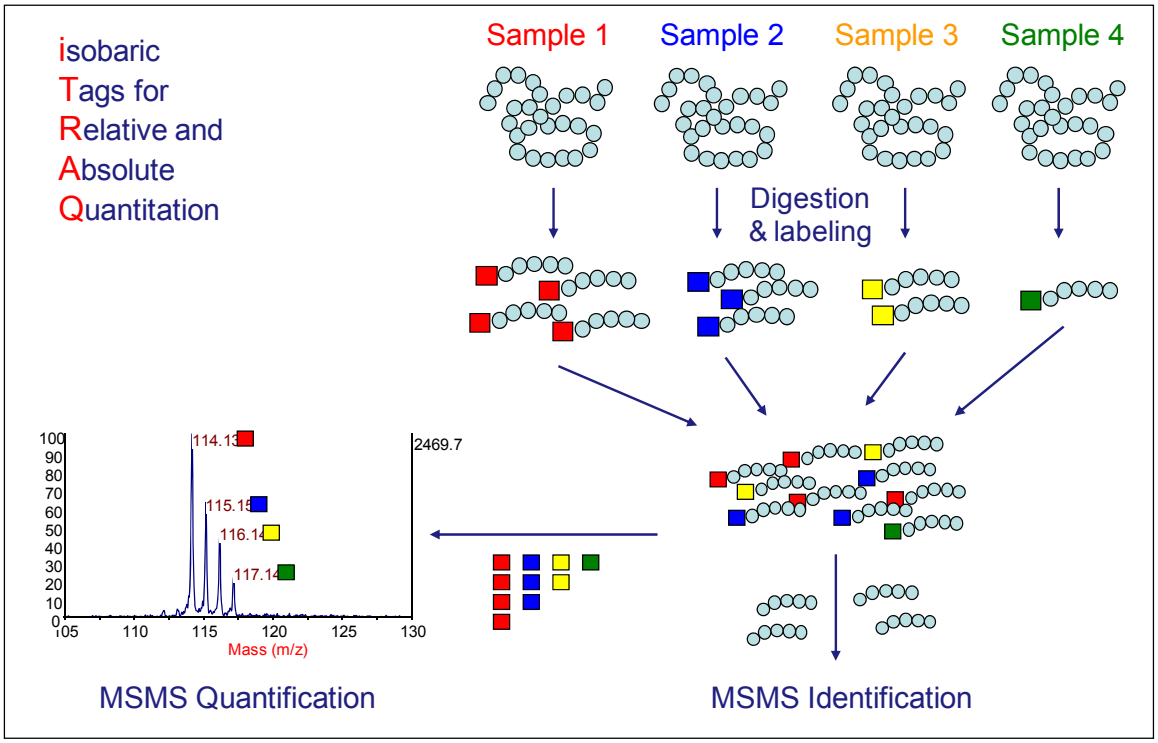


Figure 44 : La méthode iTRAQ 4-plex.

La méthode iTRAQ présente de nombreux avantages. Le tag étant isobarique, il évite la complexification du spectre MS et le risque de superposition des pics. Il augmente aussi la sensibilité de détection : chaque pic correspond à la somme du même peptide marqué pour chaque condition. Un autre avantage clé réside dans sa capacité à réaliser des analyses multiplexes (jusqu'à 8-plex). Enfin, une quantification absolue est également possible par l'ajout d'un standard interne. Cependant, son inconvénient majeur réside dans le marquage tardif des échantillons qui augmente les risques d'erreur de quantification dûs à des pertes inégales de protéines durant la préparation de l'échantillon (lyse cellulaire, extraction,…).

La dernière génération de tags iTRAQ permet de réaliser des analyses 8-plex pour comparer jusqu'à 8 conditions différentes [74]. Le groupement rapporteur du tag s'étend de 113 à 121 Da, avec la masse à 120 Da laissée libre pour les ions immonium phénylalanine potentiels (Figure 45).

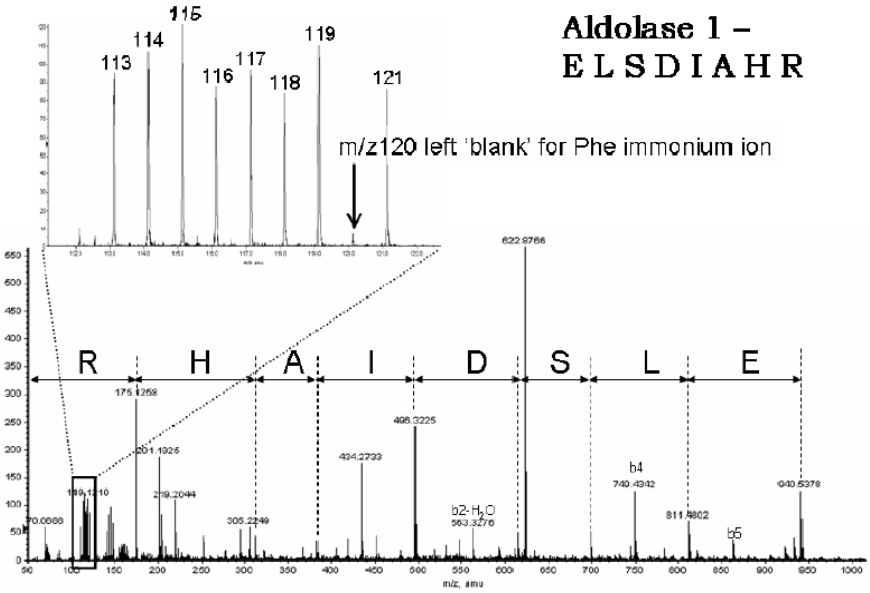


Figure 45 : Exemple de spectre MS/MS obtenu avec la méthode iTRAQ 8-plex.

L'ajout de 4 masses supplémentaires a nécessité de modifier un peu la composition chimique du tag, même si le principe reste le même avec une partie spécifique des groupements amines, un groupement perte de neutre et un rapporteur (Figure 46). Le tag est isobarique et a une masse totale de 305 Da (contre 145 Da pour le tag 4-plex). Cette augmentation de masse est « à double tranchant » : elle permet d'avoir accès à des peptides de masse faible (< 750 *m/z*) qui sont généralement perdu dans le bruit. D'un autre côté, certains peptides de masse plus importante (> 2500 *m/z*), peuvent être moins bien détectés, notamment avec une source MALDI.

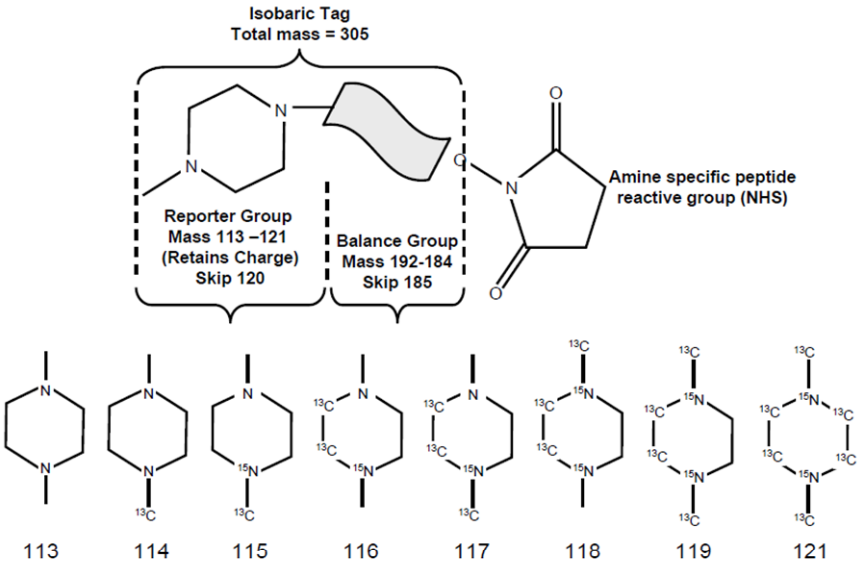


Figure 46 : Structure chimique des tags iTRAQ 8-plex.

L'analyse iTRAQ 8-plex permet de tester plus de conditions simultanément ou d'augmenter les répliquas biologiques. Cependant, la multiplication des conditions à tendance à linéariser les

profils de pics rapporteurs et donc à sous-estimer les différences réelles [75]. Cette génération de tag 8-plex est prometteuse mais encore jeune et peu de publications ont validé pour l'instant son efficacité [75, 76].

Enfin, d'autres méthodes utilisant un tag isobarique ont également été développées. Citons notamment, les approches TMT (isobaric Tandem Mass Tags) de Thermo Scientific et Exactag de PerkinElmer.

L'approche TMT permet de faire des analyses 6-plex et génère des groupements rapporteurs compris entre 126 et 131 Da [77, 78]. Son principe est très similaire à celui de l'iTRAQ avec une fixation des tags sur les fonctions amines au niveau peptidique et une quantification en MS/MS.

L'approche Exactag a été décrite récemment [79]. On a peu de recul sur son efficacité réelle puisqu'aucune étude l'utilisant n'a été publiée à ce jour. Elle permet de faire des analyses 10-plex et a l'avantage de marquer les échantillons au niveau protéique (sur les résidus Cystéine ou Lysine), donc plus précocement que l'iTRAQ. Par contre, la taille du tag semble présenter un inconvénient majeur. Celui-ci fait 972 Da et génère des groupements rapporteurs à 343, 349, 355, 400, 406, 412, 457, 463, 469 et 475 Da. Ce tag volumineux peut avoir des effets sur la séparation des échantillons et rend compliquée son utilisation avec un spectromètre de masse à source MALDI générant des ions mono-chargés.

3. Les méthodes « label-free »

Récemment, de nouvelles approches ont été développées pour obtenir des informations semi-quantitatives sans modification ou marquage préalable des échantillons. Celles-ci sont dites méthodes « label-free ».

- ***La stratégie PMSS*** (Peptide Match Score Summation) représente une première approche [80]. Elle repose sur l'hypothèse que le score attribué à une protéine par un moteur de recherche est basé sur la somme des scores d'identification de ses peptides et donc que le score protéique est proportionnel à la quantité de protéines. La comparaison des scores peut ainsi donner des informations semi-quantitatives.
- ***L'approche SpS*** (spectrum sampling) est très similaire et repose sur le décompte du nombre de spectres MS/MS associés à une même protéine [81].
- ***La méthode PAIs*** (Protein Abundance Indices) est jugée plus fiable car basée sur différents paramètres observables. Par exemple, le nombre de peptides identifiant une protéine augmente avec l'abondance de la protéine. Comme les grosses protéines génèreront

statistiquement plus de peptides mesurables que les plus petites, le PAI est calculé en normalisant le nombre de peptides observés par le nombre de peptides théoriquement observables pour la protéine d'intérêt [82, 83]. Ishihama et son équipe ont décrit une variante nommée emPAI (exponentialy modified PAI) établissant une relation logarithmique entre le nombre de peptides observés et la quantité de protéine dans un échantillon donné [84].

Malheureusement, il n'est pas possible de prédire la réponse du détecteur MS pour un peptide particulier car les propriétés d'extraction et d'ionisation sont peptides-dépendantes. En conséquence, le courant d'ions (XICs pour extracted ion currents) obtenu pour les différents peptides d'une même protéine est très différent même s'ils sont en réalité présents en même concentration. Néanmoins, ces sources d'erreurs ne s'appliquent théoriquement pas lorsque l'on compare un même peptide issu de différents runs chromatographiques réalisés selon des conditions expérimentales identiques. En conséquence, deux protéomes peuvent être comparés s'ils sont analysés l'un après l'autre selon des paramètres de séparation / analyse reproductibles [85, 86]. Des logiciels ont été développés pour extraire les intensités observées pour un même peptide dans différents runs afin de déterminer et comparer leur abondance relative. On peut citer, par exemple, MSight [87], SuperHirn [88], MapQuant [89], SpecArray [90] ou VEMS [91-93].

Les principaux avantages de ce type de méthodes sont l'absence de marquage et leur application quelque soit l'instrument de masse utilisé. Cependant, leur gros inconvénient réside dans le risque d'erreurs multiples durant les procédés d'analyse en parallèle. Elles nécessitent donc la réalisation de LC et MS très précises et reproductibles.

4. Les méthodes avec standard : vers une quantification absolue

Les approches décrites précédemment sont très utiles pour la quantification relative de protéomes complexes non caractérisés en fonction de différentes conditions. Néanmoins, elles ne sont pas ou peu adaptées à la quantification absolue des variations observées. La quantification absolue nécessite de connaître au préalable la protéine à doser. La technique de choix repose sur un dosage de type ELISA (à condition d'avoir à disposition l'anticorps dirigé contre la protéine à doser). Des techniques protéomiques basées sur l'utilisation d'un standard interne ont été développées : AQUA, QCAT ou PSAQ (Figure 47).

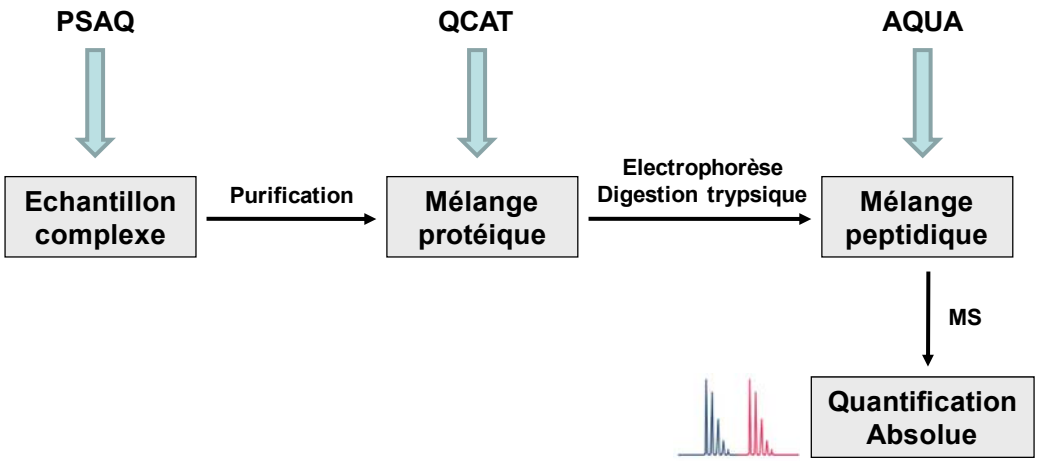


Figure 47 : La quantification absolue en protéomique.

4.1. Peptide AQUA

La méthode de quantification absolue par peptide AQUA (Absolute QUAntification) a été développée en 2003 par Gygi *et al.* [94].

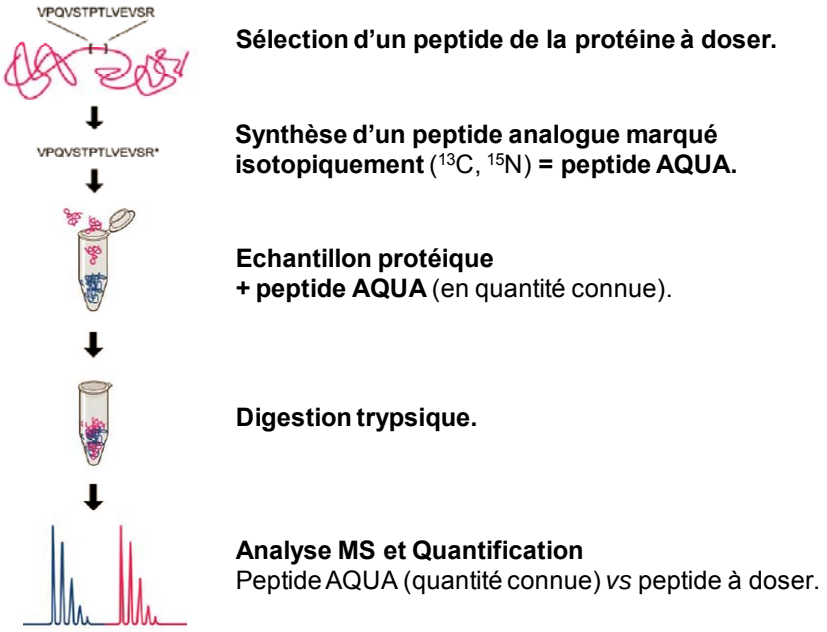


Figure 48 : Principe de l'approche AQUA (Sigma).

Son principe repose sur l'utilisation d'un peptide trypsique synthétique spécifique de la protéine d'intérêt à doser (Figure 48). Ce peptide est synthétisé en y incorporant des formes isotopiques stables de certains acides aminés (par exemple $^{13}C_6$-arginine). Il est alors ajouté en quantité connue dans l'échantillon protéique biologique. L'échantillon est protéolysé et analysé en spectrométrie de masse. Le spectre MS met en évidence une paire de pics au niveau du peptide d'intérêt. Le peptide synthétique ayant été ajouté en quantité connue, la quantité de peptide natif, donc de la protéine peut être déduite du ratio d'expression.

Les limites de cette approche sont que la quantification est faite à partir d'un seul peptide et que ce peptide AQUA est ajouté tardivement à l'échantillon. Le dosage ne tiendra donc pas compte des pertes induites par l'ensemble des étapes antérieures à l'ajout du peptide AQUA : extraction, purification, digestion des protéines,… De plus, la difficulté à obtenir de tels peptides en quantité exacte, limite cette approche.

4.2. QCAT ou QconCAT

Les techniques QCAT [95] ou QconCAT [96] ont un principe similaire à l'AQUA mais présentent l'avantage de permettre la quantification absolue de plusieurs protéines simultanément (Figure 49). Elles reposent sur la sélection d'une série de peptides trypsiques caractéristiques des protéines d'intérêt à quantifier. Un gène synthétique codant pour un concatémère de ces différents peptides est façonné. Son expression permettra la production d'une protéine chimérique constituée d'un accolement bout à bout des peptides de chaque protéine à doser. Comme pour la technique AQUA, cette protéine artificielle est synthétisée en y incorporant des formes isotopiques stables de certains acides aminés, permettant ainsi la quantification à partir du spectre MS. Cette protéine concatémère est ajoutée à l'échantillon biologique avant la digestion trypsique afin d'éliminer le biais dû, par exemple, à une digestion partielle. Par contre cette protéine étant chimérique la digestion peut-être gênée ou au contraire privilégiée par rapport aux protéines natives. De plus, il n'est pas facile de produire une telle protéine de façon stable.

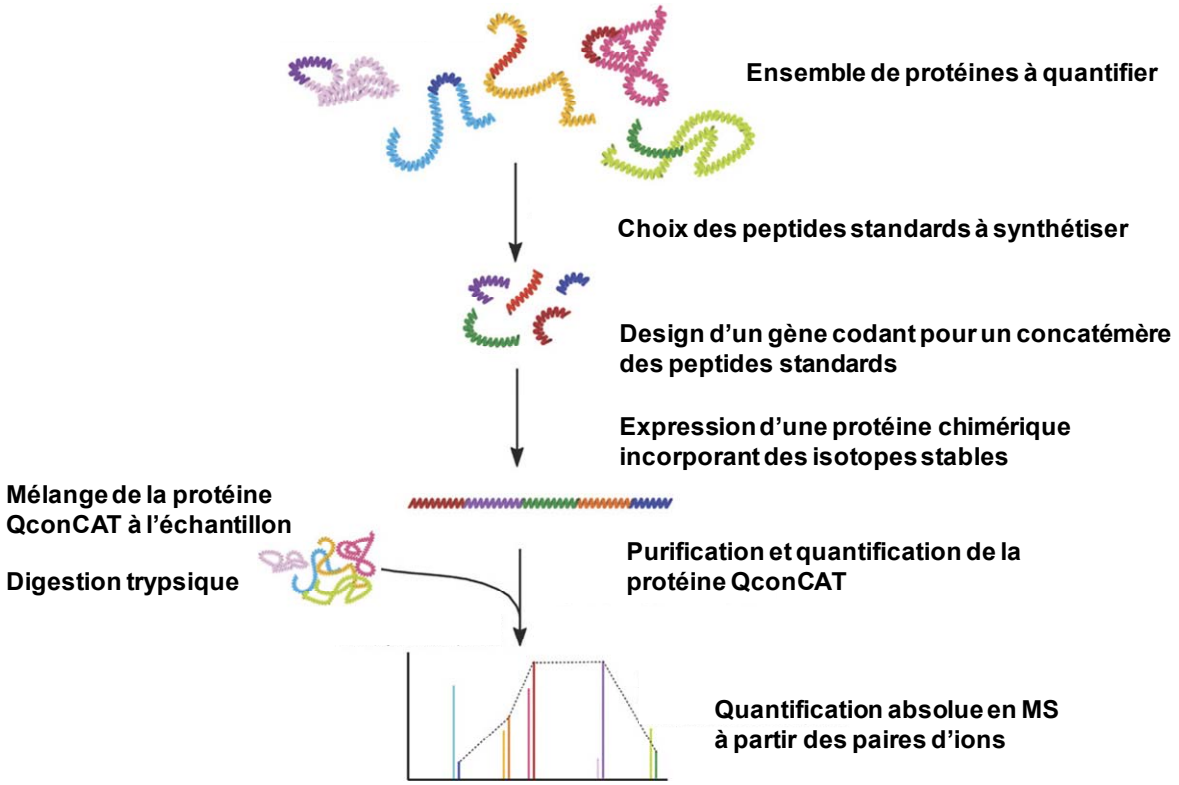


Figure 49 : Principe de l'approche QconCAT.

4.3. PSAQ

Bien que les méthodes AQUA et QconCAT aient apporté des avancées notoires dans la quantification absolue de protéines d'échantillons biologiques complexes, la quantification absolue par ces techniques souffre des limitations suivantes :

- Une incapacité à prendre en compte l'efficacité réelle de l'étape de protéolyse avant l'analyse MS.
- Une incompatibilité avec le pré-fractionnement d'échantillon qui est souvent nécessaire lorsque l'on travaille sur des échantillons biologiques.
- Une faible couverture de la séquence protéique, limitant la fiabilité de la quantification.

Pour surmonter ces inconvénients, Brun *et al.* ont proposé la méthode PSAQ (Protein Standard Absolute Quantification) qui repose sur la production *in vitro* de la protéine à quantifier sous la forme d'une protéine recombinante entière marquée isotopiquement [97, 98]. Cette protéine est ajoutée à l'échantillon très en amont dans le processus d'analyse et digérée exactement de la même façon que la protéine native. De plus, la quantification est beaucoup plus fiable car elle se fait à partir de l'ensemble des peptides trypsiques de la protéine et non pas un seul comme lors des approches précédentes. Par contre, elle nécessite la production *in vitro* de la protéine, ce qui est souvent long, compliqué voire impossible. Il se pose aussi le problème des modifications post-traductionnelles, non prises en compte généralement dans la protéine recombinante.

CHAPITRE 3 : Le cancer du poumon

1. Introduction

Le cancer du poumon est le cancer le plus fréquent au niveau mondial, à la fois en terme d'incidence (1,2 millions de nouveaux cas chaque année dans le monde, soit 12.3 % des cancers) et de mortalité (1,1 millions de morts, soit 17,8 % de la mortalité totale due aux cancers) (Figure 50). Le rapport mortalité/incidence est donc égal à 0.9, ce qui reflète la gravité de cette tumeur.

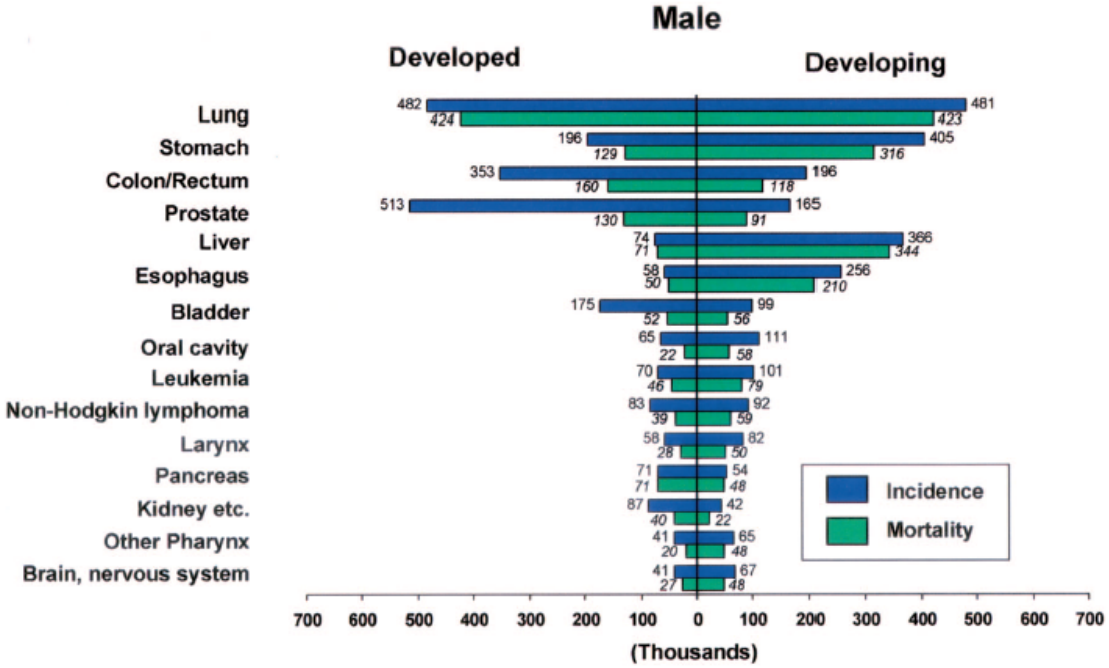


Figure 50 : Diagramme représentant le nombre estimé de nouveaux cas de cancer (incidence) et de décès (mortalité) dans le monde en 2002 (d'après Parkin et al. [99]).
Les données sont exprimées en milliers de cas pour chaque type de cancer répertoriés pour les pays développés et en voie de développement. Seuls les individus mâles sont représentés ici. Le diagramme équivalent pour les femmes est disponible dans la publication de Parkin et al [99].

Le facteur de risque principal est le tabagisme actif, responsable de 90 % des cas de cancers du poumon. Le tabac contient en effet de nombreux carcinogènes dont le principal est le benzo-α-pyrène (le goudron). D'autres facteurs environnementaux ont une incidence, tels que l'inhalation d'amiante par exemple.

Le taux d'incidence chez l'homme est de 34,9 pour 100 000, avec des taux plus élevés dans les pays développés (ce taux est en baisse dans les pays où la consommation de tabac diminue). Chez les femmes, le taux d'incidence est inférieur (11,1 pour 100 000) mais ne cesse d'augmenter (à corréler avec l'augmentation de la consommation de tabac chez les femmes depuis une quinzaine d'années). Aux Etats-Unis, le nombre de femmes décédées d'un cancer du poumon dépasse depuis peu les décès dus à un cancer du sein [99-101].

2. Classification / Histologie

L'observation au microscope de la taille et de l'apparence des cellules malignes permet de classer les cancers du poumon en deux grands groupes histologiques [102] :

- ***les cancers du poumon non à petites cellules*** ou NSCLC (pour Non-Small Cell Lung Carcinoma)
- ***les cancers du poumon à petites cellules*** ou SCLC (pour Small Cell Lung Carcinoma)

2.1. Cancers du poumon non à petites cellules

Les cancers du poumon non à petites cellules (NSCLC pour Non Small Cell Lung Carcinoma) représentent 75 à 85 % des cancers du poumon. Il existe trois sous-types de NSCLC : les cancers du poumon à cellules squameuses (ou cellules malpighiennes), les adénocarcinomes et les cancers du poumon à grandes cellules. Ces trois sous-types sont pris en charge et traités de façon similaire.

2.2. Cancers du poumon à petites cellules

Les cancers du poumon à petites cellules (SCLC pour Small Cell Lung Cancer) représentent 15 à 25 % des cancers du poumon. Ces cellules possèdent des vésicules remplies d'hormones neuroendocrines. Ce type de cancer est grandement associé au tabagisme. Il est traité par chimio ou radiothérapie mais possède généralement un mauvais pronostic avec un fort taux de métastases.

3. Signes et symptômes

Les patients atteints de cancer du poumon ne présentent pas de symptômes spécifiques, notamment à un stade très précoce de la maladie. Les principaux signes reflètent généralement un stade assez avancé du cancer. Ces symptômes sont : une toux persistante avec apparition de crachats sanglants, une difficulté à respirer, un essoufflement anormal, une douleur à la poitrine et un amaigrissement inhabituel [103]. Cette non-spécificité des symptômes nécessite la réalisation d'examens plus poussés tels que la visualisation de la tumeur par imagerie.

4. Diagnostic

Les techniques d'imagerie telles que la tomographie (PET : Positron Emission Tomography, ou CET : Compute Emission Tomography) ou l'Imagerie par Résonance Magnétique (IRM) sont couramment employées pour le diagnostic des cancers mais présentent une limite de détection insuffisante pour leur diagnostic précoce.

La tomographie de type CT (spiral computed tomography), basée sur l'utilisation de source à rayons X en rotation rapide, produit des images avec une bonne résolution spatiale et de contraste. Elle permet de détecter des tumeurs du poumon de quelques millimètres de diamètre [104, 105]. Des études pilotes rapportent la détection de tumeur de stade I avec une sensibilité de 70-85% et une taille moyenne des cancers détectés inférieure à 20 mm avec une médiane à 10 mm environ [104]. L'étude ELCAP (Early Lung Cancer Action Project), incluant 1000 individus fumeurs classés à haut risque, a démontré que la CT est beaucoup plus sensible qu'une radio des poumons pour la détection précoce de cancer du poumon [104].

Une biopsie est généralement faite pour confirmer le diagnostic et caractériser la tumeur.

Néanmoins, les limites de détection des techniques actuelles engendrent un grand intérêt pour la recherche de bio-marqueurs sériques pour le diagnostic précoce.

5. Traitements

5.1. Cancers du poumon non à petites cellules

Le traitement des cancers du poumon non à petites cellules fait appel aux résections chirurgicales, à la radiothérapie, à la chimiothérapie et aux thérapeutiques ciblées (actuellement l'Erlotinib et le Cetuximab) [106, 107].

Les formes précoces (stades I et II) sont traitées chirurgicalement (lobectomie ou pneumonectomie), avec une chimiothérapie complémentaire dans certains cas, appelée chimiothérapie adjuvante. Les cancers bronchiques plus tardifs mais non métastatiques (stade III) sont difficilement opérables en raison d'une atteinte thoracique trop importante. Ils sont donc traités par chimiothérapie et radiothérapie. Les patients avec un cancer métastatique (de stade IV ou "généralisé") peuvent bénéficier d'une chimiothérapie et de thérapeutiques ciblées. La radiothérapie peut être administrée dans certains cas, à visée palliative, pour réduire les symptômes, par exemple les douleurs d'une métastase osseuse.

Les taux de survie à 5 ans dépendent grandement du stade de la tumeur lors de la prise en charge du patient. Ainsi, les patients diagnostiqués au stade I ont une survie à 5 ans de 60 à 70 %, les

patients de stade II de 40 à 50%, les patients de stade III de 15 à 30 %. Les patients de stade IV ont un taux de survie très faible (survie de 8 mois à une année) [106].

5.2. Cancers du poumon à petites cellules

Les cancers du poumon à petites cellules sont caractérisés par une prolifération cellulaire rapide et une tendance à la métastase précoce. En clinique, deux stades de SCLC sont distingués : un stade limité où la tumeur est présente dans un seul hémithorax, et un stade étendu avec des métastases. Environ 20 à 25 % des patients ont un stade limité, traitable par résection chirurgicale. Néanmoins, les taux de survie à 5 ans restent faibles, de 15 à 25 %. Des approches couplant chimio et radiothérapies sont aussi recommandées [106-108]. Pour les stades avancés, le traitement de choix est la chimiothérapie, généralement la cisplatine, la carboplatine ou l'etoposide. Les approches actuelles utilisent aussi de nouvelles drogues telles que l'inhibiteur de la topoisomérase I et les taxans [106-108].

6. Les bio-marqueurs : une nécessité

6.1. Généralités

Les bio-marqueurs réfèrent à tous les moyens de mesure quantitative de l'homéostasie biologique, permettant de distinguer ce qui est anormal de ce qui est normal [109]. En conséquence, la définition la plus simple d'un bio-marqueur est : une molécule indiquant une altération dans la physiologie normale. En pratique, un bio-marqueur doit avoir une utilité en clinique [110]. En cancérologie, les bio-marqueurs sont non seulement utiles au diagnostic des cancers mais apportent aussi des informations importantes en thérapie telles que la vérification du stade ou la réponse thérapeutique. Les bio-marqueurs tumoraux contribuent également à la compréhension des différents stades de cette maladie. Actuellement, la classification des cancers repose essentiellement sur la classification anatomo-pathologiste développée en France par Pierre Denoix dans les années 1940 (système TNM pour Tumor, Node, metastases). Les bio-marqueurs pourraient donner plus d'informations dans certains domaines de la biologie non renseignés par le système TNM. En clinique, les bio-marqueurs tumoraux sont utiles à la sélection d'un traitement approprié. Actuellement, des thérapies tumorales sont développées en tenant compte de bio-marqueurs spécifiques, comme par exemple le « trastuzumab » pour les patientes atteintes d'un cancer du sein HER2/neu positif [111].

6.2. Types de bio-marqueurs tumoraux

Il existe plusieurs types de bio-marqueurs tumoraux rattachés à différents domaines : la génomique/transcriptomique, l'épigénétique, la protéomique, la métabolomique ou les techniques d'imagerie.

La génomique/transcriptomique utilise des puces ADN, la PCR, la RT-PCR, le séquençage d'ADN, le marquage FISH (fluorescent in situ hybridization), etc… pour détecter les altérations génétiques accompagnant l'état cancéreux.

L'analyse des modifications épigénétiques est aussi une source de bio-marqueurs. De telles modifications sont souvent rencontrées dans les îlots CpG de la région régulatrice des gènes, causant une diminution de l'expression des gènes concernés. Ces altérations peuvent être à l'origine de la sortie des cellules du cycle normal et aboutir à la formation de cancer [112, 113].

Les techniques protéomiques incluent la spectrométrie de masse, les ELISA, l'immunohistochimie, etc…, et visent à découvrir de nouveau bio-marqueurs protéiques.

La métabolomique étudie non pas les macromolécules (ADN, protéines) mais des molécules de faible poids moléculaire ou des métabolites tels que les acides aminés, les peptides, les lipides ou les carbohydrates. Le métabolome est prédit comme contenant seulement 2500 petites molécules et peut représenter une source de marqueurs.

6.3. Applications en diagnostic et thérapeutique

Les bio-marqueurs tumoraux reflètent la nature hétérogène des cancers de façon plus adéquate que le système TNM. D'un point de vue histologique, les cancers du poumon sont divisés en deux sous-types présentés précédemment (SCLC et NSCLC). D'autres critères de classification sont utilisables tels que la présence ou non de mutation de l'EGFR. Ces sous-catégories ne sont pas déterminées sans des actes invasifs pour le patient tels que les biopsies. En conséquence, le diagnostic fiable d'un sous-type de cancer par mesure de bio-marqueur spécifique serait un réel bénéfice pour les patients. Ces marqueurs sont non seulement attendus comme facteurs de prédisposition chez les patients à haut risque tels que les grands fumeurs mais aussi pour le diagnostic précoce du cancer [109]. Ces marqueurs seraient aussi utiles pour orienter la thérapie car leur niveau d'expression peut varier en réponse au traitement. Ils sont utilisables pour caractériser les cibles de drogues. Citons, par exemple, l'aminopeptidase-p et l'annexine A1 découverts au niveau des surfaces endothéliales dans un modèle de cancer du poumon chez la souris [114]. A l'inverse, ils peuvent aider au design de nouvelles drogues.

6.4. Les bio-marqueurs actuels du cancer du poumon

La forte mortalité associée au cancer du poumon est due en grande partie à l'incapacité actuelle de le diagnostiquer de façon suffisamment précoce. En effet, les outils utilisés en clinique (tomographie, bronchoscopie et analyse des crachats) ne sont pas suffisants. En conséquence, la découverte de nouveaux bio-marqueurs spécifiques du cancer du poumon est très importante pour la détection précoce et la thérapie ciblée.

Il n'existe actuellement aucun marqueur ou combinaison de marqueurs utilisés en clinique pour le diagnostic précoce du cancer du poumon dans des populations asymptomatiques ou dans un groupe de patients à risque (tels que les fumeurs). Néanmoins certains marqueurs ont pu être caractérisés [115]. On trouve ainsi, au niveau génétique, l'hyperméthylation des promoteurs et la mutation des gènes K-*ras* et p53 [116]. Les niveaux d'ARNm ne sont pas toujours directement corrélés aux niveaux des protéines, les molécules réalisant les fonctions [117]. En effet, il existe une bien plus grande variété de formes protéiques que les 20 488 gènes décrits dans le génome humain [11]. Celle-ci est notamment due à l'épissage alternatif, la protéolyse et les modifications post-traductionnelles. Cela signifie que les bio-marqueurs protéiques peuvent être plus spécifiques du type et du statut du cancer.

Les bio-marqueurs protéiques du cancer du poumon sont classés en 3 catégories en fonction de la source de protéines : sérum, tissu ou crachat [115]. Dans les crachats, les cellules cancéreuses issues des sites cancéreux sont les principales sources de protéines. Dans les biopsies de tissu pulmonaire, on trouve non seulement les cellules cancéreuses mais aussi d'autres molécules impliquées dans les défenses de l'organisme telles que les cellules de l'immunité, les cytokines et des dérivés des réponses immune et inflammatoire. Le sang est la source potentiellement la plus riche. On peut y trouver, par exemple, les bio-marqueurs trouvés dans les biopsies ainsi que plusieurs fragments protéiques générés dans le tissu environnant ou par des protéines ou cellules circulantes. Dans la mesure où les caractéristiques ultimes d'un bio-marqueur sont d'être spécifique, précoce, détectable de façon non invasive et de permettre de suivre l'évolution post-thérapie du cancer, le sang représente une source de protéines pertinente.

Des marqueurs protéiques ont ainsi été décrits comme par exemple, la neuron specific enolase (NSE), le carcinoembryonic antigen (CEA), les fragments de cytokeratine 19 (CYFRA 21-1) squamous cell carcinoma antigen (SCC), le cancer antigen CA 125 (CA 125), le tissue polypeptide antigen (TPA) ou la kallikrein B1 (KLKB1) (Tableau 4 et Tableau 5) [115, 116].

Protéine	Diagnostique	Suivi thérapeutique	Suivi pronostic	Ontologie	Détails
CEA	AdenoCA, SCLC (>10µg/L)	AdenoCA, NSCLC avancé	AdenoCA, NSCLC	Constituant de la membrane cellulaire, Superfamille des Ig à ancre lipidique	Utilisé combiné avec CYFRA. Souvent élevé chez les fumeurs
CYFRA21-1	SCC, NSCLC (NSCLC : sensibilité 23-70%)	NSCLC avancé	NSCLC, SCC	Constituant du cytosquelette	Souvent élevé chez des patients ayant des maladies bénignes du poumon
TPA	NSCLC, SCC	-	NSCLC	-	-
ProGRP	SCLC (sensibilité 47-86%) (> 200 ng/L : forte suspicion)	SCLC	-	Hormone neuropeptide	Elevé dans les maladies rénales et bénignes du poumon
NSE	SCLC (sensibilité 74%) (> 100 µg/L : forte probabilité)	SCLC	SCLC	Activité phosphoglycérate déhydrogénase	Utilisé combiné avec ProGRP. Peut corrélé avec un faible taux de survie. Elevé dans les maladies inflammatoires.
Tumor M2-pyruvate kinase	AdenoCA (sensibilité 50-71%)	-	AdenoCA	Activité pyruvate kinase, glycolyse	Elevé dans de multiple maladies malignes et inflammatoires

CEA = carcinoembryonic antigen, CYFRA 21-1 = cytokeratin 19 fragment, TPA = tissue polypeptide antigen, ProGRP = progastrin-releasing peptide, NSE = neuron-specific enolase, AdenoCA = adenocarcinoma, SCC = squamous cell carcinoma, SCLC = small cell lung cancer, NSCLC = non-small cell lung cancer

Tableau 4 : Les bio-marqueurs protéiques actuellement utilisés pour le cancer du poumon (d'après Sung et Cho, Tableau 2) [115, 116].

Protéine	Diagnostique	Ontologie	Détails
Serum amyloid A	Cancer du poumon	Activité transporteur lipidique	Elevé (jusqu'à 62,4 µg/mL contre 2 µg/mL en moyenne chez les contrôles sains)
Haptoglobin-alpha-2	AdenoCA	Protéolyse, réponse immune	-
APOA1	AdenoCA	Activité transporteur lipidique, inhibiteur d'activité lipase	Fragment de l'apolipoprotéine A-1 diminué chez les patients cancéreux
KLKB1	AdenoCA	Protéolyse	Fragment de 17-18 kDA de la kallkrein B1

Tableau 5: Les bio-marqueurs protéiques potentiels du cancer du poumon (d'après Sung et Cho) [115, 116].

Les bio-marqueurs disponibles pour le cancer du poumon ne sont pas suffisamment sensibles et spécifiques pour être utilisés cliniquement en diagnostic, stratification, pronostic ou pour la prédiction de la réponse aux drogues. Néanmoins, de gros efforts sont actuellement entrepris pour y remédier.

Trois aspects doivent être considérés. Tout d'abord, il est nécessaire d'analyser les marqueurs actuellement disponibles ou potentiels dans de grandes cohortes d'échantillons cliniques incluant d'autres cancers ou maladies (maladies inflammatoires). Deuxièmement, il est nécessaire de se focaliser sur un sous-type spécifique de cancers du poumon afin d'augmenter la spécificité et la

découverte de protéines minoritaires. Troisièmement, un seul bio-marqueur n'est pas suffisant pour prédire et suivre les cancers du poumon. Au contraire, une combinaison de bio-marqueurs fiables avec des informations quantitatives serait réellement utile en clinique.

CHAPITRE 4 : La protéine p53

Le gène *TP53* est un gène suppresseur de tumeur jouant un rôle primordial dans le maintien de l'intégrité du génome.

1. Introduction – Historique

La protéine nucléaire p53 a été mise en évidence en 1979 dans un modèle de cellules tumorales [118, 119]. Dans un premier temps, p53 est considéré comme un proto-oncogène. Ce n'est qu'en 1989 que le pouvoir transformant observé est attribué à une mutation du gène p53, la forme sauvage n'ayant quant à elle pas cette propriété [120, 121]. Bien au contraire, p53 est un anti-oncogène qui supprime donc l'effet transformant des oncogènes.
Intitulée « Molécule de l'année » en 1993 par le journal Science (Science 1993 Dec 24 ; 262(5142) :1953) et « gardien du génome » par de nombreux auteurs [122], l'intérêt qui lui est porté est tel qu'entre 1993 et 1996, 4300 articles sont publiés avec le terme « p53 » dans leur titre. Cet intérêt massif et unique pour une protéine reflète le rôle primordial qu'elle joue dans la régulation du cycle cellulaire, ainsi que l'importance et la fréquence de ses anomalies en pathologie maligne.

2. Structure de p53

Le gène *TP53* se trouve sur le bras court du chromosome 17 (17p13.1) [123]. Ce gène possède 11 exons avec plusieurs domaines grandement conservés au cours de l'évolution [124].
Il code pour une protéine p53. Il s'agit d'une phosphoprotéine de 53 kDa et de 393 acides aminés qui joue le rôle de facteur de transcription.

Transactivation (1-44)	SH³ binding (58-101)	DNA-binding (102-292)	Tetramerisation (324-355)	C-ter + NLS (363-393)

Figure 51 : Les domaines structuraux de la protéine p53.

La protéine p53 a l'architecture classique des facteurs de transcription et possède 5 domaines fonctionnels (Figure 51) :

- ***Un domaine de transactivation*** au niveau N-terminal (acides aminés 1-44). Ce site est impliqué dans l'activation de la transcription des gènes cibles de p53 , une fois la

fixation de p53 sur leur ADN [125]. Il joue aussi un rôle important dans le maintien de l'intégrité de p53 à travers son interaction avec la protéine mdm2 [126].

- ***Une région hydrophobe riche en prolines*** (acides aminés 58-101).
  Elle permet la fixation aux domaines SH^3 des molécules de transduction du signal.

- ***Un domaine central de fixation à l'ADN*** (acides aminés 102-292).
  Ce domaine, codé par les exons 5-8, permet la reconnaissance et l'interaction avec l'ADN double brins des cibles transcriptionnelles de p53 (répétitions de la séquence consensus RRRCA/TA/TGYYY) [127, 128]. Ce domaine forme le cœur de la protéine et est essentiel pour ses propriétés de transactivation. La majorité des mutations de p53 constatées dans les cancers sont situées dans cette région et induisent un changement de conformation de la protéine ou la perte de résidus indispensables à sa fixation sur l'ADN.

- ***Un domaine d'oligomérisation*** (acides aminés 324-355).
  La fixation de p53 sur l'ADN cible est optimisée lorsque 4 molécules interagissent simultanément. Cette oligomérisation se fait en deux étapes : la première consiste en la dimérisation de deux molécules et la seconde en l'association de deux dimères. L'oligomérisation de p53 est permise grâce à un domaine, très conservé au cours de l'évolution, situé juste après le domaine de fixation à l'ADN [129].

- ***Un domaine régulateur*** au niveau C-terminal (acides aminés 363-393).
  Cette région hautement basique influence la fixation séquence spécifique de p53 à l'ADN [130, 131]. Ce domaine contient aussi les trois séquences NLS (Nuclear Localisation Signal) de la protéine.

3. Rôles de la protéine p53

Une des raisons principales expliquant la forte fréquence de mutation de *TP53* dans les cancers (mutation dans environ 50 % des cancers) est que la protéine p53 joue un rôle clé dans la lutte contre la tumorogenèse.

L'activation des fonctions de p53 se fait en réponse à différents stimuli ou stress [132]. Ceux-ci incluent les agents responsables de lésions de l'ADN (les rayons UV, X ou gamma, les carcinogènes, les agents alkylants, les mycotoxines,…), l'hypoxie, la perturbation de l'adhésion cellulaire, la déplétion des ribonucléotides ou l'activation des cascades de signalisation de croissance cellulaire. La vitesse, le niveau et le mode d'activation de p53 varient selon la nature du stress, son intensité et la sensibilité du type cellulaire considéré [132]. En conséquence, il

n'existe pas de « réponse p53 » uniforme et les conséquences de l'induction de p53 varient grandement selon le stimulus [133].

La protéine p53 est exprimée de façon constitutive, mais ne s'accumule pas dans les cellules à cause de sa rapide dégradation par le protéasome. Cette dégradation est initiée par son interaction avec la protéine mdm2, principal régulateur de p53 (Figure 52). La protéine mdm2 est une ubiquitine ligase qui va ubiquitiner p53 et la transporter jusqu'au protéasome afin qu'elle y soit dégradée [126, 134, 135].

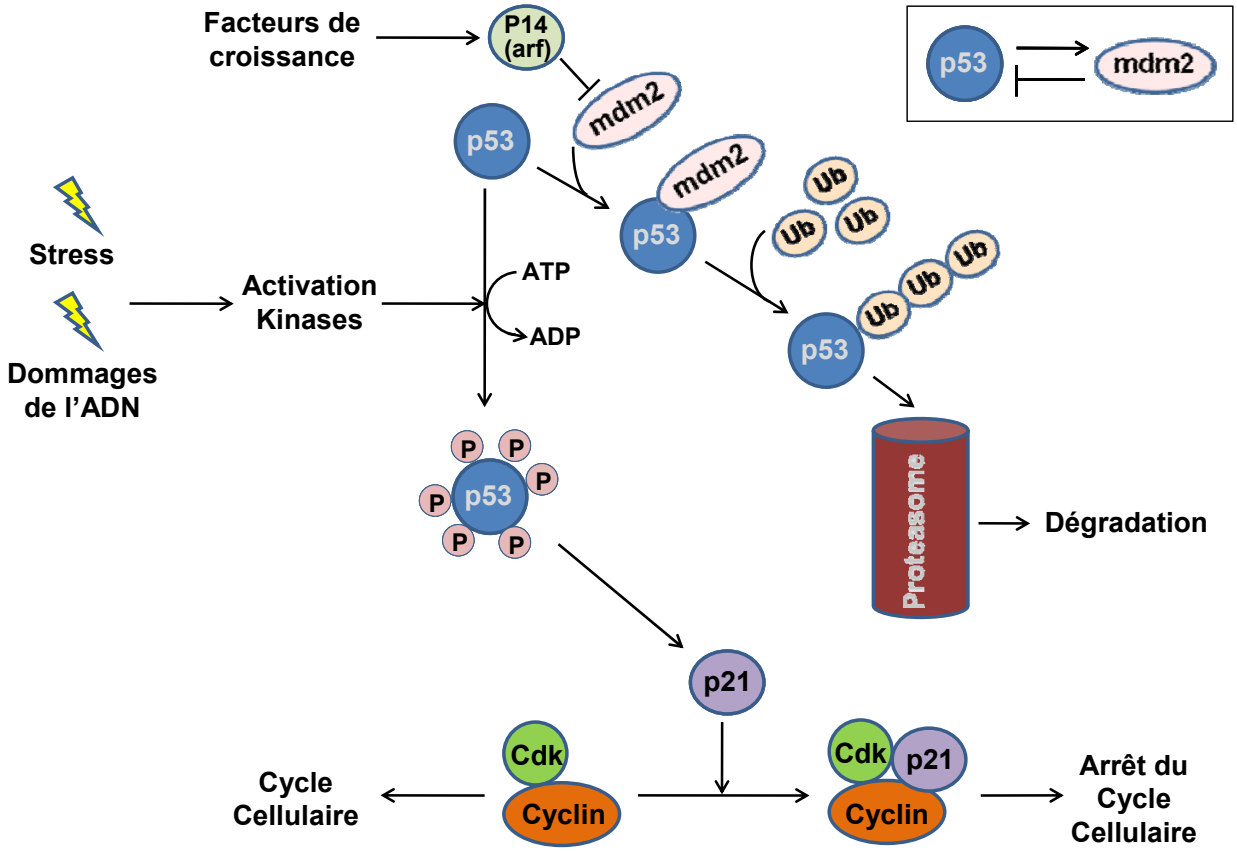


Figure 52 : Schéma simplifié de la régulation de p53.

En réponse à un stimulus, p53 est phosphorylée en de multiples sites et échappe à la dégradation par sa dissociation de mdm2. Cette dissociation résulte de deux principaux phénomènes. Le premier est la phosphorylation de p53 et/ou mdm2 sur des résidus clés par l'activation de kinases en réponse aux stress ou dommages de l'ADN [133]. Le second est la séquestration de mdm2 par la protéine p14(arf) en réponse à la stimulation par des facteurs de croissance [136]. En résumé, p53 peut être stabilisée par au moins deux phénomènes : la réponse au stress lié à des dommages de l'ADN, et le contrôle normal de la croissance cellulaire en réponse à la stimulation par des facteurs de croissance.

De façon intéressante, le promoteur du gène mdm2 possède un site de régulation par p53. La transcription de mdm2 est donc p53-dépendante. Ce phénomène est à l'origine de la mise en place

d'une boucle d'auto régulation (régulation rétro active ou « feedback ») où l'expression de p53 induit la surexpression de son inhibiteur mdm2 [137].

La stabilisation de p53 va entraîner des changements au niveau de sa conformation. La protéine va s'organiser sous forme tétramérique pour devenir un facteur de transcription actif. P53 peut alors se fixer de façon spécifique sur les séquences d'ADN de ses gènes cibles afin de réguler (positivement ou négativement) leur expression. Ces gènes cibles incluent des régulateurs du cycle cellulaire, de l'apoptose, de la réparation de l'ADN et de la différenciation [138]. L'effet de l'activation de p53 sur la cellule dépend de l'action cumulée de l'ensemble de ses cibles. En général, celles-ci forcent la cellule à sortir de son cycle de réplication de l'ADN, soit en induisant l'apoptose, soit en favorisant l'arrêt transitoire ou définitif du cycle cellulaire [139].

Dans des circonstances normales, p53 est transitoirement activée lors du « check-point » situé entre la phase G1 et S du cycle cellulaire, et participe ainsi au contrôle de la vitesse et du moment d'entrée des cellules dans la phase S. Si les cellules sont soumises à un stress, p53 va s'accumuler en grande quantité et empêcher la prolifération des cellules possédant des lésions de l'ADN. P53 induit, par exemple, l'expression de la protéine p21 qui est un inhibiteur de plusieurs cyclines-CDK nécessaires à la poursuite du cycle cellulaire.

Une vue globale non exhaustive de la voie p53 est représentée Figure 53 [140]. On distingue trois niveaux de régulation différents :

- ***Les médiateurs*** (en amont de p53) : en réponse aux différents stress, ils envoient les stimuli nécessaires à l'activation de p53.
- ***L'environnement proche de p53*** (« core regulation ») : il englobe les protéines participant à la régulation directe de p53 (mdm2, p14arf).
- ***Les cibles*** (en aval de p53) : ce sont les effecteurs permettant de réaliser les différentes fonctions de p53 : arrêt du cycle cellulaire, apoptose, réparation de l'ADN, inhibition de l'angiogenèse et de la métastase.

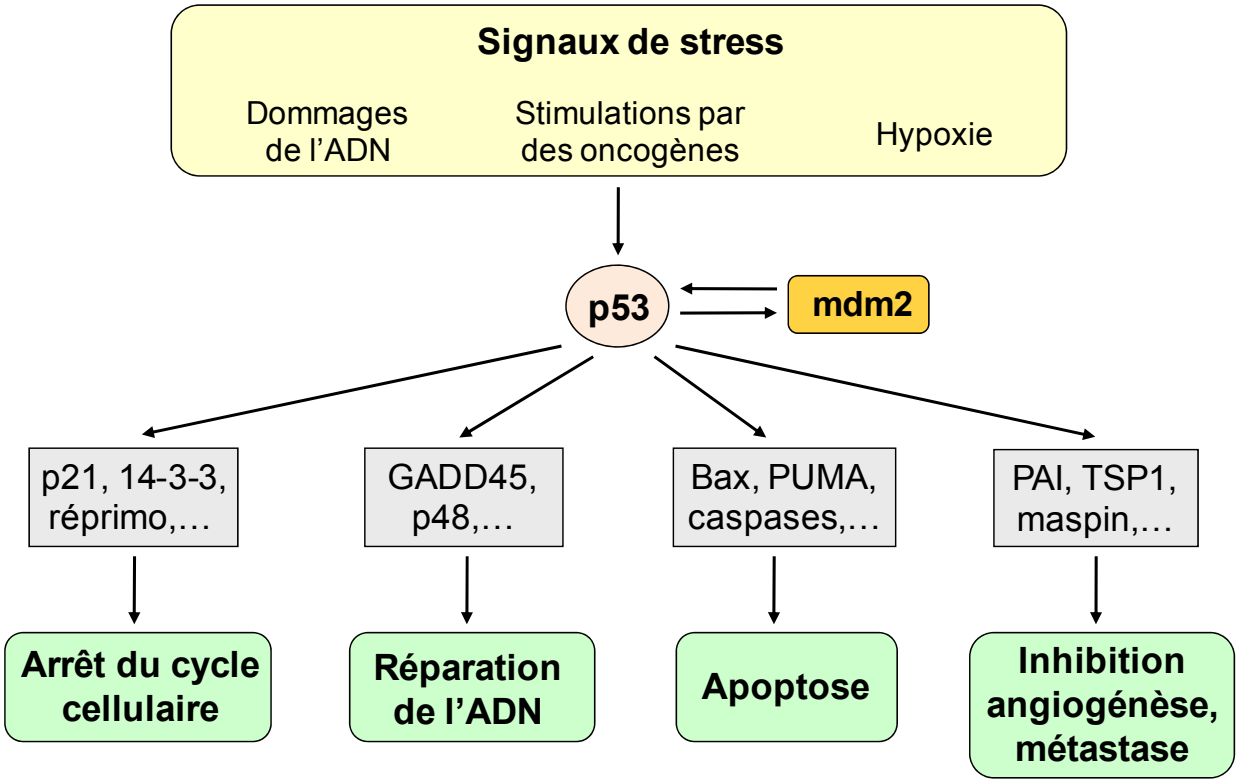


Figure 53 : Schéma synthétique de la voie de signalisation p53 (d'après Wang *et al.*, 2003) [141].

4. Mutations de p53 dans les cancers

4.1. Généralités

Le développement d'un cancer suit un processus Darwinien, sélectionnant les cellules ayant la meilleure capacité à s'adapter à leur environnement perturbé. La survie de ces cellules tumorales repose sur l'accumulation séquentielle d'altérations génétiques pour un ensemble de gènes contrôlant la prolifération et la survie cellulaire, les réponses aux stress, les interactions des cellules avec leur environnement et l'homéostasie des gènes [142].

Parmi la grande diversité des voies génétiques impliquées dans le développement d'un cancer, le gène *TP53* apparaît comme un élément clé et ce quelque soit l'origine cellulaire ou tissulaire du cancer. En effet, la voie p53 est altérée dans la majorité des cancers invasifs. Ces altérations résultent d'une perte d'allèles, d'une mutation ponctuelle (généralement « missense ») ou d'une inactivation de la protéine par séquestration [143, 144]. Des bases de données référencient l'ensemble des mutations connues de *TP53*. On peut citer, par exemple, la base de données de l'IARC International Agency for Research on Cancer (www.iarc.fr/p53/) qui référencie actuellement plus de 25 000 entrées [145].

La perte de fonction de p53 prédispose les cellules à accumuler rapidement une multitude de mutations génétiques. Une démonstration de l'importance de cette perte de fonction a été réalisée

par Donehower sur un modèle de souris [146]. Il conclut que des souris ayant une perte de fonction de la protéine p53 se développent (presque) normalement mais meurent à un âge très précoce suite à de multiples cancers.

La protéine p53 diffère des autres suppresseurs de tumeurs par le fait qu'environ 75 % des mutations rapportées sont « missense » (substitution d'un acide aminé par un autre) [138]. Cela contraste avec d'autres suppresseurs de tumeur tels que RB (protéine du rétinoblastome), APC (adenomatous polyposis coli) ou BRCA1 (breast cancer 1, early onset) qui sont généralement altérés par des délétions, des décalages du cadre de lecture ou des mutations « nonsens » (codon STOP) et ont donc une séquence protéique grandement modifiée. Les mutations de *TP53* sont présentes sur l'ensemble des exons, avec une grande prédominance sur les exons 4 à 9 codant pour le domaine central de fixation à l'ADN. L'exon 2 (le premier exon codant situé en N-terminal) et l'exon 11 (le dernier exon en C-terminal) ne portent que 0,1 % des mutations de p53 (voir base de données IARC). Cela est surprenant car ces deux régions terminales possèdent des domaines régulateurs, ainsi que des modifications post-traductionnelles très importants dans le contrôle de l'activité de cette protéine [143]. La partie N-terminale porte par exemple le site de fixation de la protéine mdm2, le principal régulateur de la stabilité de p53. Les mutations ponctuelles de la région N-terminale ont généralement peu d'effet sur la fonction de la protéine. Par contre, des doubles mutants, notamment au niveau des résidus 22 et 23, entrainent une perte de l'activité de transactivation de la protéine [147, 148]. La partie C-terminale participe à la régulation négative de la fixation de p53 à l'ADN [149].

Le domaine central de fixation à l'ADN (DNA-binding) est la région de p53 la plus fréquemment mutée. Il possède une structure complexe composée de deux feuillets bêta reliés entre eux par un ensemble de boucles et d'hélices [150]. Ces boucles sont maintenues en place grâce à la fixation d'un atome de zinc, essentiel à la stabilisation de l'ensemble de la structure [151]. Dans ce domaine, tous les résidus ne sont pas sujets de la même façon aux mutations. Environ 30 % de l'ensemble des mutations concernent 6 résidus appelés « hotspots » (codons 175, 245, 248, 249, 273, 282) [150, 152]. Ces résidus sont tous localisés dans le domaine DNA-binding et jouent un rôle important soit dans l'interaction p53-ADN (codons 245, 248, 273), soit dans la conformation de p53 (codons 175, 249, 282) (Tableau 6).

Codon	% mutations totales	% mutations codon
175Arg	**4,6**	**100**
175His	4,1	89
175Cys, Gln, Gly, Leu ou Pro	0,5	11
245Gly	**3,3**	**100**
245Ser	1,7	52
245Asp	0,7	21
245Cys	0,4	12
245Ala, Arg, Asn, Gly ou Leu	0,5	15
248Arg	**7,2**	**100**
248Gln	3,6	50
248Trp	3,0	42
248Leu	0,4	5
248Gly ou Pro	0,2	3
249Arg	**3,2**	**100**
249Ser	2,2	69
249Met	0,4	12,5
249Trp	0,2	6
249Gly, Ile, Lys ou thr	0,4	12,5
273 Arg	**6,8**	**100**
273His	3,2	47
273Cys	2,6	38
273Leu	0,8	12
273Gly, Ser ou Tyr	0,2	3
282Arg	**2,7**	**100**
282Trp	2,3	86
282Gln, Gly, His, Leu ou Pro	0,4	14

Tableau 6 : Spectre de mutation des 6 codons « hotspots » dans les cancers (d'après Vousden et Lu [153])

A l'exception du codon 249, les 5 autres « hotspots » ont des séquences CpG méthylées. La mutation de ces résidus induit une perte de la capacité de p53 à supprimer la prolifération. Cependant, des mutations sont aussi présentes sur d'autres sites, en particulier au niveau des résidus impliqués dans la fixation du zinc (codons 176, 179, 238, 242, qui représentent à eux quatre environ 5 % des mutations).

Il est finalement important de noter que certaines mutations ont pour conséquence un gain de fonction pour la protéine p53 (ou GOF pour « Gain Of Function »). Ces mutations rendent la protéine p53 plus stable et/ou lui font acquérir des activités pro-oncogéniques [154-156].

4.2. Le cas du cancer du poumon

Le cancer du poumon possède la plus grande fréquence de mutation du gène *TP53* [157]. Cette fréquence est élevée à la fois pour les cancers du poumon non à petites cellules (NSCLC) et les cancers du poumon à petites cellules (SCLC), respectivement 50 et 70 % [158]. Concernant les

NSCLC, les mutations de *TP53* sont plus souvent présentes dans les carcinomes à cellules squameuses plutôt que dans les adénocarcinomes [159].

Le tabagisme est le principal facteur de risque de développement d'un cancer du poumon, la fumée de cigarette ayant un pouvoir de mutagénèse direct sur l'ADN. Le gène *TP53* est l'une de ces cibles de mutations [160]. Certaines caractéristiques de mutation sont propres au cancer du poumon. On observe, en effet, peu de mutations au niveau du codon 175 et, au contraire, un nombre important de mutations concentrées dans la région 151-159 avec un « hotspot » au niveau des codons 157 et 158 [157]. De plus, on note une augmentation de la fréquence des transversions G:C → T:A dans les cancers du poumon de fumeurs (33%) par rapport aux non fumeurs (15%) ou aux autres cancers (15%) [161, 162]. Ces transversions G:C → T:A sont situées sur les séquences CpG méthylées des codons « hotspots ».

4.3. Les mutations de TP53 dans l'ADN circulant : une source de bio-marqueur

La détection de mutations de l'ADN dans les plasmas de patients constitue une approche non invasive prometteuse pour le diagnostic précoce des cancers.

L'exposition à l'alfatoxine et l'infection par le virus de l'hépatite B (HBV) sont, par exemple, des facteurs à risque de développement d'un HCC (hepatocellular carcinoma) [163]. L'alfatoxine est responsable de mutations au niveau de la Ser-249 de p53 détectables dans le plasma [164, 165], et peut donc constituer un bio-marqueur de diagnostic précoce des HCC [166].

Le benzo-pyrene, un carcinogène majeur présent dans le tabac, est impliqué dans les mutations des codons 157 et 158 observées dans les cancers du poumon de fumeurs [167]. La détection de ces mutations dans l'ADN circulant de population à risque constitue une source de diagnostic potentielle de cette maladie [168].

4.4. Les thérapies anti-p53 : un parcours semé d'embuches

4.4.1. La restauration par thérapie génique de la fonction de p53

La fonction de *TP53* étant perdue dans de nombreux cancers, il est pertinent d'essayer de la restaurer en remplaçant le gène mutant par une copie sauvage fonctionnelle [152]. La restauration de la fonction *TP53* dans les cellules tumorales pourrait ainsi bloquer le développement de la tumeur et sensibiliser les cellules aux stress cytotoxiques et donc améliorer la réponse thérapeutique. Traiter les cancers par thérapie génique nécessite une haute efficacité d'incorporation du gène *TP53* sauvage dans les cellules tumorales *in vivo*. Il faut aussi une expression suffisante de protéine p53 fonctionnelle pour induire la suppression de la tumeur sans

affecter les cellules saines. Deux approches de thérapie génique ont été appliquées dans ce but [152].

La première est l'utilisation de rétrovirus qui s'intègrent sous une forme stable dans le génome des cellules infectées. Cette approche a été validée sur des lignées tumorales du cancer du poumon et sur des modèles de xénogreffe chez la souris [169-171]. Un seul essai clinique de phase I a été entrepris à ce jour pour démontrer la non-toxicité de cette approche pour les patients [172]. Cependant, l'utilisation de rétrovirus présente trois limitations majeures quant à leur utilisation en thérapie génique : ils peuvent induire des dommages dans le génome, leur utilisation nécessite de forts titres de virus et leur efficacité de transduction est relativement faible.

Une seconde approche de thérapie génique du gène *TP53* repose sur l'utilisation d'adénovirus. Des études ont montré la faible toxicité de ces virus chez l'homme [173]. De plus, puisqu'ils ne s'intègrent pas dans le génome, le risque de mutagénèse est grandement diminué. Plusieurs études cliniques ont ainsi été initiées. En ce qui concerne les cancers du poumon non à petites cellules, des études cliniques de phase I [174-176] et de phase II [177, 178] ont été entreprises. Celles-ci ont donné des premiers résultats encourageants en démontrant une non-toxicité pour les patients et une efficacité d'action de la restauration de p53. En Octobre 2003, l'Ad5CMV-p53 (Gendicine) a obtenu une autorisation de mise sur le marché en Chine pour traiter les HNSCC (head and neck squamous cell carcinoma) et représente le premier médicament commercial de thérapie génique pour les cancers [179].

D'autres essais plus poussés sont en cours. Néanmoins, un des essais cliniques sur lequel reposaient beaucoup d'espoirs a été arrêté prématurément. Cet essai portait sur des patientes atteintes d'un cancer de l'ovaire de stade III, avec une maladie résiduelle limitée et une mutation du gène *TP53*. L'administration intrapéritonéale du virus constituait un élément favorable supplémentaire. L'absence d'amélioration des résultats dans le groupe ayant reçu le virus, associée à une morbidité non négligeable, a conduit les investigateurs à interrompre cet essai [180].

4.4.2. Les autres approches

Des résultats expérimentaux portant sur une classe de petites molécules inhibant la liaison entre mdm2 et p53 ont été entrepris [181]. Il s'agit de composés imidazoline, dénommés nutlins, dont la propriété est de bloquer le site de fixation de mdm2 pour p53. La structure du complexe inhibiteur-mdm2 a été finement caractérisée par diffraction X et l'activité des molécules a été

démontrée i*n vitro* et sur des modèles de tumeurs greffées [182]. Ces molécules sont utilisables dans le cas de cellules tumorales ayant conservé les deux gènes p53 sauvages.
PRIMA-1 (p53 reactivation and induction of massive apoptosis-1) et RITA (reactivation of p53 and induction of tumor cell apoptosis) sont deux petites molécules capables de réactiver les fonctions de p53 dans les cellules tumorales [183, 184].

5. p53 et les protéines sécrétées

La protéine p53 influence aussi la sécrétion de protéines par une action transcriptionnelle ou non. Parmi les gènes régulés, un certain nombre code pour des protéines sécrétées. Celles-ci jouent sur la régulation de facteurs de croissance (Insulin-like growth factor binding protein-3 ou IGFBP3), l'inhibition de protéases pouvant altérer la matrice extracellulaire (Serpin B5, Plasminogen activator inhibitor), la modification de la matrice extracellulaire pour bloquer l'angiogenèse ou la migration cellulaire (Thrombospondin) [185]. L'effet transcriptionnel de p53 sur la sécrétion peut aussi être indirect, p53 régulant une de ses cibles transcriptionnelles qui influencera dans un second temps la sécrétion d'autres protéines. Un exemple marquant de ce type de régulation a été décrit récemment par Yu *et al.* [186]. Ce dernier démontre une action de p53 sur la sécrétion de protéines par la voie des exosomes. p53 régule directement le gène TSAP6 dont l'expression est suffisante pour induire la sécrétion de protéines par la voie exosomale. Cette action de p53 a été récemment confirmée sur un modèle *in vivo* de souris [187]. On peut aussi imaginer une action de p53 sur la sécrétion de protéines par des mécanismes indépendants de la transcription en jouant sur leur stabilité, leur sécrétion ou à travers une action par des microRNA (répresseurs post-transcriptionnels) [188, 189]. En conséquence, une partie de la réponse de p53 aux stress est de produire des protéines sécrétées qui vont communiquer avec les cellules adjacentes dans le but de produire une réponse coordonnée. La protéine p53 influence aussi le micro-environnement tumoral à travers l'inhibition de processus tumoraux (angiogenèse, invasion) soit en induisant la sécrétion de protéines anti-tumorogéniques, soit en inhibant des protéines sécrétées pro-tumorogéniques [190]. Très peu de travaux étudiant l'influence de p53 sur la sécrétion des protéines ont été réalisés. Une seule étude protéomique sur des cellules de gliome, a été réalisée à ce jour, aboutissant à l'identification de 60 protéines sécrétées [191].

CHAPITRE 5 : La tumorogenèse et la modulation de l'environnement tumoral

1. Introduction

La carcinogénèse nécessite de nombreuses altérations génétiques afin de conférer aux cellules tumorales des avantages sélectifs de croissance [192]. La plupart des cellules tumorales accumule ainsi plusieurs modifications au niveau nucléotidique, génique, voire chromosomique [193]. Certains cancers acquièrent aussi des changements épigénétiques pouvant entrainer d'important effets sur l'expression des gènes [194].

Hanahan et Weinberg ont suggéré que la transformation d'une cellule normale en cellule cancéreuse nécessitait l'acquisition d'au moins six propriétés par la cellule [142] :

- Indépendance vis à vis des signaux stimulant la prolifération,
- Insensibilité aux signaux inhibiteurs,
- Abolition de l'apoptose,
- Capacité proliférative illimitée,
- Capacité de susciter l'angiogenèse,
- Acquisition d'un pouvoir invasif.

L'évolution tumorale nécessite également un changement dans l'interaction des cellules cancéreuses entre elles et avec leur environnement. Le développement d'une lésion tumorale implique donc le franchissement de barrières anatomiques : invasion du tissu normal adjacent, migration à travers le réseau vasculaire pour coloniser des organes à distance, mais aussi la néovascularisation pour se développer au delà d'un certain volume. L'invasion du tissu normal adjacent et les métastases dans d'autres organes à distance de la lésion primitive sont les facultés essentielles des lésions tumorales, susceptibles de mettre en péril le pronostic vital.

2. Invasion et métastase

2.1. Remaniement de la matrice extracellulaire

La matrice extracellulaire (MEC) correspond à l'ensemble de macromolécules extracellulaires du tissu conjonctif. Elle est constituée en grande partie de glycoprotéines (collagène, fibronectine, laminine,…), de protéines pures (élastine) et de glycosaminoglycanes / protéoglycanes. Ces macromolécules sont synthétisées et sécrétées par les cellules et la MEC est probablement propre

à chaque type tissulaire. La MEC intervient dans le soutien structural, l'adhérence, le mouvement et la régulation de la cellule. La MEC est un environnement dynamique qui joue un rôle clé dans la régulation des fonctions cellulaires durant les procédés de remodelages normaux ou pathologiques tels que le développement embryonnaire, la réparation des tissus, l'inflammation et la métastase et l'invasion tumorales.

Le remaniement de la MEC durant la tumorogenèse nécessite l'intervention de plusieurs acteurs. Le premier mécanisme implique une dégradation de la MEC par l'activation de voies protéolytiques spécifiques. Le second repose sur une migration cellulaire sans apoptose.

2.1.1. Les cascades protéolytiques

Les cellules tumorales sont capables de sécréter des protéases digérant la MEC et favorisant l'invasion du tissu [195]. Parmi les 500 à 600 protéases contenues dans le « degradome » humain [196], seule une petite partie joue un rôle important dans la protéolyse de la MEC durant la progression du cancer [197].

Il existe plusieurs cascades protéolytiques associées à la dégradation de la MEC dans les cancers.

- Le système urokinase plasminogen activator (uPA) et son récepteur (uPAR).
- Les métalloprotéases (MMP) et leurs inhibiteurs tissulaires (TIMP).
- Les cathepsines.

2.1.1.1. Le système uPA-uPAR

L'uPA est une sérine protéase sécrétée sous forme de précurseur inactif (Figure 54) [195]. Ce dernier se fixe sur son récepteur spécifique (uPAR) le rendant ainsi plus susceptible à l'activation par protéolyse. C'est l'activité résiduelle de la plasmine qui est responsable de la conversion de pro-uPA en uPA. Par la suite, l'uPA stimulant la maturation du plasminogène en plasmine, il s'installe une boucle de rétro-action positive. L'uPA active aussi directement le pro-HGF (hepatocyte growth factor) et clive la fibronectine et son propre inhibiteur PAI1 (plasminogen activator inhibitor 1). Elle participe via la plasmine à l'activation des autres cascades protéolytiques dont celle des métalloprotéinases (MMPs). L'uPAR interagit également avec les intégrines soit en cis (sur la même cellule), soit plus probablement en trans (intégrines exprimées par une cellule adjacente) par la stimulation des voies de signalisation classiques impliquant RAS ou PI3K.

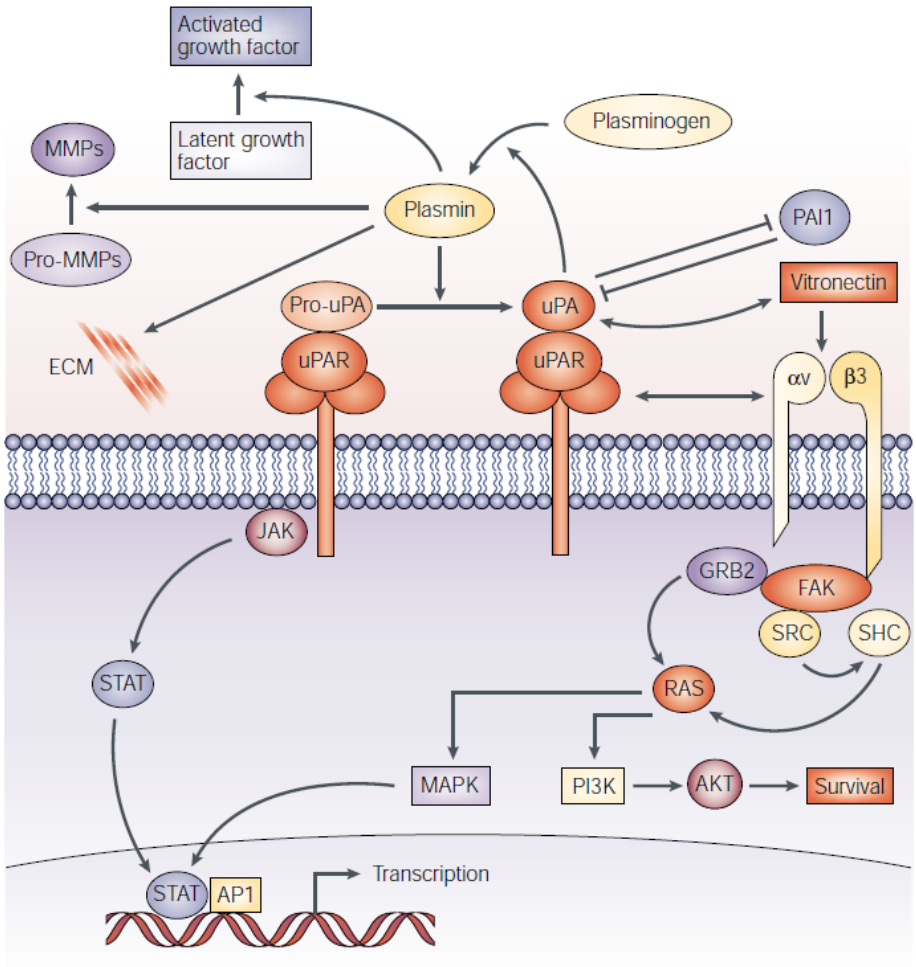


Figure 54 : La voie uPA-uPAR (d'après Rao [195]).

2.1.1.2. Les métalloprotéases (MMPs)

Les métalloprotéases (MMPs) font partie d'une vaste famille d'endopeptidases dirigées contre les différents constituants de la MEC (collagène, gélatine,...). La majorité des MMPs est sécrétée et se retrouve sous forme libre dans le tissu interstitiel. Quelques unes sont cependant des enzymes membranaires. L'activité des MMPs est régulée au niveau de l'expression génique, de l'activation de la forme pro-enzyme (pro-MMP) et par leurs inhibiteurs tissulaires spécifiques (TIMPs) [198]. Leur expression est induite par des facteurs de croissance, des cytokines, des promoteurs de tumeurs, des stress physiques, la transformation oncogénique et les interactions intercellulaires ou cellule-MEC.

Plusieurs études réalisées sur des modèles animaux ont montré l'importance des MMPs dans la pathogenèse des cancers [199, 200] (Figure 55). En règle générale, les animaux ayant un gène *MMP* invalidé ou surexprimant un TIMP ont une diminution de la tumorogenèse. Des études ont démontré que des cellules cancéreuses injectées, par voie intraveineuse chez des souris déficientes en MMP-2 ou MMP-9, colonisaient moins les poumons que chez des souris sauvages

[201, 202]. Les MMPs régulent aussi l'angiogenèse tumorale et pourraient être impliquées dans le "switch" angiogénique observé lors de la néo-vascularisation des tumeurs [203]. Elles ont également une action croisée sur les autres voies (uPA ou cathépsines). Différentes approches thérapeutiques sont développées pour inhiber l'activité des MMPs. Les inhibiteurs naturels (TIMP1 et 2) ou synthétiques (batimastat, marimastat), la calphostine (inhibiteur de la PKC), ou encore des vecteurs antisens et des ribozymes se sont montrés efficaces dans des systèmes *in vitro* ou dans des modèles animaux [204]. Des facteurs de transcription, tels que p53, régulent aussi négativement l'expression des MMPs [205].

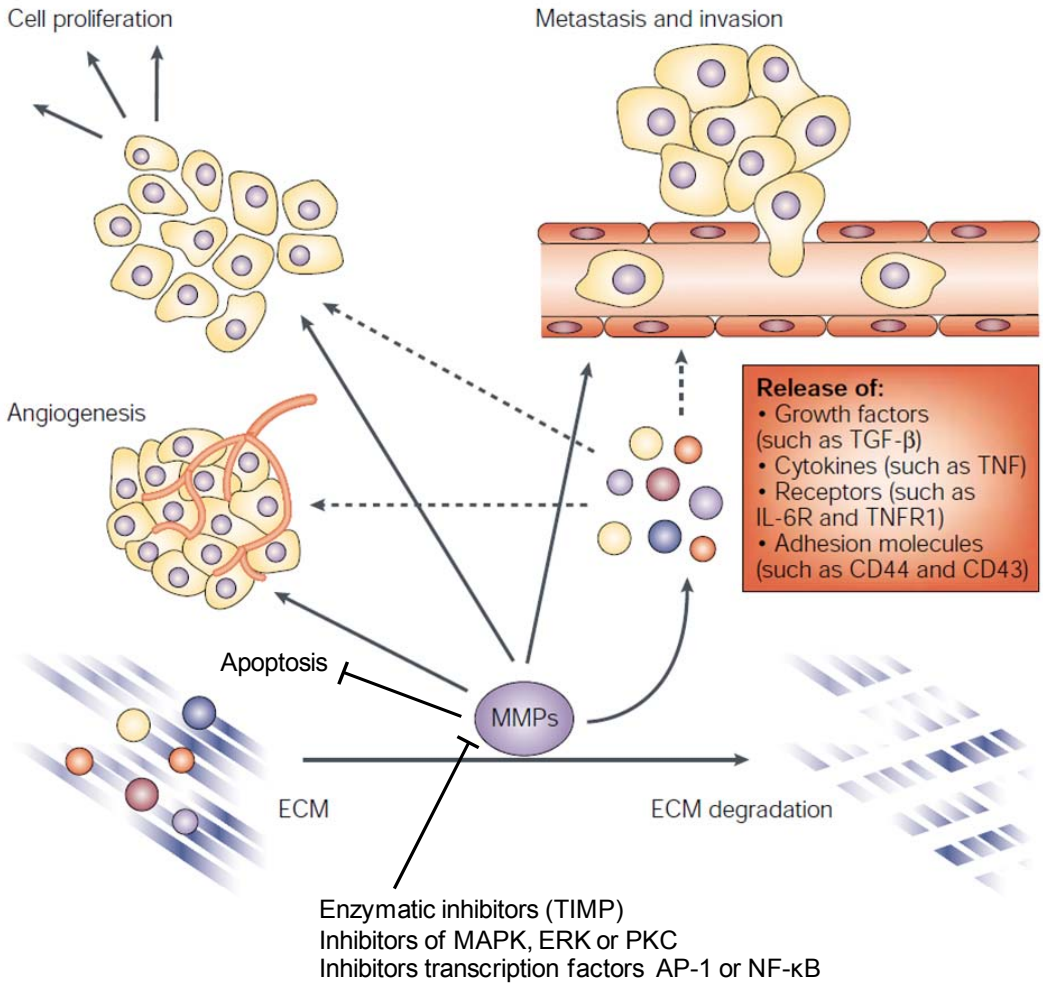


Figure 55 : Les MMPs favorisent le développement des cellules cancéreuses (d'après Rao [195]).

2.1.1.3. Les cathepsines

Les cathepsines sont des protéases à cystéine localisées dans le lysosome mais également sécrétées. La cathepsine B participe à la dissémination des cellules cancéreuses en dégradant les composés de la MEC [206] ou en activant d'autres protéases capables de le faire [207]. Les intégrines ont un rôle important dans l'activation de la cathepsine B [208]. Dans cette cascade

compliquée, la pro-cathepsine B peut être activée par la cathepsine D, l'elastase, la cathepsine G, l'uPA ou le tPA (Figure 56).

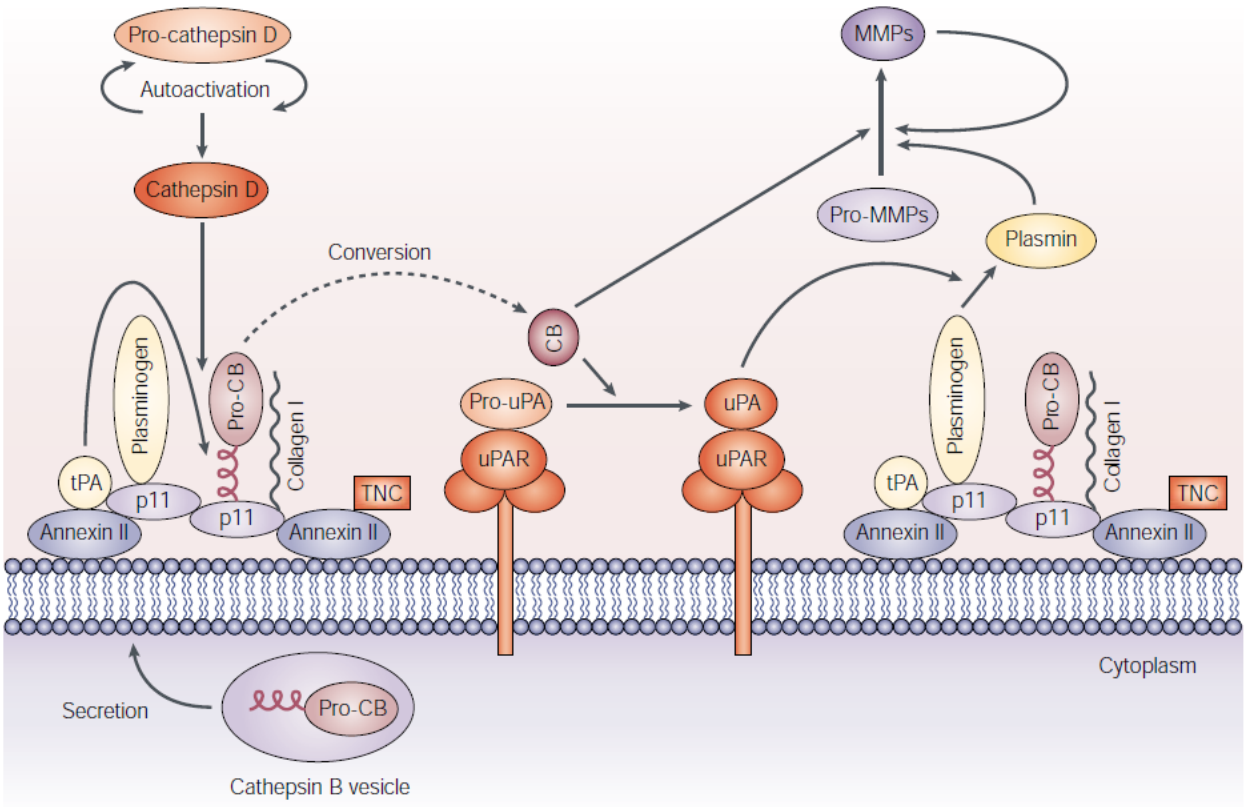


Figure 56 : La cascade protéolytique de la cathepsine B (d'après Rao [195]).

La cathepsine B a un rôle actif dans l'initiation de la cascade protéolytique impliquant uPA, le plasminogène et la plasmine, et il active le TGF-β latent [209]. Elle peut aussi activer directement les MMPs [210] ou détruire leurs inhibiteurs TIMP1 et 2 [211].

Les différentes cascades protéolytiques sont donc connectées entre elles. La plasmine participe à la conversion des pro-MMPs en MMPs. La cathepsine B sécrétée sous forme de précurseur activé par la cathepsine D ou le tPA, agit sur la conversion du pro-uPA en uPA mais aussi sur l'activation des pro-MMPs.

2.1.2. Les mécanismes d'invasion sans protéolyse

Même si les différentes cascades protéolytiques dégradant la MEC sont inactivées, les cellules tumorales conservent une capacité de migration. Celle-ci se fait par mouvements amiboïdes et implique une régulation très fine du cytosquelette (notamment des fibres d'actine) et l'attachement à des éléments extracellulaires. La voie de régulation la plus importante met en jeu la tyrosine-kinase FAK (focal adhesion kinase), un élément central des complexes d'adhérence.

Cette kinase est activée par des intégrines, des facteurs de croissances et des hormones. Les cellules, dont le gène *FAK* est invalidé, ont une capacité de migration très fortement diminuée *in vitro* [212].

2.2. La transition épithélio-mésenchymateuse

L'acquisition de propriétés invasives et migratoires par les cellules épithéliales transformées s'accompagne de changements morphologiques caractéristiques. Ce changement phénotypique est appelé transition épithélio-mésenchymateuse (TEM) et reproduit un mécanisme physiologique qui a normalement lieu durant l'embryogénèse. La TEM se manifeste par un changement d'expression des molécules d'adhérence, incluant la perte de l'adhérence cellule-cellule due à l'E-cadhérine et l'expression de protéines mésothéliales telles que la vimentine ou la N-cadhérine [213]. La perte d'E-cadhérine favorise le détachement des cellules tumorales et la N-cadhérine facilite la mobilité et migration cellulaire (Figure 57). Les membres de la famille TGFβ sont impliqués dans la diminution d'expression de l'E-cadhérine, par l'activation de la voie Raf-Mek-Erk et PI3K (probablement via Ras), afin d'induire les répresseurs ou activateurs nécessaires à la TEM [214]. L'activation de NF-κB a un effet anti-apoptotique mais induit aussi l'expression de la vimentine et de la tenascine pour favoriser la TEM [215, 216].

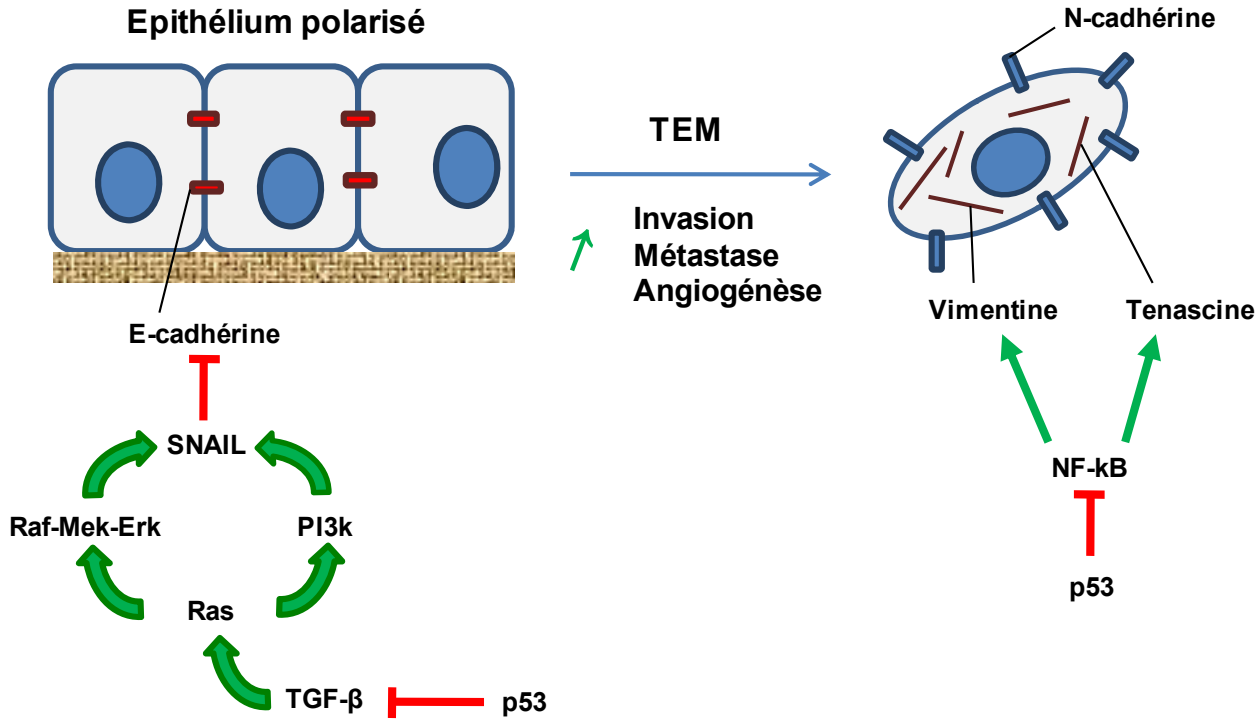


Figure 57 : La transition épithélio-mésenchymateuse (TEM) durant la tumorogenèse.
Exemples de 2 voies impliquant les protéines TGFβ (bas, gauche) et NF-κB (bas, droite)

2.3. Les régulateurs de la métastase

Différents gènes peuvent augmenter le potentiel métastatique de la cellule.
L'ostéopontine, phosphoprotéine sécrétée, a un niveau d'expression systématiquement corrélé avec le caractère métastatique [217]. Elle se lie notamment aux CD44 et intégrines, deux classes de protéines impliquées dans le processus métastatique.
Les récepteurs aux chémokines (par exemple CXCR4 - CXCL12) sont exprimés par les cellules cancéreuses [218]. Ils sont normalement impliqués dans les processus chimiotactiques permettant aux leucocytes de cibler des organes précis. La forte expression de chémokines dans certains tissus (os, poumon), combinée à l'expression de leurs récepteurs à la surface des cellules cancéreuses, constitue un mécanisme d'attraction de cs cellules vers des organes cibles.
Les gènes suppresseurs de métastase sont peu nombreux. Nous pouvons notamment citer BRMS-1, E-cadhérine et KAI-1.

3. La néo-angiogenèse et la progression tumorale

Les tumeurs peuvent rester pendant des années dans une phase dite de « dormance ». L'acquisition de la capacité à générer des néo-vaisseaux est une étape clé du développement tumoral appelée "switch" angiogénique. Le micro-environnement tumoral régule ce "switch" angiogénique de façon plus ou moins précoce selon les types de cancers [219, 220].
Le remaniement de la MEC est l'étape incontournable du processus angiogénique. Il implique la sécrétion d'effecteurs par les cellules tumorales et stromales telles que des protéases de dégradation de la MEC (MMPs) ou des facteurs de croissance de prolifération des cellules endothéliales (VEGF, bFGF, PDGF) (Figure 58). Certains facteurs de croissance sont liés à la MEC et donc libérés lors de sa dégradation pour accentuer l'angiogenèse. D'autres composés libérés lors de la dégradation de la MEC peuvent exercer une action stimulatrice ou inhibitrice. Ainsi des fragments de collagène XV (endostatine) ou de collagène IV (arrestine, tumstatine, canstatine) font partie des plus puissants inhibiteurs de prolifération et de migration des cellules endothéliales.

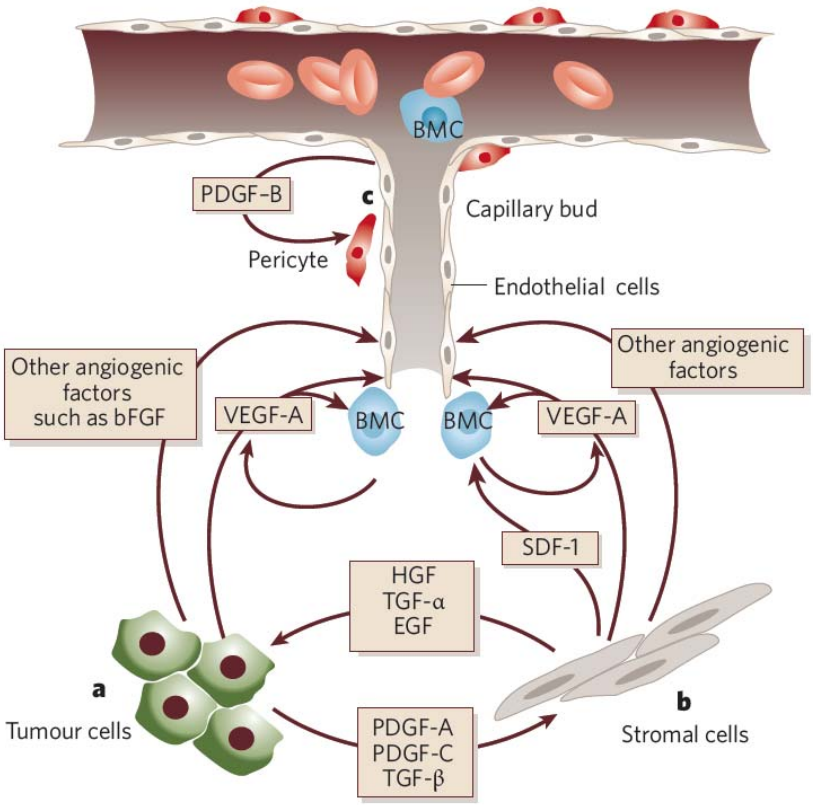


Figure 58 : Les acteurs de l'environnement tumoral impliqués dans la néo-angiogenèse (d'après Ferrara et al. [221]).

Des médiateurs directs et indirects de l'angiogenèse peuvent être distingués.

Les médiateurs directs incluent des facteurs de croissance plus ou moins spécifiques des cellules endothéliales : VEGF (Vascular endothelial growth factor A à E), bFGF (basic fibroblast growth factor), PDGF (platelet-derived growth factor), PlGF (placental growth factor), TGFα (transforming growth factor alpha) mais aussi des chimiokines, TNFα (tumour necrosis factor alpha), IL-8 et les angiopoïétines. Des inhibiteurs, tels que les thrombospondines (TSP1 et 2) ou les statines, s'opposent à l'action de ces effecteurs directs de l'angiogenèse.

Des médiateurs ayant un rôle indirect sur l'angiogenèse ont également été décrits. Quelques exemples, ainsi que leurs modes d'action sont présentés Tableau 7.

Protéines	Modes d'actions
KRAS, HRAS	Stimulation de VEGF, Inhibition de TSP1
ERBB2	Stimulation de VEGF
EGFR	Stimulation de VEGF, bFGF, IL8
FOS	Augmentation de l'expression de VEGF
HPV-16	Stimulation de la sécrétion de VEGF et IFNa
BCL2	Stimulation de VEGF
TP53	Activation de la transcription de TSP1

Tableau 7 : Les effecteurs indirects de l'angiogenèse.

4. La sénescence

Le terme « sénescence » caractérise l'arrêt irréversible des cellules dans leur cycle cellulaire [222]. Les cellules sénescentes s'arrêtent en phase G1/G0 et ne sont plus capables de rentrer en phase S en réponse à des mitogènes [223]. Les cellules sénescentes restent viables et métaboliquement actives pendant une longue période [224] et deviennent résistantes à l'apoptose [225]. Différentes voies sont capables d'induire la sénescence cellulaire.

4.1. La voie p53-p21

Le rôle joué par p53 dans la sénescence a été établi par plusieurs études [226, 227]. La protéine p21, une cible transcriptionnelle de p53, est responsable de l'arrêt du cycle cellulaire [228]. La surexpression de p21 est capable d'induire la sénescence chez de nombreuses cellules [229, 230], alors que la délétion de p21 l'empêche [231]. La protéine p53 peut aussi agir par une voie indépendante de p21 [222].

4.2. La communication SMS des cellules (Senescence-Messaging Secretome)

La sénescence cellulaire s'accompagne de changements significatifs au niveau des facteurs associés à la MEC [232]. Par exemple, l'expression des protéines sécrétées IGF1 (Insuline Growth factor 1) [233], IGFBP3 (IGF binding protein 3) [234], PAI1 (Plasminogen Activator Inhibitor 1) [235], plasmine, Wnt, TGFβ, ou des interleukines est modifiée durant la sénescence. Ces facteurs permettent aux cellules de communiquer entre elles et sont regroupés sous le nom « SMS » (Senescence-Messaging Secretome) par Kuiman et Peeper [236] (Figure 59).

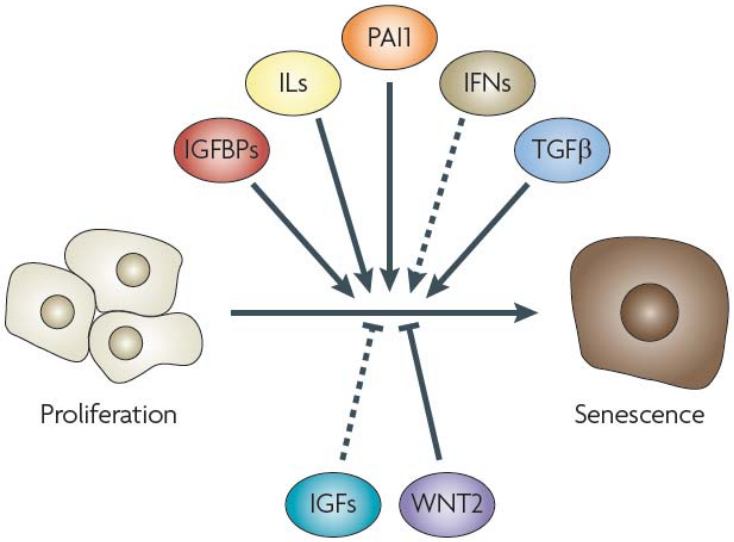


Figure 59 : Les facteurs du SMS (Senescence-Messaging Secretome) qui contribuent à l'induction et la maintenance de la sénescence (d'après Kuilman et Peeper [236]).

Les IGFBPs jouent un rôle dans l'induction de la sénescence par la séquestration des IGFs et donc la suppression de leur signalisation [237]. IGFBP3, une cible transcriptionnelle de p53, est associée à la sénescence [234]. IGFBP7 peut déclencher la sénescence en affectant la voie des Erk (MAPK) [238].
PAI1, un régulateur de la MEC, est généralement surexprimé par les cellules sénescentes [235]. PAI1 est modulé par p53 et inhibe les deux activateurs du plasminogène tPA et uPA. Il empêche ainsi la dégradation de la MEC et la libération de facteurs de croissance nécessaires à la prolifération.

L'ensemble des processus impliqués dans la tumorogenèse (métastase, invasion, angiogenèse, TEM, sénescence,…) nécessite l'intervention d'un certain nombre de facteurs sécrétés. Mieux comprendre ces phénomènes est primordial et des approches de type secretome peuvent être très utiles.

CHAPITRE 6 : Le secretome

Ce chapitre s'appuie essentiellement sur la revue publiée en 2008 par notre équipe dans la revue de médecine interne [239].

1. Définitions

La sécrétomique est la science permettant l'étude globale des protéines sécrétées (secretome) par une cellule, un tissu ou un organisme à un instant donné et sous certaines conditions. De façon plus cadrée, le secretome est défini comme le sous-protéome correspondant à l'ensemble des protéines retrouvées dans le milieu conditionné de cellules en culture [240].
Ce terme a été employé pour la première fois en 2000 par Tjalsma et son équipe lors d'une étude des protéines sécrétées par la bactérie *Bacillus Subtilis* [241], puis repris par la suite dans des études sur les cellules eucaryotes [240, 242].
Ces protéines sont des molécules actives, impliquées dans de nombreux processus physiologiques, et indispensables à la communication des cellules avec leur environnement. Certaines sont des hormones également utilisées comme agents thérapeutiques, comme par exemple l'insuline. Le secretome reflète aussi l'état fonctionnel d'une cellule dans des circonstances données. En effet, une cellule peut sécréter des protéines différentes (ou en proportions différentes) en fonction des stimuli et des messages qu'elle reçoit. Par exemple, les lymphocytes B sécrètent une grande variété d'immunoglobulines (IgM, IgG, IgA) dépendant de l'antigène ou des stimuli extérieurs [243]. Les protéines sécrétées jouent donc un rôle crucial dans de nombreux processus physiologiques et pathologiques et cela en fait une source potentielle idéale pour la recherche de cibles thérapeutiques ou de candidats bio marqueurs de nombreuses maladies tels que les cancers. Chez les organismes supérieurs, la sécrétion des protéines est un processus relativement complexe et extrêmement contrôlé faisant intervenir différentes voies.

2. Les différentes voies de sécrétion

Les protéines sont sécrétées par la cellule selon 3 grandes voies : (Figure 60) [240].

- la voie de sécrétion classique
- la voie de sécrétion non-classique
- le relargage par des exosomes

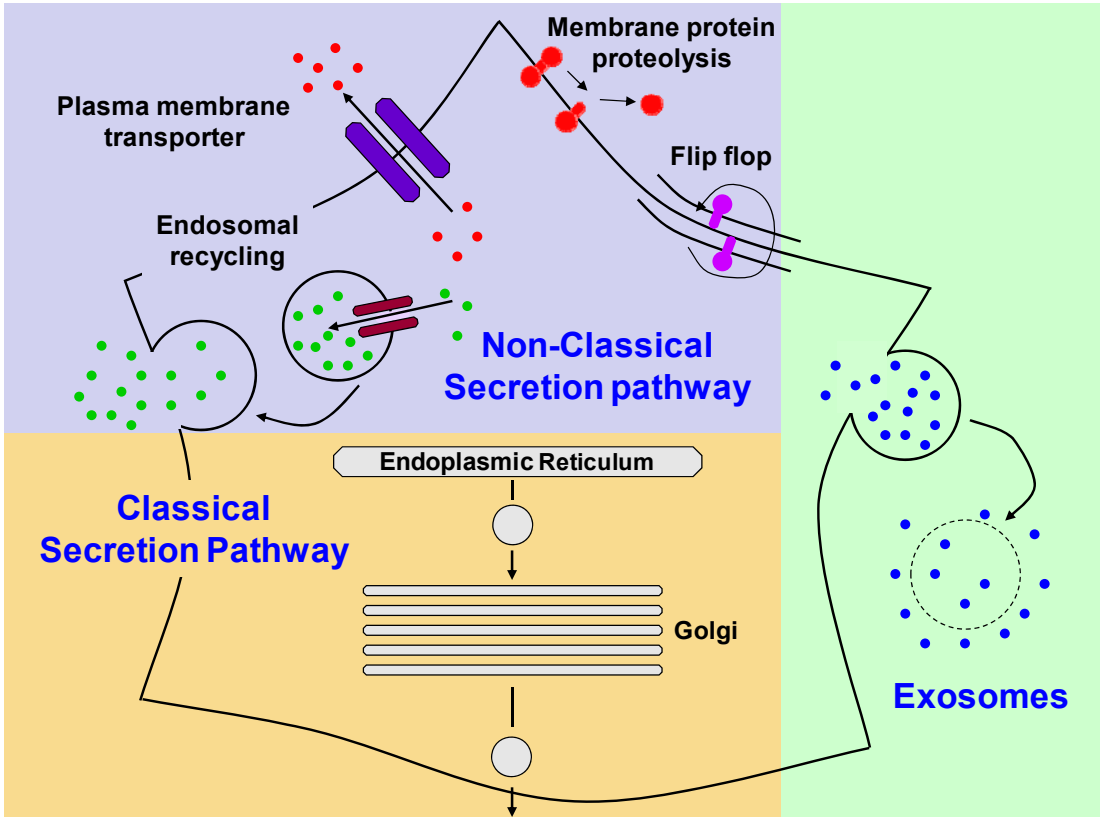


Figure 60 : Les différentes voies de sécrétion de protéines par la cellule.
D'après Chenau *et al.* Revue de Médecine Interne, 2008 [239].

Les protéines sécrétées par la voie classique possèdent un peptide signal à leur extrémité N-terminale de 15 à 30 acides aminés les ciblant comme devant être sécrétées. Ce peptide signal dirige la protéine à travers le réticulum endoplasmique, puis il est clivé. La protéine mature est alors transportée à travers l'appareil de Golgi, puis jusqu'à la membrane plasmique dans des vésicules. Ces dernières fusionnent avec la membrane et permettent ainsi le relargage par exocytose des protéines sous forme soluble dans le milieu extracellulaire.

Les protéines sécrétées selon la voie non-classique ne possèdent pas de peptide signal et sont ainsi sécrétées par différents mécanismes ne faisant intervenir ni le réticulum, ni l'appareil de Golgi. Des protéines peuvent être sécrétées par l'intermédiaire de transporteurs membranaires (ex : FGF-1 et -2), par le recyclage de vésicules endosomales préexistantes suivie d'une exocytose classique (ex : IL-1 bêta, HMGB1), ou par un mécanisme de flip-flop c'est-à-dire une translocation de protéines de la couche interne de la membrane plasmique à la couche externe (ex : HASPB) [244].

La troisième voie de sécrétion correspond au ***relargage de protéines contenues dans des exosomes***. Cette voie des exosomes correspond à l'exocytose de vésicules membranaires qui contiennent des protéines exprimées de façon ubiquitaire telles que des enzymes du métabolisme, des protéines du cytosquelette ou des protéines chaperones [245].

3. Les méthodes d'analyse du secretome

3.1. Analyse indirecte in silico

La bioinformatique, grâce à la comparaison de séquences et la recherche de la présence du peptide signal, permet de prédire les protéines potentiellement sécrétées par les cellules. Les logiciels SignalP [246] et SecretomeP [247, 248] permettent de rechercher respectivement les protéines sécrétées selon les voies classiques (*via* la recherche de la présence du peptide signal) et non-classiques. Il n'existe actuellement aucun moyen de prédire les protéines sécrétées par la voie des exosomes.

D'autres approches reposant sur la comparaison de séquences à des séquences références de protéines connues comme sécrétées sont aussi utilisées grâce à des logiciels d'alignement comme BLAST (Basic Local Alignment Search Tool) [249]. Klee *et al*. ont ainsi reconstitué le secretome prédictif humain, porcin et du poisson *Fugu* en associant l'analyse BLAST et l'identification des peptides signaux [250].

Ces analyses prédictives permettent de générer rapidement une liste de protéines potentiellement sécrétées par un tissu ou par un organisme. Néanmoins, ces données ne restent que prédictives et peuvent contenir un taux important de faux-positifs ou faux-négatifs.

3.2. Analyse directe par protéomique

Le secretome, de par sa nature, représente une source d'échantillon idéale pour être analysé en protéomique. C'est un sous-protéome, donc une fraction des protéines de la cellule bien moins complexe qu'un lysat cellulaire total. De plus, les protéines sécrétées sont, par définition, solubles, évitant ainsi les risques de précipitation souvent rencontrés avec d'autres types d'échantillons (protéines membranaires, nucléaires,…). Ces études protéomiques nécessitent néanmoins des adaptations. Les cellules doivent être cultivées en milieu conditionné en absence de sérum de veau fœtal afin de ne pas contaminer les protéines sécrétées par des protéines sériques. La préparation de l'échantillon protéique est aussi une étape cruciale. Les protéines doivent, par exemple, être généralement concentrées avant de pouvoir être étudiées.

Tout le panel des techniques protéomiques peut être utilisé pour l'analyse du secretome, que ce soient les approches de types PMF ou shotguns. Un grand intérêt est également porté sur l'analyse différentielle des secretomes, de nombreux processus pathologiques étant connus pour modifier la sécrétion de protéines. Ainsi, des études différentielles utilisant des approches DIGE [240], SILAC [251], ICAT [252] ou iTRAQ [253] ont récemment été utilisées avec succès.

4. Intérêts du secretome

L'étude du secretome permet de mieux comprendre les processus moléculaires impliqués dans certains états pathologiques. Au niveau des tumeurs, une cellule tumorale présente des modifications de son activité et de son métabolisme, responsables d'une variation de la sécrétion de certaines protéines. Ces variations induisent des changements d'interactions des cellules tumorales avec leur microenvironnement par l'activation de certains processus de tumorogenèse tels que l'invasion / métastase, la néo-angiogenèse ou la survie cellulaire [191]. L'analyse du secretome peut ainsi permettre de mieux appréhender de tels processus et ainsi donner des pistes en vue de la découverte de nouvelles cibles thérapeutiques.

La recherche de nouveaux bio-marqueurs sériques de maladies reste un challenge primordial. Cependant la découverte de telles molécules directement dans le sérum est très difficile de part la nature très complexe de ce fluide (estimée à environ 10000 protéines) et la très large gamme dynamique de concentrations des protéines (du mg/mL pour l'albumine au pg/mL pour les interleukines) [254]. En conséquence, l'analyse du secretome représente une alternative séduisante, les protéines sécrétées par une cellule ou un tissu étant généralement retrouvées dans le sang. Les tumeurs de la prostate sécrètent par exemple en abondance la protéine PSA (Prostate Specific Antigen ou Kallikrein-3). Les différences d'expression de cette protéine sont détectées dans le sérum et utilisées comme critère pour le diagnostic de ce cancer [255]. L'analyse différentielle de l'ensemble des protéines sécrétées par des tissus sains *versus* pathologiques, grâce aux techniques précédemment citées, représente donc une approche puissante pour la découverte de nouveaux candidats bio marqueurs pour le diagnostic précoce de pathologies [256].

Cependant, à ce jour, les différentes études ont été menées *in vitro* sur des systèmes de cultures cellulaires assez éloignés des conditions physiologiques réelles. Cela soulève la question de savoir si ces protéines sont réellement sécrétées *in vivo*. Peu d'études *in vivo* du secretome ont été conduites [257] de part la difficulté à extraire et préparer un tel échantillon.

OBJECTIFS DU TRAVAIL
ET CHOIX METHODOLOGIQUES

OBJECTIFS DU TRAVAIL ET CHOIX METHODOLOGIQUES

Les interactions entre les cellules tumorales et leur environnement jouent un rôle majeur dans le développement de la tumeur, son maintien, sa progression, l'invasion et la néo-angiogenèse. Ces évènements nécessitent une modification de la communication intercellulaire initiée notamment par une modulation des protéines libérées par les cellules tumorales [142]. Une meilleure connaissance de ces facteurs est indispensable car ils représentent des cibles thérapeutiques et diagnostiques potentielles. Nous nous sommes plus particulièrement intéressés au cancer du poumon. Celui-ci est un véritable problème de santé publique et la première cause de mortalité due à un cancer dans le monde. La protéine p53 est un « suppresseur de tumeur » majeur très fréquemment muté dans les cancers. Le cancer du poumon présente la plus grande fréquence de mutation (*TP53* muté dans 50 à 70 % des cas) [143, 144]. Le rôle clé de p53 dans la défense anti-tumorale, couplé à sa fonction de facteur de transcription font de sa mutation une modification majeure pour la cellule, entraînant une dérégulation directe et indirecte de l'expression d'un grand nombre de protéines. Certaines de ces protéines sont libérées dans le milieu extracellulaire et leur modulation contribue au changement d'interaction des cellules tumorales avec leur environnement. Paradoxalement, cette influence de p53 a été relativement peu étudiée.

Dans ce travail, nous nous proposons d'étudier l'influence de p53 sur la libération de protéines par des cellules du cancer du poumon. Cette étude est faite sans *à priori* quant à la découverte des cibles protéiques d'intérêt. Un procédé d'analyse protéomique différentielle robuste et adéquat a donc été établi et validé. Dans un second temps, ces protéines sont recherchées dans un modèle *in vivo* pour évaluer l'utilisation du secretome comme source potentielle de marqueurs tumoraux.

Dans la première partie de ce travail, nous avons étudié une technologie innovante de fractionnement, l'isoélectrofocalisation en solution OFFGEL, dans le but de valider son utilisation pour l'analyse des peptides trypsiques d'échantillons complexes (secretome et plasma humain). Sa compatibilité avec la méthode de marquage iTRAQ a également été observée.

Dans la seconde partie, nous avons mis en place un procédé d'analyse protéomique permettant d'étudier l'influence de p53 sur la libération de protéines par la lignée H358 (adénocarcinome de cancer du poumon non à petites cellules). L'expression conditionnelle de p53 sauvage donne des secretomes différents qui sont comparés par analyse différentielle iTRAQ. Les résultats obtenus *in vitro* sont confrontés à un modèle *in vivo* de xénogreffe tumorale chez la souris. Ce modèle permettra de mettre en valeur l'expression des protéines d'intérêt au niveau de la tumeur et leur diffusion éventuelle dans le fluide sanguin.

La dernière partie de ce travail est une étude préliminaire visant à évaluer l'intérêt des modèles « secretome » dans la recherche de bio-marqueurs potentiels. Dans ce but, un faible panel de prélèvements de patients a été analysé.

RESULTATS

Première partie :

Développement et validation de la technologie OFFGEL pour la séparation de mélanges peptidiques complexes et compatibilité avec le marquage iTRAQ.

RESULTATS

CHAPITRE 1 : Développement et validation de la technologie OFFGEL pour la séparation de mélanges peptidiques complexes et compatibilité avec le marquage iTRAQ.

1. Problématique et objectifs du travail

Les échantillons biologiques eucaryotes (extraits cellulaires, tissulaires ou fluides biologiques) possèdent un protéome très complexe, ainsi qu'une gamme dynamique de concentration protéique très large. Leur analyse en protéomique nécessite donc la mise en place préalable de techniques efficaces de décomplexification de l'échantillon.

Pour les analyses protéomiques de type LC-MS/MS, le procédé standard de fractionnement d'un mélange peptidique complexe consiste à combiner deux séparations chromatographiques d'affinité différente. Les deux types de colonnes communément utilisées sont les SCX (Strong Cation eXchange) et les RP (Reverse Phase) qui séparent les peptides respectivement selon leur charge (affinité pour les cations) et leur hydrophobicité.

Bien que ce procédé d'analyse ait fait ses preuves, l'étape de chromatographie SCX possède quelques limitations quant à son utilisation en protéomique. Les colonnes SCX utilisées (colonnes capillaires de faible diamètre interne) ont une capacité de chargement insuffisante. De plus, l'élution des peptides de ces colonnes par gradients de sels n'est pas idéale. La quantité de peptides disponible dans chaque fraction à l'issue de cette première séparation, ainsi que la composition du tampon dans lequel ils se trouvent sont deux paramètres importants à prendre en compte pour obtenir une qualité d'analyse optimale lors des étapes ultérieures (séparation en phase inverse et analyse en masse). Enfin, les profils d'élution de ces colonnes présentent généralement une mauvaise répartition de la séparation des peptides. La majorité des peptides ont des paramètres de rétention sur SCX relativement proches, le fractionnement est donc réparti sur un nombre réduit de fractions éluées (inférieur à 10). Ce pouvoir séparateur est insuffisant pour décomplexifier de façon satisfaisante des mélanges peptidiques complexes.

L'approche OFFGEL représente une alternative séduisante à la chromatographie SCX.

Notre travail s'est axé sur la question suivante :

« Le fractionnement OFFGEL est-il une approche efficace pour la séparation de mélanges peptidiques eucaryotes complexes et est-il compatible avec la technologie de marquage isotopique iTRAQ ? »

Pour répondre à cette question nous avons utilisé deux types d'échantillons protéiques complexes (secretome et plasma) et étudié les profils de séparation obtenus après leur fractionnement par OFFGEL. La compréhension de ces profils expérimentaux complétée par des informations issues d'un modèle théorique nous permettra d'évaluer l'efficacité de séparation de l'approche OFFGEL et de l'intégrer ou non à notre procédé d'analyse de protéomes complexes.

2. Moyens mis en œuvre

Nous avons choisi pour cette étude deux types d'échantillons protéiques complexes : le secretome de la lignée cellulaire H358 et un pool de plasma sain (plasma CQI issu de la plate-forme protéomique de Dijon).
L'analyse de l'échantillon secretome permettra de valider la méthode OFFGEL pour l'étude ultérieure des protéines sécrétées influencées par p53.
Le choix du plasma découle du fait qu'il s'agit du fluide biologique le plus complexe en termes de nombre de protéines et de gamme dynamique de concentration. En effet, le nombre de protéines retrouvées dans le plasma est estimé à plus de 10000 [258]. L'HUPO a initié le projet PPP (Plasma Proteome Project) qui référençait, en 2005, 3020 protéines [259] dont 697 validées avec une très haute confiance [260]. De plus, les études de protéomes plasmatiques constituent une thématique d'intérêt au laboratoire.
Après digestion des protéines par la trypsine, les peptides sont fractionnés par isoélectrofocalisation en solution grâce à l'OFFGEL Fractionator 3100 d'Agilent. Cette séparation se fait le long d'une strip IPG de 24 cm possédant une gamme linéaire de pH allant de 3 à 10. Elle permet de répartir l'ensemble des peptides du mélange en 24 fractions liquides différentes (gamme de 0.33 unités de pH par fraction). A l'issue de la focalisation, chacune des 24 fractions est collectée. L'analyse du contenu peptidique des fractions est faite par LC-MS/MS : séparation par nano-chromatographie en phase-inverse (RP-LC) et caractérisation des séquences peptidiques par MALDI-TOF/TOF. L'évaluation de la qualité de séparation des peptides est réalisée par analyse informatique. Celle-ci repose sur l'étude des séquences des peptides identifiés pour chaque fraction et la compréhension des profils de répartition obtenus. L'utilisation du logiciel

"Compute pI/MW" (http://us.expasy.org/tools/pi_tool.html) permet de calculer le pI théorique attendu pour chaque peptide identifié.

Un modèle algorithmique a aussi été développé pour expliquer les profils expérimentaux obtenus. Ce modèle repose sur la génération de l'ensemble des peptides théoriques possibles répondant aux critères suivants : peptides comprenant entre 0 et 6 acides aminés parmi ceux ayant une chaîne latérale contribuant au calcul du pI (Asp, Glu, Arg, Lys, His, Tyr).

3. Résultats

Ces travaux sont présentés dans l'article suivant :

Peptides OFFGEL electrophoresis: a suitable pre-analytical step for complex eukaryotic samples fractionation compatible with quantitative iTRAQ labelling.

Jérôme Chenau, Sylvie Michelland, Jonathan Sidibé and Michel Seve.

Cet article a été publié dans le journal Proteome Science (*Proteome Science* 2008, **6**:9).

Résumé des principaux résultats :

Les peptides issus du secretome et du plasma possèdent un pI moyen par fraction concordant avec les pI théoriques attendus (Figure 1 de l'article).

L'OFFGEL présente une bonne qualité de séparation des peptides. Celle-ci est évaluée en étudiant la répartition d'un peptide unique dans des fractions OFFGEL successives. En ce qui concerne le secretome, 76,9 % des peptides sont focalisés dans une fraction unique et 93,7 % dans une ou deux fractions successives. Pour le plasma, 62,6 % des peptides sont focalisés dans une fraction unique et 81,5 % dans une ou deux fractions successives (Figure 2 et Table 1 de l'article).

La répartition des peptides n'est pas homogène le long de la gamme de pH de la strip (Figure 3 de l'article). Ce phénomène est à corréler avec les résultats issus de notre modèle algorithmique montrant que certaines valeurs de pI sont plus probables que d'autres (Figure 5 de l'article).

La répartition des peptides le long de la gamme de pH de la strip est en adéquation avec leur composition en acides aminés. Les peptides des fractions de pH acide sont, par exemple, riches en acides aminés à chaîne latérale acide (Glutamate et Aspartate) au contraire des fractions de pH plus basique (Figure 4 de l'article).

Le marquage iTRAQ ne perturbe pas la focalisation lors de la séparation OFFGEL (Table 2 de l'article).

4. Conclusion

Notre étude démontre l'efficacité de l'approche OFFGEL comme première étape de séparation d'un mélange complexe de peptides issus d'échantillons eucaryotes (secretome et plasma), en vue de leur analyse ultérieure en LC-MS/MS.

L'OFFGEL est compatible avec la technique de marquage isotopique iTRAQ et peut donc être utilisé dans des protocoles d'analyse différentielle.

Ces résultats font de l'OFFGEL une technique de décomplexification de choix à intégrer dans notre procédé d'analyse protéomique. Cette approche, de par sa compatibilité avec l'iTRAQ, sera très utile pour l'étude différentielle des protéines sécrétées par des cellules du cancer du poumon en fonction du statut p53.

5. Publication I

Proteome Science

BioMed Central

Research

Open Access

Peptides OFFGEL electrophoresis: a suitable pre-analytical step for complex eukaryotic samples fractionation compatible with quantitative iTRAQ labeling

Jérôme Chenau[1], Sylvie Michelland[1,3], Jonathan Sidibe[2] and Michel Seve*[1,3]

Address: [1]Centre de Recherche INSERM/UJF U823, Cibles Diagnostiques ou Thérapeutiques et Vectorisation des Drogues dans le Cancer du Poumon Institut Albert Bonniot, 38706 La Tronche, France, [2]Université Joseph Fourier, 38041 Grenoble Cedex 09, France and [3]Centre d'Innovation en Biologie, Pavillon B, CHU de la Tronche, 38043 Grenoble Cedex 9, France

Email: Jérôme Chenau - jerome.chenau@gmail.com; Sylvie Michelland - smichelland@chu-grenoble.fr; Jonathan Sidibe - jonathan.sidibe@wanadoo.fr; Michel Seve* - Michel.Seve@ujf-grenoble.fr

* Corresponding author

Published: 26 February 2008

Proteome Science 2008, **6**:9 doi:10.1186/1477-5956-6-9

Received: 1 December 2007
Accepted: 26 February 2008

This article is available from: http://www.proteomesci.com/content/6/1/9

© 2008 Chenau et al; licensee BioMed Central Ltd.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/2.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Abstract

Background: The proteomes of mammalian biological fluids, cells and tissues are complex and composed of proteins with a wide dynamic range. The effective way to overcome the complexity of these proteomes is to combine several fractionation steps. OFFGEL fractionation, recently developed by Agilent Technologies, provides the ability to pre-fractionate peptides into discrete liquid fractions and demonstrated high efficiency and repeatability necessary for the analysis of such complex proteomes.

Results: We evaluated OFFGEL fractionator technology to separate peptides from two complex proteomes, human secretome and human plasma, using a 24-wells device encompassing the pH range 3–10. In combination with reverse phase liquid chromatography, peptides from these two samples were separated and identified by MALDI TOF-TOF. The repartition profiles of the peptides in the different fractions were analyzed and explained by their content in charged amino acids using an algorithmic model based on the possible combinations of amino acids. We also demonstrated for the first time the compatibility of OFFGEL separation technology with the quantitative proteomic labeling technique iTRAQ allowing inclusion of this technique in complex samples comparative proteomic workflow.

Conclusion: The reported data showed that OFFGEL system provides a highly valuable tool to fractionate peptides from complex eukaryotic proteomes (plasma and secretome) and is compatible with iTRAQ labeling quantitative studies. We therefore consider peptides OFFGEL fractionation as an effective addition to our strategy and an important system for quantitative proteomics studies.

Background

The proteomes of mammalian cells, tissues and body fluids are complex and display a wide dynamic range of proteins concentration. In order to overcome the human proteome complexity and determine the proteome content, it is necessary to use sample fractionation steps.

The recent introduction of commercially available OFFGEL fractionator system by Agilent Technologies, provides an efficient and reproducible separation technique [1,2]. This separation is based on immobilized pH gradient (IPG) strips and permits to separate peptides and proteins according to their isoelectric point (pI), but is realized in solution [1,3]. Therefore, its micropreparative scale provides fraction volumes large enough to perform subsequent analyses as reverse phase (RP) – liquid chromatography (LC) – MALDI MS/MS.

The secretome, first introduced by Tjalsma *et al* in 2000, describes the global study of secreted proteins by a cell, tissue or organism at any given time or under certain conditions [4,5]. Secretome is the origin of circulating proteins in the body and a very promising source for discovery of new biomarkers candidates. Human plasma is the most complex body fluid and contains a large number of proteins with a dynamic range of at least 9–10 orders of magnitude [6]. This complexity is a problem for proteomic analysis and it is necessary to develop efficient separation techniques to determine its precise protein composition.

In this study, we applied OFFGEL fractionation in combination with RP LC-MALDI MS/MS analysis to separate peptides of complex samples as secretome and plasma and we analyzed and explained the specific repartition profiles obtained. We also evaluated the efficiency of OFFGEL fractionation with peptides previously labeled by the isobaric amine-specific tags used in recently developed iTRAQ™ (isobaric tags for relative and absolute quantification) technology [7]. The validation of OFFGEL electrophoresis on separation of eukaryote tryptic digests will permit to include it in proteomic workflow of complex samples.

Results and discussion

The OFFGEL fractionator system was used in order to separate secretome and depleted plasma peptides according to their pI using a 24-wells device encompassing the pH range 3–10. The peptides are recovered in liquid phase, which is much more convenient for the others subsequent separation experiments like liquid chromatography. Expected pH ranges by fraction were calculated according to the IPG strip supplier data and the fractions size (Table 1). The attempted resolution was at least 0.3 pH units [1]. The average experimental pI for each fraction was computed from the peptides lists after filtering for false positives and replicates (Figure 1A and Table 1). Experimental pI values were globally similar for both secretome and plasma samples. The average experimental pI deviated from the average theoretical pI values with an average error of 6.36% for plasma and 6.69% for secretome. Error was < 6% in fractions 3–11 (pH 4.0–6.1) and was > 10% in fractions 14–17 (pH 6.8–7.7) (Figure 1B). The observed error increase was correlated with the low number of peptides found in this pH range and previously reported by Horth *et al* [2]. Standard deviations were weak for fractions 1–9 (pH 3.5–5.5) and relatively high for fractions 18–24 (pH 7.81–9.65). Globally, standard deviation increased from the neutral pH region to the basic pH region. This increase is correlated with the focalization level assessed by the ratio of unique peptides/total peptides, smaller for fractions in neutral and basic pH regions. Indeed, the most unique peptides to each fraction were found in fractions 2–5 (pH 3.6–4.7). This is in agreement with previously published studies reported that unique peptides to each fraction are more found in acidic region rather than neutral and basic regions [2,8-10]. The OFFGEL good quality separation can be validated looking at the number of unique peptides, identified only in a single fraction. 62.6% and 76.9% of peptides were found in a single fraction in plasma and secretome samples respectively (Figure 2). We observed that 81.5% and 93.7% of peptides were found in one or two fractions in plasma and secretome samples respectively. This result correlates with the findings of Hörth et al, who found that 74% of tryptic peptides of *E. Coli* focus in one well and 90% focus in two wells [2]. On the other hand, the worst separated peptide is found in 15 distinct fractions for plasma sample and in 8 distinct fractions for secretome sample. Although plasma sample has been depleted before OFFGEL fractionation, the sample remained complex because of the large dynamic range of protein concentration and the difficulty to remove the totality of the higher abundant proteins.

Peptides were unevenly distributed along the IPG strip scale (Figure 3A and 3B). Gaps were observed in fractions 6–7 (pH 4.66–5.19), 14 (pH 6.76–7.03) and 17–19 (pH 7.55–8.34). To more understand the factors determining peptides pI distribution, we digested in silico the 9504 plasma proteins identified from the Human Plasma Project [11] and calculated the pI distribution of the 14800 resulting peptides, after elimination of replicate peptides. Interestingly, graphical representation of the pI distribution of these peptides fits fairly with those obtained from our experimental data (Figure 3C). Gaps were found at fractions 7 and 15–17. For gaps in neutral and basic region, in silico distribution is slightly shifted on the left and could be explained by trypsin miss-cleavages in experimental data. These results correspond to Hörth's study using OFFGEL fractionation of total E.Coli peptides [2] and to theoretical results of Lam et al [12]. So the repartition across the OFFGEL IPG strip of proteome tryptic digest is not dependent on the sample nature or organism (eukaryote or prokaryote). Some molecular considerations have been introduced to explain the unevenly distribution of peptides. We calculated the average

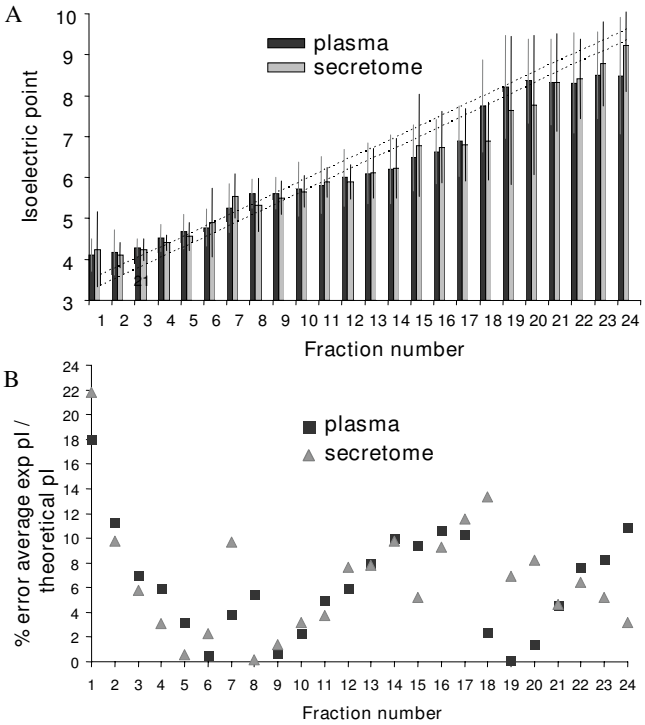


Figure 1
Comparison of experimental and theoretical pI for each OFFGEL fraction. (A) **Average experimental pI values for all peptides identified per fraction**. The dark area relates to the plasma peptides and the white area to the secretome peptides. Standard deviation is represented by errors bars. The broken lines are based on the theoretical pI values calculated according to the supplier's specifications of the IPG gel strip of 24 cm length (pH 3–10). (B) **Percentage error between the average experimental pI and the average theoretical pI per fraction.** Black squares correspond to the error for plasma and triangles correspond to the error for secretome.

number of each amino acid by peptide and by fraction. We then focused on the 7 charged amino acids aspartate (D), glutamate (E), arginine (R), lysine (K), histidine (H), tyrosine (Y), cysteine (C) which mainly contribute to the overall charge of the peptides and used for the theoretical calculation of their pI [13]. As expected, the number of acidic amino acids D and E by peptide globally decreases from acidic to basic fractions (Figure 4A). Acidic peptides (Fraction 1–3) have an average of 2 D or E amino acids while basic peptides have less than 0.2 acidic amino acid. The number of the basic amino acids R and K is globally constant in each fraction, explained by the action of trypsin, cutting proteins after these residues. Expecting no miss-cleavage, each peptide should contain an arginine or a lysine residue at the c-terminus. Arginine is more frequently observed: 80% versus 20% of the peptides containing a lysine residue. In our experimental conditions, cysteine residues do not contribute to the charge of the peptides as they were reduced and alkylated prior the OFFGEL separation. However, cysteine repartition is not homogenous along the strip with less cysteine in the peptides migrating in fractions 6 to 8, which contain a weak number of peptides (Figure 4B). Conversely, fractions 19 and 20 which are also low abundant fractions have peptides with a high content in cysteine residues. Tyrosine is present in all fractions but is more abundant in the basic fractions because of its basic side-chain. Histidine is present mainly in two regions: fractions 6–9 and 15–18. It is also important to note that the mean size of the peptides decreases from 15 amino acids in acidic fractions to 10 amino acids in basic fractions.

We built a calculated model of the OFFGEL peptide separation based on the generation of all possible peptides containing a combination of 6 amino acids out of the 6 charged amino acids contributing to the pI calculation (D, E, R, K, H, Y) with the condition that only one arginine or lysine is present in a given peptide. 96% of the peptides sequenced from plasma or secretome have 0 to 6 charged amino acids. We generated 252 theoretical combinations and calculated their respective pI. We obtained a set of possible pI with a profile globally homogenous up to the 7th fraction and the presence of steps in the pI scale after this fraction (Figure 5A). These steps explain the higher standard deviation obtained for the basic fractions. From this set of pI, peptides were reattributed to the different fractions (Figure 5B). It does not exist any combination resulting in a peptide with a pI below 3.66. That explains why the average pI observed for the first fraction is higher than expected by the strip specifications and why the first fraction standard deviation is so high. The same problem was observed for fractions 16 to 18. In our model, it does not exist any combination of charged amino acids resulting in a peptide with a pI in the range 7.29 to 8.08 corresponding to these fractions. Last, no combination generated a pI over 9.78. However, we observed experimentally a lot of peptides with pI between 10 and 14, accumulating in the fraction 24. The analysis of their sequence revealed miss-cleavages with two adjacent lysine or arginine residues. By comparison of the profile obtained from the model and experimental data (Figure 3), we observed that the fractions with few peptides, *i.e.* fractions 6 and 7 (pH 4.66–5.19) and fractions 16 and 19 (pH 7.29–8.34) correspond with a lack of possible combinations of charged amino acids resulting in peptides in these ranges of pH (Figure 5B). A previous study has shown the compatibility of the recently developed isotopic labeling technique iTRAQ with isoelectrofocalisation separation (in-gel) [14]. We wanted here to study the compatibility of OFFGEL fractionation with iTRAQ. We labeled the secretome sample with iTRAQ reagent and separated the peptides with OFFGEL following the same

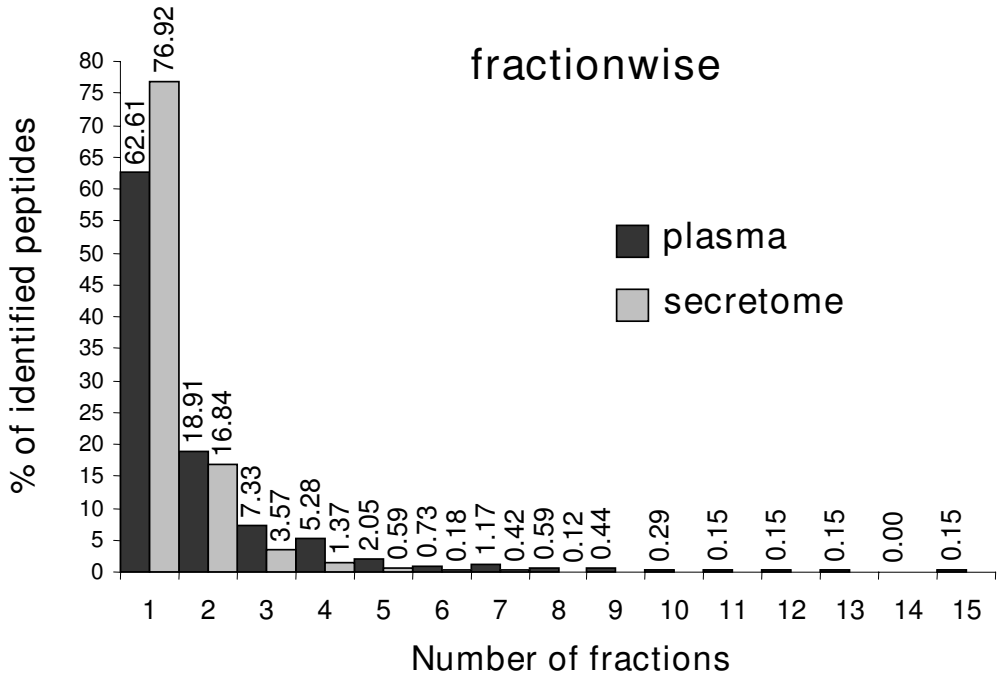


Figure 2
Fractionwise distribution of identified peptides in plasma sample (grey bars) and secretome sample (white bars).

Table 1: The expected pH ranges per fraction were calculated according to the IPG strip supplier data and the OFFGEL well dimensions and compared with experimental pI obtained for plasma and secretome samples.

			Plasma		Secretome	
Fraction	**Expected pH range**		**Peptide pI Average**	**s.d.**	**Peptide pI Average**	**s.d**
1	3.35	3.61	4.11	0.40	4.24	0.92
2	3.61	3.88	4.17	0.56	4.11	0.31
3	3.88	4.14	4.29	0.22	4.24	0.26
4	4.14	4.40	4.53	0.32	4.40	0.19
5	4.40	4.66	4.67	0.42	4.56	0.33
6	4.66	4.93	4.77	0.45	4.90	0.84
7	4.93	5.19	5.25	0.59	5.55	0.54
8	5.19	5.45	5.61	0.36	5.33	0.65
9	5.45	5.71	5.61	0.39	5.50	0.41
10	5.71	5.98	5.71	0.67	5.66	0.39
11	5.98	6.24	5.81	0.70	5.88	0.36
12	6.24	6.50	5.99	0.70	5.88	0.42
13	6.50	6.76	6.10	0.75	6.11	0.61
14	6.76	7.03	6.20	0.85	6.22	0.72
15	7.03	7.29	6.48	0.80	6.78	1.24
16	7.29	7.55	6.63	0.78	6.73	0.87
17	7.55	7.81	6.89	0.87	6.79	0.88
18	7.81	8.08	7.75	1.13	6.88	0.95
19	8.08	8.34	8.20	1.26	7.64	1.81
20	8.34	8.60	8.36	1.03	7.78	1.70
21	8.60	8.86	8.33	1.05	8.33	1.20
22	8.86	9.13	8.31	1.23	8.42	0.97
23	9.13	9.39	8.49	1.06	8.77	1.04
24	9.39	9.65	8.48	1.42	9.22	1.13

(s.d. = standard deviation)

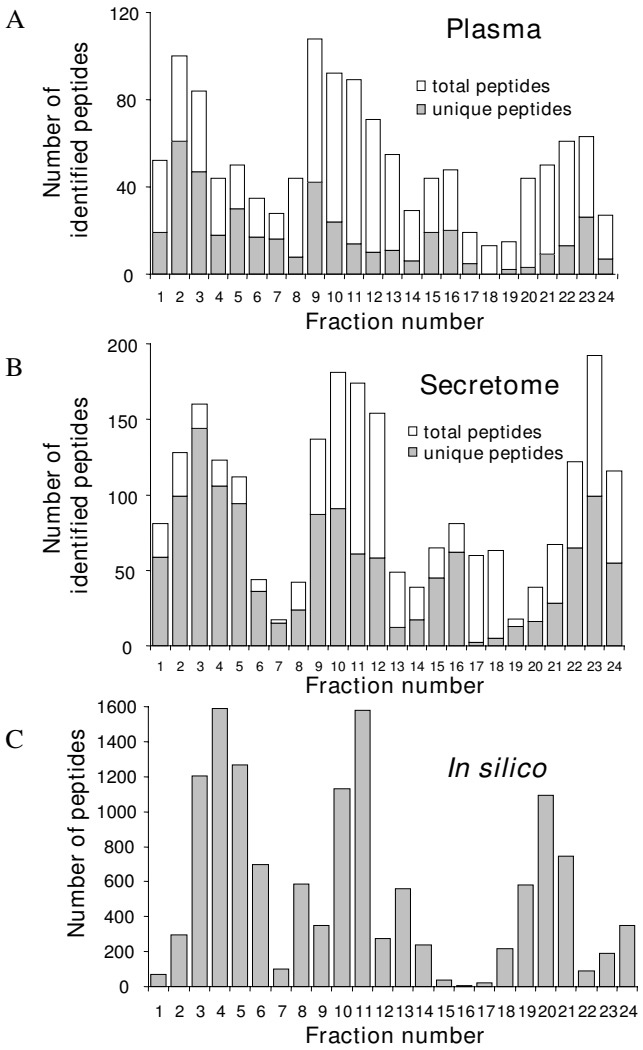


Figure 3
Total number of peptides identified per fraction: (A) in plasma sample and **(B)** in secretome sample. The grey area relates to the unique peptides in each fraction. **(C)** I*n silico* repartition of all the peptides identified for the 9504 plasma proteins from the Human Plasma Project found on the HUPO website.

protocol as the unlabelled sample for the separation. For this purpose, the peptide composition of three different fractions in different pH range was compared (fractions 2: acidic pH; 11: neutral pH and 24: basic pH) (Table 2). iTRAQ labeling did not affect the separation for the fractions 2 and 11, as the mean pI was identical for these fractions with or without labeling and a majority of peptides were found present in the same fraction in both conditions (68.1 % and 74.1 % respectively). For fraction 24, we observed a slight modification of the pI by iTRAQ labeling ($\Delta pI = 0.48$) and a lower percentage of coverage. It could be explained by a higher number of miss-cleaved peptides recovered from this fraction. To conclude, we showed for the first time the compatibility of iTRAQ labeling with OFFGEL separation. OFFGEL fractionation is particularly interesting for quantitative proteomic analysis because higher amounts of sample can be loaded and separated, allowing an optimization of the subsequent analytical steps. These results were also validated for plasma samples (data not shown), a complex sample that may require the coupling of two or three separation stages in order to access to the less abundant proteins. For example, a combination of OFFGEL, Strong Cation Exchange (SCX) and nano RP chromatographies seems to be an interesting approach to apprehend the complexity of the plasma. An another advantage of OFFGEL is its ability to desalt the sample during separation which can avoid in certain processes a supernumerary desalting step [15].

Conclusion

Sample fractionation is an essential step for proteomics analysis. The data reported here show that OFFGEL system provides a highly valuable tool to fractionate peptides from complex eukaryotic samples like plasma or secretome. We also demonstrated the compatibility of OFFGEL fractionation with the iTRAQ labeling reagents used for quantitative proteomics. Finally, we consider peptides OFFGEL separation as a highly valuable technique to integrate in a proteomic workflow for complex sample analysis.

Methods

Plasma samples

For quality and standardization issues, the plasma samples used were IQC samples (Internal Quality Control) provided by the proteomic laboratory of Dijon (Plateform Proteomic IFR- santé-STIC, France) as calibrated and reference samples. IQC was a pool of inactivate plasma packaged in controlled conditions. The samples were received in 50 μL aliquots and stored at -130°C until use. The samples were depleted using PROTEOPREP 20S spin column technology, according to the procedure recommended by Sigma-Aldrich, which remove the 20 highest abundant proteins. After 10 depletion cycles, corresponding to 80 μL of crude plasma, depleted solutions are recovered and then concentrated by ultra filtration using a 5 kDa molecular mass cut-off spin column (Amicon Ultra, Millipore).

Secretome samples

H358 human non-small lung adenocarcinoma cell line [16] were grown in RPMI-1640 medium with glutamine, 10% heat-inactivated fetal bovine serum and Penicillin/Streptomycin 5000 U in 175-cm^2 dishes until they reached a confluence state of approximately 60–70 %. They were then gently washed four times with phosphate buffered saline and two times with serum free medium to eliminate serum contaminants and left in serum free

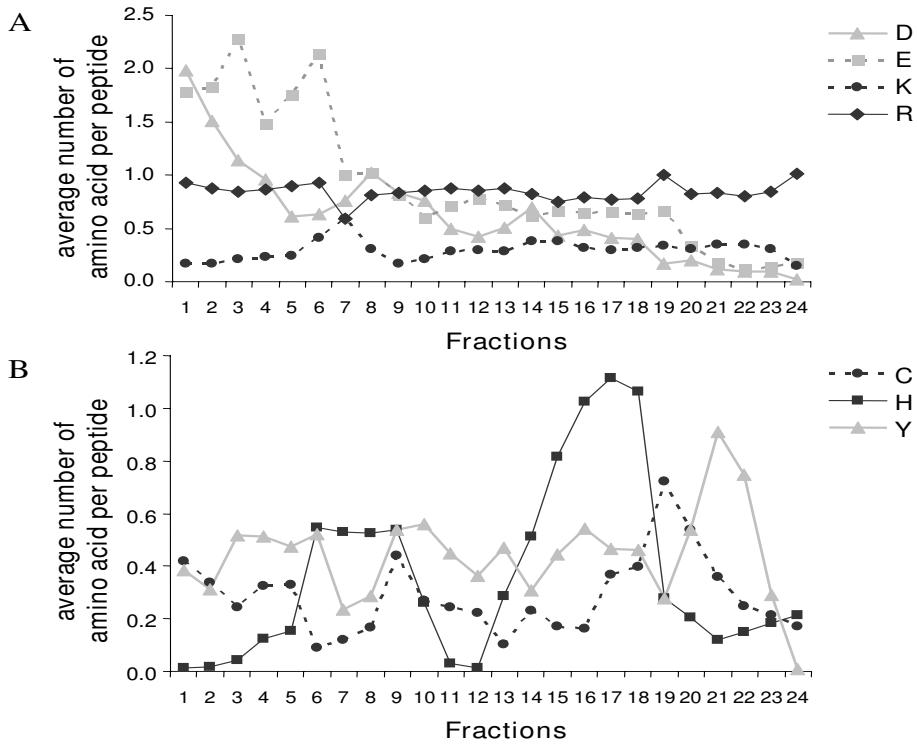


Figure 4

Average distribution of charged amino acids per peptide in each OFFGEL fraction obtained for secretome separation. **(A)** Average number of acidic amino acids (aspartic acid (D) and glutamic acid (E)), and basic amino acids (lysine (K) and arginine (R)).**(B)** Average number of cysteine (C), histidine (H) and tyrosine (Y) charged amino acids.

medium for 72 hours. The conditioned medium was collected and cooled down on ice. Floating cells and cellular debris were removed by centrifugation (200 × g, 10 min) followed by sterile filtration (pore size: 0.2 μm). Proteins were then concentrated and desalted by ultra-filtration using a 5 kDa molecular mass cut-off spin column (Amicon Ultra, Millipore) according to the manufacturer instructions. The total protein amount was determined using a standard Bradford protein assay (Bio-Rad).

Tryptic digestion

Secretome and plasma samples were treated simultaneously and with the same protocol. 200 μg of each sample were reduced with 45 mM dithiothreitol at 50°C for 35 min and alkylated with 100 mM iodoacetamide at room temperature for 45 min. Trypsin (Promega) was added at an enzyme: protein ratio of 3:100 w/w and incubated overnight at 37°C. The digests were dried by vacuum centrifugation prior to the OFFGEL peptides fractionation.

iTRAQ labelling

200 μg of proteins from secretome sample were resuspended, reduced, alkylated, and digested according to the standard protocol supplied by the manufacturer (Applied Biosystems). Then, 100 μg of each digest were labeled either with iTRAQ reagent 114 or iTRAQ reagent 117. After labeling, samples were pooled in a ratio 1:1 (v/v) (for a total of 200 μg of peptide digests) and dried by vacuum centrifugation prior to the OFFGEL peptides fractionation.

OFFGEL peptides fractionation

To perform peptide fractionation according to their pI, the 3100 OFFGEL Fractionator and the OFFGEL Kit 3–10 (both from Agilent Technology) were used following the user protocol. The device was set up for the 24 fractions separation by using 24-cm-long IPG gel strip with a linear pH gradient ranging at 3–10. The peptides are separated in a two-phases system: liquid upper phase (focusing buffer provided by the supplier) separated in wells and lower IPG gel strip phase. The wells are isolated from each others. There is no direct fluidic connection between the wells. The peptides migrate through the IPG gel that plays the role of 'bridge" between each well and are retrieved in the solution at the IPG region where pH is peptides pI. 200 μg of secretome (with or without iTRAQ labelling) or plasma tryptic digests were resuspended with focusing buffer to a final volume of 3.6 mL. 150 μL of this sample

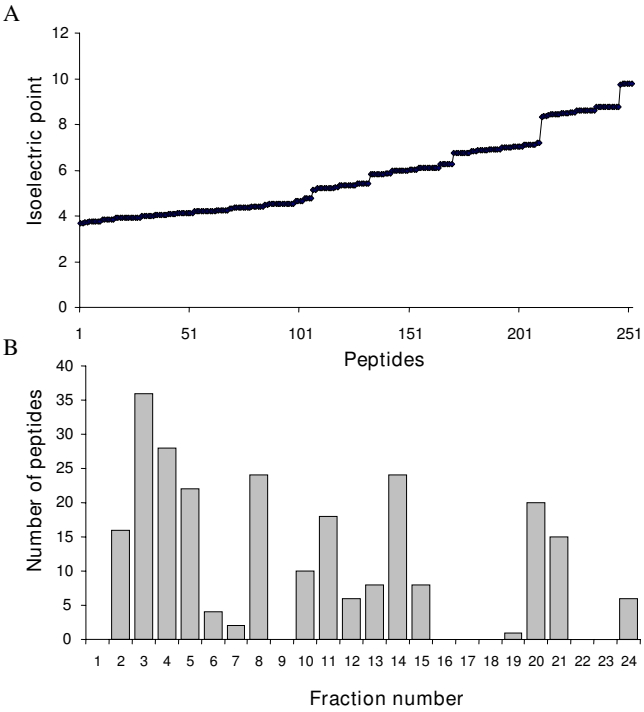


Figure 5
Analysis of peptides generated by a calculated model. Our model is based on the generation of all possible peptides containing a combination of 6 amino acids out of the 6 charged amino acids contributing to the pI calculation (D, E, R, K, H, Y). Only one arginine or lysine is theoretically present in a given peptide after tryptic digestion. **(A) Isoelectric point repartition of peptides obtained by the combination model**. Generated peptides are classified by increasing theoretical pI. **(B) Distribution of theoretical peptides per fraction.** Generated peptides are classified in function of their theoretical pI correspondence to the pH range of each fraction.

was loaded in each of the 24 wells. The sample was focused using the recommended method for OFFGEL peptides 24 wells fractionation with a maximum current of 50 μA. The focusing was stopped after total voltage reaches 50 kVh. During the focusing, oil was added to the electrodes to prevent any evaporation effect. After focusing, 50 to 150 μl of sample was recovered for each well and transferred in individual micro tubes. To recover as much as possible the focusing peptides, 150 μl of methanol was added to each well, incubated for 15 min without voltage [3]. Corresponding peptides fractions were pooled and concentrated by vacuum centrifugation prior to LC-MALDI MS/MS analysis.

Nano Reversed-phased LC-MALDI MS/MS analysis

Peptides were re-dissolved in 20 μl 0.2% trifluoroacetic acid. Peptides separation was performed on an Ultimate nanoHPLC System (Dionex/LC Packings, France) equipped with a PepMapC18 column (Dionex/LC Packings; 3-μm particles, 10 nm pore size, 75-μm *i.d.*), an autosampler and a Probot microfraction collector. The mobile phase consisted of a gradient of solvents A (0.05% trifluoroacetic acid; 2% acetonitrile in water) and B (0.05% trifluoroacetic acid; 80% acetonitrile in water). Injection was performed with 100% solvent A. The peptides were separated with a linear gradient of solvent B from 0–5% in 5 min, followed by an increase until to 40% of solvent B in 30 min and to 55% in 10 min at a flow rate of 0.3 μL/min. The column was washed and regenerated with 90% solvent B for 10 min and with 100% solvent A. For MALDI MS/MS analysis, column effluent was mixed in a 1:3 ratio with MALDI matrix (2 mg/mL α-cyano-4-hydroxycinnamic acid in 0.1% trifluoroacetic acid/70% acetonitrile) (v/v) and deposed on an Opti-tof LC/MALDI Insert 123 × 81 mm plate (Applied Biosystems) at a frequency of one spot/15s.

MALDI plates were analyzed by MALDI TOF TOF 4800 proteomics Analyzer mass spectrometer (Applied Biosystems) in positive reflector ion mode. MS spectra from m/z 700–3500 were acquired for each spot using 1500 laser shots. The ten most intense peaks in each MS spectrum above an S/N threshold of 100 were selected for MS/MS analysis.

Data analysis

Peptides and proteins identification were performed using the GPS (Global Proteome Server) Explorer software V3.6 (Applied Biosystems) with Mascot (Matrix Science) as the database search engine (V2.0). Each MS/MS spectrum was searched against a database of human protein

Table 2: iTRAQ labeling influence on peptides repartition for OFFGEL fractions 2, 11 and 24 for secretome sample.

	+ iTRAQ		- iTRAQ			
Fraction	**Peptide pI Average**	**s.d.**	**Peptide pI Average**	**s.d.**	**ΔpI**	**% of Coverage**
2	4.21	0.66	4.11	0.31	0.10	68.12
11	5.88	0.53	5.88	0.36	0.00	74.14
24	8.74	1.35	9.22	1.13	-0.48	56.90

(s.d. = standard deviation; ΔpI = difference of the peptide pI average for a same fraction with and without iTRAQ labeling)

sequences (Swiss-Prot, downloaded January 2006), resulting in a set of tryptic peptides matches with confidence values. Only the peptides with C.I.% > 85%, for any MS/MS spectrum were retained for further analysis. These peptide identifications were then combined using the MASCOT search engine to yield a set of human protein identifications with confidence values. The MASCOT searches were run using the following parameters: methionine oxidation, cystein carbamidomethylation modifications were selected as variable; 1 missed cleavage allowed; precursor error tolerance at < 50 ppm; MS/MS fragment tolerance set to 0.2 Da and charge set to +1; full trypsin specificity (N- and C-terminal also applied). Only proteins with at least two specific peptides matched were considered positively identified.

The theoretical pI of the peptides, were calculated using "Compute pI/Mw" tool accessible on the Expasy website [13,17]. The identified peptides with an ion score C.I.% higher than 95% were always kept in the peptide list even if theoretical pI did not match with pI of the relevant individual OFFGEL fraction. For peptide ion score C.I.% between 85% and 95%, peptides were kept if calculated theoretical pI was correlated with pH range ± 2.

Abbreviations

LC, Liquid Chromatography; pI, Isoelectric point; MALDI, Matrix Assisted Laser Desorption Ionization; MS, Mass Spectrometry; IPG, Isoelectric Point Gel; IEF, IsoElectric Focusing; IQC, Internal Quality Control; HUPO, Human Proteome Organisation.

Competing interests

The author(s) declare that they have no competing interests.

Authors' contributions

JC performed the laboratory experiments on secretome samples, carried out the LC separation of peptides samples and the MALDI-TOF/TOF analysis, and assisted in the writing of the final manuscript. JS performed the plasma samples depletion and participated in LC separation and MALDI-TOF/TOF analysis. SM participated in the design of the study and wrote the manuscript. MS supervised and coordinated the project. All authors read and approved the final manuscript.

Acknowledgements

The medical proteomic platform of the Center for Innovation in Biology (CIB) benefits from financial supports from the Lyon-Auvergne-Rhône-Alpes Canceropole (CLARA), Nanobio project, Grenoble university hospital and the French research minister. This work was supported by a grant from the French research minister to JC. This work was supported by a Grant from INCA (National Institute of Cancer) free projects 2006 PL06_025, "Analysis of secreted and targeted proteins from non-small cell lung cancers for the identification of new seric biomarkers". We would like to thank Patrick Ducoroy (Platform Proteomic IFR-Sante-STIC) for providing Internal Quality Control Plasma samples and Pascal Bertolino for developing specific algorithms.

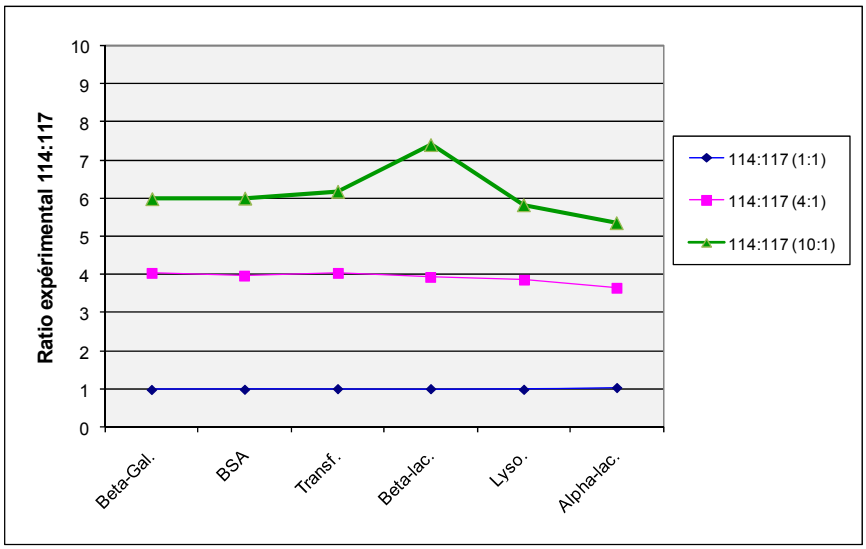
References

1. Michel PE, Reymond F, Arnaud IL, Josserand J, Girault HH, Rossier JS: **Protein fractionation in a multicompartment device using Off-Gel isoelectric focusing.** *Electrophoresis* 2003, **24(1-2):**3-11.
2. Horth P, Miller CA, Preckel T, Wenz C: **Efficient fractionation and improved protein identification by peptide OFFGEL electrophoresis.** *Mol Cell Proteomics* 2006, **5(10):**1968-1974.
3. Heller M, Michel PE, Morier P, Crettaz D, Wenz C, Tissot JD, Reymond F, Rossier JS: **Two-stage Off-Gel isoelectric focusing: protein followed by peptide fractionation and application to proteome analysis of human plasma.** *Electrophoresis* 2005, **26(6):**1174-1188.
4. Hathout Y: **Approaches to the study of the cell secretome.** *Expert Rev Proteomics* 2007, **4(2):**239-248.
5. Tjalsma H, Bolhuis A, Jongbloed JD, Bron S, van Dijl JM: **Signal peptide-dependent protein transport in Bacillus subtilis: a genome-based survey of the secretome.** *Microbiol Mol Biol Rev* 2000, **64(3):**515-547.
6. Anderson NL, Polanski M, Pieper R, Gatlin T, Tirumalai RS, Conrads TP, Veenstra TD, Adkins JN, Pounds JG, Fagan R, Lobley A: **The human plasma proteome: a nonredundant list developed by combination of four separate sources.** *Mol Cell Proteomics* 2004, **3(4):**311-326.
7. Ross PL, Huang YN, Marchese JN, Williamson B, Parker K, Hattan S, Khainovski N, Pillai S, Dey S, Daniels S, Purkayastha S, Juhasz P, Martin S, Bartlet-Jones M, He F, Jacobson A, Pappin DJ: **Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents.** *Mol Cell Proteomics* 2004, **3(12):**1154-1169.
8. Fraterman S, Zeiger U, Khurana TS, Rubinstein NA, Wilm M: **Combination of peptide OFFGEL fractionation and label-free quantitation facilitated proteomics profiling of extraocular muscle.** *Proteomics* 2007, **7(18):**3404-3416.
9. Cargile BJ, Talley DL, Stephenson JL Jr.: **Immobilized pH gradients as a first dimension in shotgun proteomics and analysis of the accuracy of pI predictability of peptides.** *Electrophoresis* 2004, **25(6):**936-945.
10. Essader AS, Cargile BJ, Bundy JL, Stephenson JL Jr.: **A comparison of immobilized pH gradient isoelectric focusing and strong-cation-exchange chromatography as a first dimension in shotgun proteomics.** *Proteomics* 2005, **5(1):**24-34.
11. **Human Proteome Organisation** [http://www.hupo.org]
12. Lam HT, Josserand J, Lion N, Girault HH: **Modeling the isoelectric focusing of peptides in an OFFGEL multicompartment cell.** *J Proteome Res* 2007, **6(5):**1666-1676.
13. Bjellqvist B, Hughes GJ, Pasquali C, Paquet N, Ravier F, Sanchez JC, Frutiger S, Hochstrasser D: **The focusing positions of polypeptides in immobilized pH gradients can be predicted from their amino acid sequences.** *Electrophoresis* 1993, **14(10):**1023-1031.
14. Lengqvist J, Uhlen K, Lehtio J: **iTRAQ compatibility of peptide immobilized pH gradient isoelectric focusing.** *Proteomics* 2007, **7(11):**1746-1752.
15. Arnaud IL, Josserand J, Jensen H, Lion N, Roussel C, Girault HH: **Salt removal during Off-Gel electrophoresis of protein samples.** *Electrophoresis* 2005, **26(9):**1650-1658.
16. Dubrez L, Coll JL, Hurbin A, Solary E, Favrot MC: **Caffeine sensitizes human H358 cell line to p53-mediated apoptosis by inducing mitochondrial translocation and conformational change of BAX protein.** *The Journal of biological chemistry* 2001, **276(42):**38980-38987.
17. **Compute pI/Mw tool on Expasy website** [http://us.expasy.org/tools/pi_tool.html]

6. Résultat complémentaire : Validation de l'approche iTRAQ

La méthode de marquage iTRAQ a été validée avant son utilisation pour l'analyse différentielle des protéines sécrétées influencées par p53. Ce test est effectué sur un mélange de 6 protéines connues que l'on a digéré à la trypsine (mix6). Le digestat de mix6 est réparti en différents aliquots et chacun est marqué avec un tag iTRAQ spécifique. Ces aliquots marqués sont alors mélangés selon des ratios de concentrations définis.

A



B

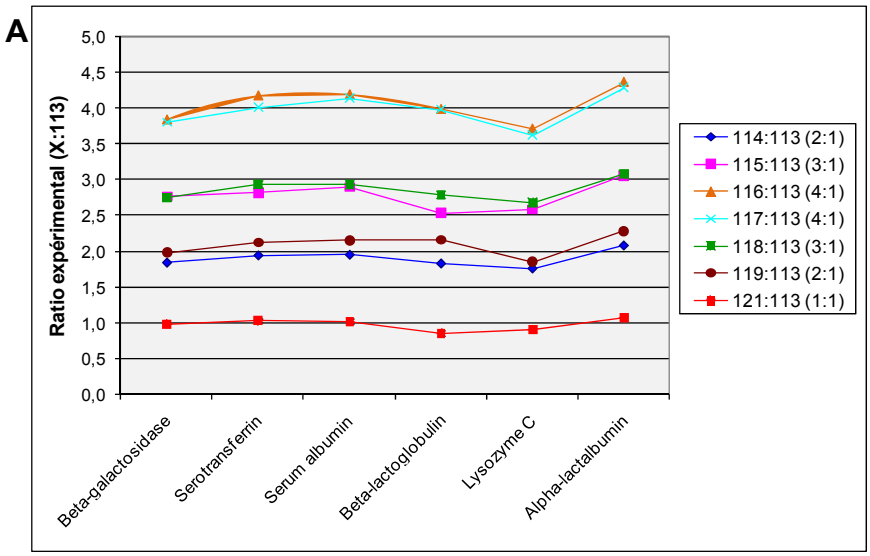
Ratio imposé	Ratio moyen expérimental	Std. Dev.
1:1	1,01	0,02
4:1	3,92	0,14
10:1	6,13	0,69

Figure 61 : Validation de la quantification relative des protéines par marquage iTRAQ 2-plex.

Exemple d'un essai 2-plex comparant 2 aliquots d'un même échantillon marqués spécifiquement par un tag iTRAQ et mélangés selon 3 ratios 114:117 différents : ratios 1:1 ; 4:1 et 10:1. L'échantillon pour chaque marquage est constitué d'un mélange de 6 protéines standards digérées à la trypsine (mix6).

A. Graphique des ratios 114:117 expérimentaux obtenus pour chacune des 6 protéines du mix6. Le ratio théorique appliqué est noté entre parenthèse pour chaque ratio.

B. Tableau des ratios iTRAQ obtenus.
Ratio imposé : Ratio 114:117 appliqué. Ratio moyen expérimental : Moyenne des valeurs obtenues pour les 6 protéines du mix6 pour un ratio imposé donné. Std. Dev. : Déviation standard inter-protéines pour un ratio imposé donné.



B

Ratio imposé	Ratio moyen expérimental	Std. Dev.
114:113 (2:1)	1,89	0,11
115:113 (3:1)	2,77	0,20
116:113 (4:1)	4,04	0,24
117:113 (4:1)	3,97	0,24
118:113 (3:1)	2,85	0,15
119:113 (2:1)	2,09	0,16
121:113 (1:1)	0,97	0,08

Figure 62 : Validation de la quantification relative des protéines par marquage iTRAQ 8-plex.

Exemple d'un essai 8-plex comparant 8 aliquots d'un même échantillon marqués spécifiquement par un tag iTRAQ et mélangés selon le ratio 1:2:3:4:4:3:2:1. L'échantillon pour chaque marquage est constitué d'un mélange de 6 protéines standards digérées à la trypsine (mix6).

A. Graphique des différents ratios obtenus pour chacune des 6 protéines du mix6. Ces ratios sont présentés par rapport à la valeur 113. Le ratio théorique appliqué est noté entre parenthèse pour chaque ratio.

B. Tableau des ratios iTRAQ obtenus par rapport à la valeur 113.
Ratio imposé : Ratio X:113 appliqué. Ratio moyen expérimental : Moyenne des valeurs obtenues pour les 6 protéines du mix6 pour un ratio imposé donné. Std. Dev. : Déviation standard inter-protéines pour un ratio imposé donné.

La fiabilité de la quantification est jugée en comparant les ratios attendus et les ratios obtenus expérimentalement. Ce test est effectué pour les kits iTRAQ 2-plex et 8-plex (Figure 61 et Figure 62). Dans les deux cas, les ratios obtenus sont très proches de ceux attendus. Ces résultats sont en accord avec ceux précédemment publiés [66, 74] et valident la fiabilité de l'iTRAQ pour la quantification relative de mélanges peptidiques.

Nous notons néanmoins que les ratios d'expression importants (ratio de 10 par exemple) ont tendance à être atténués expérimentalement (voir ratio 1:10 de la Figure 61). Ainsi, certains ratios d'expression obtenus lors de notre analyse du secretome modulé par p53 dans les cellules H358 sont sûrement sous-estimés. Par exemple, la quinone oxidoreductase, identifiée comme 5,7 fois plus concentrée dans le secretome en présence de p53, est peut-être modulée de façon plus conséquente (voir Chapitre 2 des résultats). Ce biais reste limité car il n'empêche pas d'identifier les protéines modulées par p53 mais entraine une légère sous estimation des ratios d'expression les plus conséquents.

A partir de ces résultats, nous avons également fixé les seuils de ratios jugés significatifs, c'est-à-dire différents de 1. Ainsi, les ratios sont considérés comme différents de 1 lorsqu'ils seront inférieurs à 0,8 ou supérieur à 1,25. Ces valeurs seuils sont en accord avec celles utilisées lors de précédentes études utilisant l'iTRAQ [66, 74].

Deuxième partie :

Analyse différentielle des protéines sécrétées modulées par p53 dans des cellules du Cancer du Poumon Non à Petites Cellules

CHAPITRE 2 : Analyse différentielle des protéines sécrétées modulées par p53 dans des cellules du Cancer du Poumon Non à Petites Cellules.

1. Problématique et objectifs du travail

Notre travail s'est axé autour de la question suivante :

« Quelles sont les protéines sécrétées modulées par p53 et impliquées dans la tumorogenèse des cellules tumorales pulmonaires ? »

La protéine p53 joue un rôle crucial dans la progression tumorale. Elle est très souvent mutée dans les cancers du poumon et cette mutation a des répercussions au niveau de l'interaction des cellules tumorales avec leur micro-environnement par une modulation de la sécrétion de protéines. Une meilleure connaissance de ces facteurs est primordiale pour mieux comprendre la communication des cellules tumorales entre elles et avec leur environnement. De plus, ces protéines représentent des pistes pour la recherche de cibles thérapeutiques ou diagnostiques potentielles.

Afin de répondre à notre question, nous avons mis en place un modèle cellulaire permettant une expression inductible de la protéine p53 sauvage. Ce modèle *in vitro* donne l'opportunité d'étudier les protéines influencées par p53 qui sont libérées par des cellules tumorales pulmonaires dans le milieu extra-cellulaire.
La réalisation de cette étude a nécessité le développement d'un procédé d'analyse performant. Notre choix s'est porté sur une approche protéomique, très efficace pour l'analyse globale des protéines sans *à priori*, ainsi que leur quantification relative. Cette étude permettra de mieux comprendre le rôle joué par p53 dans la modulation des interactions intercellulaires par la régulation de la sécrétion de protéines. De plus, la mise en place d'un modèle de xénogreffe de la lignée H358 chez la souris permettra d'appréhender la relation entre le secretome *in vitro* et la libération *in vivo* de protéines par la tumeur.

2. Moyens mis en œuvre

Pour réaliser cette étude, nous avons choisi comme modèle cellulaire la lignée H358 (ATCC CRL5807), issue d'un adénocarcinome du Cancer du Poumon Non à Petites Cellules (NSCLC). Ce type de cancer du poumon est d'un point de vue histologique le plus fréquemment rencontré (entre 75 et 85 % des cas).

Cette lignée est résistante à l'apoptose due à la privation de sérum pendant la culture [261, 262]. Les cellules peuvent ainsi être cultivées en milieu sans sérum durant le temps nécessaire à la sécrétion d'une quantité suffisante de protéines (72 heures). Ce paramètre est indispensable pour ne pas contaminer le milieu conditionné des cellules (contenant les protéines sécrétées) par des protéines du sérum de veau fœtal. Les cellules H358 ont une délétion homozygote du gène *TP53*. Elles ont été précédemment transfectées par un plasmide codant pour le gène *TP53* sauvage placé sous le contrôle positif d'un promoteur tetracycline-dépendant (système TetOn) [263, 264]. L'ajout de doxycycline dans le milieu de culture des cellules à une concentration finale non toxique de 1 µg/mL déclenche l'expression de la protéine p53 par ces cellules. Cette induction de p53 commence 4 heures après l'ajout de doxycycline, atteint son maximum d'expression au bout de 24 heures et reste élevée durant 13 jours. L'expression de p53 dans ces cellules déclenche un arrêt du cycle cellulaire en phase G_2/M mais pas d'augmentation d'apoptose.
De par sa capacité à pousser en milieu sans sérum et son expression inductible de p53 sauvage, la lignée H358/TetOn/p53 est un modèle idéal pour notre étude.

Les protéines sécrétées par les cellules sont récupérées en prélevant le milieu de culture. Celui-ci est centrifugé et filtré afin d'éliminer les éventuelles cellules flottantes ou débris cellulaires. Le milieu est finalement concentré et dessalé pour obtenir une quantité suffisante de protéines analysables.

Un procédé d'analyse protéomique semi-quantitative adapté à l'étude différentielle de l'influence de p53 sur le secretome des cellules H358 a été développé.
Une première approche utilise un marquage iTRAQ différentiel des échantillons et leur analyse par LC-MS/MS. Les échantillons protéiques sont digérés en solution par la trypsine. Chaque digestat est alors marqué par un tag iTRAQ spécifique. Une première étude iTRAQ 4-plex est réalisée sur deux préparations d'échantillons indépendantes pour chacune des deux conditions (p53- et p53+). Les peptides spécifiquement marqués sont mélangés et séparés simultanément par la méthode OFFGEL. Chacune des 24 fractions OFFGEL est ensuite analysée par LC-MS/MS (chromatographie en phase inverse et MALDI-TOF/TOF en off-line) afin d'identifier et quantifier chacun des peptides. Une étude 2-plex est effectuée indépendamment sur la première préparation de chaque échantillon avec une inversion des tags pour éliminer le risque de différences d'expression tag dépendante. Le choix de développer l'approche iTRAQ résulte de ses nombreux avantages. En effet, elle permet une analyse multiplex, sensible, sans complexification du spectre MS (marquage isobarique). Par contre, elle repose sur un marquage tardif des échantillons (après la digestion des protéines). Afin de minimiser le risque expérimental

induit par les étapes de préparation précédant le marquage, nous avons décidé de réaliser les réplicas définis précédemment et de ne juger comme fiable que les résultats concordants.

Une approche « gel-based » utilisant une séparation des protéines par électrophorèse bidimensionnelle a aussi été appliquée. Celle-ci a pour but d'analyser l'influence de p53 sur la modulation des modifications post-traductionnelles des protéines sécrétées. Cette analyse permet aussi de recouper les résultats obtenus avec ceux de l'approche iTRAQ. Chaque condition est analysée en triplicat de façon à être statistiquement fiable. Les images de gels sont analysées par le logiciel ImageMaster (GE Healthcare). Celui-ci va détecter et associer les spots des différents gels, puis réaliser leur quantification relative. Les protéines d'intérêt sont alors digérées "in-gel" et identifiées par MALDI-TOF/TOF en PMF et/ou MS/MS.

Le procédé d'analyse protéomique mis en place pour cette étude est présenté Figure 63.

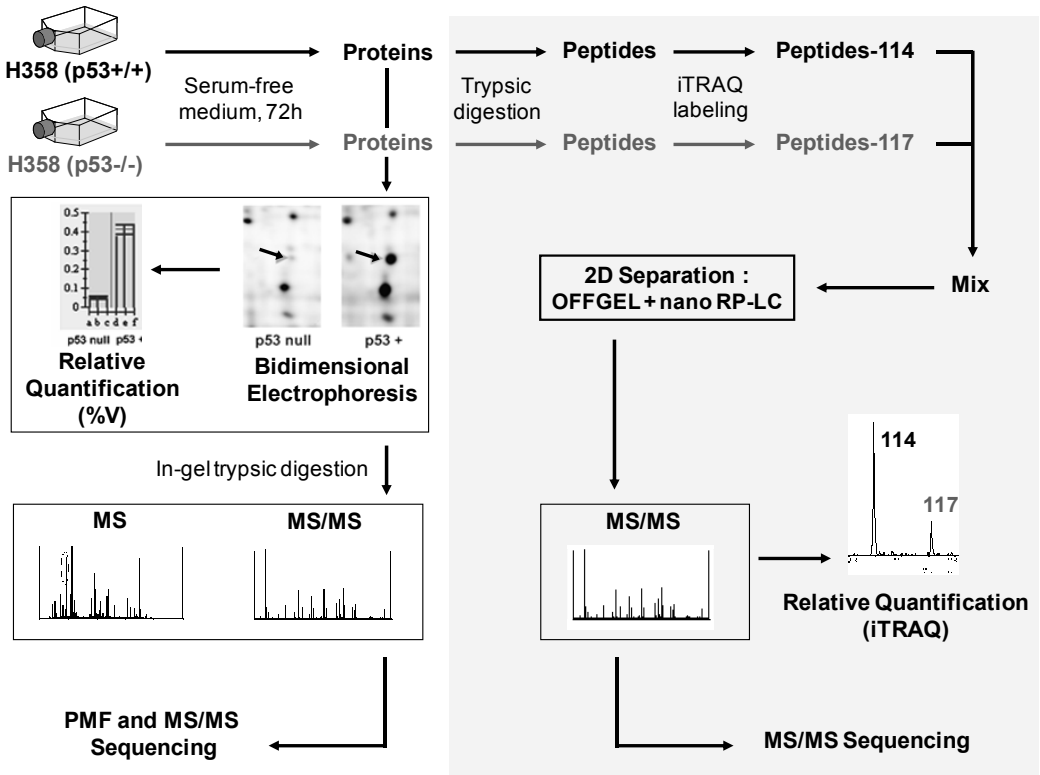


Figure 63 : Représentation du procédé d'analyse protéomique utilisé pour l'analyse différentielle des protéines p53-modulées relarguées par la lignée H358.
Gauche : Analyse par électrophorèse bidimensionnelle.
Droite : Analyse par marquage iTRAQ.

L'étude *in vivo* est réalisée sur un modèle de xénogreffe des cellules H358/TetOn/p53 chez des souris nude. Une fois que la tumeur s'est implantée, les souris sont réparties en deux groupes. L'un des groupes de souris reçoit de la doxycycline dans son eau de boisson afin d'induire l'expression de p53 au niveau du greffon. Après 72 heures, les souris sont sacrifiées, et la tumeur et le plasma collectés. L'analyse des protéines d'intérêt est réalisée pour ces deux types d'échantillons par ELISA et Western-blot.

3. Résultats

Ces travaux sont présentés dans l'article suivant :

The cell line secretome, a suitable tool for investigating proteins released in vivo by tumors: Application to the study of p53-modulated proteins secreted in lung cancer cells.
Jérôme CHENAU, Sylvie MICHELLAND, Florence de FRAIPONT, Véronique JOSSERAND, Jean-Luc COLL, Marie-Christine FAVROT, Michel SEVE.

Cet article a été soumis dans « Journal of Proteome research ».

Résumé des principaux résultats :

Nous avons développé un procédé d'analyse protéomique efficace pour l'analyse différentielle des protéines sécrétées par des cellules NSCLC en fonction du statut p53 (Figure 63).
L'analyse LC-MS/MS avec marquage différentiel iTRAQ a permis d'identifier 909 protéines du secretome de la lignée H358. (Tableau supplémentaire S1 de la publication). Le taux de faux positifs est de 0,14 %
Parmi ces protéines, 91 ont un niveau de sécrétion modulé de façon significative par l'expression de la protéine p53 sauvage. Quarante-sept sont surexprimées et 44 sous-exprimées dans le secretome lorsque p53 est exprimé (Tableaux 1 et 2 de la publication).
L'électrophorèse bidimensionnelle a permis l'identification de 184 protéines dont 182 sont recoupées avec les 909 identifiées par l'approche LC-MS/MS. Treize protéines ont une modulation d'expression dépendante de p53 en accord avec les résultats obtenus par l'approche iTRAQ. Par contre, même si plusieurs états de PTMs sont résolus sur les gels, aucune altération p53-dépendante de ces PTMs n'a pu être caractérisée (Tableaux 1 et 2 et Figure 1 de la publication).
L'étude de la distribution des protéines modulées par p53 selon les 3 différentes voies de sécrétion (classique, non-classique et exosome) a été effectuée (Figure 2 et tableaux 1 et 2 de la publication).
P53 est un facteur de transcription. Quinze protéines sur les 47 surexprimées et 12 sur les 44 sous-exprimées dans le secretome en présence de p53 sont des cibles transcriptionnelles de p53 préalablement démontrées (Tableaux 1 et 2 de la publication).
L'analyse de la littérature montre une relation inverse entre la variation des protéines modulées par p53 et les procédés de tumorogenèse (Figure 3 et tableau supplémentaire S2 de la publication).

Les résultats protéomiques ont été validés pour 16 protéines par Western-blot (Figure 4A de la publication).

Un Western-blot de contrôle a validé l'action spécifique de p53 sur ces variations (élimination d'un effet dox) (Figure 4B de la publication).

L'expression de p53 chez des souris xénogreffées par des cellules H358/TetOn/p53 a été validée par Western-blot. Celui-ci confirme la fonctionnalité de p53 par l'observation de l'activation d'une de ses cibles transcriptionnelles, p21 (Figure 5 de la publication).

L'analyse par ELISA des variations d'expression des protéines GDF-15, VEGF et FGF-19 a été validée sur les extraits totaux de tumeur des souris. GDF-15 est également détectée dans le plasma des souris et sa modulation d'expression par p53 a été démontrée (Figure 6 de la publication).

4. Conclusion

Nous avons développé et validé un procédé efficace pour l'analyse différentielle des secretomes.

Cette approche a permis d'établir la liste de protéines d'un secretome la plus complète à ce jour dans laquelle 91 protéines sont influencées par p53.

Ces résultats sont utiles pour mieux comprendre le rôle joué par p53 dans la tumorogenèse à travers sa régulation de facteurs sécrétés.

La modulation par p53 est confirmée *in vivo* chez la souris pour trois protéines (GDF-15, FGF-19 et VEGF).

La protéine GDF-15 est également modulée par p53 dans les plasmas des souris xénogreffées. C'est la première fois qu'une protéine exprimée de façon différentielle dans une étude secretome est retrouvée et quantifiée dans le plasma de souris xénogreffées.

Le secretome d'une lignée cellulaire reflète la libération de protéines par les tumeurs *in vivo* et représente donc un modèle fiable pour la caractérisation des marqueurs potentiels de tumeur.

5. Publication II

The cell line secretome, a suitable tool for investigating proteins released in vivo by tumors: Application to the study of p53-modulated proteins secreted in lung cancer cells

Journal:	*Journal of Proteome Research*
Manuscript ID:	pr-2009-00383g.R1
Manuscript Type:	Article
Date Submitted by the Author:	
Complete List of Authors:	CHENAU, Jerome; Université Joseph Fourier - Grenoble 1, INSERM, Institut Albert Bonniot U823 MICHELLAND, Sylvie; Grenoble Hospital, Centre d'Innovation en Biologie DE FRAIPONT, Florence; Grenoble Hospital, Département de Biologie Intégrée Josserand, Veronique; IAB Coll, Jean-Luc; INSERM FAVROT, Marie-Christine; Grenoble Hospital, Centre d'Innovation en Biologie SEVE, Michel; Grenoble Hospital, Centre d'Innovation en Biologie

scholarONE™
Manuscript Central

The cell line secretome, a suitable tool for investigating proteins released *in vivo* by tumors: Application to the study of p53-modulated proteins secreted in lung cancer cells

Jérôme CHENAU[1], Sylvie MICHELLAND[1, 2], Florence de FRAIPONT[1, 3], Véronique JOSSERAND[1], Jean-Luc COLL[1], Marie-Christine FAVROT[1, 2], Michel SEVE[1, 2 *]

[1] Université Joseph Fourier - Grenoble 1, INSERM, Institut Albert Bonniot U823, Grenoble, France

[2] CHU Grenoble, Centre d'Innovation en Biologie, Plate-forme de Protéomique Médicale, Grenoble, France

[3] CHU Grenoble, Cancérologie Biologique et Biothérapies, département de Biologie Intégrée, Grenoble, France

* Correspondence

Prof. Michel SEVE
Plate-forme Protéomique Médicale - Centre d'Innovation en Biologie
CHU Grenoble
BP 217
38043 Grenoble cedex 09
FRANCE
Tel: 33 (0)4 76 76 54 84
Fax: 33 (0)4 76 76 56 64
Email: Michel.Seve@ujf-grenoble.fr

Abbreviations

CM: Conditioned medium; Dox: Doxycycline; MMTS: Methyl methanethiosulfonate; NSCLC: Non-Small Cell Lung Cancers; SILAC: Stable Isotope Labeling with Amino acid in Culture; TCEP: Tris-(2-carboxyethyl) phosphine; TEAB: Triethylammonium bicarbonate; WB: Western blot; wt: wild type

Keywords

Proteomics, iTRAQ, lung cancer, p53, secretome, xenograft mouse model

Abstract

Malignant processes such as metastasis, invasion, or angiogenesis are tightly dependent on the composition of the extracellular medium, which is itself affected by the release of proteins by the tumor cells. p53, a major tumor suppressor protein very frequently mutated and/or inactivated in cancer cells, is known to modulate the release of proteins by the tumor cells; however , while p53-modulated intracellular proteins have been extensively studied, little is known concerning their extracellular counterparts. Here, we characterized the p53-dependent secretome of a lung tumor model *in vitro* (H358 human non-small cell lung adenocarcinoma cell line with a homozygous deletion of p53) and demonstrate that the modulation of exported proteins can also be detected *in vivo* in the plasma of tumor-bearing mice. We used a clone of H358, stably transfected with a tetracycline-inducible wild-type p53-expressing vector. Using iTRAQ labeling and LC-MALDI-MS/MS analysis, we identified 909 proteins released in vitro by the cells, among which 91 are p53-modulated. Three proteins (GDF-15, FGF-19, and VEGF) were also investigated in H358/TetOn/p53 xenograft mice. The ELISA dosage on total tumor protein extracts confirmed the influence of p53 on the release of these proteins *in vivo*. Moreover, the GDF-15 concentration was measured in the plasma and its p53-dependent modulation was confirmed. To our knowledge, this is the first report establishing that the *in vitro* cell line secretome is reliable and reflects the extracellular release of proteins from tumor cells *in vivo* and could be used to identify putative tumor markers.

Introduction

Lung cancer is a major public health issue and the leading cause of cancer-related deaths worldwide [1, 2]. Non-small cell lung cancers (NSCLC) account for approximately 80% of all types of lung cancers [3, 4]. Tumor–stroma interactions play a major role in tumor development, maintenance, progression, invasion, and neo-angiogenesis. Modulation of the stromal composition by tumor-produced proteins plays a pivotal role in the intercellular cross-talk [5].

In this study, we examined the role of wild-type (wt) p53 protein on the tumor microenvironment. Indeed, p53 is a major tumor suppressor with a DNA-binding activity known to control the transcription of a large panel of genes, including some encoding for secreted factors [6]. It plays an important role in maintaining cell integrity by controlling cell cycle progression and cell survival in response to DNA damage [7, 8]. *TP53* gene is the most frequently mutated gene in human cancer. Indeed, a mutational inactivation of the *TP53* gene is observed in around 50% of human cancers and over 60% of lung cancers [8-11]. Loss of p53 function predisposes cells to a rapid accumulation of multiple genetic changes. As a consequence, mice without the functional *TP53* gene breed and develop normally but die at an early age from multiple cancers [12]. Additionally, restoration of p53 function in tumor cells can prevent tumor development [13]. p53 also influences tumor–stroma interactions by modulating secreted pro- or anti-tumorigenic proteins [14, 15]. While p53-regulated intracellular proteins have been thoroughly studied [16-18], there are few data on the influence of p53 on the release of proteins by tumor cells. In a recent study on glioma cells, 60 secreted proteins modulated by p53 expression were described [19].

The term “secretome” refers to the whole pool of proteins released by cells in culture medium at a given time and under defined conditions [20, 21]. These secreted proteins can be growth factors, extracellular matrix-degrading proteinases, cell motility factors, immunoregulatory cytokines, or other bioactive molecules. They are essential in the processes of differentiation, invasion, metastasis, and angiogenesis by regulating cell-to-cell and cell-to-extracellular matrix interactions. These tumor-secreted proteins are also thought to be released in body fluids such as blood or urine

and can be measured by noninvasive assays [22]. Thus, cancer secretome analysis could be a promising tool supporting the identification of cancer biomarkers. Recent differential secretome studies have been performed in colon cancer using DIGE [21], in pancreatic and lung cancers using SILAC [23, 24], or in prostate cancer using ICAT [25]. To our knowledge, however, the influence of p53 on the secretome of the lung cancer cells has not been described.

The secreted proteins modulated by p53 and involved in the tumor–stroma interactions of the lung cancer cells were studied from the H358 cell line, a human non-small cell lung adenocarcinoma cell line of bronchioalveolar origin with a homozygous deletion of p53 [26]. This cell line was stably transfected with a wt-p53-expressing plasmid under the positive control of a tetracycline-dependent promoter (H358/TetOn/p53) [27, 28]. The expression of wt p53 triggers G_2/M cell cycle arrest without inducing apoptosis [27] and is sufficient to induce tumor regression *in vivo* [28]. Serum removal from culture medium significantly decreased the number of cells in the G_2/M phase and increased the number of cells in the G_0/G_1 phase, both in p53-null and in p53-expressing cells, but it did not induce a significant apoptosis in any condition [27, 29].

The release of p53-modulated proteins in the H358/TetOn/p53 cell line was analyzed by means of proteomics using LC-MALDI-MS/MS differential analysis with specific iTRAQ labeling. Moreover, we conducted a two-dimensional gel electrophoresis analysis to study the influence of p53 expression on the modulation of post-translational modifications (PTMs) of secreted proteins. Detailed analyses of the known function of these p53-modulated proteins indicate that this in vitro model contains the hallmark of the main biological processes that favor tumor progression. In order to test whether the secretome could be a suitable tool for investigating proteins released *in vivo* by tumors, H358/TetOn/p53 xenografts were established in nude mice. The relative levels of the proteins of interest were determined in crude tumor protein extracts and in the plasma of mice to investigate a potential use of the secretome in the discovery of tumor markers.

Materials and Methods

Cell culture and preparation of secreted protein samples

The human non-small cell lung adenocarcinoma cell line H358 of bronchioalveolar origin (ATCC CRL5807) has a homozygous deletion of p53. This cell line was transfected with a tetracycline-inducible wt p53-expressing vector (TetOn system) [28]. Cells were grown in RPMI-1640 medium with 2 mM L-glutamine, 10% heat-inactivated fetal calf serum, and penicillin (100 U/mL) / streptomycin (100 µg/mL) in 175-cm^2 dishes. The H358/TetOn/p53 clone was maintained by adding hygromycin (50 µg/mL) and geniticin (G418: 100 µg/mL). All products were purchased from Gibco. When cells reached a confluence state of approximately 60–70%, they were gently washed four times with PBS and twice with serum-free medium to eliminate serum contaminants and then left in serum-free medium for 72 h with or without the tetracycline analog doxycycline. Doxycycline induces p53 expression and was added to serum-free medium to a final concentration of 1 µg/mL, which is not toxic. The analysis of apoptosis in H358/TetON/p53 cells incubated in serum-free medium for 72 h in the presence or absence of 1 µg/mL doxycycline was realized by flow cytometry (FACSCalibur, Becton Dickinson). The cells were also treated for 12 h with 1 µM staurosporine as a positive control for apoptosis. Proteolytic activation of procaspase-3 was measured with PE Caspase-3 Active Apoptosis Kit (BD Pharmingen) following the manufacturer's instructions. The different conditioned media were then collected and cooled on ice. Floating cells and cellular debris were removed by centrifugation (200 *g*, 10 min) followed by sterile filtration (pore size: 0.2 µm). Proteins were then concentrated and desalted by ultra-filtration using a 5,000-Da molecular mass cut-off spin column (Amicon Ultra-15, Millipore) according to the manufacturer's instructions. The total protein amount was determined using a standard Bradford protein assay (Thermo Scientific).

Isobaric Tags for Relative and Absolute Quantitation (iTRAQ)

Peptide isobaric labeling with iTRAQ reagents

Three different cultures of the H358/TetOn/p53 cell line were realized. In each case, one part of the cells was induced for wt-p53 expression and the other not. Six secretome samples were retained and dried by vacuum centrifugation. These secretome samples were analyzed by a 2-plex or a 4-plex iTRAQ experiment. For iTRAQ labeling, samples were treated in parallel, essentially as described by Ross et al. [30]. Stock reagents and buffers (TEAB, SDS, TCEP, MMTS, and the four isobaric tagging reagents) were obtained from Applied Biosystems. Briefly, 100 µg of each protein sample was reduced with 2.5 mM TCEP (60°C for 1 h) and cystein residues were blocked with 10 mM MMTS at room temperature for 20 min. Trypsin solution (Porcine modified, Promega) was added in a 1:20 ratio (w/w) and incubated for 20 h at 37°C. Protein digests were dried by vacuum centrifugation and suspended in 30 µL of dissolution buffer (5 M TEAB, 0.1% SDS). iTRAQ reagent (1 unit in ethanol) was added to tryptic digests and the mixture was incubated at room temperature for 1 h. In order to eliminate a potential tag-dependent effect, one replicate was conducted with inverted iTRAQ tags.

OFFGEL isoelectrofocusing

To perform peptide fractionation according to p*I*, the 3100 OFFGEL Fractionator and the OFFGEL Kit 3-10 (both from Agilent Technology) were used following the user protocol [31]. We have previously validated that OFFGEL is an efficient tool for fractionating peptides from complex proteomes (such as secretome) and that it is compatible with iTRAQ labeling [32]. The device was set up to separate peptides in 24 liquid fractions using a 24-cm IPG gel strip with a linear pH gradient ranging from 3 to 10. Two hundred micrograms of iTRAQ-labeled peptide mix was dried by vacuum centrifugation and suspended with a focusing buffer (1.2% ampholytes 3–10 in water (v/v)) as described previously [32]. This sample solution was loaded in each of the 24 wells placed above the IPG strip and peptides were focused with a constant current of 50 µA until the total voltage reached 50 kVh. After focusing, 50–150 µL of sample was recovered from each well,

transferred in individual tubes, and dried in a vacuum concentrator prior to LC-MALDI-MS/MS analysis.

Reversed-phase Liquid Chromatography (RP-LC)

Peptide separation was performed on an Ultimate-300 chromatography system (Dionex-LC Packings) equipped with a Probot MALDI spotting device. Each dried OFFGEL fraction was dissolved in 15 µL 0.1% TFA (v/v) and 5 µL was injected and captured onto a trap column (nano-Precolumn ™, PepMap ™, C18, 300 µm i.d. × 5 mm, LC Packings). Peptides were eluted on to an analytical column (PepMap ™, C18, 75 µm i.d. × 15 cm, LC Packings) using an automated binary gradient (300 nL/min) from 100% buffer A (2% (v/v) ACN ; 0.05% (v/v) TFA) to 55% buffer B (80% (v/v) ACN; 0.05% (v/v) TFA) over 45 min. For MALDI-MS/MS analysis, column effluent was mixed in a 1:3 ratio with MALDI matrix (2 mg/mL α-Cyano-4-hydroxy-cinamic acid) through a 25-nL mixing tee and spotted in 32 × 50 spot arrays (Opti-tof LC/MALDI Insert 123 × 81 mm plate, Applied Biosystems) at a frequency of one spot per 15 s.

MALDI MS/MS analysis

MS and MS/MS analyses were performed using the 4800 MALDI-TOF/TOF Analyzer (Applied Biosystems/MDS Sciex). After screening all LC-MALDI sample positions in MS-positive reflector mode using 2,000 laser shots, the fragmentation of automatically selected precursors was performed at a collision energy of 1 kV using air as collision gas (pressure, ~2×10^{-6} Torr). MS spectra were acquired between *m/z* 750 and 3500. Up to 12 of the most intense ion signals per spot position having an S/N greater than 80 were selected as precursors for MS/MS acquisition.

Peptide and protein identification and iTRAQ quantification

Peptides and proteins were identified using the ProteinPilot™ Software V 2.0 (Applied Biosystems) [33]. It is based on two algorithms: Paragon™ as a database search engine, and ProGroup™ to analyze database results and identify the minimum set of defensible proteins by removing the redundancy search for each unique peptide. Each MS/MS spectrum was searched for against a database of human protein sequences (UniProtKB, release 12.0 of 24-07-07, 276 256

sequence entries). Only proteins with a ProteinPilot score greater than 4 (99% confidence and two distinct peptides matched) were considered positively identified. To determine levels of false-positive peptide identifications, MS/MS spectra were also searched against the corresponding randomized database generated from a Perl script downloaded from the Matrix Science website (www.matrixscience.com/downloads/decoy.pl.gz). As shown by Peng *et al*., the false-positive rate can be estimated using the following formula: $\%fal = 2\ [n_{rev} / (n_{rev} + n_{real})]$ with %fal being the estimated false-positive rate, and n_{rev} and n_{real} the number of unique peptides identified from the reverse and the real databases, respectively [34]. For each MS/MS spectrum acquired, signature-ion peak areas at 114.1, 115.1, 116.1, and 117.1 *m/z* were used by the software for relative quantification. The software performed an automatic experimental bias correction that accounted for unequal mixing during the combination of the different labeled samples. For each iTRAQ protein ratio, the software calculates a *p*-value (between 0 and 1) that reflects the iTRAQ ratio variation between the different peptides matched with the same protein. Only proteins with a *p*-value protein ratio less than 0.05 were considered to have a real change of expression.

Two-dimensional gel electrophoresis (2-DE)

Protein separation by two-dimensional gel electrophoresis

The proteins from conditioned medium of H358/TetOn/p53 cells induced or not induced for wt-p53 were analyzed in triplicate using 2-DE as described elsewhere [35]. Isoelectric focusing was performed with the Ettan IPGphor II system (GE Biosciences) according to the manufacturer's instructions. One hundred micrograms of secreted proteins was solubilized in 350 µL IEF sample buffer (9 M urea, 65 mM DTT, 60 mM CHAPS, 5% Orange G (v/v)). Prior to the IEF step, ampholines 3-10 (if IPG strip pH 3–10 was used) or ampholines 4-7 (if IPG strip pH 4–7 was used) were added to a final concentration of 2% (v/v). The solubilized proteins were loaded onto a dehydrated linear IPG strip, pH 3–10 or pH 4–7 (GE Biosciences). Focusing was performed at 20°C, 50 µA/strip, as follows: 10–12 h passive rehydration, 1 h at 500 V, 2-h gradient 500–1,000 V, 3-h gradient 1,000–8,000 V, 1.5 h at 8,000 V. After focusing, the strips were incubated with 52 mM

DTT in equilibration buffer (6 M urea, 50 mM Tris-HCl, 2.5% SDS, 50% glycerol, pH 8.8) for 20 min to reduce disulfide bonds of the proteins. Proteins were then carbamidomethylated for a further 20 min (220 mM iodoacetamide in equilibration buffer). The IPG strips were transferred to a 10–12.5% SDS polyacrylamide gel. SDS-PAGE was carried out vertically in a protean II chamber (Bio-Rad) according to the manufacturer's instructions, with a current of 20 mA per gel at 10°C. After electrophoresis, protein spots were visualized by the silver staining technique, as described previously [36]. The 2-DE stained gels were scanned using ImageScanner and LabScan software (v. 5.0; GE Biosciences) and the digitalized gel images were normalized and comparatively analyzed (three gels per condition) using ImageMaster software (v. 5.0; GE Biosciences).

In-gel digestion

For protein identification, silver-stained protein spots were excised from the gel and washed with destaining solution (30 mM potassium ferricyanide, 100 mM sodium thiosulfate), then washed once each with water, digestion buffer (25 mM NH_4HCO_3), and modified digestion buffer (25 mM NH_4HCO_3/ACN 1:1 (v:v)). Subsequently, gel pieces were shrunk with 100% ACN and swollen with 20 µL trypsin solution (6 ng/µL trypsin from Promega, in digestion buffer). Digestion was performed for 12–16 h at 37°C. The supernatant was then transferred into a fresh tube and the gel pieces were extracted with a cycle of 50% ACN (v/v), 0.1% TFA (v/v), and 100% ACN (v/v). Each step lasted 20 min. After that, gel pieces were sonicated and the supernatant was pooled with the supernatant of the digestion. These extracted peptides were dried with vacuum centrifugation and resuspended in 10 µL 0.1% TFA (v/v). Finally, extracted peptides were desalted with C18 reverse-phase Zip-Tip (Millipore) according to the manufacturer's instructions. Peptides were eluted with 3 µL of 50% ACN, 0.1% TFA (v/v), mixed at a 1:1 ratio with MALDI matrix (10 mg/mL α-cyano-4-hydroxy-cinnamic acid (CHCA) matrix in 50% ACN, 0.1% TFA (v/v)), and spotted on a stainless steel MALDI sample plate (Applied Biosystems).

Mass spectrometry analysis

MS and MS/MS analyses of tryptic peptides were performed using the 4800 MALDI-TOF/TOF Analyzer (Applied Biosystems). MS spectra were acquired in positive reflector ion mode using 4000 Series Explorer software (V. 3.5). MS spectra were acquired between *m/z* 750 and 4000 using 2000 laser shots. To look for less abundant proteins, the ten least intense peaks in each MS spectrum above an S/N threshold of 80 were selected for MS/MS analysis. Fragmentation of automatically selected precursors was performed at a collision energy of 1 kV using air as collision gas (pressure of ~1×10^{-6} Torr). For peak detection, 5-point Gaussian smoothing and an S/N of 20 were applied using cluster area S/N optimization.

Protein identification

Protein identification was performed by peptide mass fingerprint (PMF) and/or specific MS/MS peptide sequencing. The combined MS and MS/MS spectra from each spot were processed using GPS Explorer V3.6 (Applied Biosystems) with Mascot (Matrix Science) as the database search engine (V 2.0). Each MS and MS/MS spectrum was searched against a database of human protein sequences (UniProtKB, release 12.0 of 24-07-07, 276.256 sequence entries), resulting in a set of tryptic peptides matched with confidence values. A search was also performed against a database of mammalian protein sequences to eliminate possible residual calf serum contaminants. The Mascot searches were run using the following parameters: methionine oxidation, cysteine carbamidomethylation as variable; one missed cleavage allowed; precursor tolerance at 50 ppm, MS/MS fragment tolerance set to 0.2 Da, and charge set to +1. Identification was assigned to a protein spot feature if the protein score was calculated to be greater than 60, correlating to a confidence interval greater than 95%.

Western blot analysis

The conditioned media or total protein extracts (2–20 µg) were subjected to standard SDS-PAGE and immunoblot analysis. The primary antibodies used were: anti-kallikrein-6 KLK6 (sc-20624; 1:500), anti-NPC2 (sc-30346; 1:100), anti-PIG3 (sc-16327; 1:100) (all from Santa Cruz

Biotechnologies), anti-FGF19 (ab25728; 1:1000), anti-TPA (ab28219; 1:1000), anti-VEGF (ab9544; 1:100), anti-GGH (ab37802; 1:1500), anti-clusterin (ab39991; 1:1000), anti-IL8 (ab16223; 1:2500), anti-IGFII (ab9574; 1:2000), anti-PLTP (ab18990; 1:250), anti-GDF15 (ab39999; 1:2500), anti-IGFBP3 (ab36529; 1:2500) (all from Abcam), anti-FGFBP1 (AF1593; R&D Systems; 1:1000), anti-p21 (OP64; Calbiochem; 1:250), and anti-p53 (DO-7; DakoCytomation; 1:1000). Anti-actin (ab1801; Abcam; 1:1000), anti-alpha-enolase ENO1 (ab49191; Abcam; 1:750), α-SOD1 (sc-11407; Santa Cruz Biotechnologies) antibodies were used as loading controls. The secondary antibodies used were: ECL anti-mouse igG horseradish peroxidase (HRP)-conjugated (NA9310; GE Biosciences; 1:7500), anti-rabbit igG HRP-conjugated (NA9340; GE Biosciences; 1:7500), anti-goat igG HRP-conjugated (A9452; Sigma; 1:7500) and anti-chicken igG HRP-conjugated (62-3120; Zymed; 1:3000). Western blot analyses were performed using a Hybond ECL membrane (GE Biosciences) and profiles were detected using the ECL detection system (GE Biosciences) according to standard procedures.

H358/TetOn/p53 xenografted mouse model

Following the guidelines of the Institutional Animal Care and Use Committee, the animals were regularly monitored for signs of illness or suffering (lack of normal grooming and avoidance behaviors). Human lung adenocarcinoma H358/TetOn/p53 cells were harvested from culture, and 20×10^6 cells (in 200 µL sterile PBS) were injected subcutaneously in 6-week-old female NMRI nude mice (Janvier breeding, Le Genest Saint Isle, France) (s = 18 mice). Eight weeks later, when tumor diameters ranged from 5 to 10 mm, the mice were randomized into two groups (n=9) according to tumor volume. Three control mice (without tumor xenograft) were added to each group. As previously described, one group of mice was exposed to doxycycline (2 mg/mL) added to the drinking water for 3 days; the other group was untreated during the same period [28].

Three days after the beginning of treatment, the mice were injected intravenously with heparin (50 UI, 10µL), blood was collected by intra cardiac sampling and was centrifuged at 500 g, 4°C for 20

min. Plasma was sampled and aliquots were stored at −80°C. Mice were sacrificed by cervical dislocation and tumors were excised, weighed, rinsed with PBS, and stored at −80°C.

Animals were weighed and examined daily for signs of tumor growth. The tumors were measured once a week in two perpendicular diameters. The volume was calculated as follows: $a \times b^2 \times 0.5$, where 'a' and 'b' are the largest and smallest diameters, respectively.

Quantitative assays for human GDF-15, FGF-19, and VEGF using ELISA

Proteins from the tumors were extracted in pre-chilled RIPA buffer (Mbiotech) containing a protease inhibitor cocktail (Complete, Roche) using a dispersing instrument Ultra-Turrax T8 (Ika Labortechnik) on ice, vortexed, and sonicated. The homogenate was kept on ice for 10 min and centrifuged at 12,000 *g*, 4°C for 20 min. The supernatant was transferred into a new tube and stored at −20°C. The total protein amount was determined using a standard Bradford protein assay (Thermo Scientific). For ELISA, total protein extracts were adjusted to 1 mg/mL. Plasma and protein extracts from the tumors were used for the analysis of GDF-15, FGF-19, and VEGF secretion from the human NSCLC H358/TetOn/p53 cells using the DuoSet Human GDF-15, Quantikine Human FGF-19, and VEGF Immunoassay Kits (R&D Systems) following the manufacturer's protocols. The assays are based on the quantitative sandwich enzyme immunoassay employing monoclonal antibody, specific for human GDF-15, FGF-19, and VEGF pre-coated onto a microplate for solid-phase ELISA. The protein concentrations were extrapolated from the standard curve generated using recombinant human proteins. ELISA results are presented as the mean ± SD of several biological replicates ($n \geq 6$). Statistical significance was determined using the nonparametric Mann-Whitney test.

Bioinformatics and annotation

The predicted secretion pathways of proteins were studied with the SignalP software (www.cbs.dtu.dk/services/SignalP/) for the classical secretion pathway [37, 38] and SecretomeP software (www.cbs.dtu.dk/services/SecretomeP/) for the non-classical secretion pathway [39].

Interactions between p53 and p53-influenced proteins, and the relationship between these proteins and tumorigenic processes, were obtained from a review of the literature.

Results

Effect of p53 expression in the H358/TetOn/p53 cell line

The secreted proteins modulated by p53 and involved in the tumor-stroma interactions of the lung cancer cells were studied from the p53-null human non-small cell lung adenocarcinoma H358 cell line. This cell line was stably transfected with a wild type p53 cDNA construct under the positive control of a tetracycline-dependent promoter (H358/TetOn/p53) [27, 28]. Doxycycline induces the expression of wt-p53 in the H358/TetOn/p53 cell line and the biological activity of p53 was assessed by its ability to up-regulate its transcriptional target p21 (Figure 1A). To verify that serum-free medium combined to p53 expression did not induce H358 cells apoptosis, we measured the proteolytic activation of procaspase-3 by flow cytometric analysis (Figure 1B). We did not observe any significant difference between control and doxycyclin treated cells.

P53-modulated secretome analysis by iTRAQ labeling

Tryptic digests from secretome samples of H358/TetOn/p53 cells induced or not for p53 expression were specifically labeled with iTRAQ tags. We identified 1,430 proteins with a ProteinPilot Score greater than 2 (one or more peptides matched with > 99% confidence) and 909 proteins with a score greater than 4 (two or more distinct peptides matched with > 99% confidence) in the secretome of the H358/TetOn/p53 cells (see Supplemental Table S1). The criteria used were optimal because the estimated peptide false-positive rate was 0.14%. Ninety-one proteins have a secretion level significantly influenced by wt p53 expression. Forty seven proteins have enhanced (Table 1) and 44 proteins have reduced (Table 2) accumulation in conditioned medium (CM) upon p53 induction. Replicates were used to achieve accurate statistical analysis. The ratios listed are an average between three replicates and standard deviations are rather low (SD ≤ 15%) (Tables 1 and 2).

p53-modulated secretome analysis by two-dimensional electrophoresis (2-DE)

We studied the influence of p53 on PTMs in the H358/TetOn/p53 secretome by 2-DE analysis of secreted proteins. Triplicates for each condition were used to ensure reproducibility [40]. We found an average of more than 600 protein spots on each gel; 316 spots were successfully identified for a total of 184 individual proteins. Only two proteins were not identified by the iTRAQ approach. A semi-quantitative analysis of the differential secretome between p53-deficient and wt p53 re-expressing H358/TetOn/p53 cells was performed by comparing spot volume (%V) using the ImageMaster software. An example of an accumulated protein upon p53 expression is given with Serpin B5 (Figure 2A). Although many PTM states were resolved in 2D gels, we did not identify any p53-dependent alterations of these PTMs. For instance, Kallikrein-6 glycoprotein was observed under different glycosylation states, but all these forms were underexpressed in the same way upon p53 expression (Figure 2B). We identified 15 proteins with p53-modulated expression. These proteins were also found to be modulated by p53 in the iTRAQ analysis (Tables 1, 2).

Mode of release of p53-modulated proteins in CM

The primary sequences from the secreted proteins identified in CM upon p53 induction were analyzed with SignalP and SecretomeP. Among the 47 proteins with enhanced accumulation, 34 (72.3%) were predicted to have a signal peptide specific to the classical secretion pathway or to be secreted by the non-classical pathway. Thirteen proteins (27.7%) were secreted by an unknown or exosome pathway. Among the 44 proteins with reduced accumulation, 41 (93.2%) were predicted to have a signal peptide specific to the classical secretion pathway or to be secreted by a non-classical pathway. Three proteins (6.8%) were secreted by an unknown or exosome pathway (Figure 3, Tables 1 and 2). These results are algorithm predictions. Indeed, the presence of a signal peptide is not necessarily a sufficient condition for the secretion of proteins. For instance, the proteins targeted to the plasma membrane usually have a signal peptide or a signal anchor responsible for their insertion in the membrane of the reticulum endoplasmic. Nevertheless, the plasma membrane proteins could be released in the extracellular medium, for example, by proteolysis.

Functional analysis of p53-modulated secreted proteins

We investigated the relationship between the p53-modulated secreted proteins, p53, and tumorigenic processes. Among the 47 proteins with enhanced accumulation in CM upon p53 induction, 15 were previously described to be induced by p53, such as serpin B5, insulin-like growth factor-binding protein 3 (IGF-BP3), putative quinone oxidoreductase, and growth/differentiation factor 15 (GDF-15) [41-48]. On the other hand, among the 44 proteins with reduced accumulation in CM upon p53 induction, 12 were previously described to be repressed by p53, such as tissue and urokinase plasminogen activator, vascular endothelial growth factor (VEGF), and osteopontin [49-52] (Tables 1 and 2).

The relationship between p53-modulated secreted proteins and tumorigenic processes has been also studied. Among the 47 proteins with enhanced accumulation in CM upon p53 induction, 21 were previously described to be implicated in apoptosis and/or cell cycle arrest, three in inhibition of angiogenesis, and nine in inhibition of metastasis/invasion processes (Figure 4, Supplemental Table S2). Conversely, among the 44 proteins with reduced accumulation in CM upon p53 induction, 16 are implicated in cell proliferation and survival, 12 in enhanced angiogenesis, and 26 in enhanced metastasis/invasion processes (Figure 4, Supplemental Table S2).

Validation of proteomic results by Western blot analysis

To validate our results, 16 identified proteins were analyzed by Western blot (Figure 5A). Alpha-enolase (ENO1) and superoxide dismutase 1 (SOD1) were used as a loading control since their levels were not found to be increased by wt p53. Four proteins, in which putative quinone oxidoreductase (PIG3) and GDF-15 were found to have enhanced accumulation in CM upon p53 induction, were effectively confirmed by Western blot. Ten proteins, in which VEGF, FGF-19, and kallikrein-6 (KLK6) were found to have reduced accumulation, were also validated by Western blot. These data show that our proteomic analysis with iTRAQ and 2-DE can be used to reliably identify differential expression of proteins released in the CM (Tables 1 and 2). We also conducted

a control experiment to confirm that the modulation of proteins released depended only on the expression of p53 and did not result from a direct effect of doxycycline (Figure 5B).

Expression levels of GDF-15, VEGF, and FGF-19 proteins in H358/TetOn/p53 tumor xenograft mice tissues and in plasma

p53 expression in H358 xenograft mice was induced or not induced by the addition of doxycycline, and the biological activity of restored p53 was assessed by its ability to up regulate its main transcriptional target $p21^{waf1\text{-}cip1}$ (Figure 6). GDF-15, VEGF, and FGF-19 proteins were measured in total tumor protein extracts using human-specific ELISA kits. For these proteins, we confirmed the influence of p53 on their expression. Indeed, GDF-15 was expressed 2.9-fold higher in total protein extracts from xenograft mice without induction of p53 (Figure 7A). In contrast, VEGF and FGF-19 had a ratio of 0.40 and 0.26, respectively (Figure 7B and 7C). Since proteins secreted by tumor cells could enter the blood stream, we also analyzed their levels in blood plasma using human-specific ELISA kits. The levels of human GDF-15 in the plasma of mice were much higher in the doxycycline-treated xenograft-bearing mice than in the non-induced xenograft-bearing mice (Figure 7D). Interestingly, control mice without H358 xenografts had no signal on ELISA, with or without doxycycline treatment. Thus, the antibodies used did not cross-react with the mice plasma proteins. The levels of human VEGF and FGF-19 in the plasma were too low to be detected and measured.

Discussion

In this study, we used a human p53-deficient non-small cell lung adenocarcinoma cell line transfected by a TetOn/p53 plasmid to allow wt p53-inducible expression. Using iTRAQ and 2-DE semi-quantitative proteomics approaches, we studied the variation of the H358/TetOn/p53 secretome in reaction to p53 expression. The identification of proteins by 2-DE and iTRAQ processes matched. Indeed, 182 out of 184 proteins identified with the 2-DE process were also found with the iTRAQ approach. With the 2-DE approach, we identified many fewer proteins than with the iTRAQ approach (184 proteins *versus* 909). This was probably due to the limitations of the 2-DE technique: the proteins with low or high p*I* and molecular weights are difficult to visualize and the sensitivity of the staining is also a limiting factor. Our analysis identified 91 secreted proteins whose extracellular level is modulated by p53, 47 being over expressed and 44 under expressed in CM in response to wt-p53 expression. The modulation of secreted proteins by p53 was validated for some targets to abolish a potential direct effect of doxycycline. Recent proteomics data obtained on this cell line in the laboratory from the analysis of different cellular compartments confirmed that doxycycline it-self does not have any effect on the expression of proteins modified by p53.

The mechanism through which p53 might regulate the secretion of proteins is currently unknown. p53 has transcription-dependent effects that could be direct or indirect, i.e., mediated by p53 itself or by one or several transcriptional targets of p53. It is also conceivable that p53 regulates specific proteins directly in a transcription-independent manner by affecting their stability or secretion. Analysis of the p53-modulated secretome provides information on both transcription-dependent and independent regulations.

Among the 91 p53-modulated secreted proteins, 15 are transcriptional targets previously shown to be induced and 12 to be repressed by p53 (see references in Tables 1 and 2). To examine whether the CM levels of the other secreted proteins identified were regulated by p53 at the gene expression

level, it would be interesting to study the differential expression of their mRNAs by microarrays or quantitative PCR in the H358 cell line. This kind of analysis has not been previously reported for the p53-dependent gene targets in the H358 cell line. Such a transcriptomic study was conducted on the 2024 glioma cell line [19]. The comparison of these transcriptomic results with the proteomics results obtained for the same cell line did not correlate and the authors concluded that p53 could act essentially by non-transcriptional regulation.

The p53-modulated secreted proteins may be regulated indirectly through transcription-independent mechanisms such as changes in protein stability and localization, or the modulation of proteolytic degradation. p53 could potentially affect protein stability and/or secretion through regulation of PTMs, for example, phosphorylation, glycosylation, acetylation, and hydroxylation, which may mark some proteins either for degradation or for localization [53]. To investigate the influence of p53 on PTMs, we used 2-DE. Nevertheless, no p53-dependent alteration of PTMs in the secretome was identified. Alternatively, p53 may be also involved in the regulation of secretory pathways. Proteins could be released to the extracellular compartment *via* three different secretion pathways: a classical secretion pathway, a non-classical secretion pathway, and an exosome pathway [21]. The classical secretion pathway is mediated by a signal peptide at the N-terminus of the protein. The non-classical pathway corresponds to the secretion of proteins without signal peptide that can be released by a flip-flop mechanism, endosomal recycling, proteolysis, through a membrane transporter, or by other unknown processes. The third secretion pathway corresponds to the release of exosomes that is to say the exocytosis of membranous vesicles. Ubiquitously expressed molecules such as enzymes of the intracellular metabolism, cytoskeletal and chaperones proteins are exosome constituents. A previous study reported 295 proteins secreted by the exosome pathway [54]. We identified 129 of them, such as alpha-enolase, heat shock 70 kDa protein, destrin, or triosephosphate isomerase. Most of secreted proteins negatively regulated by p53 were released through classical pathway (86.4%). In the other hand, 48.9% of proteins positively regulated were released through this pathway. p53 could alter the secretion rate of intracellular proteins through

either an increased release of specific proteins or through up regulation of a particular non-classical secretory pathway. Recent evidences suggest that p53 may act as a general regulator of the exosome pathway [55]. The authors demonstrated the ability of p53 to transcribe the *TSAP6* gene whose product is sufficient to induce the secretion of exosomes *in vitro*, and they considered p53 as a general regulator of this pathway. They showed, for instance, that the majority of serpin B5 reaches the extracellular space *via* exosomes. Hence, the increase in proteins secreted by an unknown or a non-classical pathway observed when p53 is expressed could be explained by a role of p53 on exosome secretion mechanisms. Moreover, some p53-modulated proteins found are membrane proteins or bound to membrane proteins, so they could be partially released in CM by proteolysis.

The tumor-suppressive protein p53 plays a central role by down regulating many tumorigenic processes such as metastasis, invasion, angiogenesis, cell survival, and cell proliferation [56]. Metastasis and invasion processes utilize similar strategies, involving changes in the cell-to-cell and cell-to-extracellular matrix (ECM) interactions, and the activation of extracellular proteases [5]. The ability of tumor cells to digest the ECM by secreting proteolytic enzymes is recognized to correlate with their tissue invasiveness [57]. The proteolytic network implicates the serine-protease system of urokinase plasminogen activator (uPA), uPA receptor (uPAR) and plasminogen, which results in the activation of latent matrix metalloproteinases (MMPs) and is mainly responsible of ECM degradation. Interestingly, this pathway seems to be implicated when p53 is not expressed in H358/TetOn/p53 cell line. Indeed, uPA, tissue plasminogen activator (tPA) tenascin and TGF- β, all implicated in the ECM degradation, are over-expressed in absence of p53. Even if we did not identify MMPs in the secretome, their general inhibitors (TIMPs 1 and 2) are over-expressed in absence of p53, probably to counterbalance the up regulation of MMPs. We identified others proteases up-regulated in secretome in absence of p53, for example kallikrein-6 that belongs to the kallikrein superfamily of serine proteases family and have been shown to play a role in ECM degradation [58]. On the other hand, two protease inhibitors (leukocyte elastase inhibitor and serpin B5) are up regulated in presence of p53. Some proteins are involved in cell adhesion and may

participate to the cell adhesion changes. For example, E-cadherin, cadherin-6 (K-cadherin), cadherin-related tumor suppressor, catenins or ephrin A1 are up-regulated, and cadherin-2 (N-cadherin) is down-regulated when p53 was expressed. The invasion processes involve also morphological changes called epithelial-mesenchymal transition (EMT) [59]. We noted some characteristic of EMT in the H358/TetOn/p53 cell line in absence of p53 such as the down regulation of E-cadherin and the up regulation of N-cadherin that promote tumor cells migration and mobility. Consequently, our results support the role of p53 in the inhibition of the metastasis/invasion processes by the modulation of proteolytic cascades and cells interactions.

p53 plays also a role in the modulation of angiogenesis process. Indeed, VEGF and TGF-β, 2 growth factors implicated in angiogenesis, are down-regulated when p53 is expressed. We also observed the down-regulation of interleukin-8 and IGF-II, two pro-angiogenic factors. At the opposite, the anti-angiogenic protein serpin-B5 is up-regulated when p53 is expressed.

p53 affects cell survival and proliferation too. Moreover, some proteins involved in tumor proliferation arrest or apoptosis are up regulated in the CM from H358/TetOn/p53 cells after wt-p53 induction (*e.g.,* 14-3-3 proteins, IGFBP3, GDF-15, and putative quinone oxidoreductase), whereas proteins increasing cell proliferation and survival are down regulated (*e.g.,* amphiregulin and interleukin-8). This cell cycle arrest by p53 is consistent with our previously published results [27].

This current proteomic analysis gives a comprehensive and global view of the p53-modulated secreted proteins correlated with p53 tumor suppressor function. Overall, many up regulated secreted proteins upon p53 expression prevent tumorigenesis. Conversely, in p53-deficient cells, we observed over secretion of proteins involved in the progression of tumorigenesis. For most of them, this is the first study in which a relationship with p53 has been shown.

Some proteins modulated by p53 are poorly described in the literature and could be interesting targets for further analysis. For instance, FGF-19 that belongs to the fibroblast growth factors family is a particular protein that binds only one receptor, the FGF receptor 4 [60]. Desnoyers *and al.* have recently demonstrated that this protein is over-expressed in many tumors, in which lung

tumors and that it is implicated in tumor progression. The FGFs are closely regulated by the chaperones FGF-binding proteins (FGF-BP) [61] that are often up regulated in cancers [62, 63]. We found that FGF-BP1 is greatly down-regulated in presence of p53. This protein is poorly studied and could play important roles, especially by an interaction with FGF-19. A last example is the ZIP10 zinc transporter. Studies have demonstrated the role of zinc flow in cancer progression. ZIP10 has been poorly described but Kagara *and al.* found that it could be implicated in breast tumor progression [64]. Further studies on these proteins could be consequently very promising.

The influence of p53 on the secretome was demonstrated *in vivo* with a xenograft mouse model for three proteins: GDF-15, VEGF, and FGF-19. Moreover, we noted that the proteins modulation by p53 were very similar with the iTRAQ ratios obtained *in vitro*. GDF-15 has a ratio of 2.9 (*vs* 2.8 using iTRAQ), VEGF of 0.40 (*vs* 0.34), and FGF-19 of 0.26 (*vs* 0.30). As tumor-secreted proteins could enter the blood stream, we also analyzed their levels in blood plasma. Interestingly, the levels of GDF-15 expression in blood plasma correlated with their *in vitro* expression. FGF-19 and VEGF concentrations were below detection levels. This is the first report to establish a connection between the *in vitro* cell line secretome and the presence of these proteins *in vivo* in the plasma.

This study may be helpful for the discovery of new potential lung cancer markers or therapeutic targets. The discovery of such molecules directly in blood is difficult because of their very low concentrations. We showed here that secretome obtained *in vitro* may reflect the release of proteins from tumors *in vivo*, as suggested by our xenograft mouse model. Consequently, such a differential "secretomics" approach could be used as an alternative methodology for identifying putative tumor markers.

In conclusion, we developed an original analytical process efficient for the differential study of the p53-influenced secretome. This approach allowed us to identify the most complete secretome list to date, in which 91 are p53-modulated secreted proteins. This study provides more informations on the influence of p53 in the modulation of cell-to-cell and cell-to-stroma interactions by the

differential release of proteins. Finally, we demonstrated for the first time that proteins of tumor cells secretome could be released into the blood stream.

Acknowledgements

The medical proteomic platform of the Center for Innovation in Biology (CIB) receives financial support from the Lyon-Auvergne-Rhône-Alpes Canceropole (CLARA), Grenoble University Hospital, INSERM and the French Research Ministry. This work was supported by a doctoral studentship from the French Research Ministry to J. Chenau. This work was supported by a Grant from the INCA (French National Institute of Cancer) free projects 2006 PL06_025, "Analysis of secreted and targeted proteins from non-small cell lung cancers for the identification of new seric biomarkers". We thank Nadège Mondello, Jacqueline Méo, Valérie Konik-Mathevet and Corine Tenaud for technical assistance.

This information is available free of charge via the Internet at http://pubs.acs.org.

Conflict of interest

The authors declare no conflict of interest

References

1. Jemal, A.; Siegel, R.; Ward, E.; Hao, Y.; Xu, J.; Murray, T.; Thun, M. J., Cancer statistics, 2008. *CA Cancer J Clin* **2008,** 58, (2), 71-96.
2. Smith, R. A.; Glynn, T. J., Epidemiology of lung cancer. *Radiol Clin North Am* **2000,** 38, (3), 453-70.
3. Gridelli, C.; Rossi, A.; Maione, P., Treatment of non-small-cell lung cancer: state of the art and development of new biologic agents. *Oncogene* **2003,** 22, (42), 6629-38.
4. Hoffman, P. C.; Mauer, A. M.; Vokes, E. E., Lung cancer. *Lancet* **2000,** 355, (9202), 479-85.
5. Hanahan, D.; Weinberg, R. A., The hallmarks of cancer. *Cell* **2000,** 100, (1), 57-70.
6. Guimaraes, D. P.; Hainaut, P., TP53: a key gene in human cancer. *Biochimie* **2002,** 84, (1), 83-93.
7. Robles, A. I.; Linke, S. P.; Harris, C. C., The p53 network in lung carcinogenesis. *Oncogene* **2002,** 21, (45), 6898-907.
8. Szymanska, K.; Hainaut, P., TP53 and mutations in human cancer. *Acta Biochim Pol* **2003,** 50, (1), 231-8.
9. Levine, A. J.; Momand, J.; Finlay, C. A., The p53 tumour suppressor gene. *Nature* **1991,** 351, (6326), 453-6.
10. Lewis, P. D.; Parry, J. M., In silico p53 mutation hotspots in lung cancer. *Carcinogenesis* **2004,** 25, (7), 1099-107.
11. Brambilla, E.; Negoescu, A.; Gazzeri, S.; Lantuejoul, S.; Moro, D.; Brambilla, C.; Coll, J. L., Apoptosis-related factors p53, Bcl2, and Bax in neuroendocrine lung tumors. *Am J Pathol* **1996,** 149, (6), 1941-52.
12. Donehower, L. A.; Harvey, M.; Slagle, B. L.; McArthur, M. J.; Montgomery, C. A., Jr.; Butel, J. S.; Bradley, A., Mice deficient for p53 are developmentally normal but susceptible to spontaneous tumours. *Nature* **1992,** 356, (6366), 215-21.
13. Coll, J. L.; Negoescu, A.; Louis, N.; Sachs, L.; Tenaud, C.; Girardot, V.; Demeinex, B.; Brambilla, E.; Brambilla, C.; Favrot, M., Antitumor activity of bax and p53 naked gene transfer in lung cancer: in vitro and in vivo analysis. *Hum Gene Ther* **1998,** 9, (14), 2063-74.
14. Zigrino, P.; Loffek, S.; Mauch, C., Tumor-stroma interactions: their role in the control of tumor cell invasion. *Biochimie* **2005,** 87, (3-4), 321-8.
15. Van Meir, E. G.; Roemer, K.; Diserens, A. C.; Kikuchi, T.; Rempel, S. A.; Haas, M.; Huang, H. J.; Friedmann, T.; de Tribolet, N.; Cavenee, W. K., Single cell monitoring of growth arrest and morphological changes induced by transfer of wild-type p53 alleles to glioblastoma cells. *Proc Natl Acad Sci U S A* **1995,** 92, (4), 1008-12.
16. Gu, S.; Liu, Z.; Pan, S.; Jiang, Z.; Lu, H.; Amit, O.; Bradbury, E. M.; Hu, C. A.; Chen, X., Global investigation of p53-induced apoptosis through quantitative proteomic profiling using comparative amino acid-coded tagging. *Mol Cell Proteomics* **2004,** 3, (10), 998-1008.
17. Sun, Y., p53 and its downstream proteins as molecular targets of cancer. *Mol Carcinog* **2006,** 45, (6), 409-15.
18. Rahman-Roblick, R.; Roblick, U. J.; Hellman, U.; Conrotto, P.; Liu, T.; Becker, S.; Hirschberg, D.; Jornvall, H.; Auer, G.; Wiman, K. G., p53 targets identified by protein expression profiling. *Proc Natl Acad Sci U S A* **2007,** 104, (13), 5401-6.
19. Khwaja, F. W.; Svoboda, P.; Reed, M.; Pohl, J.; Pyrzynska, B.; Van Meir, E. G., Proteomic identification of the wt-p53-regulated tumor cell secretome. *Oncogene* **2006**.
20. Hathout, Y., Approaches to the study of the cell secretome. *Expert Rev Proteomics* **2007,** 4, (2), 239-48.
21. Volmer, M. W.; Stuhler, K.; Zapatka, M.; Schoneck, A.; Klein-Scory, S.; Schmiegel, W.; Meyer, H. E.; Schwarte-Waldhoff, I., Differential proteome analysis of conditioned media to detect Smad4 regulated secreted biomarkers in colon cancer. *Proteomics* **2005,** 5, (10), 2587-601.
22. Xue, H.; Lu, B.; Lai, M., The cancer secretome: a reservoir of biomarkers. *J Transl Med* **2008,** 6, 52.

23. Gronborg, M.; Kristiansen, T. Z.; Iwahori, A.; Chang, R.; Reddy, R.; Sato, N.; Molina, H.; Jensen, O. N.; Hruban, R. H.; Goggins, M. G.; Maitra, A.; Pandey, A., Biomarker discovery from pancreatic cancer secretome using a differential proteomic approach. *Mol Cell Proteomics* **2006,** 5, (1), 157-71.
24. Zhong, L.; Roybal, J.; Chaerkady, R.; Zhang, W.; Choi, K.; Alvarez, C. A.; Tran, H.; Creighton, C. J.; Yan, S.; Strieter, R. M.; Pandey, A.; Kurie, J. M., Identification of secreted proteins that mediate cell-cell interactions in an in vitro model of the lung cancer microenvironment. *Cancer Res* **2008,** 68, (17), 7237-45.
25. Martin, D. B.; Gifford, D. R.; Wright, M. E.; Keller, A.; Yi, E.; Goodlett, D. R.; Aebersold, R.; Nelson, P. S., Quantitative proteomic analysis of proteins released by neoplastic prostate epithelium. *Cancer Res* **2004,** 64, (1), 347-55.
26. Takahashi, T.; Carbone, D.; Takahashi, T.; Nau, M. M.; Hida, T.; Linnoila, I.; Ueda, R.; Minna, J. D., Wild-type but not mutant p53 suppresses the growth of human lung cancer cells bearing multiple genetic lesions. *Cancer Res* **1992,** 52, (8), 2340-3.
27. Dubrez, L.; Coll, J. L.; Hurbin, A.; Solary, E.; Favrot, M. C., Caffeine sensitizes human H358 cell line to p53-mediated apoptosis by inducing mitochondrial translocation and conformational change of BAX protein. *J Biol Chem* **2001,** 276, (42), 38980-7.
28. Dubrez, L.; Coll, J. L.; Hurbin, A.; de Fraipont, F.; Lantejoul, S.; Favrot, M. C., Cell cycle arrest is sufficient for p53-mediated tumor regression. *Gene Ther* **2001,** 8, (22), 1705-12.
29. Hurbin, A.; Dubrez, L.; Coll, J. L.; Favrot, M. C., Inhibition of apoptosis by amphiregulin via an insulin-like growth factor-1 receptor-dependent pathway in non-small cell lung cancer cell lines. *J Biol Chem* **2002,** 277, (51), 49127-33.
30. 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.; Pappin, D. J., Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. *Mol Cell Proteomics* **2004,** 3, (12), 1154-69.
31. Horth, P.; Miller, C. A.; Preckel, T.; Wenz, C., Efficient fractionation and improved protein identification by peptide OFFGEL electrophoresis. *Mol Cell Proteomics* **2006,** 5, (10), 1968-74.
32. Chenau, J.; Michelland, S.; Sidibe, J.; Seve, M., Peptides OFFGEL electrophoresis: a suitable pre-analytical step for complex eukaryotic samples fractionation compatible with quantitative iTRAQ labeling. *Proteome Sci* **2008,** 6, 9.
33. Shilov, I. V.; Seymour, S. L.; Patel, A. A.; Loboda, A.; Tang, W. H.; Keating, S. P.; Hunter, C. L.; Nuwaysir, L. M.; Schaeffer, D. A., The Paragon Algorithm, a next generation search engine that uses sequence temperature values and feature probabilities to identify peptides from tandem mass spectra. *Mol Cell Proteomics* **2007,** 6, (9), 1638-55.
34. Peng, J.; Elias, J. E.; Thoreen, C. C.; Licklider, L. J.; Gygi, S. P., Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS) for large-scale protein analysis: the yeast proteome. *J Proteome Res* **2003,** 2, (1), 43-50.
35. Goldman, D.; Merril, C. R.; Ebert, M. H., Two-dimensional gel electrophoresis of cerebrospinal fluid proteins. *Clin Chem* **1980,** 26, (9), 1317-22.
36. Shevchenko, A.; Wilm, M.; Vorm, O.; Mann, M., Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Anal Chem* **1996,** 68, (5), 850-8.
37. Bendtsen, J. D.; Nielsen, H.; von Heijne, G.; Brunak, S., Improved prediction of signal peptides: SignalP 3.0. *J Mol Biol* **2004,** 340, (4), 783-95.
38. Nielsen, H.; Engelbrecht, J.; Brunak, S.; von Heijne, G., Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. *Protein Eng* **1997,** 10, (1), 1-6.
39. Bendtsen, J. D.; Jensen, L. J.; Blom, N.; Von Heijne, G.; Brunak, S., Feature-based prediction of non-classical and leaderless protein secretion. *Protein Eng Des Sel* **2004,** 17, (4), 349-56.
40. Gorg, A.; Weiss, W.; Dunn, M. J., Current two-dimensional electrophoresis technology for proteomics. *Proteomics* **2004,** 4, (12), 3665-85.

41. Flatt, P. M.; Polyak, K.; Tang, L. J.; Scatena, C. D.; Westfall, M. D.; Rubinstein, L. A.; Yu, J.; Kinzler, K. W.; Vogelstein, B.; Hill, D. E.; Pietenpol, J. A., p53-dependent expression of PIG3 during proliferation, genotoxic stress, and reversible growth arrest. *Cancer Lett* **2000,** 156, (1), 63-72.
42. Contente, A.; Dittmer, A.; Koch, M. C.; Roth, J.; Dobbelstein, M., A polymorphic microsatellite that mediates induction of PIG3 by p53. *Nat Genet* **2002,** 30, (3), 315-20.
43. Wong, J.; Li, P. X.; Klamut, H. J., A novel p53 transcriptional repressor element (p53TRE) and the asymmetrical contribution of two p53 binding sites modulate the response of the placental transforming growth factor-beta promoter to p53. *J Biol Chem* **2002,** 277, (29), 26699-707.
44. Yang, H.; Filipovic, Z.; Brown, D.; Breit, S. N.; Vassilev, L. T., Macrophage inhibitory cytokine-1: a novel biomarker for p53 pathway activation. *Mol Cancer Ther* **2003,** 2, (10), 1023-9.
45. Buckbinder, L.; Talbott, R.; Velasco-Miguel, S.; Takenaka, I.; Faha, B.; Seizinger, B. R.; Kley, N., Induction of the growth inhibitor IGF-binding protein 3 by p53. *Nature* **1995,** 377, (6550), 646-9.
46. Grimberg, A.; Liu, B.; Bannerman, P.; El-Deiry, W. S.; Cohen, P., IGFBP-3 mediates p53-induced apoptosis during serum starvation. *Int J Oncol* **2002,** 21, (2), 327-35.
47. Spiesbach, K.; Tannapfel, A.; Mossner, J.; Engeland, K., TAp63gamma can substitute for p53 in inducing expression of the maspin tumor suppressor. *Int J Cancer* **2005,** 114, (4), 555-62.
48. Ito, R.; Nakayama, H.; Yoshida, K.; Oda, N.; Yasui, W., Loss of maspin expression is associated with development and progression of gastric carcinoma with p53 abnormality. *Oncol Rep* **2004,** 12, (5), 985-90.
49. Kunz, C.; Pebler, S.; Otte, J.; von der Ahe, D., Differential regulation of plasminogen activator and inhibitor gene transcription by the tumor suppressor p53. *Nucleic Acids Res* **1995,** 23, (18), 3710-7.
50. Shetty, S.; Gyetko, M. R.; Mazar, A. P., Induction of p53 by urokinase in lung epithelial cells. *J Biol Chem* **2005,** 280, (30), 28133-41.
51. Lee, J. S.; Kim, H. S.; Park, J. T.; Lee, M. C.; Park, C. S., Expression of vascular endothelial growth factor in the progression of cervical neoplasia and its relation to angiogenesis and p53 status. *Anal Quant Cytol Histol* **2003,** 25, (6), 303-11.
52. Tataria, M.; Quarto, N.; Longaker, M. T.; Sylvester, K. G., Absence of the p53 tumor suppressor gene promotes osteogenesis in mesenchymal stem cells. *J Pediatr Surg* **2006,** 41, (4), 624-32; discussion 624-32.
53. Kamemura, K.; Hart, G. W., Dynamic interplay between O-glycosylation and O-phosphorylation of nucleocytoplasmic proteins: a new paradigm for metabolic control of signal transduction and transcription. *Prog Nucleic Acid Res Mol Biol* **2003,** 73, 107-36.
54. Pisitkun, T.; Shen, R. F.; Knepper, M. A., Identification and proteomic profiling of exosomes in human urine. *Proc Natl Acad Sci U S A* **2004,** 101, (36), 13368-73.
55. Yu, X.; Harris, S. L.; Levine, A. J., The regulation of exosome secretion: a novel function of the p53 protein. *Cancer Res* **2006,** 66, (9), 4795-801.
56. Jin, S.; Levine, A. J., The p53 functional circuit. *J Cell Sci* **2001,** 114, (Pt 23), 4139-40.
57. Rao, J. S., Molecular mechanisms of glioma invasiveness: the role of proteases. *Nat Rev Cancer* **2003,** 3, (7), 489-501.
58. Borgono, C. A.; Diamandis, E. P., The emerging roles of human tissue kallikreins in cancer. *Nat Rev Cancer* **2004,** 4, (11), 876-90.
59. Guarino, M.; Rubino, B.; Ballabio, G., The role of epithelial-mesenchymal transition in cancer pathology. *Pathology* **2007,** 39, (3), 305-18.
60. Xie, M. H.; Holcomb, I.; Deuel, B.; Dowd, P.; Huang, A.; Vagts, A.; Foster, J.; Liang, J.; Brush, J.; Gu, Q.; Hillan, K.; Goddard, A.; Gurney, A. L., FGF-19, a novel fibroblast growth factor with unique specificity for FGFR4. *Cytokine* **1999,** 11, (10), 729-35.
61. Abuharbeid, S.; Czubayko, F.; Aigner, A., The fibroblast growth factor-binding protein FGF-BP. *Int J Biochem Cell Biol* **2006,** 38, (9), 1463-8.

62. Liaudet-Coopman, E. D.; Berchem, G. J.; Wellstein, A., In vivo inhibition of angiogenesis and induction of apoptosis by retinoic acid in squamous cell carcinoma. *Clin Cancer Res* **1997,** 3, (2), 179-84.
63. Czubayko, F.; Liaudet-Coopman, E. D.; Aigner, A.; Tuveson, A. T.; Berchem, G. J.; Wellstein, A., A secreted FGF-binding protein can serve as the angiogenic switch in human cancer. *Nat Med* **1997,** 3, (10), 1137-40.
64. Kagara, N.; Tanaka, N.; Noguchi, S.; Hirano, T., Zinc and its transporter ZIP10 are involved in invasive behavior of breast cancer cells. *Cancer Sci* **2007,** 98, (5), 692-7.
65. Ruiz de Almodovar, C.; Ruiz-Ruiz, C.; Rodriguez, A.; Ortiz-Ferron, G.; Redondo, J. M.; Lopez-Rivas, A., Tumor necrosis factor-related apoptosis-inducing ligand (TRAIL) decoy receptor TRAIL-R3 is up-regulated by p53 in breast tumor cells through a mechanism involving an intronic p53-binding site. *J Biol Chem* **2004,** 279, (6), 4093-101.
66. Todorovicc, V.; Chen, C. C.; Hay, N.; Lau, L. F., The matrix protein CCN1 (CYR61) induces apoptosis in fibroblasts. *J Cell Biol* **2005,** 171, (3), 559-68.
67. Yoon, H.; Liyanarachchi, S.; Wright, F. A.; Davuluri, R.; Lockman, J. C.; de la Chapelle, A.; Pellegata, N. S., Gene expression profiling of isogenic cells with different TP53 gene dosage reveals numerous genes that are affected by TP53 dosage and identifies CSPG2 as a direct target of p53. *Proc Natl Acad Sci U S A* **2002,** 99, (24), 15632-7.
68. Collavin, L.; Lazarevic, D.; Utrera, R.; Marzinotto, S.; Monte, M.; Schneider, C., wt p53 dependent expression of a membrane-associated isoform of adenylate kinase. *Oncogene* **1999,** 18, (43), 5879-88.
69. Ito, S.; Fukusato, T.; Nemoto, T.; Sekihara, H.; Seyama, Y.; Kubota, S., Coexpression of glucose transporter 1 and matrix metalloproteinase-2 in human cancers. *J Natl Cancer Inst* **2002,** 94, (14), 1080-91.
70. Jin, Y. J.; Wang, J.; Qiao, C.; Hei, T. K.; Brandt-Rauf, P. W.; Yin, Y., A novel mechanism for p53 to regulate its target gene ECK in signaling apoptosis. *Mol Cancer Res* **2006,** 4, (10), 769-78.
71. Levina, E.; Oren, M.; Ben-Ze'ev, A., Downregulation of beta-catenin by p53 involves changes in the rate of beta-catenin phosphorylation and Axin dynamics. *Oncogene* **2004,** 23, (25), 4444-53.
72. Hara, S.; Nakashima, S.; Kiyono, T.; Sawada, M.; Yoshimura, S.; Iwama, T.; Banno, Y.; Shinoda, J.; Sakai, N., p53-Independent ceramide formation in human glioma cells during gamma-radiation-induced apoptosis. *Cell Death Differ* **2004,** 11, (8), 853-61.
73. Atencio, I. A.; Ramachandra, M.; Shabram, P.; Demers, G. W., Calpain inhibitor 1 activates p53-dependent apoptosis in tumor cell lines. *Cell Growth Differ* **2000,** 11, (5), 247-53.
74. Thomas, A.; Giesler, T.; White, E., p53 mediates bcl-2 phosphorylation and apoptosis via activation of the Cdc42/JNK1 pathway. *Oncogene* **2000,** 19, (46), 5259-69.
75. Gadea, G.; Lapasset, L.; Gauthier-Rouviere, C.; Roux, P., Regulation of Cdc42-mediated morphological effects: a novel function for p53. *Embo J* **2002,** 21, (10), 2373-82.
76. Okuyama, T.; Kurata, S.; Tomimori, Y.; Fukunishi, N.; Sato, S.; Osada, M.; Tsukinoki, K.; Jin, H. F.; Yamashita, A.; Ito, M.; Kobayashi, S.; Hata, R. I.; Ikawa, Y.; Katoh, I., p63(TP63) elicits strong trans-activation of the MFG-E8/lactadherin/BA46 gene through interactions between the TA and DeltaN isoforms. *Oncogene* **2008,** 27, (3), 308-17.
77. Orre, L. M.; Pernemalm, M.; Lengqvist, J.; Lewensohn, R.; Lehtio, J., Up-regulation, modification, and translocation of S100A6 induced by exposure to ionizing radiation revealed by proteomics profiling. *Mol Cell Proteomics* **2007,** 6, (12), 2122-31.
78. Oshiro, M. M.; Futscher, B. W.; Lisberg, A.; Wozniak, R. J.; Klimecki, W. T.; Domann, F. E.; Cress, A. E., Epigenetic regulation of the cell type-specific gene 14-3-3sigma. *Neoplasia* **2005,** 7, (9), 799-808.
79. Gu, Y. M.; Jin, Y. H.; Choi, J. K.; Baek, K. H.; Yeo, C. Y.; Lee, K. Y., Protein kinase A phosphorylates and regulates dimerization of 14-3-3 epsilon. *FEBS Lett* **2006,** 580, (1), 305-10.

80. Dohn, M.; Jiang, J.; Chen, X., Receptor tyrosine kinase EphA2 is regulated by p53-family proteins and induces apoptosis. *Oncogene* **2001,** 20, (45), 6503-15.
81. Criswell, T.; Klokov, D.; Beman, M.; Lavik, J. P.; Boothman, D. A., Repression of IR-inducible clusterin expression by the p53 tumor suppressor protein. *Cancer Biol Ther* **2003,** 2, (4), 372-80.
82. Pugacheva, E. N.; Ivanov, A. V.; Kravchenko, J. E.; Kopnin, B. P.; Levine, A. J.; Chumakov, P. M., Novel gain of function activity of p53 mutants: activation of the dUTPase gene expression leading to resistance to 5-fluorouracil. *Oncogene* **2002,** 21, (30), 4595-600.
83. Rother, K.; Dengl, M.; Lorenz, J.; Tschop, K.; Kirschner, R.; Mossner, J.; Engeland, K., Gene expression of cyclin-dependent kinase subunit Cks2 is repressed by the tumor suppressor p53 but not by the related proteins p63 or p73. *FEBS Lett* **2007,** 581, (6), 1166-72.
84. Yun, J.; Chae, H. D.; Choy, H. E.; Chung, J.; Yoo, H. S.; Han, M. H.; Shin, D. Y., p53 negatively regulates cdc2 transcription via the CCAAT-binding NF-Y transcription factor. *J Biol Chem* **1999,** 274, (42), 29677-82.
85. Harada, H.; Nakagawa, K.; Saito, M.; Kohno, S.; Nagato, S.; Furukawa, K.; Kumon, Y.; Hamada, K.; Ohnishi, T., Introduction of wild-type p53 enhances thrombospondin-1 expression in human glioma cells. *Cancer Lett* **2003,** 191, (1), 109-19.
86. Peterson, E. J.; Bogler, O.; Taylor, S. M., p53-mediated repression of DNA methyltransferase 1 expression by specific DNA binding. *Cancer Res* **2003,** 63, (20), 6579-82.
87. Zhang, L.; Zhan, Q.; Zhan, S.; Kashanchi, F.; Fornace, A. J., Jr.; Seth, P.; Helman, L. J., p53 regulates human insulin-like growth factor II gene expression through active P4 promoter in rhabdomyosarcoma cells. *DNA Cell Biol* **1998,** 17, (2), 125-31.
88. Guida, T.; Salvatore, G.; Faviana, P.; Giannini, R.; Garcia-Rostan, G.; Provitera, L.; Basolo, F.; Fusco, A.; Carlomagno, F.; Santoro, M., Mitogenic effects of the up-regulation of minichromosome maintenance proteins in anaplastic thyroid carcinoma. *J Clin Endocrinol Metab* **2005,** 90, (8), 4703-9.
89. Grimberg, A.; Coleman, C. M.; Shi, Z.; Burns, T. F.; MacLachlan, T. K.; Wang, W.; El-Deiry, W. S., Insulin-like growth factor factor binding protein-2 is a novel mediator of p53 inhibition of insulin-like growth factor signaling. *Cancer Biol Ther* **2006,** 5, (10), 1408-14.
90. Sembritzki, O.; Hagel, C.; Lamszus, K.; Deppert, W.; Bohn, W., Cytoplasmic localization of wild-type p53 in glioblastomas correlates with expression of vimentin and glial fibrillary acidic protein. *Neuro Oncol* **2002,** 4, (3), 171-8.
91. Shou, J.; Ali-Osman, F.; Multani, A. S.; Pathak, S.; Fedi, P.; Srivenugopal, K. S., Human Dkk-1, a gene encoding a Wnt antagonist, responds to DNA damage and its overexpression sensitizes brain tumor cells to apoptosis following alkylation damage of DNA. *Oncogene* **2002,** 21, (6), 878-89.
92. Hidaka, H.; Ishiko, T.; Ishikawa, S.; Ikeda, O.; Mita, S.; Iwamura, T.; Chijiiwa, K.; Ogawa, M., Constitutive IL-8 expression in cancer cells is associated with mutation of p53. *J Exp Clin Cancer Res* **2005,** 24, (1), 127-33.
93. Jonson, T.; Albrechtsson, E.; Axelson, J.; Heidenblad, M.; Gorunova, L.; Johansson, B.; Hoglund, M., Altered expression of TGFB receptors and mitogenic effects of TGFB in pancreatic carcinomas. *Int J Oncol* **2001,** 19, (1), 71-81.

FIGURES LEGENDS

Figure 1: Effect of p53 expression in the H358/TetOn/p53 cell line.

(A) Western-blot was realized on total proteins extracts of p53-null H358 cell line cultivated in serum-free medium with or without doxycycline treatment during 72h. The Western blot showed wt-p53 induction following doxycycline treatment and the downstream activation of the cell cycle inhibitor p21 protein.

(B) Flow cytometry analysis of apoptosis in H358/TetON/p53 cells incubated for 72 h in the presence or absence of 1 µg/mL doxycycline in serum-free medium. H358 cells are also treated for 72 h with 1 µM staurosporine (positive control for apoptosis). Proteolytic activation of procaspase-3 is observed with PE Caspase-3 Active Apoptosis Kit of BD Pharmingen.

Figure 2: Enlargement of two-dimensional electrophoresis profiles obtained for examples of p53-modulated secreted proteins.

Proteins of the conditioned medium from H358/TetOn/p53 cells with or without doxycycline induction were analyzed by 2-DE using linear IPG strips pH 4–7 and 11% SDS-PAGE. In these examples, the protein spots indicated show **(A)** the upregulation of Serpin B5 in response to wt-p53 induction and **(B)** upregulation of the quinone oxidoreductase (black arrow) and the downregulation of kallikrein-6 (black circle) in response to wt-p53 induction.

Figure 3: Distribution of p53-modulated secreted proteins according to their secretion pathways: classical secretion, nonclassical secretion, and unknown or exosome pathways.

Secretion pathways of proteins with enhanced accumulation in conditioned medium upon p53 induction are shown on the left. Secretion pathways of proteins with reduced accumulation in conditioned medium upon p53 induction are shown on the right. The mode of secretion of the proteins was predicted by SignalP software for the classical secretion pathway and SecretomeP

software for the nonclassical secretion pathway. *: 10 of these 13 proteins secreted by an unknown pathway are membrane proteins or are bound to membrane proteins.

Figure 4: Diagram representing the relationship between p53-modulated secreted proteins and tumorigenic processes. Proteins with enhanced or reduced accumulation in the conditioned medium upon p53 induction are shown in black and gray, respectively. Proteins that promote tumorigenic processes are represented with positive values and proteins that inhibit these processes have negative values.

Figure 5: Verification of secretome protein expression by Western blotting.

(A) Western blots were performed using concentrated serum-free CM collected after 72 h from H358/TetOn/p53 cells with or without doxycycline addition. Differential expression of fibroblast growth factor-binding protein 1 (FGFBP1), fibroblast growth factor 19 (FGF19), tissue-type plasminogen activator (TPA), gamma-glutamyl hydrolase (GGH), vascular endothelial growth factor A (VEGF), clusterin (CLUS), kallikrein-6 (KLK6), insulin-like growth factor II (IGF-II), epididymal secretory protein E1 (NPC2), interleukin-8 (IL8), quinone oxidoreductase (PIG3), growth/differentiation factor 15 (GDF15), phospholipid transfer protein (PLTP), insulin-like growth factor-binding protein 3 (IGFBP3) in response to wt-p53 expression was examined. Alpha-enolase (ENO1) and superoxide dismutase 1 (SOD1) served as loading control and remained unchanged.

(B) Western blot controls for doxycycline effect were performed using concentrated serum-free CM collected after 72 h of doxycycline treatment from H358 (control for dox) and H358/TetOn/p53 cells with or without doxycycline addition. Differential expression of GDF-15, PLTP, TPA, and FGF-19 in response to wt-p53 expression was examined. ENO1 served as a loading control and remained unchanged.

Figure 6: p53 expression in the H358/TetOn/p53 tumor xenograft mice tissues

Western blots were performed using total tumor protein extracts for all the mice in each group. The Western blots showed wt-p53 induction following doxycycline treatment and the downstream activation of the cell cycle inhibitor $p21^{waf1\text{-}cip1}$ protein (example for four mice of each group).

Figure 7: Expression levels of GDF-15, VEGF, and FGF-19 proteins in H358/TetOn/p53 tumor xenograft mice tissues and in plasma using ELISA.

(A) Levels of GDF-15 protein in total tumor proteins extracts (in ng per mg of total proteins). Mean ± SD (n=9), $p < 0.001$, versus control. **(B)** Levels of GDF-15 protein in plasma samples (in pg/mL plasma). Mean ± SD (n=7 for xenograft plasma samples and n=3 for control plasma samples), $p < 0.001$, versus control. **(C)** Levels of VEGF protein in total tumor protein extracts (in ng per mg of total proteins). Mean ± SD (n=9), $p < 0.01$, versus control. **(D)** Levels of FGF-19 protein in total tumor protein extracts (in pg per mg of total proteins). Mean ± SD (n=6), $p < 0.01$, versus control. H358 p53+ and H358 p53−: tumor or plasma samples of H358/TetOn/p53 xenograft mice with or without doxycycline treatment. Control dox+ and dox−: control mice with or without doxycycline treatment.

Table 1: Secreted proteins with enhanced accumulation in culture medium upon p53 induction.
Proteins with enhanced levels in the CM in response to wt-p53 are represented with their UniProt SwissProt accession number (No.), total number of unique peptides matched (Pep.), molecular weight (MW; kDa), isoelectric point (p*I*), mode of secretion (Secr.) and known p53 interactions (p53). Only proteins with a ProteinPilot score > 2 (> 99% confidence) were considered. Listed ratios are an average of the three different ratios obtained by 2-plex and 4-plex iTRAQ experiments. The iTRAQ (+/−) ratio refers to the relative secretion of proteins under wt-p53 compared to p53-null conditions. Proteins upregulated by wt-p53 have a ratio > 1.25. The ratios are followed by standard deviation (SD) inter-replicates. The 2-DE column indicates whether the proteins were also found by 2-DE analyses and if their secretion levels were found concordant (C) or unclear (U) with iTRAQ. The WB column indicates the proteins confirmed by Western blot analyses and if their secretion levels were found concordant (C). An empty space indicates that proteins were not found or not searched for in that particular experiment. The mode of secretion of the proteins was obtained from SignalP and SecretomeP software (C: classical; NC: nonclassical; U: unknown; mb: membrane proteins). Interactions between p53 and p53-modulated secreted proteins were obtained from a review of the literature (I: p53-induced protein; R: p53-repressed protein; L: Protein linked to p53; D: protein degrading p53). Proteins with * were identified with only one peptide (99% confidence) and were kept because they were confirmed by Western blot or have a previously known interaction with p53.

Protein name	No.	Pep.	MW	p*I*	2-DE	WB	Ratio +/-	SD	Secr.	p53
Putative quinone oxidoreductase	Q53FA7	4	35.51	6.67	C	C	5.6	1.6	NC	I [41, 42]
Growth/differentiation factor 15	Q99988	6	34.17	9.79	C	C	2.8	1.1	C	I [43, 44]
Phospholipid transfer protein	P55058	13	54.74	6.53		C	2.5	0.2	C	
Cadherin-6	P55285	2	88.31	4.77			2.5	0.7	U (mb)	
Tumor necrosis factor receptor superfamily member 10C *	O14798	1	27.38	4.79			2.41	-	C (mb)	I [65]
Tissue alpha-L-fucosidase	P04066	7	53.69	6.37			2.1	0.4	C	
Syndecan-4	P31431	8	21.64	4.39			2.0	0.4	C (mb)	L [66]
Insulin-like growth factor-binding protein 3	P17936	2	31.67	9.03		C	1.8	-	C	I [45, 46]
Versican core protein	P13611	10	372.82	4.43			1.8	0.2	C	I [67]
Lysosomal alpha-mannosidase	O00754	4	113.74	6.84			1.75	0.04	C	
EH domain-containing protein 1	Q9H4M9	7	60.63	6.35			1.8	0.1	U (mb)	
Adenylate kinase isoenzyme 1	P00568	3	21.64	8.73			1.7	0.2	NC	I [68]

Solute carrier family 2, facilitated glucose transporter member 1	P11166	3	54.08	8.93		1.7	0.3	U (mb)	L [69]
Ephrin type-A receptor 2	P29317	7	108.25	5.81		1.6	0.2	C (mb)	I [70]
Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta 2	P62879	2	37.33	5.60		1.6	0.3	U (mb)	
Glypican-1	P35052	13	61.65	7.07		1.6	0.2	C (mb)	
Serpin B5	P36952	8	42.11	5.72	C	1.6	0.1	C	I [47, 48]
Junctional adhesion molecule A	Q9Y624	3	32.58	8.09		1.6	0.1	C (mb)	
Catenin alpha -1	P35221	14	100.07	5.95		1.57	0.05	U (mb)	
Catenin beta-1	P35222	14	85.50	5.53		1.54	0.01	NC	R [71]
Leukocyte elastase inhibitor	P30740	7	42.74	5.90	U	1.53	0.05	NC	
Acid ceramidase	Q13510	8	44.65	7.52		1.5	0.2	C	I [72]
Calnexin	P27824	3	67.57	4.47		1.5	0.1	C (mb)	
Lysosomal protective protein	P10619	8	54.47	6.16		1.5	0.2	C	
Plastin-3	P13797	3	70.44	5.52		1.5	0.3	U	
Catenin delta-1	O60716	19	108.77	5.86		1.50	0.06	U (mb)	
Glutathione reductase, mitochondrial	P00390	7	56.22	8.74	C	1.5	0.2	NC	
Calpain small subunit 1	P04632	3	28.32	5.05		1.5	0.2	NC (mb)	D [73]
Plectin-1	Q15149	180	531.73	5.73		1.5	0.1	U	
Protein FAM3C	Q92520	18	24.68	8.52	C	1.5	0.2	C	
Proactivator polypeptide	P07602	23	58.11	5.06		1.5	0.1	C	
Cadherin-related tumor suppressor homolog	Q14517	3	506.28	4.84		1.5	0.2	C (mb)	
MARCKS-related protein	P49006	2	19.53	4.68		1.4	0.2	U (mb)	
Brain acid soluble protein 1	P80723	11	22.69	4.64		1.4	0.1	U (mb)	
Tyrosine-protein kinase-like 7	Q13308	7	118.39	6.67		1.4	0.3	C (mb)	

Cell division control protein 42 homolog	P60953	2	21.31	5.76		1.4	0.1	NC (mb)	I [74, 75]
Endothelial protein C receptor	Q9UNN8	3	26.67	6.70		1.4	0.1	C (mb)	
Gamma-synuclein	O76070	3	13.30	4.89	U	1.4	0.2	U	
Lactadherin	Q08431	8	43.12	8.47		1.4	0.1	C (mb)	I [76]
Suppressor of tumorigenicity protein 14	Q9Y5Y6	3	94.77	6.11		1.4	0.1	NC (mb)	
Protein S100-A6	P06703	4	10.17	5.33	C	1.38	0.05	NC	I [77]
Puromycin-sensitive aminopeptidase	P55786	9	103.28	5.49		1.35	0.05	U (mb)	
14-3-3 protein sigma	P31947	2	27.77	4.68	C	1.3	0.3	NC	I [78]
Epithelial-cadherin	P12830	12	97.46	4.58		1.3	0.2	C (mb)	
14-3-3 protein zeta/delta	P63104	23	27.73	4.73	C	1.3	0.1	U (mb)	I [79]
Ephrin-A1	P20827	4	23.79	6.34		1.28	0.09	C (mb)	I [80]
NAD(P)H dehydrogenase [quinone] 1	P15559	9	30.87	8.91		1.28	0.02	NC	

Table 2: Secreted proteins with reduced accumulation in conditioned medium upon p53 induction.
See Table 1 legend. Proteins downregulated by wt-p53 have a ratio < 0.8.

Protein name	No.	Pep.	MW	p/	2-DE	WB	Ratio +/-	SD	Secr.	p53
Fibroblast growth factor-binding protein 1	Q14512	7	26.26	9.28		C	0.23	0.04	C (mb)	
Zinc transporter ZIP10	Q9ULF5	4	94.13	6.25			0.30	0.03	C (mb)	
Fibroblast growth factor 19	O95750	2	24.00	6.55		C	0.30	0.07	C	
Tissue-type plasminogen activator	P00750	8	62.92	8.14		C	0.33	0.04	C	R [49]
Gamma-glutamyl hydrolase	Q92820	12	35.96	6.66		C	0.34	0.06	C	
Vascular endothelial growth factor A	P15692	5	27.04	9.21		C	0.34	0.04	C	R [51]
Clusterin	P10909	17	52.46	5.89	C	C	0.38	0.02	C	R [81]
Kallikrein-6	Q92876	18	26.84	7.15	C	C	0.38	0.05	C	
Deoxyuridine 5'-triphosphate nucleotidohydrolase, mitochondrial	P33316	4	26.71	9.65			0.44	0.05	NC	R [82]
Beta-2-microglobulin	P61769	10	13.71	6.06	C		0.46	0.05	C	
Fibulin-2	P98095	13	126.54	4.71			0.47	0.03	C	
Cadherin-2	P19022	4	99.79	4.64	C		0.47	0.06	C (mb)	
Fibulin-1	P23142	6	77.26	5.11			0.49	0.05	C	
Cyclin-dependent kinases regulatory subunit 2	P33552	2	9.86	8.07			0.49	-	U	R [83, 84]
Osteopontin *	P10451	1	35.42	4.37			0.5	0.1	C	R [52]
Insulin-like growth factor-binding protein 4	P22692	2	27.93	6.81			0.5	0.1	C	
Laminin subunit beta-1	P07942	23	198.07	4.84			0.51	0.06	C (mb)	
Polypeptide N-acetylgalactosaminyltransferase 6	Q8NCL4	5	71.16	8.47			0.51	0.09	C (mb)	
Tenascin	P24821	38	240.87	4.79			0.51	0.04	C (mb)	
Transferrin receptor protein 1	P02786	6	84.87	6.18			0.52	0.05	C (mb)	

Gamma-interferon-inducible lysosomal thiol reductase	P13284	2	29.15	4.88			0.53	0.09	C	
Thrombospondin-1	P07996	48	129.38	4.71			0.53	0.06	C (mb)	I [85]
DNA (cytosine-5)-methyltransferase 1	P26358	2	183.17	7.99			0.54	0.07	U	R [86]
Insulin-like growth factor II *	P01344	1	20.14	9.50		C	0.56	0.04	C	R [87]
EGF-containing fibulin-like extracellular matrix protein 1	Q12805	20	54.64	4.95			0.56	0.03	C	
Epididymal secretory protein E1	P61916	12	16.56	7.56	C	C	0.56	0.06	C	
DNA replication licensing factor MCM6	Q14566	8	92.89	5.29			0.58	0.01	U	R [88]
Insulin-like growth factor-binding protein 2	P18065	18	35.14	7.48	C		0.58	0.04	C	I [89]
Vimentin	P08670	2	53.65	5.06			0.6	0.1	NC	L [90]
Amphiregulin	P15514	7	27.90	7.01			0.60	0.07	C (mb)	
Dickkopf-related protein 1	O94907	5	28.67	8.80			0.6	0.2	C	I [91]
Metalloproteinase inhibitor 2	P16035	19	24.38	7.45	C		0.61	0.06	C	
Legumain	Q99538	15	49.41	6.07			0.62	0.07	C	
Dipeptidyl-peptidase 1	P53634	26	51.84	6.53			0.62	0.07	C	
Interleukin-8	P10145	2	11.10	9.10		C	0.63	0.04	C	R [92]
Beta-1.3-galactosyl-O-glycosyl-glycoprotein beta-1.6-N-acetylglucosaminyltransferase	Q02742	5	49.79	8.65			0.63	0.07	C (mb)	
Urokinase-type plasminogen activator	P00749	34	48.53	8.78	U		0.63	0.06	C	R [49, 50]
Stanniocalcin-2	O76061	9	33.25	6.93			0.64	0.06	C	
Lysyl oxidase homolog 2	Q9Y4K0	16	86.73	5.95			0.68	0.08	C	
Transforming growth factor beta-2	P61812	2	47.75	8.82			0.7	0.1	C	R [93]
Beta-1.3-N-acetylglucosaminyltransferase lunatic fringe	Q8NES3	10	41.77	9.35			0.70	0.04	C (mb)	
Cathepsin L2	O60911	2	37.33	8.98			0.7	0.2	C	
Metalloproteinase inhibitor 1	P01033	9	23.17	8.46	U		0.74	0.05	C	
Transforming growth factor-beta-induced protein ig-h3	Q15582	50	74.68	7.62			0.75	0.08	C	

Supplemental data

Supplemental Table S1: Proteins released in conditioned medium of the H358 cell line.

Proteins in the conditioned medium of H358 cell line are listed with their UniProt accession number (Access. No.) and total number of peptides matched (Pep.). Only proteins with ProteinPilot score > 2 (one peptide matched with 99% confidence) were considered.

Supplemental Table S2: List of p53-regulated secreted proteins having a relationship with tumorigenic processes that were used in the diagram representation of Figure 5B.

(A) Secreted proteins with enhanced accumulation in culture medium upon p53 induction. **(B)** Secreted proteins with reduced accumulation in conditioned medium upon p53 induction. Tumorigenic processes: Proliferation/Survival (P/S), Angiogenesis (A) and Metastasis/Invasion (M/I). +: enhance this tumorigenic process. −: inhibit this tumorigenic process.

Figure 1

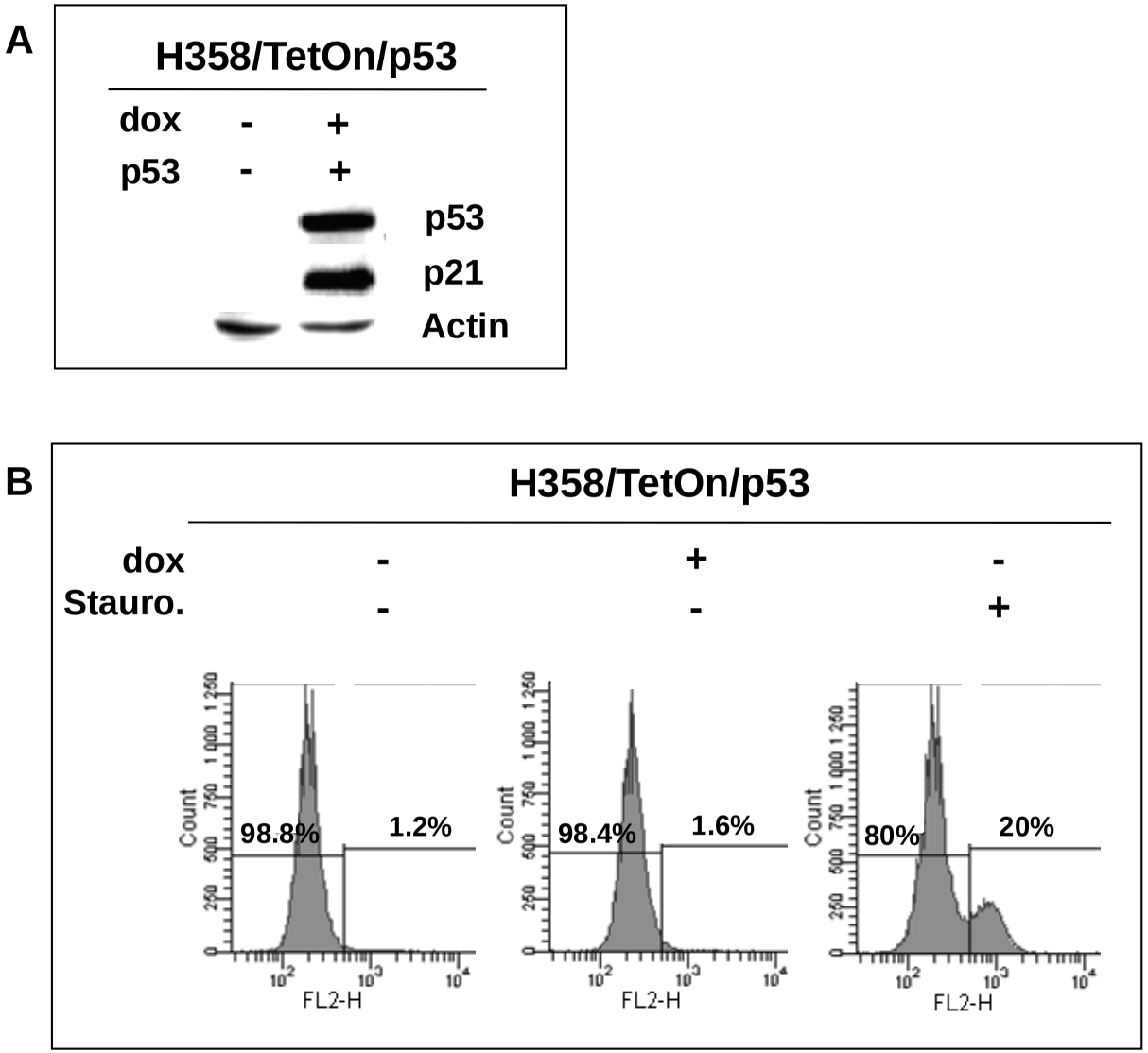

A
H358/TetOn/p53
dox - +
p53 - +
p53
p21
Actin
B
H358/TetOn/p53
dox - + -
Stauro. - - +
98.8%
1.2%
98.4%
1.6%
80%
20%
Count
FL2-H

Figure 2

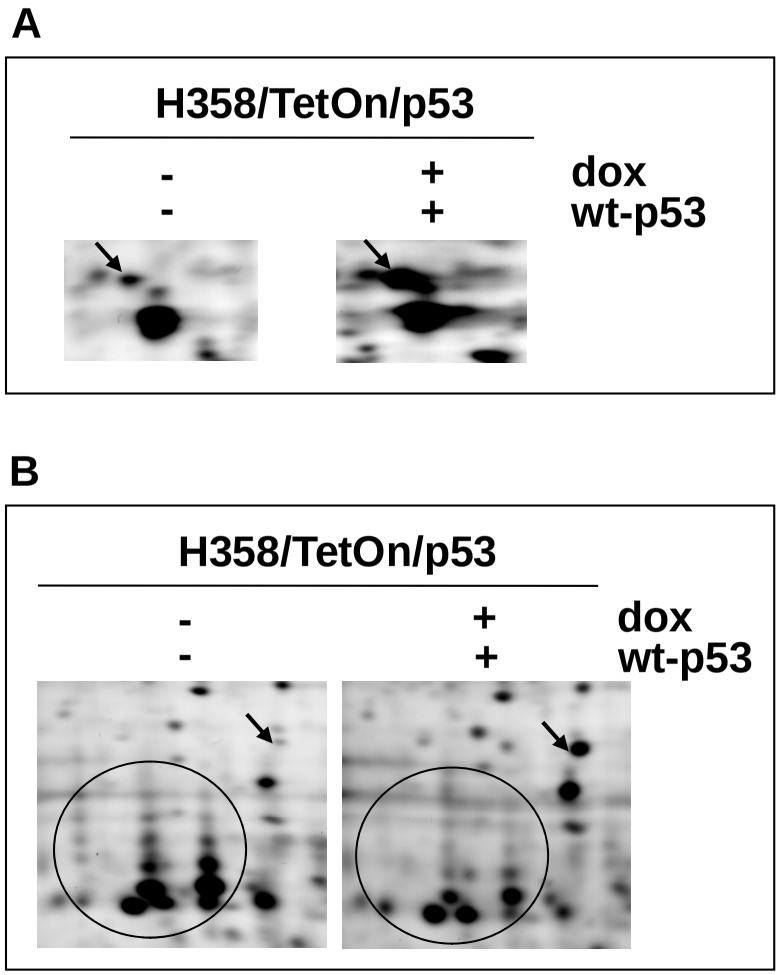


Figure 3

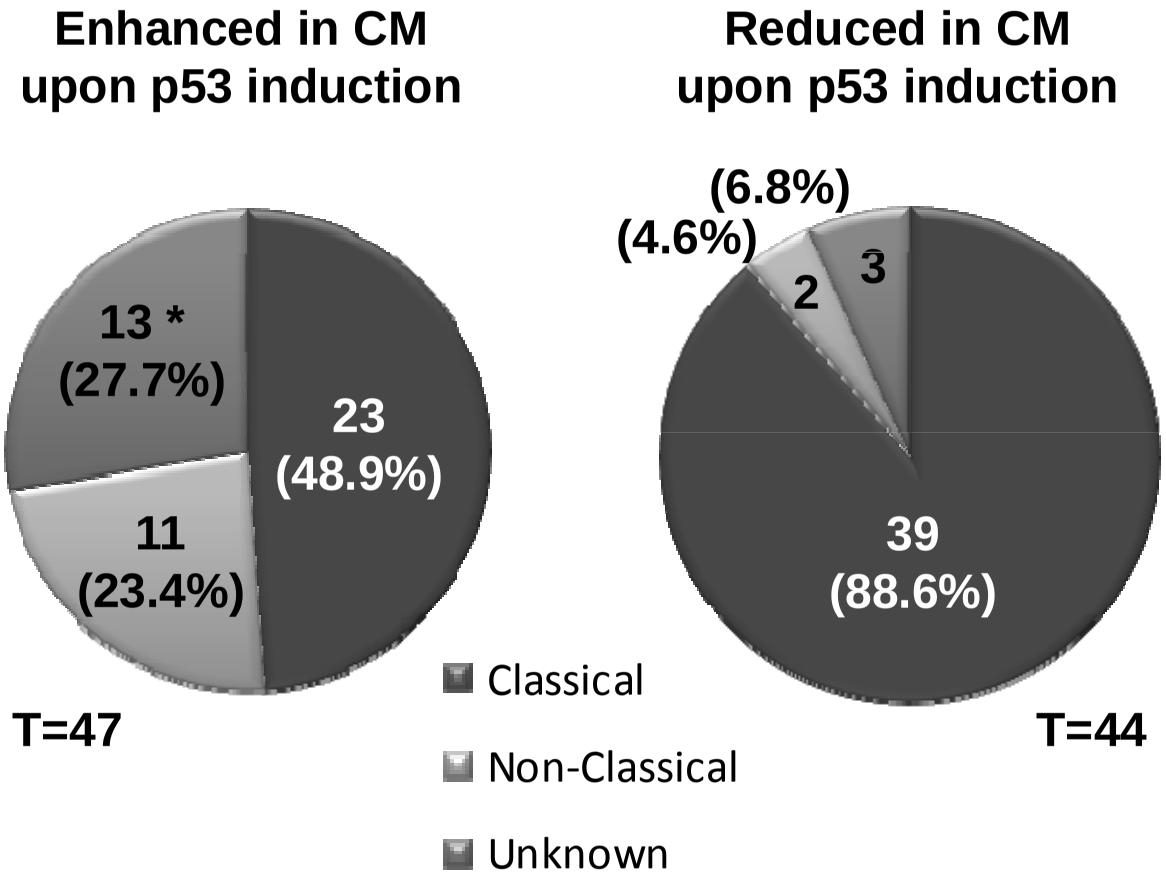

Enhanced in CM upon p53 induction
13 * (27.7%)
23 (48.9%)
11 (23.4%)
T=47
Reduced in CM upon p53 induction
(6.8%)
(4.6%)
3
2
39 (88.6%)
T=44
Classical
Non-Classical
Unknown

Figure 4

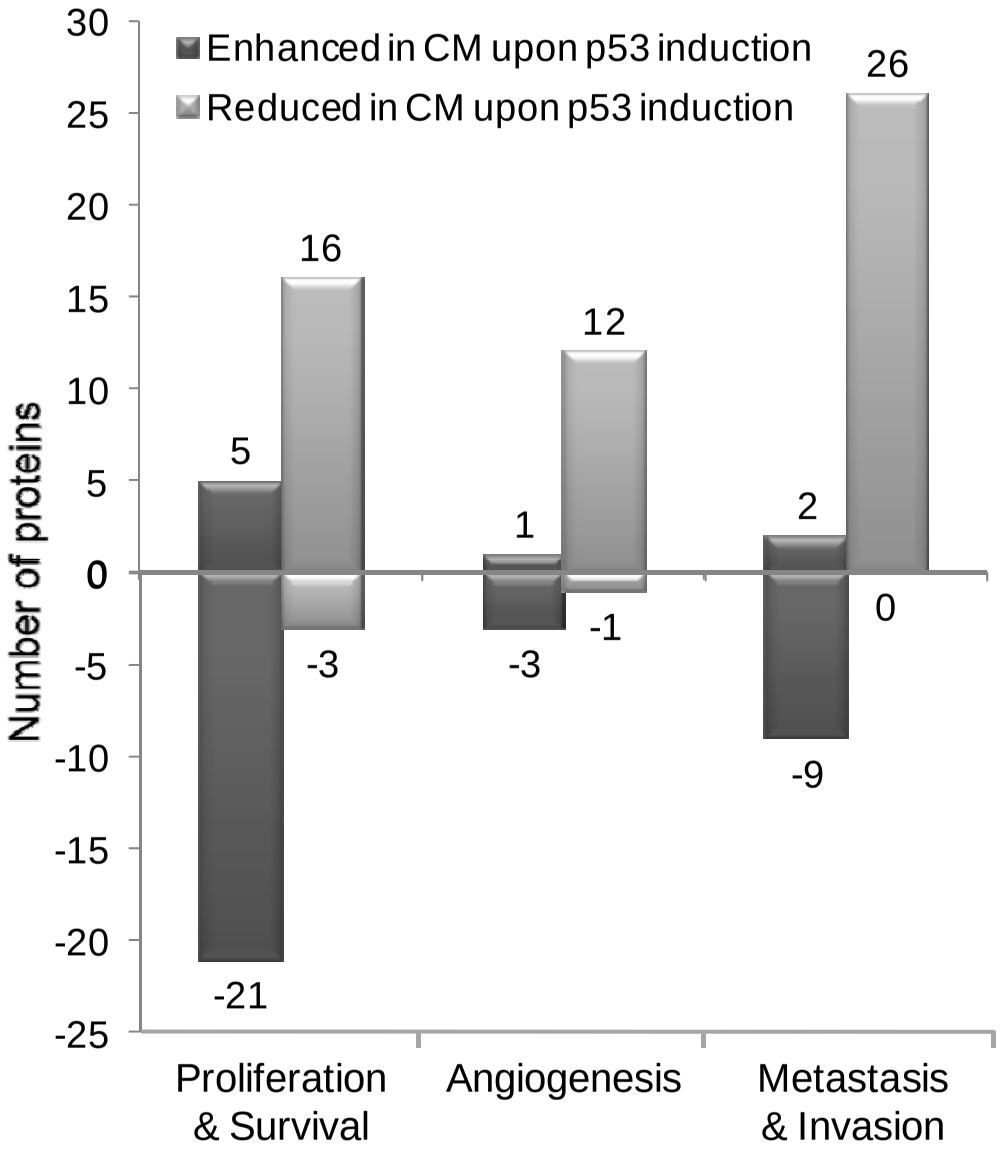

Enhanced in CM upon p53 induction
Reduced in CM upon p53 induction
Number of proteins
30
25
20
15
10
5
0
-5
-10
-15
-20
-25
5
16
-21
-3
1
12
-3
-1
2
26
-9
0
Proliferation & Survival
Angiogenesis
Metastasis & Invasion

Figure 5

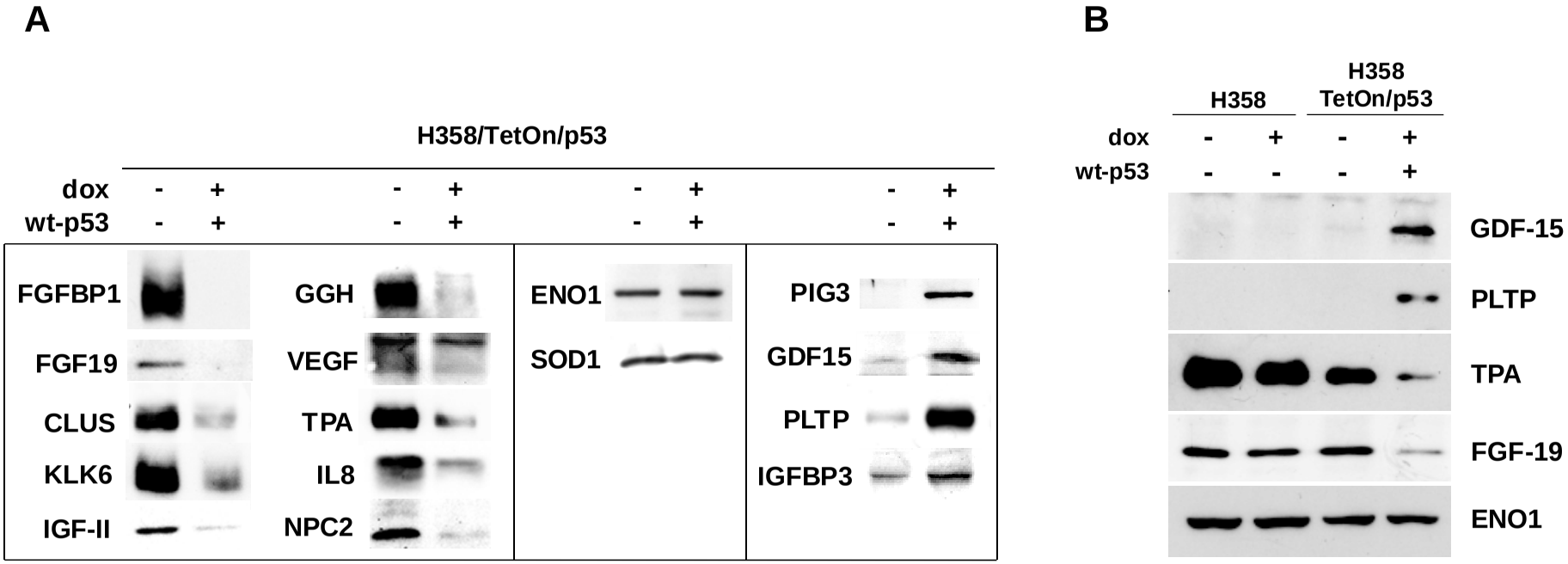

A
H358/TetOn/p53
dox - + - + - + - +
wt-p53 - + - + - + - +
FGFBP1
FGF19
CLUS
KLK6
IGF-II
GGH
VEGF
TPA
IL8
NPC2
ENO1
SOD1
PIG3
GDF15
PLTP
IGFBP3
B
H358
H358 TetOn/p53
dox - + - +
wt-p53 - - - +
GDF-15
PLTP
TPA
FGF-19
ENO1

Figure 6

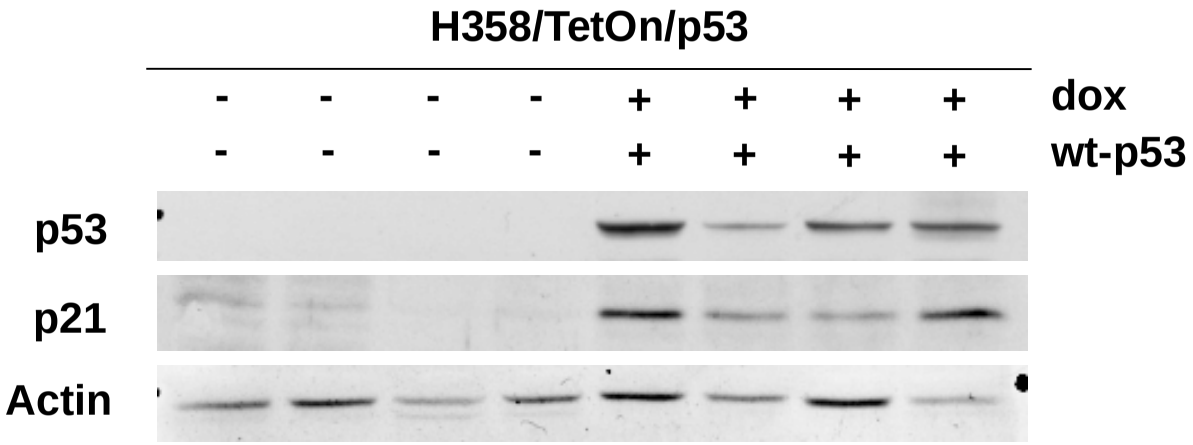


Figure 7

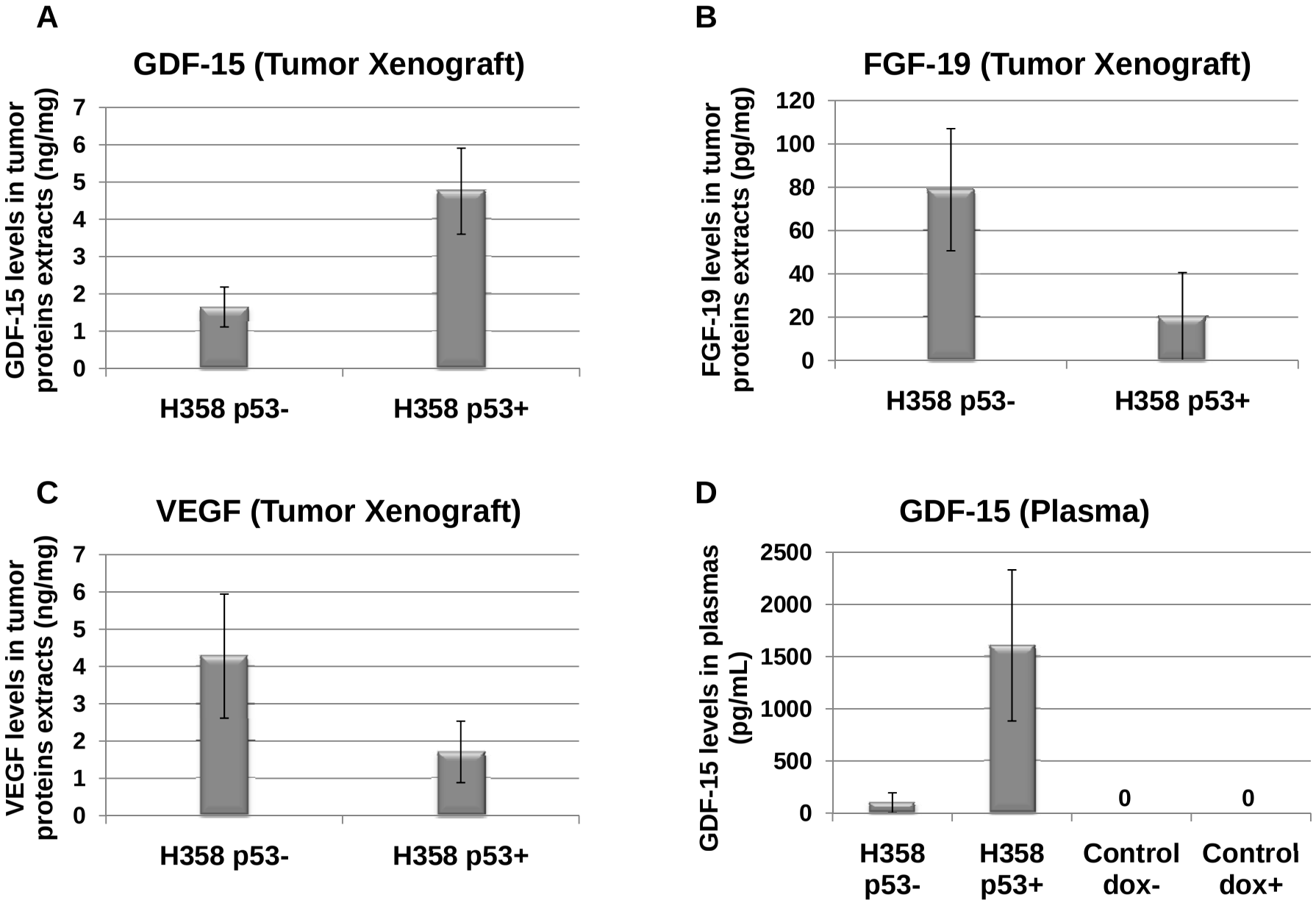
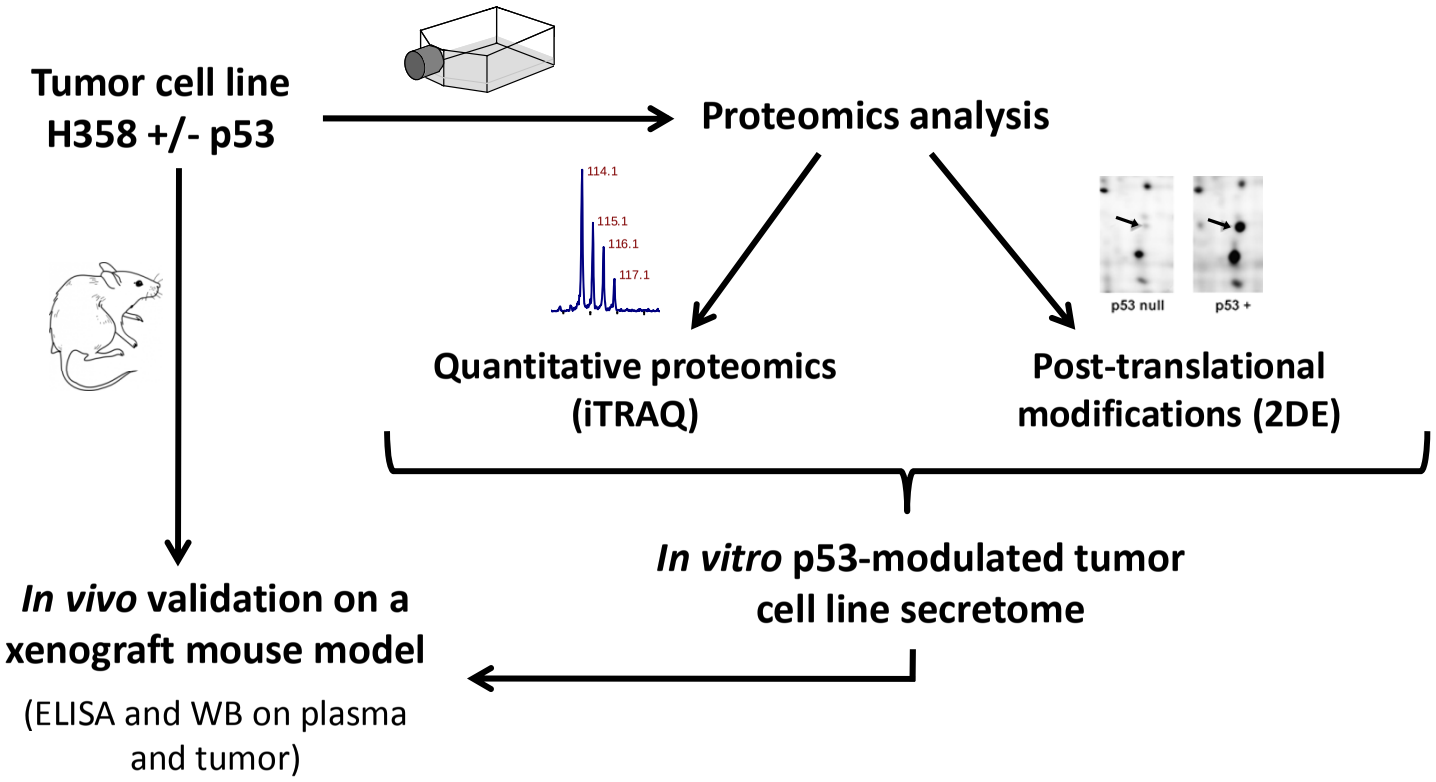

A
GDF-15 (Tumor Xenograft)
GDF-15 levels in tumor proteins extracts (ng/mg)
0
1
2
3
4
5
6
7
H358 p53-
H358 p53+
B
FGF-19 (Tumor Xenograft)
FGF-19 levels in tumor proteins extracts (pg/mg)
0
20
40
60
80
100
120
H358 p53-
H358 p53+
C
VEGF (Tumor Xenograft)
VEGF levels in tumor proteins extracts (ng/mg)
0
1
2
3
4
5
6
7
H358 p53-
H358 p53+
D
GDF-15 (Plasma)
GDF-15 levels in plasmas (pg/mL)
0
500
1000
1500
2000
2500
0
0
H358 p53-
H358 p53+
Control dox-
Control dox+

Figure for table of contents – a graphic and a 70 words synopsis



Synopsis: We characterized the p53-modulated secretome of a lung tumor model *in vitro* and demonstrated that the modulation of exported proteins can also be detected *in vivo* in the plasma of tumor-bearing mice. To our knowledge, this is the first report establishing that the *in vitro* cell line secretome is reliable and reflects the extracellular release of proteins from tumor cells *in vivo,* and could be used to identify putative tumor markers.

Supplemental table S1

No.	Accession No.	Names	Pep.
1	Q15149|PLEC1_HUMAN	Plectin-1	180
2	P98160|PGBM_HUMAN	Basement membrane-specific heparan sulfate proteoglycan core protein precursor	103
3	O15230|LAMA5_HUMAN	Laminin subunit alpha-5 precursor	79
4	O00468|AGRIN_HUMAN	Agrin precursor	76
5	P21333|FLNA_HUMAN	Filamin-A	65
6	P49327|FAS_HUMAN	Fatty acid synthase	63
7	Q13813|SPTA2_HUMAN	Spectrin alpha chain, brain	57
8	P02545|LMNA_HUMAN	Lamin-A/C	56
9	O43707|ACTN4_HUMAN	Alpha-actinin-4	55
10	Q09666|AHNK_HUMAN	Neuroblast differentiation-associated protein AHNAK	55
11	Q15582|BGH3_HUMAN	Transforming growth factor-beta-induced protein ig-h3 precursor	50
12	P07900|HS90A_HUMAN	Heat shock protein HSP 90-alpha	49
13	P63261|ACTG_HUMAN	Actin, cytoplasmic 2	49
14	P07996|TSP1_HUMAN	Thrombospondin-1 precursor	48
15	P14618|KPYM_HUMAN	Pyruvate kinase isozymes M1/M2	46
16	P11142|HSP7C_HUMAN	Heat shock cognate 71 kDa protein	46
17	P05067|A4_HUMAN	Amyloid beta A4 protein precursor	46
18	P02751|FINC_HUMAN	Fibronectin precursor	46
19	O00391|QSOX1_HUMAN	Sulfhydryl oxidase 1 precursor	43
20	Q13753|LAMC2_HUMAN	Laminin subunit gamma-2 precursor	43
21	P06733|ENOA_HUMAN	Alpha-enolase	41
22	P07237|PDIA1_HUMAN	Protein disulfide-isomerase precursor	40
23	Q08380|LG3BP_HUMAN	Galectin-3-binding protein precursor	40
24	P15311|EZRI_HUMAN	Ezrin	39
25	O94985|CSTN1_HUMAN	Calsyntenin-1 precursor	39
26	P05787|K2C8_HUMAN	Keratin, type II cytoskeletal 8	39
27	P07355|ANXA2_HUMAN	Annexin A2	39
28	Q16787|LAMA3_HUMAN	Laminin subunit alpha-3 precursor	39
29	P08727|K1C19_HUMAN	Keratin, type I cytoskeletal 19	38
30	P24821|TENA_HUMAN	Tenascin precursor	38
31	P11047|LAMC1_HUMAN	Laminin subunit gamma-1 precursor	37
32	Q01082|SPTB2_HUMAN	Spectrin beta chain, brain 1	37
33	P11021|GRP78_HUMAN	78 kDa glucose-regulated protein precursor	36
34	P08603|CFAH_HUMAN	Complement factor H precursor	36
35	P00749|UROK_HUMAN	Urokinase-type plasminogen activator precursor	34
36	P29401|TKT_HUMAN	Transketolase	33
37	P30101|PDIA3_HUMAN	Protein disulfide-isomerase A3 precursor	32
38	P02787|TRFE_HUMAN	Serotransferrin precursor	32
39	P08238|HS90B_HUMAN	Heat shock protein HSP 90-beta	31
40	P16870|CBPE_HUMAN	Carboxypeptidase E precursor	31
41	P19338|NUCL_HUMAN	Nucleolin	30
42	P35579|MYH9_HUMAN	Myosin-9	30
43	Q06481|APLP2_HUMAN	Amyloid-like protein 2 precursor	30
44	Q00610|CLH1_HUMAN	Clathrin heavy chain 1	29
45	P07339|CATD_HUMAN	Cathepsin D precursor	29
46	P04075|ALDOA_HUMAN	Fructose-bisphosphate aldolase A	29
47	P12814|ACTN1_HUMAN	Alpha-actinin-1	28
48	P08729|K2C7_HUMAN	Keratin, type II cytoskeletal 7	28
49	P08107|HSP71_HUMAN	Heat shock 70 kDa protein 1	28
50	P09651|ROA1_HUMAN	Heterogeneous nuclear ribonucleoprotein A1	27
51	Q6YHK3|CD109_HUMAN	CD109 antigen precursor	27
52	P00558|PGK1_HUMAN	Phosphoglycerate kinase 1	27
53	P22626|ROA2_HUMAN	Heterogeneous nuclear ribonucleoproteins A2/B1	27
54	Q13751|LAMB3_HUMAN	Laminin subunit beta-3 precursor	27
55	Q14697|GANAB_HUMAN	Neutral alpha-glucosidase AB precursor	27
56	P53634|CATC_HUMAN	Dipeptidyl-peptidase 1 precursor	26
57	P14625|ENPL_HUMAN	Endoplasmin precursor	26
58	P04083|ANXA1_HUMAN	Annexin A1	24
59	Q8NBJ4|GP73_HUMAN	Golgi phosphoprotein 2	23
60	P05783|K1C18_HUMAN	Keratin, type I cytoskeletal 18	23
61	P07942|LAMB1_HUMAN	Laminin subunit beta-1 precursor	23
62	Q02818|NUCB1_HUMAN	Nucleobindin-1 precursor	23
63	P63104|1433Z_HUMAN	14-3-3 protein zeta/delta	23
64	P07602|SAP_HUMAN	Proactivator polypeptide precursor [Contains: Saposin-A]	23
65	P68371|TBB2C_HUMAN	Tubulin beta-2C chain	23
66	P60174|TPIS_HUMAN	Triosephosphate isomerase	22
67	P01034|CYTC_HUMAN	Cystatin-C precursor	21
68	P00338|LDHA_HUMAN	L-lactate dehydrogenase A chain	21
69	P13667|PDIA4_HUMAN	Protein disulfide-isomerase A4 precursor	21
70	P31948|STIP1_HUMAN	Stress-induced-phosphoprotein 1	21
71	O43278|SPIT1_HUMAN	Kunitz-type protease inhibitor 1 precursor	21
72	P23246|SFPQ_HUMAN	Splicing factor, proline- and glutamine-rich	20
73	P12956|KU70_HUMAN	ATP-dependent DNA helicase 2 subunit 1	20
74	P27797|CALR_HUMAN	Calreticulin precursor	20
75	P13010|KU86_HUMAN	ATP-dependent DNA helicase 2 subunit 2	20
76	P13639|EF2_HUMAN	Elongation factor 2	20
77	Q12905|ILF2_HUMAN	Interleukin enhancer-binding factor 2	20
78	Q12805|FBLN3_HUMAN	EGF-containing fibulin-like extracellular matrix protein 1 precursor	20
79	P61978|HNRPK_HUMAN	Heterogeneous nuclear ribonucleoprotein K	20
80	P16035|TIMP2_HUMAN	Metalloproteinase inhibitor 2 precursor	19
81	Q86X29|LSR_HUMAN	Lipolysis-stimulated lipoprotein receptor	19
82	P13497|BMP1_HUMAN	Bone morphogenetic protein 1 precursor	19
83	P06748|NPM_HUMAN	Nucleophosmin	19
84	Q02809|PLOD1_HUMAN	Procollagen-lysine,2-oxoglutarate 5-dioxygenase 1 precursor	19
85	P11216|PYGB_HUMAN	Glycogen phosphorylase, brain form	19
86	O60716|CTND1_HUMAN	Catenin delta-1	19
87	P55072|TERA_HUMAN	Transitional endoplasmic reticulum ATPase	18
88	Q92520|FAM3C_HUMAN	Protein FAM3C precursor	18
89	P18065|IBP2_HUMAN	Insulin-like growth factor-binding protein 2 precursor	18
90	O75369|FLNB_HUMAN	Filamin-B	18
91	Q92876|KLK6_HUMAN	Kallikrein-6 precursor	18
92	P55268|LAMB2_HUMAN	Laminin subunit beta-2 precursor	18
93	P09429|HMGB1_HUMAN	High mobility group protein B1	17
94	P26038|MOES_HUMAN	Moesin	17
95	O60437|PEPL_HUMAN	Periplakin	17
96	P45452|MMP13_HUMAN	Collagenase 3 precursor	17
97	Q14204|DYHC_HUMAN	Dynein heavy chain, cytosolic	17
98	P68104|EF1A1_HUMAN	Elongation factor 1-alpha 1	17
99	P51858|HDGF_HUMAN	Hepatoma-derived growth factor	17
100	P10909|CLUS_HUMAN	Clusterin precursor	17
101	P04406|G3P_HUMAN	Glyceraldehyde-3-phosphate dehydrogenase	17
102	P08581|MET_HUMAN	Hepatocyte growth factor receptor precursor	17
103	Q9Y4K0|LOXL2_HUMAN	Lysyl oxidase homolog 2 precursor	16
104	P11940|PABP1_HUMAN	Polyadenylate-binding protein 1	16
105	Q96AE4|FUBP1_HUMAN	Far upstream element-binding protein 1	16
106	P23284|PPIB_HUMAN	Peptidyl-prolyl cis-trans isomerase B precursor	16
107	Q13098|CSN1_HUMAN	COP9 signalosome complex subunit 1	16
108	P12277|KCRB_HUMAN	Creatine kinase B-type	16
109	P39687|AN32A_HUMAN	Acidic leucine-rich nuclear phosphoprotein 32 family member A	15
110	Q16531|DDB1_HUMAN	DNA damage-binding protein 1	15
111	P58107|EPIPL_HUMAN	Epiplakin	15
112	Q99538|LGMN_HUMAN	Legumain precursor	15
113	P09960|LKHA4_HUMAN	Leukotriene A-4 hydrolase	15
114	Q9GZM7|TINAL_HUMAN	Tubulointerstitial nephritis antigen-like precursor	15
115	Q14103|HNRPD_HUMAN	Heterogeneous nuclear ribonucleoprotein D0	14
116	P05455|LA_HUMAN	Lupus La protein	14
117	P02768|ALBU_HUMAN	Serum albumin precursor	14

118	P22223|CADH3_HUMAN	Cadherin-3 precursor	14
119	P35221|CTNA1_HUMAN	Catenin alpha-1	14
120	P23528|COF1_HUMAN	Cofilin-1	14
121	P13796|PLSL_HUMAN	Plastin-2	14
122	P20810|ICAL_HUMAN	Calpastatin	14
123	P11413|G6PD_HUMAN	Glucose-6-phosphate 1-dehydrogenase	14
124	P35222|CTNB1_HUMAN	Catenin beta-1	14
125	P10586|PTPRF_HUMAN	Receptor-type tyrosine-protein phosphatase F precursor	14
126	Q13838|UAP56_HUMAN	Spliceosome RNA helicase BAT1	14
127	O60687|SRPX2_HUMAN	Sushi repeat-containing protein SRPX2 precursor	14
128	P26641|EF1G_HUMAN	Elongation factor 1-gamma	13
129	Q15233|NONO_HUMAN	Non-POU domain-containing octamer-binding protein	13
130	P34932|HSP74_HUMAN	Heat shock 70 kDa protein 4	13
131	Q14974|IMB1_HUMAN	Importin subunit beta-1	13
132	P62937|PPIA_HUMAN	Peptidyl-prolyl cis-trans isomerase A	13
133	P55058|PLTP_HUMAN	Phospholipid transfer protein precursor	13
134	P62258|1433E_HUMAN	14-3-3 protein epsilon	13
135	P35052|GPC1_HUMAN	Glypican-1 precursor	13
136	Q92945|FUBP2_HUMAN	Far upstream element-binding protein 2	13
137	P98095|FBLN2_HUMAN	Fibulin-2 precursor	13
138	P06744|G6PI_HUMAN	Glucose-6-phosphate isomerase	13
139	O95084|PRS23_HUMAN	Serine protease 23 precursor	13
140	Q14247|SRC8_HUMAN	Src substrate cortactin	13
141	P06753|TPM3_HUMAN	Tropomyosin alpha-3 chain	12
142	Q92820|GGH_HUMAN	Gamma-glutamyl hydrolase precursor	12
143	P61916|NPC2_HUMAN	Epididymal secretory protein E1 precursor	12
144	Q9NPR2|SEM4B_HUMAN	Semaphorin-4B precursor	12
145	P61981|1433G_HUMAN	14-3-3 protein gamma	12
146	P12830|CADH1_HUMAN	Epithelial-cadherin precursor	12
147	Q9Y4L1|HYOU1_HUMAN	Hypoxia up-regulated protein 1 precursor	12
148	P19021|AMD_HUMAN	Peptidyl-glycine alpha-amidating monooxygenase precursor	12
149	Q13435|SF3B2_HUMAN	Splicing factor 3B subunit 2	12
150	P50990|TCPQ_HUMAN	T-complex protein 1 subunit theta	12
151	Q92598|HS105_HUMAN	Heat shock protein 105 kDa	12
152	P07910|HNRPC_HUMAN	Heterogeneous nuclear ribonucleoproteins C1/C2	11
153	P80723|BASP_HUMAN	Brain acid soluble protein 1	11
154	P51991|ROA3_HUMAN	Heterogeneous nuclear ribonucleoprotein A3	11
155	Q16881|TRXR1_HUMAN	Thioredoxin reductase 1, cytoplasmic precursor	11
156	Q13740|CD166_HUMAN	CD166 antigen precursor	11
157	Q14152|IF3A_HUMAN	Eukaryotic translation initiation factor 3 subunit 10	11
158	Q8NC51|PAIRB_HUMAN	Plasminogen activator inhibitor 1 RNA-binding protein	11
159	P08758|ANXA5_HUMAN	Annexin A5	11
160	P29692|EF1D_HUMAN	Elongation factor 1-delta	11
161	P17931|LEG3_HUMAN	Galectin-3	11
162	P14314|GLU2B_HUMAN	Glucosidase 2 subunit beta precursor	11
163	P04439|1A03_HUMAN	HLA class I histocompatibility antigen, A-3 alpha chain precursor	11
164	Q99497|PARK7_HUMAN	Protein DJ-1	11
165	P50454|SERPH_HUMAN	Serpin H1 precursor	11
166	Q13263|TIF1B_HUMAN	Transcription intermediary factor 1-beta	11
167	P61160|ARP2_HUMAN	Actin-like protein 2	11
168	P09382|LEG1_HUMAN	Galectin-1	11
169	P20700|LMNB1_HUMAN	Lamin-B1	11
170	P22392|NDKB_HUMAN	Nucleoside diphosphate kinase B	11
171	Q06830|PRDX1_HUMAN	Peroxiredoxin-1	11
172	P00491|PNPH_HUMAN	Purine nucleoside phosphorylase	11
173	P61769|B2MG_HUMAN	Beta-2-microglobulin precursor	10
174	Q01105|SET_HUMAN	Protein SET	10
175	Q07955|SFRS1_HUMAN	Splicing factor, arginine/serine-rich 1	10
176	Q8NES3|LFNG_HUMAN	Beta-1,3-N-acetylglucosaminyltransferase lunatic fringe	10
177	P06865|HEXA_HUMAN	Beta-hexosaminidase alpha chain precursor	10
178	Q04637|IF4G1_HUMAN	Eukaryotic translation initiation factor 4 gamma 1	10
179	P40925|MDHC_HUMAN	Malate dehydrogenase, cytoplasmic	10
180	P37802|TAGL2_HUMAN	Transgelin-2	10
181	O75083|WDR1_HUMAN	WD repeat protein 1	10
182	Q16706|MA2A1_HUMAN	Alpha-mannosidase 2	10
183	O43852|CALU_HUMAN	Calumenin precursor	10
184	Q16658|FSCN1_HUMAN	Fascin	10
185	P07195|LDHB_HUMAN	L-lactate dehydrogenase B chain	10
186	Q15365|PCBP1_HUMAN	Poly(rC)-binding protein 1	10
187	Q15084|PDIA6_HUMAN	Protein disulfide-isomerase A6 precursor	10
188	P13611|CSPG2_HUMAN	Versican core protein precursor	10
189	Q15046|SYK_HUMAN	Lysyl-tRNA synthetase	9
190	P43490|NAMPT_HUMAN	Nicotinamide phosphoribosyltransferase	9
191	Q9UL46|PSME2_HUMAN	Proteasome activator complex subunit 2	9
192	P40926|MDHM_HUMAN	Malate dehydrogenase, mitochondrial precursor	9
193	P21741|MK_HUMAN	Midkine precursor	9
194	Q86SF2|GALT7_HUMAN	N-acetylgalactosaminyltransferase 7	9
195	P26599|PTBP1_HUMAN	Polypyrimidine tract-binding protein 1	9
196	P07737|PROF1_HUMAN	Profilin-1	9
197	P46940|IQGA1_HUMAN	Ras GTPase-activating-like protein IQGAP1	9
198	P78371|TCPB_HUMAN	T-complex protein 1 subunit beta	9
199	P22314|UBE1_HUMAN	Ubiquitin-activating enzyme E1	9
200	P23526|SAHH_HUMAN	Adenosylhomocysteinase	9
201	P31939|PUR9_HUMAN	Bifunctional purine biosynthesis protein PURH	9
202	Q14126|DSG2_HUMAN	Desmoglein-2 precursor	9
203	P38159|HNRPG_HUMAN	Heterogeneous nuclear ribonucleoprotein G	9
204	P01033|TIMP1_HUMAN	Metalloproteinase inhibitor 1 precursor	9
205	P15559|NQO1_HUMAN	NAD(P)H dehydrogenase [quinone] 1	9
206	Q06323|PSME1_HUMAN	Proteasome activator complex subunit 1	9
207	P55786|PSA_HUMAN	Puromycin-sensitive aminopeptidase	9
208	Q15393|SF3B3_HUMAN	Splicing factor 3B subunit 3	9
209	O76061|STC2_HUMAN	Stanniocalcin-2 precursor	9
210	P18206|VINC_HUMAN	Vinculin	9
211	P26583|HMGB2_HUMAN	High mobility group protein B2	8
212	P31431|SDC4_HUMAN	Syndecan-4 precursor	8
213	P61604|CH10_HUMAN	10 kDa heat shock protein, mitochondrial	8
214	P62906|RL10A_HUMAN	60S ribosomal protein L10a	8
215	Q08945|SSRP1_HUMAN	FACT complex subunit SSRP1	8
216	Q15424|SAFB1_HUMAN	Scaffold attachment factor B	8
217	Q9UHD8|SEPT9_HUMAN	Septin-9	8
218	P37837|TALDO_HUMAN	Transaldolase	8
219	P30085|KCY_HUMAN	UMP-CMP kinase	8
220	P53999|TCP4_HUMAN	Activated RNA polymerase II transcriptional coactivator p15	8
221	Q14019|COTL1_HUMAN	Coactosin-like protein	8
222	Q14566|MCM6_HUMAN	DNA replication licensing factor MCM6	8
223	P60842|IF4A1_HUMAN	Eukaryotic initiation factor 4A-I	8
224	Q09028|RBBP4_HUMAN	Histone-binding protein RBBP4	8
225	Q14703|MBTP1_HUMAN	Membrane-bound transcription factor site-1 protease precursor	8
226	P30041|PRDX6_HUMAN	Peroxiredoxin-6	8
227	Q96G03|PGM2_HUMAN	Phosphoglucomutase-2	8
228	Q15293|RCN1_HUMAN	Reticulocalbin-1 precursor	8
229	Q9P2E9|RRBP1_HUMAN	Ribosome-binding protein 1	8
230	P30153|2AAA_HUMAN	Serine/threonine-protein phosphatase 2A 65 kDa regulatory subunit A alpha isoform	8
231	P62136|PP1A_HUMAN	Serine/threonine-protein phosphatase PP1-alpha catalytic subunit	8
232	P36952|SPB5_HUMAN	Serpin B5 precursor	8
233	P48643|TCPE_HUMAN	T-complex protein 1 subunit epsilon	8
234	P00750|TPA_HUMAN	Tissue-type plasminogen activator precursor	8
235	Q6EMK4|VASN_HUMAN	Vasorin precursor	8
236	Q13510|ASAH1_HUMAN	Acid ceramidase precursor	8
237	P53396|ACLY_HUMAN	ATP-citrate synthase	8
238	Q9Y376|CAB39_HUMAN	Calcium-binding protein 39	8
239	Q96HE7|ERO1A_HUMAN	ERO1-like protein alpha precursor	8

240	Q9Y653|GPR56_HUMAN	G-protein coupled receptor 56 precursor	8
241	P63244|GBLP_HUMAN	Guanine nucleotide-binding protein subunit beta 2-like 1	8
242	P04792|HSPB1_HUMAN	Heat shock protein beta-1	8
243	O60506|HNRPQ_HUMAN	Heterogeneous nuclear ribonucleoprotein Q	8
244	Q08431|MFGM_HUMAN	Lactadherin precursor	8
245	P10619|PPGB_HUMAN	Lysosomal protective protein precursor	8
246	P60660|MYL6_HUMAN	Myosin light polypeptide 6	8
247	Q10471|GALT2_HUMAN	Polypeptide N-acetylgalactosaminyltransferase 2	8
248	Q9UHG2|PCSK1_HUMAN	ProSAAS precursor	8
249	P31949|S10AB_HUMAN	Protein S100-A11	8
250	P26639|SYTC_HUMAN	Threonyl-tRNA synthetase, cytoplasmic	8
251	P54727|RD23B_HUMAN	UV excision repair protein RAD23 homolog B	8
252	Q15631|TSN_HUMAN	Translin	7
253	P52209|6PGD_HUMAN	6-phosphogluconate dehydrogenase, decarboxylating	7
254	P30040|ERP29_HUMAN	Endoplasmic reticulum protein ERp29 precursor	7
255	P30740|ILEU_HUMAN	Leukocyte elastase inhibitor	7
256	P50897|PPT1_HUMAN	Palmitoyl-protein thioesterase 1 precursor	7
257	P10155|RO60_HUMAN	60 kDa SS-A/Ro ribonucleoprotein	7
258	Q08211|DHX9_HUMAN	ATP-dependent RNA helicase A	7
259	Q9Y5B9|SPT16_HUMAN	FACT complex subunit SPT16	7
260	Q14512|FGFP1_HUMAN	Fibroblast growth factor-binding protein 1 precursor	7
261	Q15181|IPYR_HUMAN	Inorganic pyrophosphatase	7
262	P22234|PUR6_HUMAN	Multifunctional protein ADE2	7
263	P19105|MLRM_HUMAN	Myosin regulatory light chain 2, nonsarcomeric	7
264	P67809|YBOX1_HUMAN	Nuclease sensitive element-binding protein 1	7
265	Q13162|PRDX4_HUMAN	Peroxiredoxin-4	7
266	Q86SR1|GLT10_HUMAN	Polypeptide N-acetylgalactosaminyltransferase 10	7
267	P35637|FUS_HUMAN	RNA-binding protein FUS	7
268	P08195|4F2_HUMAN	4F2 cell-surface antigen heavy chain	7
269	O15144|ARPC2_HUMAN	Actin-related protein 2/3 complex subunit 2	7
270	O14672|ADA10_HUMAN	ADAM 10 precursor	7
271	P54819|KAD2_HUMAN	Adenylate kinase isoenzyme 2, mitochondrial	7
272	P15514|AREG_HUMAN	Amphiregulin precursor	7
273	P12429|ANXA3_HUMAN	Annexin A3	7
274	P02649|APOE_HUMAN	Apolipoprotein E precursor	7
275	P15291|B4GT1_HUMAN	Beta-1,4-galactosyltransferase 1	7
276	P29373|RABP2_HUMAN	Cellular retinoic acid-binding protein 2	7
277	P39060|COIA1_HUMAN	Collagen alpha-1(XVIII) chain precursor [Contains: Endostatin]	7
278	P27695|APEX1_HUMAN	DNA-(apurinic or apyrimidinic site) lyase	7
279	Q9H4M9|EHD1_HUMAN	EH domain-containing protein 1	7
280	P29317|EPHA2_HUMAN	Ephrin type-A receptor 2 precursor	7
281	Q16394|EXT1_HUMAN	Exostosin-1	7
282	P00390|GSHR_HUMAN	Glutathione reductase, mitochondrial precursor	7
283	O95757|HS74L_HUMAN	Heat shock 70 kDa protein 4L	7
284	P31943|HNRH1_HUMAN	Heterogeneous nuclear ribonucleoprotein H	7
285	P14923|PLAK_HUMAN	Junction plakoglobin	7
286	Q9NS15|LTBP3_HUMAN	Latent-transforming growth factor beta-binding protein 3 precursor	7
287	P12004|PCNA_HUMAN	Proliferating cell nuclear antigen	7
288	Q15459|SF3A1_HUMAN	Splicing factor 3 subunit 1	7
289	O75533|SF3B1_HUMAN	Splicing factor 3B subunit 1	7
290	Q8NBS9|TXND5_HUMAN	Thioredoxin domain-containing protein 5 precursor	7
291	P04066|FUCO_HUMAN	Tissue alpha-L-fucosidase precursor	7
292	Q13308|PTK7_HUMAN	Tyrosine-protein kinase-like 7 precursor	7
293	P62988|UBIQ_HUMAN	Ubiquitin	7
294	Q96FW1|OTUB1_HUMAN	Ubiquitin thioesterase OTUB1	7
295	Q8NFL0|B3GN7_HUMAN	UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 7	7
296	P67936|TPM4_HUMAN	Tropomyosin alpha-4 chain	6
297	P62191|PRS4_HUMAN	26S protease regulatory subunit 4	6
298	Q14240|IF4A2_HUMAN	Eukaryotic initiation factor 4A-II	6
299	P04264|K2C1_HUMAN	Keratin, type II cytoskeletal 1	6
300	P12270|TPR_HUMAN	Nucleoprotein TPR	6
301	P07686|HEXB_HUMAN	Beta-hexosaminidase beta chain precursor	6
302	Q9BQT9|CSTN3_HUMAN	Calsyntenin-3 precursor	6
303	O00299|CLIC1_HUMAN	Chloride intracellular channel protein 1	6
304	Q9UHL4|DPP2_HUMAN	Dipeptidyl-peptidase 2 precursor	6
305	Q08345|DDR1_HUMAN	Epithelial discoidin domain-containing receptor 1 precursor	6
306	P28799|GRN_HUMAN	Granulins precursor	6
307	Q04760|LGUL_HUMAN	Lactoylglutathione lyase	6
308	Q03252|LMNB2_HUMAN	Lamin-B2	6
309	P15586|GNS_HUMAN	N-acetylglucosamine-6-sulfatase precursor	6
310	P25786|PSA1_HUMAN	Proteasome subunit alpha type 1	6
311	P60900|PSA6_HUMAN	Proteasome subunit alpha type 6	6
312	Q9NUQ9|FA49B_HUMAN	Protein FAM49B	6
313	Q9P258|RCC2_HUMAN	Protein RCC2	6
314	P52565|GDIR_HUMAN	Rho GDP-dissociation inhibitor 1	6
315	Q9NVA2|SEP11_HUMAN	Septin-11	6
316	P16422|TACD1_HUMAN	Tumor-associated calcium signal transducer 1 precursor	6
317	Q9C0C2|TB182_HUMAN	182 kDa tankyrase 1-binding protein	6
318	Q13200|PSMD2_HUMAN	26S proteasome non-ATPase regulatory subunit 2	6
319	Q9BRK5|CAB45_HUMAN	45 kDa calcium-binding protein precursor	6
320	Q13443|ADAM9_HUMAN	ADAM 9 precursor	6
321	Q92499|DDX1_HUMAN	ATP-dependent RNA helicase DDX1	6
322	Q9GZN4|BSSP4_HUMAN	Brain-specific serine protease 4 precursor	6
323	P07711|CATL_HUMAN	Cathepsin L precursor	6
324	Q9Y2E5|MA2B2_HUMAN	Epididymis-specific alpha-mannosidase precursor	6
325	Q99613|IF38_HUMAN	Eukaryotic translation initiation factor 3 subunit 8	6
326	P55884|IF39_HUMAN	Eukaryotic translation initiation factor 3 subunit 9	6
327	Q93063|EXT2_HUMAN	Exostosin-2	6
328	P52907|CAZA1_HUMAN	F-actin-capping protein subunit alpha-1	6
329	P23142|FBLN1_HUMAN	Fibulin-1 precursor	6
330	P09211|GSTP1_HUMAN	Glutathione S-transferase P	6
331	Q99988|GDF15_HUMAN	Growth/differentiation factor 15 precursor	6
332	P69905|HBA_HUMAN	Hemoglobin subunit alpha	6
333	Q00839|HNRPU_HUMAN	Heterogeneous nuclear ribonucleoprotein U	6
334	P62805|H4_HUMAN	Histone H4	6
335	P10253|LYAG_HUMAN	Lysosomal alpha-glucosidase precursor	6
336	P40121|CAPG_HUMAN	Macrophage-capping protein	6
337	P15531|NDKA_HUMAN	Nucleoside diphosphate kinase A	6
338	O60568|PLOD3_HUMAN	Procollagen-lysine,2-oxoglutarate 5-dioxygenase 3 precursor	6
339	Q9UQ80|PA2G4_HUMAN	Proliferation-associated protein 2G4	6
340	P35241|RADI_HUMAN	Radixin	6
341	O75326|SEM7A_HUMAN	Semaphorin-7A precursor	6
342	Q9H5Y7|SLIK6_HUMAN	SLIT and NTRK-like protein 6 precursor	6
343	Q07157|ZO1_HUMAN	Tight junction protein ZO-1	6
344	P02786|TFR1_HUMAN	Transferrin receptor protein 1	6
345	Q12907|LMAN2_HUMAN	Vesicular integral-membrane protein VIP36 precursor	6
346	O60841|IF2P_HUMAN	Eukaryotic translation initiation factor 5B	5
347	P63241|IF5A1_HUMAN	Eukaryotic translation initiation factor 5A-1	5
348	P26885|FKBP2_HUMAN	FK506-binding protein 2 precursor	5
349	P25398|RS12_HUMAN	40S ribosomal protein S12	5
350	P00505|AATM_HUMAN	Aspartate aminotransferase, mitochondrial precursor	5
351	Q16643|DREB_HUMAN	Drebrin	5
352	P49368|TCPG_HUMAN	T-complex protein 1 subunit gamma	5
353	P31946|1433B_HUMAN	14-3-3 protein beta/alpha	5
354	P27348|1433T_HUMAN	14-3-3 protein theta	5
355	P46777|RL5_HUMAN	60S ribosomal protein L5	5
356	O95994|AGR2_HUMAN	Anterior gradient protein 2 homolog precursor	5
357	Q9BZZ5|API5_HUMAN	Apoptosis inhibitor 5	5
358	Q02742|GCNT1_HUMAN	Beta-1,3-galactosyl-O-glycosyl-glycoprotein beta-1,6-N-acetylglucosaminyltransferase	5
359	O43809|CPSF5_HUMAN	Cleavage and polyadenylation specificity factor 5	5
360	O94907|DKK1_HUMAN	Dickkopf-related protein 1 precursor	5
361	P04222|1C03_HUMAN	HLA class I histocompatibility antigen, Cw-3 alpha chain precursor	5

362	P23229|ITA6_HUMAN	Integrin alpha-6 precursor	5
363	P05556|ITB1_HUMAN	Integrin beta-1 precursor	5
364	Q14847|LASP1_HUMAN	LIM and SH3 domain protein 1	5
365	Q9Y617|SERC_HUMAN	Phosphoserine aminotransferase	5
366	Q8NCL4|GALT6_HUMAN	Polypeptide N-acetylgalactosaminyltransferase 6	5
367	Q8N163|K1967_HUMAN	Protein KIAA1967	5
368	P48723|STCH_HUMAN	Stress 70 protein chaperone microsome-associated 60 kDa protein precursor	5
369	P13693|TCTP_HUMAN	Translationally-controlled tumor protein	5
370	P07437|TBB5_HUMAN	Tubulin beta chain	5
371	P59998|ARPC4_HUMAN	Actin-related protein 2/3 complex subunit 4	5
372	P02765|FETUA_HUMAN	Alpha-2-HS-glycoprotein precursor	5
373	P16278|BGAL_HUMAN	Beta-galactosidase precursor	5
374	P62158|CALM_HUMAN	Calmodulin	5
375	P07858|CATB_HUMAN	Cathepsin B precursor	5
376	P45973|CBX5_HUMAN	Chromobox protein homolog 5	5
377	Q96CT7|CC124_HUMAN	Coiled-coil domain-containing protein 124	5
378	P12109|CO6A1_HUMAN	Collagen alpha-1(VI) chain precursor	5
379	Q15828|CYTM_HUMAN	Cystatin-M precursor	5
380	Q02487|DSC2_HUMAN	Desmocollin-2 precursor	5
381	P60981|DEST_HUMAN	Destrin	5
382	Q9UKM7|MA1B1_HUMAN	Endoplasmic reticulum mannosyl-oligosaccharide 1,2-alpha-mannosidase	5
383	P62826|RAN_HUMAN	GTP-binding nuclear protein Ran	5
384	Q99880|H2B1L_HUMAN	Histone H2B type 1-L	5
385	P09529|INHBB_HUMAN	Inhibin beta B chain precursor	5
386	P26006|ITA3_HUMAN	Integrin alpha-3 precursor	5
387	O75874|IDHC_HUMAN	Isocitrate dehydrogenase [NADP] cytoplasmic	5
388	P78504|JAG1_HUMAN	Jagged-1 precursor	5
389	Q86UP2|KTN1_HUMAN	Kinectin	5
390	Q8NHP8|LAML2_HUMAN	LAMA-like protein 2 precursor	5
391	P42785|PCP_HUMAN	Lysosomal Pro-X carboxypeptidase precursor	5
392	P33908|MA1A1_HUMAN	Mannosyl-oligosaccharide 1,2-alpha-mannosidase IA	5
393	O95631|NET1_HUMAN	Netrin-1 precursor	5
394	O60462|NRP2_HUMAN	Neuropilin-2 precursor	5
395	P05114|HMGN1_HUMAN	Nonhistone chromosomal protein HMG-14	5
396	Q96HC4|PDLI5_HUMAN	PDZ and LIM domain protein 5	5
397	P32119|PRDX2_HUMAN	Peroxiredoxin-2	5
398	Q8WW12|PCNP_HUMAN	PEST proteolytic signal-containing nuclear protein	5
399	P18669|PGAM1_HUMAN	Phosphoglycerate mutase 1	5
400	P25787|PSA2_HUMAN	Proteasome subunit alpha type 2	5
401	O14818|PSA7_HUMAN	Proteasome subunit alpha type 7	5
402	P49721|PSB2_HUMAN	Proteasome subunit beta type 2	5
403	O43143|DHX15_HUMAN	Putative pre-mRNA-splicing factor ATP-dependent RNA helicase DHX15	5
404	P50395|GDIB_HUMAN	Rab GDP dissociation inhibitor beta	5
405	Q13283|G3BP1_HUMAN	Ras GTPase-activating protein-binding protein 1	5
406	O75787|RENR_HUMAN	Renin receptor precursor	5
407	Q13247|SFRS6_HUMAN	Splicing factor, arginine/serine-rich 6	5
408	Q16629|SFRS7_HUMAN	Splicing factor, arginine/serine-rich 7	5
409	Q9BRA2|TXNL5_HUMAN	Thioredoxin-like protein 5	5
410	O14773|TPP1_HUMAN	Tripeptidyl-peptidase 1 precursor	5
411	P07477|TRY1_HUMAN	Trypsin-1 precursor	5
412	Q9Y2A9|B3GN3_HUMAN	UDP-GlcNAc:betaGal beta-1,3-N-acetylglucosaminyltransferase 3	5
413	P26640|SYV_HUMAN	Valyl-tRNA synthetase	5
414	P15692|VEGFA_HUMAN	Vascular endothelial growth factor A precursor	5
415	P38919|IF4A3_HUMAN	Eukaryotic initiation factor 4A-III	4
416	P29034|S10A2_HUMAN	Protein S100-A2	4
417	P09493|TPM1_HUMAN	Tropomyosin alpha-1 chain	4
418	O43242|PSMD3_HUMAN	26S proteasome non-ATPase regulatory subunit 3	4
419	P47756|CAPZB_HUMAN	F-actin-capping protein subunit beta	4
420	O14979|HNRDL_HUMAN	Heterogeneous nuclear ribonucleoprotein D-like	4
421	P17096|HMGA1_HUMAN	High mobility group protein HMG-I/HMG-Y	4
422	Q8WWI1|LMO7_HUMAN	LIM domain only protein 7	4
423	Q15185|TEBP_HUMAN	Prostaglandin E synthase 3	4
424	P45974|UBP5_HUMAN	Ubiquitin carboxyl-terminal hydrolase 5	4
425	Q9UNM6|PSD13_HUMAN	26S proteasome non-ATPase regulatory subunit 13	4
426	P08865|RSSA_HUMAN	40S ribosomal protein SA	4
427	Q9H4A4|AMPB_HUMAN	Aminopeptidase B	4
428	Q9NYF8|BCLF1_HUMAN	Bcl-2-associated transcription factor 1	4
429	P19022|CADH2_HUMAN	Cadherin-2 precursor	4
430	P00450|CERU_HUMAN	Ceruloplasmin precursor	4
431	P09496|CLCA_HUMAN	Clathrin light chain A	4
432	Q9ULV4|COR1C_HUMAN	Coronin-1C	4
433	O00273|DFFA_HUMAN	DNA fragmentation factor subunit alpha	4
434	P33991|MCM4_HUMAN	DNA replication licensing factor MCM4	4
435	O75937|DNJC8_HUMAN	DnaJ homolog subfamily C member 8	4
436	P06396|GELS_HUMAN	Gelsolin precursor	4
437	O14792|OST1_HUMAN	Heparan sulfate glucosamine 3-O-sulfotransferase 1 precursor	4
438	Q96QV6|H2A1A_HUMAN	Histone H2A type 1-A	4
439	P29966|MARCS_HUMAN	Myristoylated alanine-rich C-kinase substrate	4
440	Q96TA1|NIBL_HUMAN	Niban-like protein	4
441	Q15223|PVRL1_HUMAN	Poliovirus receptor-related protein 1 precursor	4
442	P06703|S10A6_HUMAN	Protein S100-A6	4
443	P06454|PTMA_HUMAN	Prothymosin alpha [Contains: Thymosin alpha-1]	4
444	Q15019|SEPT2_HUMAN	Septin-2	4
445	P16949|STMN1_HUMAN	Stathmin	4
446	Q9Y490|TLN1_HUMAN	Talin-1	4
447	Q9Y2W1|TR150_HUMAN	Thyroid hormone receptor-associated protein 3	4
448	P10646|TFPI1_HUMAN	Tissue factor pathway inhibitor precursor	4
449	P08621|RU17_HUMAN	U1 small nuclear ribonucleoprotein 70 kDa	4
450	P61086|UBC1_HUMAN	Ubiquitin-conjugating enzyme E2-25 kDa	4
451	Q04446|GLGB_HUMAN	1,4-alpha-glucan-branching enzyme	4
452	P43686|PRS6B_HUMAN	26S protease regulatory subunit 6B	4
453	Q99460|PSMD1_HUMAN	26S proteasome non-ATPase regulatory subunit 1	4
454	Q15008|PSMD6_HUMAN	26S proteasome non-ATPase regulatory subunit 6	4
455	P23396|RS3_HUMAN	40S ribosomal protein S3	4
456	P21589|5NTD_HUMAN	5'-nucleotidase precursor	4
457	P05388|RLA0_HUMAN	60S acidic ribosomal protein P0	4
458	P05387|RLA2_HUMAN	60S acidic ribosomal protein P2	4
459	O95433|AHSA1_HUMAN	Activator of 90 kDa heat shock protein ATPase homolog 1	4
460	P07108|ACBP_HUMAN	Acyl-CoA-binding protein	4
461	Q10567|AP1B1_HUMAN	AP-1 complex subunit beta-1	4
462	Q9UKV3|ACINU_HUMAN	Apoptotic chromatin condensation inducer in the nucleus	4
463	Q12797|ASPH_HUMAN	Aspartyl/asparaginyl beta-hydroxylase	4
464	Q13938|CAYP1_HUMAN	Calcyphosin	4
465	P17655|CAN2_HUMAN	Calpain-2 catalytic subunit precursor	4
466	P53618|COPB_HUMAN	Coatomer subunit beta	4
467	Q96CG8|CTHR1_HUMAN	Collagen triple helix repeat-containing protein 1 precursor	4
468	Q86VP6|CAND1_HUMAN	Cullin-associated NEDD8-dissociated protein 1	4
469	P14854|CX6B1_HUMAN	Cytochrome c oxidase subunit VIb isoform 1	4
470	P33316|DUT_HUMAN	Deoxyuridine 5'-triphosphate nucleotidohydrolase, mitochondrial precursor	4
471	P25205|MCM3_HUMAN	DNA replication licensing factor MCM3	4
472	P78527|PRKDC_HUMAN	DNA-dependent protein kinase catalytic subunit	4
473	Q14118|DAG1_HUMAN	Dystroglycan precursor	4
474	O43854|EDIL3_HUMAN	EGF-like repeat and discoidin I-like domain-containing protein 3 precursor	4
475	Q92817|EVPL_HUMAN	Envoplakin	4
476	P20827|EFNA1_HUMAN	Ephrin-A1 precursor	4
477	P52803|EFNA5_HUMAN	Ephrin-A5 precursor	4
478	P41091|IF2G_HUMAN	Eukaryotic translation initiation factor 2 subunit 3	4
479	Q13347|IF32_HUMAN	Eukaryotic translation initiation factor 3 subunit 2	4
480	P48637|GSHB_HUMAN	Glutathione synthetase	4
481	Q99729|ROAA_HUMAN	Heterogeneous nuclear ribonucleoprotein A/B	4
482	P50502|F10A1_HUMAN	Hsc70-interacting protein	4
483	O00410|IMB3_HUMAN	Importin beta-3	4

484	P16144|ITB4_HUMAN	Integrin beta-4 precursor	4
485	P21399|IREB1_HUMAN	Iron-responsive element-binding protein 1	4
486	P42167|LAP2B_HUMAN	Lamina-associated polypeptide 2, isoforms beta/gamma	4
487	P50895|LU_HUMAN	Lutheran blood group glycoprotein precursor	4
488	O00754|MA2B1_HUMAN	Lysosomal alpha-mannosidase precursor	4
489	P15529|MCP_HUMAN	Membrane cofactor protein precursor	4
490	P80297|MT1X_HUMAN	Metallothionein-1X	4
491	P34059|GALNS_HUMAN	N-acetylgalactosamine-6-sulfatase precursor	4
492	P48163|MAOX_HUMAN	NADP-dependent malic enzyme	4
493	P49321|NASP_HUMAN	Nuclear autoantigenic sperm protein	4
494	Q14980|NUMA1_HUMAN	Nuclear mitotic apparatus protein 1	4
495	P68402|PA1B2_HUMAN	Platelet-activating factor acetylhydrolase IB subunit beta	4
496	Q15366|PCBP2_HUMAN	Poly(rC)-binding protein 2	4
497	P25788|PSA3_HUMAN	Proteasome subunit alpha type 3	4
498	P28066|PSA5_HUMAN	Proteasome subunit alpha type 5	4
499	Q9Y224|CN166_HUMAN	Protein C14orf166	4
500	O75629|CREG1_HUMAN	Protein CREG1 precursor	4
501	Q9UNF0|PACN2_HUMAN	Protein kinase C and casein kinase substrate in neurons protein 2	4
502	Q99584|S10AD_HUMAN	Protein S100-A13	4
503	Q96FQ6|S10AG_HUMAN	Protein S100-A16	4
504	P25815|S100P_HUMAN	Protein S100-P	4
505	Q53FA7|QORX_HUMAN	Putative quinone oxidoreductase	4
506	Q9Y383|LC7L2_HUMAN	Putative RNA-binding protein Luc7-like 2	4
507	Q01844|EWS_HUMAN	RNA-binding protein EWS	4
508	P14678|RSMB_HUMAN	Small nuclear ribonucleoprotein-associated proteins B and B'	4
509	Q14515|SPRL1_HUMAN	SPARC-like protein 1 precursor	4
510	Q15637|SF01_HUMAN	Splicing factor 1	4
511	Q01130|SFRS2_HUMAN	Splicing factor, arginine/serine-rich 2	4
512	Q7KZF4|SND1_HUMAN	Staphylococcal nuclease domain-containing protein 1	4
513	P00441|SODC_HUMAN	Superoxide dismutase [Cu-Zn]	4
514	P17987|TCPA_HUMAN	T-complex protein 1 subunit alpha	4
515	Q99832|TCPH_HUMAN	T-complex protein 1 subunit eta	4
516	P10599|THIO_HUMAN	Thioredoxin	4
517	P22102|PUR2_HUMAN	Trifunctional purine biosynthetic protein adenosine-3	4
518	Q08J23|NSUN2_HUMAN	tRNA	4
519	Q9BQE3|TBA1C_HUMAN	Tubulin alpha-1C chain	4
520	P09012|SNRPA_HUMAN	U1 small nuclear ribonucleoprotein A	4
521	Q14157|UBP2L_HUMAN	Ubiquitin-associated protein 2-like	4
522	P68036|UB2L3_HUMAN	Ubiquitin-conjugating enzyme E2 L3	4
523	O60701|UGDH_HUMAN	UDP-glucose 6-dehydrogenase	4
524	Q96QK1|VPS35_HUMAN	Vacuolar protein sorting-associated protein 35	4
525	Q9ULF5|ZIP10_HUMAN	Zinc transporter ZIP10 precursor	4
526	Q9UBS4|DJB11_HUMAN	DnaJ homolog subfamily B member 11 precursor	3
527	O60869|EDF1_HUMAN	Endothelial differentiation-related factor 1	3
528	Q14320|FA50A_HUMAN	Protein FAM50A	3
529	Q9BQ61|CS043_HUMAN	Uncharacterized protein C19orf43	3
530	P17980|PRS6A_HUMAN	26S protease regulatory subunit 6A	3
531	P35998|PRS7_HUMAN	26S protease regulatory subunit 7	3
532	P62195|PRS8_HUMAN	26S protease regulatory subunit 8	3
533	Q13442|HAP28_HUMAN	28 kDa heat- and acid-stable phosphoprotein	3
534	P78536|ADA17_HUMAN	ADAM 17 precursor	3
535	Q9BXP5|ARS2_HUMAN	Arsenite-resistance protein 2	3
536	Q14517|FATH_HUMAN	Cadherin-related tumor suppressor homolog precursor	3
537	P27824|CALX_HUMAN	Calnexin precursor	3
538	P07384|CAN1_HUMAN	Calpain-1 catalytic subunit	3
539	P83916|CBX1_HUMAN	Chromobox protein homolog 1	3
540	P47813|IF1AX_HUMAN	Eukaryotic translation initiation factor 1A	3
541	P20042|IF2B_HUMAN	Eukaryotic translation initiation factor 2 subunit 2	3
542	O75822|IF31_HUMAN	Eukaryotic translation initiation factor 3 subunit 1	3
543	O75821|IF34_HUMAN	Eukaryotic translation initiation factor 3 subunit 4	3
544	P14324|FPPS_HUMAN	Farnesyl pyrophosphate synthetase	3
545	Q96GU1|GGEE1_HUMAN	G antigen family E member 1	3
546	Q14767|LTBP2_HUMAN	Latent-transforming growth factor beta-binding protein 2 precursor	3
547	Q15599|NHRF2_HUMAN	Na(+)/H(+) exchange regulatory cofactor NHE-RF2	3
548	P13797|PLST_HUMAN	Plastin-3	3
549	O00469|PLOD2_HUMAN	Procollagen-lysine,2-oxoglutarate 5-dioxygenase 2 precursor	3
550	P25789|PSA4_HUMAN	Proteasome subunit alpha type 4	3
551	Q99873|ANM1_HUMAN	Protein arginine N-methyltransferase 1	3
552	O94979|SC31A_HUMAN	Protein transport protein Sec31A	3
553	Q9Y230|RUVB2_HUMAN	RuvB-like 2	3
554	Q16181|SEPT7_HUMAN	Septin-7	3
555	P49591|SYSC_HUMAN	Seryl-tRNA synthetase, cytoplasmic	3
556	O75368|SH3L1_HUMAN	SH3 domain-binding glutamic acid-rich-like protein	3
557	P62308|RUXG_HUMAN	Small nuclear ribonucleoprotein G	3
558	P62314|SMD1_HUMAN	Small nuclear ribonucleoprotein Sm D1	3
559	P11166|GTR1_HUMAN	Solute carrier family 2, facilitated glucose transporter member 1	3
560	P63279|UBC9_HUMAN	SUMO-conjugating enzyme UBC9	3
561	P18827|SDC1_HUMAN	Syndecan-1 precursor	3
562	Q9NXG2|THUM1_HUMAN	THUMP domain-containing protein 1	3
563	P61586|RHOA_HUMAN	Transforming protein RhoA precursor	3
564	Q13428|TCOF_HUMAN	Treacle protein	3
565	O43399|TPD54_HUMAN	Tumor protein D54	3
566	O00231|PSD11_HUMAN	26S proteasome non-ATPase regulatory subunit 11	3
567	Q16401|PSMD5_HUMAN	26S proteasome non-ATPase regulatory subunit 5	3
568	P10809|CH60_HUMAN	60 kDa heat shock protein, mitochondrial precursor	3
569	P68133|ACTS_HUMAN	Actin, alpha skeletal muscle	3
570	Q13444|ADA15_HUMAN	ADAM 15 precursor	3
571	P00568|KAD1_HUMAN	Adenylate kinase isoenzyme 1	3
572	P49588|SYAC_HUMAN	Alanyl-tRNA synthetase, cytoplasmic	3
573	P11766|ADHX_HUMAN	Alcohol dehydrogenase class 3	3
574	P54802|ANAG_HUMAN	Alpha-N-acetylglucosaminidase precursor	3
575	O43747|AP1G1_HUMAN	AP-1 complex subunit gamma-1	3
576	P54136|SYRC_HUMAN	Arginyl-tRNA synthetase, cytoplasmic	3
577	P14868|SYDC_HUMAN	Aspartyl-tRNA synthetase, cytoplasmic	3
578	Q86XP3|DDX42_HUMAN	ATP-dependent RNA helicase DDX42	3
579	P35613|BASI_HUMAN	Basigin precursor	3
580	O00462|MANBA_HUMAN	Beta-mannosidase precursor	3
581	P55957|BID_HUMAN	BH3-interacting domain death agonist	3
582	P04632|CPNS1_HUMAN	Calpain small subunit 1	3
583	P21964|COMT_HUMAN	Catechol O-methyltransferase	3
584	P09668|CATH_HUMAN	Cathepsin H precursor	3
585	Q5ZPR3|CD276_HUMAN	CD276 antigen precursor	3
586	Q9BWS9|CHID1_HUMAN	Chitinase domain-containing protein 1 precursor	3
587	P09497|CLCB_HUMAN	Clathrin light chain B	3
588	P33240|CSTF2_HUMAN	Cleavage stimulation factor 64 kDa subunit	3
589	Q9Y678|COPG_HUMAN	Coatomer subunit gamma	3
590	P04080|CYTB_HUMAN	Cystatin-B	3
591	P50238|CRIP1_HUMAN	Cysteine-rich protein 1	3
592	O00115|DNS2A_HUMAN	Deoxyribonuclease-2-alpha precursor	3
593	Q9NY33|DPP3_HUMAN	Dipeptidyl-peptidase 3	3
594	P49736|MCM2_HUMAN	DNA replication licensing factor MCM2	3
595	P55265|DSRAD_HUMAN	Double-stranded RNA-specific adenosine deaminase	3
596	Q9UNN8|EPCR_HUMAN	Endothelial protein C receptor precursor	3
597	P84090|ERH_HUMAN	Enhancer of rudimentary homolog	3
598	Q96AY3|FKB10_HUMAN	FK506-binding protein 10 precursor	3
599	Q00688|FKBP3_HUMAN	FK506-binding protein 3	3
600	Q6MZW2|FSTL4_HUMAN	Follistatin-related protein 4 precursor	3
601	P78423|X3CL1_HUMAN	Fractalkine precursor	3
602	O76070|SYUG_HUMAN	Gamma-synuclein	3
603	Q13630|FCL_HUMAN	GDP-L-fucose synthetase	3
604	Q06210|GFPT1_HUMAN	Glucosamine--fructose-6-phosphate aminotransferase 1	3
605	P41250|SYG_HUMAN	Glycyl-tRNA synthetase	3

606	P62993|GRB2_HUMAN	Growth factor receptor-bound protein 2	3
607	Q96MM7|H6ST2_HUMAN	Heparan-sulfate 6-O-sulfotransferase 2	3
608	P31942|HNRH3_HUMAN	Heterogeneous nuclear ribonucleoprotein H3	3
609	P52272|HNRPM_HUMAN	Heterogeneous nuclear ribonucleoprotein M	3
610	Q9BUJ2|HNRL1_HUMAN	Heterogeneous nuclear ribonucleoprotein U-like protein 1	3
611	P24592|IBP6_HUMAN	Insulin-like growth factor-binding protein 6 precursor	3
612	Q9Y287|ITM2B_HUMAN	Integral membrane protein 2B	3
613	Q12906|ILF3_HUMAN	Interleukin enhancer-binding factor 3	3
614	Q9Y624|JAM1_HUMAN	Junctional adhesion molecule A precursor	3
615	Q9UHB6|LIMA1_HUMAN	LIM domain and actin-binding protein 1	3
616	Q8WWY8|LIPH_HUMAN	Lipase member H precursor	3
617	P01130|LDLR_HUMAN	Low-density lipoprotein receptor precursor	3
618	P43243|MATR3_HUMAN	Matrin-3	3
619	Q14696|MESD2_HUMAN	Mesoderm development candidate 2	3
620	Q9UPN3|MACF1_HUMAN	Microtubule-actin cross-linking factor 1, isoforms 1/2/3/5	3
621	Q6WCQ1|MRIP_HUMAN	Myosin phosphatase Rho-interacting protein	3
622	Q8TBC4|UBA3_HUMAN	NEDD8-activating enzyme E1 catalytic subunit	3
623	O14786|NRP1_HUMAN	Neuropilin-1 precursor	3
624	Q9UNZ2|NSF1C_HUMAN	NSFL1 cofactor p47	3
625	P55209|NP1L1_HUMAN	Nucleosome assembly protein 1-like 1	3
626	O00151|PDLI1_HUMAN	PDZ and LIM domain protein 1	3
627	P30044|PRDX5_HUMAN	Peroxiredoxin-5, mitochondrial precursor	3
628	Q9Y285|SYFA_HUMAN	Phenylalanyl-tRNA synthetase alpha chain	3
629	P30086|PEBP1_HUMAN	Phosphatidylethanolamine-binding protein 1	3
630	Q92841|DDX17_HUMAN	Probable ATP-dependent RNA helicase DDX17	3
631	P20618|PSB1_HUMAN	Proteasome subunit beta type 1 precursor	3
632	P28070|PSB4_HUMAN	Proteasome subunit beta type 4 precursor	3
633	O60610|DIAP1_HUMAN	Protein diaphanous homolog 1	3
634	P61326|MGN_HUMAN	Protein mago nashi homolog	3
635	P60903|S10AA_HUMAN	Protein S100-A10	3
636	Q15436|SC23A_HUMAN	Protein transport protein Sec23A	3
637	Q08174|PCDH1_HUMAN	Protocadherin-1 precursor	3
638	Q9UN86|G3BP2_HUMAN	Ras GTPase-activating protein-binding protein 2	3
639	P62491|RB11A_HUMAN	Ras-related protein Rab-11A	3
640	Q14257|RCN2_HUMAN	Reticulocalbin-2 precursor	3
641	Q9NTZ6|RBM12_HUMAN	RNA-binding protein 12	3
642	Q9Y5S9|RBM8A_HUMAN	RNA-binding protein 8A	3
643	Q14563|SEM3A_HUMAN	Semaphorin-3A precursor	3
644	Q15257|PTPA_HUMAN	Serine/threonine-protein phosphatase 2A regulatory subunit B'	3
645	P62318|SMD3_HUMAN	Small nuclear ribonucleoprotein Sm D3	3
646	Q92673|SORL_HUMAN	Sortilin-related receptor precursor	3
647	Q9UMY4|SNX12_HUMAN	Sorting nexin-12	3
648	O60749|SNX2_HUMAN	Sorting nexin-2	3
649	P63208|SKP1_HUMAN	S-phase kinase-associated protein 1A	3
650	P26368|U2AF2_HUMAN	Splicing factor U2AF 65 kDa subunit	3
651	Q9Y5Y6|ST14_HUMAN	Suppressor of tumorigenicity protein 14	3
652	P50991|TCPD_HUMAN	T-complex protein 1 subunit delta	3
653	P40227|TCPZ_HUMAN	T-complex protein 1 subunit zeta	3
654	Q9BQ16|TICN3_HUMAN	Testican-3 precursor	3
655	Q9BS26|TXND4_HUMAN	Thioredoxin domain-containing protein 4 precursor	3
656	P20290|BTF3_HUMAN	Transcription factor BTF3	3
657	Q24JP5|T132A_HUMAN	Transmembrane protein 132A precursor	3
658	Q92973|TNPO1_HUMAN	Transportin-1	3
659	P68363|TBA1B_HUMAN	Tubulin alpha-1B chain	3
660	P68366|TBA4A_HUMAN	Tubulin alpha-4A chain	3
661	Q13509|TBB3_HUMAN	Tubulin beta-3 chain	3
662	Q6IBS0|TWF2_HUMAN	Twinfilin-2	3
663	Q9UMX0|UBQL1_HUMAN	Ubiquilin-1	3
664	Q93009|UBP7_HUMAN	Ubiquitin carboxyl-terminal hydrolase 7	3
665	Q16851|UGPA_HUMAN	UTP--glucose-1-phosphate uridylyltransferase	3
666	Q71U36|TBA1A_HUMAN	Tubulin alpha-1A chain	3
667	P31947|1433S_HUMAN	14-3-3 protein sigma	2
668	P46782|RS5_HUMAN	40S ribosomal protein S5	2
669	P26373|RL13_HUMAN	60S ribosomal protein L13	2
670	P39023|RL3_HUMAN	60S ribosomal protein L3	2
671	P47895|AL1A3_HUMAN	Aldehyde dehydrogenase 1A3	2
672	Q9UII2|ATIF1_HUMAN	ATPase inhibitor, mitochondrial precursor	2
673	O75934|BCAS2_HUMAN	Breast carcinoma amplified sequence 2	2
674	P56211|ARP19_HUMAN	cAMP-regulated phosphoprotein 19	2
675	Q8IUL8|CILP2_HUMAN	Cartilage intermediate layer protein 2 precursor	2
676	P52943|CRIP2_HUMAN	Cysteine-rich protein 2	2
677	Q9HC35|EMAL4_HUMAN	Echinoderm microtubule-associated protein-like 4	2
678	Q13541|4EBP1_HUMAN	Eukaryotic translation initiation factor 4E-binding protein 1	2
679	Q15056|IF4H_HUMAN	Eukaryotic translation initiation factor 4H	2
680	P55060|XPO2_HUMAN	Exportin-2	2
681	P62942|FKB1A_HUMAN	FK506-binding protein 1A	2
682	Q02790|FKBP4_HUMAN	FK506-binding protein 4	2
683	P15170|GSPT1_HUMAN	G1 to S phase transition protein 1 homolog	2
684	Q1KMD3|HNRL2_HUMAN	Heterogeneous nuclear ribonucleoprotein U-like protein 2	2
685	P12268|IMDH2_HUMAN	Inosine-5'-monophosphate dehydrogenase 2	2
686	P82970|NSBP1_HUMAN	Nucleosome-binding protein 1	2
687	O75400|PRP40_HUMAN	Pre-mRNA-processing factor 40 homolog A	2
688	Q9BUQ8|DDX23_HUMAN	Probable ATP-dependent RNA helicase DDX23	2
689	P55145|ARMET_HUMAN	Protein ARMET precursor	2
690	Q96DB5|FA82B_HUMAN	Protein FAM82B	2
691	Q9Y570|PPME1_HUMAN	Protein phosphatase methylesterase 1	2
692	Q9NTK5|GTPB9_HUMAN	Putative GTP-binding protein 9	2
693	Q9Y265|RUVB1_HUMAN	RuvB-like 1	2
694	O95171|SCEL_HUMAN	Sciellin	2
695	Q8IYB3|SRRM1_HUMAN	Serine/arginine repetitive matrix protein 1	2
696	Q9Y3A5|SBDS_HUMAN	Shwachman-Bodian-Diamond syndrome protein	2
697	P62316|SMD2_HUMAN	Small nuclear ribonucleoprotein Sm D2	2
698	O75347|TBCA_HUMAN	Tubulin-specific chaperone A	2
699	P61088|UBE2N_HUMAN	Ubiquitin-conjugating enzyme E2 N	2
700	Q9NQG5|CT077_HUMAN	Uncharacterized protein C20orf77	2
701	P08670|VIME_HUMAN	Vimentin	2
702	Q9NRX4|PHP14_HUMAN	14 kDa phosphohistidine phosphatase	2
703	Q04917|1433F_HUMAN	14-3-3 protein eta	2
704	Q01970|PLCB3_HUMAN	1-phosphatidylinositol-4,5-bisphosphate phosphodiesterase beta 3	2
705	O00232|PSD12_HUMAN	26S proteasome non-ATPase regulatory subunit 12	2
706	P48556|PSMD8_HUMAN	26S proteasome non-ATPase regulatory subunit 8	2
707	O95861|BPNT1_HUMAN	3'(2'),5'-bisphosphate nucleotidase 1	2
708	P46783|RS10_HUMAN	40S ribosomal protein S10	2
709	P39019|RS19_HUMAN	40S ribosomal protein S19	2
710	P63220|RS21_HUMAN	40S ribosomal protein S21	2
711	P62857|RS28_HUMAN	40S ribosomal protein S28	2
712	P62241|RS8_HUMAN	40S ribosomal protein S8	2
713	P30050|RL12_HUMAN	60S ribosomal protein L12	2
714	P61313|RL15_HUMAN	60S ribosomal protein L15	2
715	Q02878|RL6_HUMAN	60S ribosomal protein L6	2
716	P17858|K6PL_HUMAN	6-phosphofructokinase, liver type	2
717	O15145|ARPC3_HUMAN	Actin-related protein 2/3 complex subunit 3	2
718	P13798|ACPH_HUMAN	Acylamino-acid-releasing enzyme	2
719	P07741|APT_HUMAN	Adenine phosphoribosyltransferase	2
720	P30520|PURA2_HUMAN	Adenylosuccinate synthetase isozyme 2	2
721	Q01518|CAP1_HUMAN	Adenylyl cyclase-associated protein 1	2
722	Q9NZ08|ARTS1_HUMAN	Adipocyte-derived leucine aminopeptidase precursor	2
723	P14550|AK1A1_HUMAN	Alcohol dehydrogenase [NADP+]	2
724	P55008|AIF1_HUMAN	Allograft inflammatory factor 1	2
725	Q9BYC5|FUT8_HUMAN	Alpha-(1,6)-fucosyltransferase	2
726	P01023|A2MG_HUMAN	Alpha-2-macroglobulin precursor	2
727	Q03154|ACY1_HUMAN	Aminoacylase-1	2

728	Q86SJ2|AMGO2_HUMAN	Amphoterin-induced protein 2 precursor	2
729	Q9Y6Q5|AP1M2_HUMAN	AP-1 complex subunit mu-2	2
730	Q8NCW5|AIBP_HUMAN	Apolipoprotein A-I-binding protein precursor	2
731	P15289|ARSA_HUMAN	Arylsulfatase A precursor	2
732	Q8WWM7|ATX2L_HUMAN	Ataxin-2-like protein	2
733	Q8NE71|ABCF1_HUMAN	ATP-binding cassette sub-family F member 1	2
734	O95817|BAG3_HUMAN	BAG family molecular chaperone regulator 3	2
735	O75531|BAF_HUMAN	Barrier-to-autointegration factor	2
736	O60512|B4GT3_HUMAN	Beta-1,4-galactosyltransferase 3	2
737	P07814|SYEP_HUMAN	Bifunctional aminoacyl-tRNA synthetase	2
738	P43251|BTD_HUMAN	Biotinidase precursor	2
739	P27708|PYR1_HUMAN	CAD protein	2
740	P55285|CADH6_HUMAN	Cadherin-6 precursor	2
741	Q9HB71|CYBP_HUMAN	Calcyclin-binding protein	2
742	P13861|KAP2_HUMAN	cAMP-dependent protein kinase type II-alpha regulatory subunit	2
743	Q14444|CAPR1_HUMAN	Caprin-1	2
744	Q8NCH0|CHSTE_HUMAN	Carbohydrate sulfotransferase D4ST1	2
745	P16152|CBR1_HUMAN	Carbonyl reductase [NADPH] 1	2
746	O60911|CATL2_HUMAN	Cathepsin L2 precursor	2
747	Q9UBR2|CATZ_HUMAN	Cathepsin Z precursor	2
748	P16070|CD44_HUMAN	CD44 antigen precursor	2
749	P21926|CD9_HUMAN	CD9 antigen	2
750	P60953|CDC42_HUMAN	Cell division control protein 42 homolog precursor	2
751	P62633|CNBP_HUMAN	Cellular nucleic acid-binding protein	2
752	O75503|CLN5_HUMAN	Ceroid-lipofuscinosis neuronal protein 5	2
753	O95232|CROP_HUMAN	Cisplatin resistance-associated overexpressed protein	2
754	Q16630|CPSF6_HUMAN	Cleavage and polyadenylation specificity factor 6	2
755	P27658|CO8A1_HUMAN	Collagen alpha-1(VIII) chain precursor	2
756	P08174|DAF_HUMAN	Complement decay-accelerating factor precursor	2
757	Q99829|CPNE1_HUMAN	Copine-1	2
758	O75131|CPNE3_HUMAN	Copine-3	2
759	Q13620|CUL4B_HUMAN	Cullin-4B	2
760	P33552|CKS2_HUMAN	Cyclin-dependent kinases regulatory subunit 2	2
761	P99999|CYC_HUMAN	Cytochrome c	2
762	O43175|SERA_HUMAN	D-3-phosphoglycerate dehydrogenase	2
763	Q96EP5|DAZP1_HUMAN	DAZ-associated protein 1	2
764	Q16555|DPYL2_HUMAN	Dihydropyrimidinase-related protein 2	2
765	P26358|DNMT1_HUMAN	DNA	2
766	P33993|MCM7_HUMAN	DNA replication licensing factor MCM7	2
767	P11387|TOP1_HUMAN	DNA topoisomerase 1	2
768	Q9UJU6|DBNL_HUMAN	Drebrin-like protein	2
769	Q7Z6Z7|HUWE1_HUMAN	E3 ubiquitin-protein ligase HUWE1	2
770	Q9BUP0|EFHD1_HUMAN	EF-hand domain-containing protein 1	2
771	P24534|EF1B_HUMAN	Elongation factor 1-beta	2
772	P50402|EMD_HUMAN	Emerin	2
773	P05198|IF2A_HUMAN	Eukaryotic translation initiation factor 2 subunit 1	2
774	O00303|IF35_HUMAN	Eukaryotic translation initiation factor 3 subunit 5	2
775	Q9Y262|IF36I_HUMAN	Eukaryotic translation initiation factor 3 subunit 6-interacting protein	2
776	O15371|IF37_HUMAN	Eukaryotic translation initiation factor 3 subunit 7	2
777	P23588|IF4B_HUMAN	Eukaryotic translation initiation factor 4B	2
778	P06730|IF4E_HUMAN	Eukaryotic translation initiation factor 4E	2
779	P56537|IF6_HUMAN	Eukaryotic translation initiation factor 6	2
780	O14980|XPO1_HUMAN	Exportin-1	2
781	Q01469|FABPE_HUMAN	Fatty acid-binding protein, epidermal	2
782	O95750|FGF19_HUMAN	Fibroblast growth factor 19 precursor	2
783	Q8N441|FGRL1_HUMAN	Fibroblast growth factor receptor-like 1 precursor	2
784	P30043|BLVRB_HUMAN	Flavin reductase	2
785	P09972|ALDOC_HUMAN	Fructose-bisphosphate aldolase C	2
786	P13284|GILT_HUMAN	Gamma-interferon-inducible lysosomal thiol reductase precursor	2
787	P17900|SAP3_HUMAN	Ganglioside GM2 activator precursor	2
788	O60763|VDP_HUMAN	General vesicular transport factor p115	2
789	P47897|SYQ_HUMAN	Glutaminyl-tRNA synthetase	2
790	Q9H8Y8|GORS2_HUMAN	Golgi reassembly-stacking protein 2	2
791	Q14344|GNA13_HUMAN	Guanine nucleotide-binding protein alpha-13 subunit	2
792	P04899|GNAI2_HUMAN	Guanine nucleotide-binding protein G(i), alpha-2 subunit	2
793	P62879|GBB2_HUMAN	Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta 2	2
794	Q9UK76|HN1_HUMAN	Hematological and neurological expressed 1 protein	2
795	P52597|HNRPF_HUMAN	Heterogeneous nuclear ribonucleoprotein F	2
796	P14866|HNRPL_HUMAN	Heterogeneous nuclear ribonucleoprotein L	2
797	P37235|HPCL1_HUMAN	Hippocalcin-like protein 1	2
798	Q92769|HDAC2_HUMAN	Histone deacetylase 2	2
799	Q8IUE6|H2A2B_HUMAN	Histone H2A type 2-B	2
800	P68431|H31_HUMAN	Histone H3.1	2
801	P30464|1B15_HUMAN	HLA class I histocompatibility antigen, B-15 alpha chain precursor	2
802	P00492|HPRT_HUMAN	Hypoxanthine-guanine phosphoribosyltransferase	2
803	P22304|IDS_HUMAN	Iduronate 2-sulfatase precursor	2
804	Q8TEX9|IPO4_HUMAN	Importin-4	2
805	O95373|IPO7_HUMAN	Importin-7	2
806	P17936|IBP3_HUMAN	Insulin-like growth factor-binding protein 3 precursor	2
807	P22692|IBP4_HUMAN	Insulin-like growth factor-binding protein 4 precursor	2
808	P05161|UCRP_HUMAN	Interferon-induced 17 kDa protein precursor [Contains: Ubiquitin cross-reactive protein	2
809	P10145|IL8_HUMAN	Interleukin-8 precursor	2
810	Q96CN7|ISOC1_HUMAN	Isochorismatase domain-containing protein 1	2
811	P35527|K1C9_HUMAN	Keratin, type I cytoskeletal 9	2
812	Q07666|SAM68_HUMAN	KH domain-containing, RNA-binding, signal transduction-associated protein 1	2
813	P33176|KINH_HUMAN	Kinesin heavy chain	2
814	O43291|SPIT2_HUMAN	Kunitz-type protease inhibitor 2 precursor	2
815	P02788|TRFL_HUMAN	Lactotransferrin precursor	2
816	O00515|LAD1_HUMAN	Ladinin 1	2
817	Q9NRN7|ADPPT_HUMAN	L-aminoadipate-semialdehyde dehydrogenase-phosphopantetheinyl transferase	2
818	Q01650|LAT1_HUMAN	Large neutral amino acids transporter small subunit 1	2
819	Q14766|LTB1L_HUMAN	Latent-transforming growth factor beta-binding protein, isoform 1L precursor	2
820	P14151|LYAM1_HUMAN	L-selectin precursor	2
821	P13473|LAMP2_HUMAN	Lysosome-associated membrane glycoprotein 2 precursor	2
822	P14174|MIF_HUMAN	Macrophage migration inhibitory factor	2
823	Q14764|MVP_HUMAN	Major vault protein	2
824	O60664|M6PBP_HUMAN	Mannose-6-phosphate receptor-binding protein 1	2
825	Q9UJH8|METRN_HUMAN	Meteorin precursor	2
826	P27816|MAP4_HUMAN	Microtubule-associated protein 4	2
827	Q15691|MARE1_HUMAN	Microtubule-associated protein RP/EB family member 1	2
828	Q9Y2B0|MSAP_HUMAN	MIR-interacting saposin-like protein precursor	2
829	O15427|MOT4_HUMAN	Monocarboxylate transporter 4	2
830	Q9UKD2|MRT4_HUMAN	mRNA turnover protein 4 homolog	2
831	Q9HC84|MUC5B_HUMAN	Mucin-5B precursor	2
832	Q13155|MCA2_HUMAN	Multisynthetase complex auxiliary component p38	2
833	P20933|ASPG_HUMAN	N(4)-(beta-N-acetylglucosaminyl)-L-asparaginase precursor	2
834	Q9HB63|NET4_HUMAN	Netrin-4 precursor	2
835	Q9GZT8|NIF3L_HUMAN	NIF3-like protein 1	2
836	Q9BZM5|N2DL2_HUMAN	NKG2D ligand 2 precursor	2
837	P61970|NTF2_HUMAN	Nuclear transport factor 2	2
838	Q9H1E3|NUCKS_HUMAN	Nuclear ubiquitous casein and cyclin-dependent kinases substrate	2
839	Q14978|NOLC1_HUMAN	Nucleolar phosphoprotein p130	2
840	Q13232|NDK3_HUMAN	Nucleoside diphosphate kinase 3	2
841	Q8WXF1|PSPC1_HUMAN	Paraspeckle component 1	2
842	P01127|PDGFB_HUMAN	Platelet-derived growth factor B chain precursor	2
843	Q13310|PABP4_HUMAN	Polyadenylate-binding protein 4	2
844	P35080|PROF2_HUMAN	Profilin-2	2
845	Q8WUM4|PDC6I_HUMAN	Programmed cell death 6-interacting protein	2
846	O14737|PDCD5_HUMAN	Programmed cell death protein 5	2
847	O75340|PDCD6_HUMAN	Programmed cell death protein 6	2
848	P61289|PSME3_HUMAN	Proteasome activator complex subunit 3	2
849	P49720|PSB3_HUMAN	Proteasome subunit beta type 3	2

850	P28074|PSB5_HUMAN	Proteasome subunit beta type 5 precursor	2
851	P35659|DEK_HUMAN	Protein DEK	2
852	Q9C0B1|FATSO_HUMAN	Protein fatso	2
853	Q96C90|PP14B_HUMAN	Protein phosphatase 1 regulatory subunit 14B	2
854	P21980|TGM2_HUMAN	Protein-glutamine gamma-glutamyltransferase 2	2
855	Q08257|QOR_HUMAN	Quinone oxidoreductase	2
856	P31150|GDIA_HUMAN	Rab GDP dissociation inhibitor alpha	2
857	P61026|RAB10_HUMAN	Ras-related protein Rab-10	2
858	Q13332|PTPRS_HUMAN	Receptor-type tyrosine-protein phosphatase S precursor	2
859	P11908|PRPS2_HUMAN	Ribose-phosphate pyrophosphokinase II	2
860	P98175|RBM10_HUMAN	RNA-binding protein 10	2
861	Q15413|RYR3_HUMAN	Ryanodine receptor 3	2
862	P31153|METK2_HUMAN	S-adenosylmethionine synthetase isoform type-2	2
863	Q14151|SAFB2_HUMAN	Scaffold attachment factor B2	2
864	Q96C86|DCPS_HUMAN	Scavenger mRNA-decapping enzyme DcpS	2
865	Q9UQ35|SRRM2_HUMAN	Serine/arginine repetitive matrix protein 2	2
866	P60510|PP4C_HUMAN	Serine/threonine-protein phosphatase 4 catalytic subunit	2
867	P53041|PPP5_HUMAN	Serine/threonine-protein phosphatase 5	2
868	P62140|PP1B_HUMAN	Serine/threonine-protein phosphatase PP1-beta catalytic subunit	2
869	Q9NR45|SIAS_HUMAN	Sialic acid synthase	2
870	Q99519|NEUR1_HUMAN	Sialidase-1 precursor	2
871	P63165|SUMO1_HUMAN	Small ubiquitin-related modifier 1 precursor	2
872	P61956|SUMO2_HUMAN	Small ubiquitin-related modifier 2 precursor	2
873	Q8WVQ1|CANT1_HUMAN	Soluble calcium-activated nucleotidase 1	2
874	Q99523|SORT_HUMAN	Sortilin precursor	2
875	Q15428|SF3A2_HUMAN	Splicing factor 3A subunit 2	2
876	Q9BWJ5|SF3B5_HUMAN	Splicing factor 3B subunit 5	2
877	Q13243|SFRS5_HUMAN	Splicing factor, arginine/serine-rich 5	2
878	Q6ZRP7|QSOX2_HUMAN	Sulfhydryl oxidase 2 precursor	2
879	Q99536|VAT1_HUMAN	Synaptic vesicle membrane protein VAT-1 homolog	2
880	Q13148|TADBP_HUMAN	TAR DNA-binding protein 43	2
881	P13686|PPA5_HUMAN	Tartrate-resistant acid phosphatase type 5 precursor	2
882	Q9UGI8|TES_HUMAN	Testin	2
883	P52888|MEPD_HUMAN	Thimet oligopeptidase	2
884	O95881|TXD12_HUMAN	Thioredoxin domain-containing protein 12 precursor	2
885	Q86V81|THOC4_HUMAN	THO complex subunit 4	2
886	P23193|TCEA1_HUMAN	Transcription elongation factor A protein 1	2
887	Q13595|TRA2A_HUMAN	Transformer-2 protein homolog	2
888	P61812|TGFB2_HUMAN	Transforming growth factor beta-2 precursor	2
889	O43493|TGON2_HUMAN	Trans-Golgi network integral membrane protein 2 precursor	2
890	Q8WUA8|TSK_HUMAN	Tsukushi precursor	2
891	O75509|TNR21_HUMAN	Tumor necrosis factor receptor superfamily member 21 precursor	2
892	P09758|TACD2_HUMAN	Tumor-associated calcium signal transducer 2 precursor	2
893	O75643|U520_HUMAN	U5 small nuclear ribonucleoprotein 200 kDa helicase	2
894	P07919|UCRH_HUMAN	Ubiquinol-cytochrome c reductase complex 11 kDa protein, mitochondrial precursor	2
895	P15374|UCHL3_HUMAN	Ubiquitin carboxyl-terminal hydrolase isozyme L3	2
896	A0AVT1|UB1L2_HUMAN	Ubiquitin-activating enzyme E1-like protein 2	2
897	Q13404|UB2V1_HUMAN	Ubiquitin-conjugating enzyme E2 variant 1	2
898	Q9NYU2|UGGG1_HUMAN	UDP-glucose:glycoprotein glucosyltransferase 1 precursor	2
899	Q9H3U1|UN45A_HUMAN	UNC45 homolog A	2
900	Q9NWV4|CA123_HUMAN	Uncharacterized protein C1orf123	2
901	Q96HY6|CT116_HUMAN	Uncharacterized protein C20orf116 precursor	2
902	P38606|VATA_HUMAN	Vacuolar ATP synthase catalytic subunit A	2
903	P21281|VATB2_HUMAN	Vacuolar ATP synthase subunit B, brain isoform	2
904	Q15904|VAS1_HUMAN	Vacuolar ATP synthase subunit S1 precursor	2
905	Q9UBQ0|VPS29_HUMAN	Vacuolar protein sorting-associated protein 29	2
906	Q9C0J8|WDR33_HUMAN	WD repeat protein 33	2
907	P12955|PEPD_HUMAN	Xaa-Pro dipeptidase	2
908	Q9H773|XTP3A_HUMAN	XTP3-transactivated gene A protein	2
909	P49006|MRP_HUMAN	MARCKS-related protein	2

Supplemental table S2

A

Protein name	P/S	A	M/I
Putative quinone oxidoreductase	- [1]		
Growth/differentiation factor 15	- [2, 3]		
Phospholipid transfer protein	- [4]		
Cadherin-6			- [5]
Histone H2A type 1-A			
Tumor necrosis factor receptor superfamily member 10C *	+ [6]		
Histone H4			
Tissue alpha-L-fucosidase			
Histone H2B type 1-L			
Syndecan-4	- [7]		
Insulin-like growth factor-binding protein 3	- [8, 9]		
EH domain-containing protein 1			
Lysosomal alpha-mannosidase			
Versican core protein	- [10]		- [11]
Adenylate kinase isoenzyme 1			
Solute carrier family 2. facilitated glucose transporter member 1			
Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta 2			
Catenin alpha-1	- [12]		
Serpin B5	- [13]	- [13]	- [13, 14]
Acid ceramidase	+ [15]		
Junctional adhesion molecule A	- [16]		
Ephrin type-A receptor 2		- [17]	- [17]
Calnexin			
Catenin beta-1	- [18, 19]		
Lysosomal protective protein			
Plastin-3	- [20]		
Leukocyte elastase inhibitor			- [21, 22]
Glypican-1			- [23]
Suppressor of tumorigenicity protein 14			+ [24]
Glutathione reductase. mitochondrial			
Calpain small subunit 1	- [25, 26]		
Proactivator polypeptide			- [27]
Tyrosine-protein kinase-like 7			
Cadherin-related tumor suppressor homolog	- [28, 29]		
Protein FAM3C			+ [30]
Plectin-1			
MARCKS-related protein	- [31]		
Catenin delta-1	- [32-34]		
Brain acid soluble protein 1			
Lactadherin		+ [35]	
Gamma-synuclein	+ [36]		
Puromycin-sensitive aminopeptidase	- [37]		
Cell division control protein 42 homolog	- [38]		
Endothelial protein C receptor	+ [39]		
Epithelial-cadherin	- [18, 40]		- [41]
14-3-3 protein sigma	- [42, 43]		
Protein S100-A6	- [44]		- [45]
14-3-3 protein zeta/delta	+ [46]		
NAD(P)H dehydrogenase [quinone] 1			
Ephrin-A1	- [47]	- [48]	

B

Protein name	P/S	A	M/I
Fibroblast growth factor-binding protein 1		+ [49, 50]	
Zinc transporter ZIP10			+ [51, 52]
Fibroblast growth factor 19	+ [53, 54]		
Tissue-type plasminogen activator	+ [55, 56]		+ [56]
Gamma-glutamyl hydrolase			
Vascular endothelial growth factor A		+ [57-59]	
Clusterin			+ [60, 61]
Kallikrein-6			+ [62]
Osteopontin *	+ [63]	+ [64]	+ [65]
Beta-2-microglobulin	+ [66]		
Deoxyuridine 5'-triphosphate nucleotidohydrolase. mitochondrial	+ [67]		
Fibulin-2			+ [68]
Cyclin-dependent kinases regulatory subunit 2	+ [69]		
Fibulin-1		+ [70]	+ [71]
Insulin-like growth factor II *		+ [72]	+ [73]
Gamma-interferon-inducible lysosomal thiol reductase			
Polypeptide N-acetylgalactosaminyltransferase 6			
Thrombospondin-1		+ [74, 75]	
Tenascin	+ [76]	+ [77]	+ [78]
Laminin subunit beta-1		+ [79]	+ [79, 80]
Insulin-like growth factor-binding protein 4	- [81]	- [81]	
Transferrin receptor protein 1			
DNA (cytosine-5)-methyltransferase 1	+ [82]		+ [83]
EGF-containing fibulin-like extracellular matrix protein 1			
Epididymal secretory protein E1			
Vimentin			+ [84, 85]
Metalloproteinase inhibitor 2	+ [86, 87]		
Insulin-like growth factor-binding protein 2	+ [88]		+ [89]
Cadherin-2		+ [90]	+ [90, 91]
DNA replication licensing factor MCM6			
Amphiregulin	+ [92, 93]		+ [94]
Legumain			+ [95]
Dipeptidyl-peptidase 1	+ [96]	+ [96]	+ [96]
Beta-1.3-galactosyl-O-glycosyl-glycoprotein beta-1.6-N-acetylglucosaminyltransferase			+ [97, 98]
Interleukin-8	+ [99]	+ [100]	+ [99, 101]
Urokinase-type plasminogen activator	+ [102]		+ [103]
Stanniocalcin-2			+ [104]
Cathepsin L2			+ [105]
Dickkopf-related protein 1	- [106]		+ [107]
Lysyl oxidase homolog 2	+ [108]		+ [108]
Beta-1.3-N-acetylglucosaminyltransferase lunatic fringe			
Transforming growth factor beta-2		+ [109]	+ [110]
Metalloproteinase inhibitor 1	- / + [111, 112]		
Transforming growth factor-beta-induced protein ig-h3			+ [113, 114]

References for supplemental table S2

1. P. M. Flatt, K. Polyak, L. J. Tang, C. D. Scatena, M. D. Westfall, L. A. Rubinstein, J. Yu, K. W. Kinzler, B. Vogelstein, D. E. Hill and J. A. Pietenpol, 'p53-dependent expression of PIG3 during proliferation, genotoxic stress, and reversible growth arrest', *Cancer Lett*, 156 (2000), pp. 63-72.
2. M. K. Agarwal, K. Hastak, M. W. Jackson, S. N. Breit, G. R. Stark and M. L. Agarwal, 'Macrophage inhibitory cytokine 1 mediates a p53-dependent protective arrest in S phase in response to starvation for DNA precursors', *Proc Natl Acad Sci U S A*, 103 (2006), pp. 16278-83.
3. T. Liu, A. R. Bauskin, J. Zaunders, D. A. Brown, S. Pankhurst, P. J. Russell and S. N. Breit, 'Macrophage inhibitory cytokine 1 reduces cell adhesion and induces apoptosis in prostate cancer cells', *Cancer Res*, 63 (2003), pp. 5034-40.
4. A. Wehinger, I. Tancevski, W. Schgoer, P. Eller, K. Hochegger, M. Morak, A. Hermetter, A. Ritsch, J. R. Patsch and B. Foeger, 'Phospholipid transfer protein augments apoptosis in THP-1-derived macrophages induced by lipolyzed hypertriglyceridemic plasma', *Arterioscler Thromb Vasc Biol*, 27 (2007), pp. 908-15.
5. T. Shimazui, E. Oosterwijk, H. Akaza, P. Bringuier, E. Ruijter, H. van Berkel, J. O. Wakka, A. van Bokhoven, F. M. Debruyne and J. A. Schalken, 'Expression of cadherin-6 as a novel diagnostic tool to predict prognosis of patients with E-cadherin-absent renal cell carcinoma', *Clin Cancer Res*, 4 (1998), pp. 2419-24.
6. H. Mitsuyama, W. Matsuyama, M. Watanabe, Y. Shirahama, I. Higashimoto, T. Wada, M. Osame and K. Arimura, 'Increased expression of TRAIL receptor 3 on eosinophils in Churg-Strauss syndrome', *Arthritis Rheum*, 56 (2007), pp. 662-73.
7. V. Todorovicc, C. C. Chen, N. Hay and L. F. Lau, 'The matrix protein CCN1 (CYR61) induces apoptosis in fibroblasts', *J Cell Biol*, 171 (2005), pp. 559-68.
8. A. Grimberg, C. M. Coleman, T. F. Burns, B. P. Himelstein, C. J. Koch, P. Cohen and W. S. El-Deiry, 'p53-Dependent and p53-independent induction of insulin-like growth factor binding protein-3 by deoxyribonucleic acid damage and hypoxia', *J Clin Endocrinol Metab*, 90 (2005), pp. 3568-74.
9. P. Cohen, 'Insulin-like growth factor binding protein-3: insulin-like growth factor independence comes of age', *Endocrinology*, 147 (2006), pp. 2109-11.
10. Y. Wu, J. Wu, D. Y. Lee, A. Yee, L. Cao, Y. Zhang, C. Kiani and B. B. Yang, 'Versican protects cells from oxidative stress-induced apoptosis', *Matrix Biol*, 24 (2005), pp. 3-13.
11. Y. Wu, L. Chen, P. S. Zheng and B. B. Yang, 'beta 1-Integrin-mediated glioma cell adhesion and free radical-induced apoptosis are regulated by binding to a C-terminal domain of PG-M/versican', *J Biol Chem*, 277 (2002), pp. 12294-301.
12. S. Matsubara and M. Ozawa, 'Expression of alpha-catenin in alpha-catenin-deficient cells results in a reduced proliferation in three-dimensional multicellular spheroids but not in two-dimensional monolayer cultures', *Oncogene*, 23 (2004), pp. 2694-702.
13. N. Maass, K. Nagasaki, M. Ziebart, C. Mundhenke and W. Jonat, 'Expression and regulation of tumor suppressor gene maspin in breast cancer', *Clin Breast Cancer*, 3 (2002), pp. 281-7.
14. E. I. Chen and J. R. Yates, 'Maspin and tumor metastasis', *IUBMB Life*, 58 (2006), pp. 25-9.
15. A. Strelow, K. Bernardo, S. Adam-Klages, T. Linke, K. Sandhoff, M. Kronke and D. Adam, 'Overexpression of acid ceramidase protects from tumor necrosis factor-induced cell death', *J Exp Med*, 192 (2000), pp. 601-12.
16. C. Jacquot, X. Lanco, D. Carbonnelle, O. Sevestre, C. Tomasoni, G. Briad, M. Juget, V. Roussis and C. Roussakis, 'Effect of four genes (ALDH1, NRF1, JAM and KBL) on proliferation arrest in a non-small cell bronchopulmonary cancer line', *Anticancer Res*, 22 (2002), pp. 2229-35.
17. D. M. Brantley, N. Cheng, E. J. Thompson, Q. Lin, R. A. Brekken, P. E. Thorpe, R. S. Muraoka, D. P. Cerretti, A. Pozzi, D. Jackson, C. Lin and J. Chen, 'Soluble Eph A receptors inhibit tumor angiogenesis and progression in vivo', *Oncogene*, 21 (2002), pp. 7011-26.
18. A. M. Lowy, J. Knight and J. Groden, 'Restoration of E-cadherin/beta-catenin expression in pancreatic cancer cells inhibits growth by induction of apoptosis', *Surgery*, 132 (2002), pp. 141-8.
19. D. Olmeda, S. Castel, S. Vilaro and A. Cano, 'Beta-catenin regulation during the cell cycle: implications in G2/M and apoptosis', *Mol Biol Cell*, 14 (2003), pp. 2844-60.
20. Y. Sasaki, F. Itoh, T. Kobayashi, T. Kikuchi, H. Suzuki, M. Toyota and K. Imai, 'Increased expression of T-fimbrin gene after DNA damage in CHO cells and inactivation of T-fimbrin by CpG methylation in human colorectal cancer cells', *Int J Cancer*, 97 (2002), pp. 211-6.
21. J. Cooley, T. K. Takayama, S. D. Shapiro, N. M. Schechter and E. Remold-O'Donnell, 'The serpin MNEI inhibits elastase-like and chymotrypsin-like serine proteases through efficient reactions at two active sites', *Biochemistry*, 40 (2001), pp. 15762-70.
22. T. M. Hong, P. C. Yang, K. Peck, J. J. Chen, S. C. Yang, Y. C. Chen and C. W. Wu, 'Profiling the downstream genes of tumor suppressor PTEN in lung cancer cells by complementary DNA microarray', *Am J Respir Cell Mol Biol*, 23 (2000), pp. 355-63.
23. W. Liu, E. D. Litwack, M. J. Stanley, J. K. Langford, A. D. Lander and R. D. Sanderson, 'Heparan sulfate proteoglycans as adhesive and anti-invasive molecules. Syndecans and glypican have distinct functions', *J Biol Chem*, 273 (1998), pp. 22825-32.
24. K. F. Ding, L. F. Sun, W. T. Ge, H. G. Hu, S. Z. Zhang and S. Zheng, 'Effect of SNC19/ST14 gene overexpression on invasion of colorectal cancer cells', *World J Gastroenterol*, 11 (2005), pp. 5651-4.
25. Y. Tan, C. Wu, T. De Veyra and P. A. Greer, 'Ubiquitous calpains promote both apoptosis and survival signals in response to different cell death stimuli', *J Biol Chem*, 281 (2006), pp. 17689-98.
26. M. H. Kubbutat and K. H. Vousden, 'Proteolytic cleavage of human p53 by calpain: a potential regulator of protein stability', *Mol Cell Biol*, 17 (1997), pp. 460-8.
27. Y. Sun, X. Qi and G. A. Grabowski, 'Saposin C is required for normal resistance of acid beta-glucosidase to proteolytic degradation', *J Biol Chem*, 278 (2003), pp. 31918-23.
28. T. Tanoue and M. Takeichi, 'Mammalian Fat1 cadherin regulates actin dynamics and cell-cell contact', *J Cell Biol*, 165 (2004), pp. 517-28.
29. R. Hou, L. Liu, S. Anees, S. Hiroyasu and N. E. Sibinga, 'The Fat1 cadherin integrates vascular smooth muscle cell growth and migration signals', *J Cell Biol*, 173 (2006), pp. 417-29.
30. T. Waerner, M. Alacakaptan, I. Tamir, R. Oberauer, A. Gal, T. Brabletz, M. Schreiber, M. Jechlinger and H. Beug, 'ILEI: a cytokine essential for EMT, tumor formation, and late events in metastasis in epithelial cells', *Cancer Cell*, 10 (2006), pp. 227-39.
31. J. J. Ramsden, 'MARCKS: a case of molecular exaptation?' *Int J Biochem Cell Biol*, 32 (2000), pp. 475-9.
32. M. A. Thoreson and A. B. Reynolds, 'Altered expression of the catenin p120 in human cancer: implications for tumor progression', *Differentiation*, 70 (2002), pp. 583-9.
33. A. Ferber, C. Yaen, E. Sarmiento and J. Martinez, 'An octapeptide in the juxtamembrane domain of VE-cadherin is important for p120ctn binding and cell proliferation', *Exp Cell Res*, 274 (2002), pp. 35-44.
34. C. Z. Nong, L. L. Pan, W. S. He, X. L. Zha, H. H. Ye and H. Y. Huang, 'P120ctn overexpression enhances beta-catenin-E-cadherin binding and down regulates expression of survivin and cyclin D1 in BEL-7404 hepatoma cells', *World J Gastroenterol*, 12 (2006), pp. 1187-91.
35. J. S. Silvestre, C. Thery, G. Hamard, J. Boddaert, B. Aguilar, A. Delcayre, C. Houbron, R. Tamarat, O. Blanc-Brude, S. Heeneman, M. Clergue, M. Duriez, R. Merval, B. Levy, A. Tedgui, S. Amigorena and Z. Mallat, 'Lactadherin promotes VEGF-dependent neovascularization', *Nat Med*, 11 (2005), pp. 499-506.
36. Z. Z. Pan, W. Bruening, B. I. Giasson, V. M. Lee and A. K. Godwin, 'Gamma-synuclein promotes cancer cell survival and inhibits stress- and chemotherapy drug-induced apoptosis by modulating MAPK pathways', *J Biol Chem*, 277 (2002), pp. 35050-60.
37. J. M. de Gandarias, J. Irazusta, J. Gil, D. Fernandez, A. Varona and L. Casis, 'Ontogeny of puromycin-sensitive and insensitive aminopeptidase activities in several subcellular fractions of the rat brain', *Brain Res Bull*, 50 (1999), pp. 283-90.
38. C. F. Deroanne, D. Hamelryckx, T. T. Ho, C. A. Lambert, P. Catroux, C. M. Lapiere and B. V. Nusgens, 'Cdc42 downregulates MMP-1 expression by inhibiting the ERK1/2 pathway', *J Cell Sci*, 118 (2005), pp. 1173-83.
39. T. Cheng, D. Liu, J. H. Griffin, J. A. Fernandez, F. Castellino, E. D. Rosen, K. Fukudome and B. V. Zlokovic, 'Activated protein C blocks p53-mediated apoptosis in ischemic human brain endothelium and is neuroprotective', *Nat Med*, 9 (2003), pp. 338-42.
40. T. Ohira, R. M. Gemmill, K. Ferguson, S. Kusy, J. Roche, E. Brambilla, C. Zeng, A. Baron, L. Bemis, P. Erickson, E. Wilder, A. Rustgi, J. Kitajewski, E. Gabrielson, R. Bremnes, W. Franklin and H. A. Drabkin, 'WNT7a induces E-cadherin in lung cancer cells', *Proc Natl Acad Sci U S A*, 100 (2003), pp. 10429-34.
41. A. S. Wong and B. M. Gumbiner, 'Adhesion-independent mechanism for suppression of tumor cell invasion by E-cadherin', *J Cell Biol*, 161 (2003), pp. 1191-203.
42. H. Y. Yang, Y. Y. Wen, C. H. Chen, G. Lozano and M. H. Lee, '14-3-3 sigma positively regulates p53 and suppresses tumor growth', *Mol Cell Biol*, 23 (2003), pp. 7096-107.
43. H. Yang, R. Zhao and M. H. Lee, '14-3-3sigma, a p53 regulator, suppresses tumor growth of nasopharyngeal carcinoma', *Mol Cancer Ther*, 5 (2006), pp. 253-60.
44. J. H. Joo, S. Y. Yoon, J. H. Kim, S. G. Paik, S. R. Min, J. S. Lim, I. S. Choe, I. Choi and J. W. Kim, 'S100A6 (calcyclin) enhances the sensitivity to apoptosis via the upregulation of caspase-3 activity in Hep3B cells', *J Cell Biochem*, 103 (2008), pp. 1183-97.
45. H. H. Luu, L. Zhou, R. C. Haydon, A. T. Deyrup, A. G. Montag, D. Huo, R. Heck, C. W. Heizmann, T. D. Peabody, M. A. Simon and T. C. He, 'Increased expression of S100A6 is associated with decreased metastasis and inhibition of cell migration and anchorage independent growth in human osteosarcoma', *Cancer Lett*, 229 (2005), pp. 135-48.
46. T. Fan, R. Li, N. W. Todd, Q. Qiu, H. B. Fang, H. Wang, J. Shen, R. Y. Zhao, N. P. Caraway, R. L. Katz, S. A. Stass and F. Jiang, 'Up-regulation of 14-3-3zeta in lung cancer and its implication as prognostic and therapeutic target', *Cancer Res*, 67 (2007), pp. 7901-6.
47. D. P. Liu, Y. Wang, H. P. Koeffler and D. Xie, 'Ephrin-A1 is a negative regulator in glioma through down-regulation of EphA2 and FAK', *Int J Oncol*, 30 (2007), pp. 865-71.
48. T. Ojima, H. Takagi, K. Suzuma, H. Oh, I. Suzuma, H. Ohashi, D. Watanabe, E. Suganami, T. Murakami, M. Kurimoto, Y. Honda and N. Yoshimura, 'EphrinA1 inhibits vascular endothelial growth factor-induced intracellular signaling and suppresses retinal neovascularization and blood-retinal barrier breakdown', *Am J Pathol*, 168 (2006), pp. 331-9.
49. F. Czubayko, E. D. Liaudet-Coopman, A. Aigner, A. T. Tuveson, G. J. Berchem and A. Wellstein, 'A secreted FGF-binding protein can serve as the angiogenic switch in human cancer', *Nat Med*, 3 (1997), pp. 1137-40.
50. E. Tassi, A. Al-Attar, A. Aigner, M. R. Swift, K. McDonnell, A. Karavanov and A. Wellstein, 'Enhancement of fibroblast growth factor (FGF) activity by an FGF-binding protein', *J Biol Chem*, 276 (2001), pp. 40247-53.
51. N. Kagara, N. Tanaka, S. Noguchi and T. Hirano, 'Zinc and its transporter ZIP10 are involved in invasive behavior of breast cancer cells', *Cancer Sci*, 98 (2007), pp. 692-7.
52. L. Zhao, W. Chen, K. M. Taylor, B. Cai and X. Li, 'LIV-1 suppression inhibits HeLa cell invasion by targeting ERK1/2-Snail/Slug pathway', *Biochem Biophys Res Commun*, 363 (2007), pp. 82-8.
53. W. Engstrom and M. Granerus, 'Effects of fibroblast growth factors 19 and 20 on cell multiplication and locomotion in a human embryonal carcinoma cell line (Tera-2) in vitro', *Anticancer Res*, 26 (2006), pp. 3307-10.
54. L. Fu, L. M. John, S. H. Adams, X. X. Yu, E. Tomlinson, M. Renz, P. M. Williams, R. Soriano, R. Corpuz, B. Moffat, R. Vandlen, L. Simmons, J. Foster, J. P. Stephan, S. P. Tsai and T. A. Stewart, 'Fibroblast growth factor 19 increases metabolic rate and reverses dietary and leptin-deficient diabetes', *Endocrinology*, 145 (2004), pp. 2594-603.
55. E. Ortiz-Zapater, S. Peiro, O. Roda, J. M. Corominas, S. Aguilar, C. Ampurdanes, F. X. Real and P. Navarro, 'Tissue plasminogen activator induces pancreatic cancer cell proliferation by a non-catalytic mechanism that requires extracellular signal-regulated kinase 1/2 activation through epidermal growth factor receptor and annexin A2', *Am J Pathol*, 170 (2007), pp. 1573-84.
56. R. Paciucci, M. Tora, V. M. Diaz and F. X. Real, 'The plasminogen activator system in pancreas cancer: role of t-PA in the invasive potential in vitro', *Oncogene*, 16 (1998), pp. 625-33.
57. H. Kojima, N. Shijubo and S. Abe, 'Thymidine phosphorylase and vascular endothelial growth factor in patients with Stage I lung adenocarcinoma', *Cancer*, 94 (2002), pp. 1083-93.
58. J. S. Lee, H. S. Kim, J. J. Jung, Y. B. Kim, M. C. Lee and C. S. Park, 'Expression of vascular endothelial growth factor in invasive ductal carcinoma of the breast and the relation to angiogenesis and p53 and HER-2/neu protein expression', *Appl Immunohistochem Mol Morphol*, 10 (2002), pp. 289-95.
59. F. Tas, D. Duranyildiz, H. Oguz, H. Camlica, V. Yasasever and E. Topuz, 'Serum vascular endothelial growth factor (VEGF) and bcl-2 levels in advanced stage non-small cell lung cancer', *Cancer Invest*, 24 (2006), pp. 576-80.
60. D. Xie, J. S. Sham, W. F. Zeng, L. H. Che, M. Zhang, H. X. Wu, H. L. Lin, J. M. Wen, S. H. Lau, L. Hu and X. Y. Guan, 'Oncogenic role of clusterin overexpression in multistage colorectal tumorigenesis and progression', *World J Gastroenterol*, 11 (2005), pp. 3285-9.
61. C. Cao, E. T. Shinohara, H. Li, K. J. Niermann, K. W. Kim, K. R. Sekhar, M. Gleave, M. Freeman and B. Lu, 'Clusterin as a therapeutic target for radiation sensitization in a lung cancer model', *Int J Radiat Oncol Biol Phys*, 63 (2005), pp. 1228-36.
62. M. C. Ghosh, L. Grass, A. Soosaipillai, G. Sotiropoulou and E. P. Diamandis, 'Human kallikrein 6 degrades extracellular matrix proteins and may enhance the metastatic potential of tumour cells', *Tumour Biol*, 25 (2004), pp. 193-9.

63. T. Standal, M. Borset and A. Sundan, 'Role of osteopontin in adhesion, migration, cell survival and bone remodeling', *Exp Oncol*, 26 (2004), pp. 179-84.
64. H. Tang, J. Wang, F. Bai, L. Hong, J. Liang, J. Gao, H. Zhai, M. Lan, F. Zhang, K. Wu and D. Fan, 'Inhibition of osteopontin would suppress angiogenesis in gastric cancer', *Biochem Cell Biol*, 85 (2007), pp. 103-10.
65. V. Takafuji, M. Forgues, E. Unsworth, P. Goldsmith and X. W. Wang, 'An osteopontin fragment is essential for tumor cell invasion in hepatocellular carcinoma', *Oncogene*, 26 (2007), pp. 6361-71.
66. T. Nomura, W. C. Huang, S. Seo, H. E. Zhau, H. Mimata and L. W. Chung, 'Targeting beta2-microglobulin mediated signaling as a novel therapeutic approach for human renal cell carcinoma', *J Urol*, 178 (2007), pp. 292-300.
67. B. F. Romeike, A. Bockeler, E. Kremmer, P. Sommer, C. Krick and F. Grasser, 'Immunohistochemical detection of dUTPase in intracranial tumors', *Pathol Res Pract*, 201 (2005), pp. 727-32.
68. A. Strom, A. I. Olin, A. Aspberg and A. Hultgardh-Nilsson, 'Fibulin-2 is present in murine vascular lesions and is important for smooth muscle cell migration', *Cardiovasc Res*, 69 (2006), pp. 755-63.
69. J. Wesierska-Gadek, M. Gueorgueva and M. Horky, 'Dual action of cyclin-dependent kinase inhibitors: induction of cell cycle arrest and apoptosis. A comparison of the effects exerted by roscovitine and cisplatin', *Pol J Pharmacol*, 55 (2003), pp. 895-902.
70. C. Morrissey, L. D. True, M. P. Roudier, I. M. Coleman, S. Hawley, P. S. Nelson, R. Coleman, Y. C. Wang, E. Corey, P. H. Lange, C. S. Higano and R. L. Vessella, 'Differential expression of angiogenesis associated genes in prostate cancer bone, liver and lymph node metastases', *Clin Exp Metastasis*, 25 (2008), pp. 377-88.
71. M. Du, X. Fan, E. Hong and J. J. Chen, 'Interaction of oncogenic papillomavirus E6 proteins with fibulin-1', *Biochem Biophys Res Commun*, 296 (2002), pp. 962-9.
72. Y. W. Kwon, K. S. Kwon, H. E. Moon, J. A. Park, K. S. Choi, Y. S. Kim, H. S. Jang, C. K. Oh, Y. M. Lee, Y. G. Kwon, Y. S. Lee and K. W. Kim, 'Insulin-like growth factor-II regulates the expression of vascular endothelial growth factor by the human keratinocyte cell line HaCaT', *J Invest Dermatol*, 123 (2004), pp. 152-8.
73. D. C. Martin, J. L. Fowlkes, B. Babic and R. Khokha, 'Insulin-like growth factor II signaling in neoplastic proliferation is blocked by transgenic expression of the metalloproteinase inhibitor TIMP-1', *J Cell Biol*, 146 (1999), pp. 881-92.
74. D. D. Roberts, J. S. Isenberg, L. A. Ridnour and D. A. Wink, 'Nitric oxide and its gatekeeper thrombospondin-1 in tumor angiogenesis', *Clin Cancer Res*, 13 (2007), pp. 795-8.
75. R. T. Poon, K. K. Chung, S. T. Cheung, C. P. Lau, S. W. Tong, K. L. Leung, W. C. Yu, G. P. Tuszynski and S. T. Fan, 'Clinical significance of thrombospondin 1 expression in hepatocellular carcinoma', *Clin Cancer Res*, 10 (2004), pp. 4150-7.
76. C. Herold-Mende, M. M. Mueller, M. M. Bonsanto, H. P. Schmitt, S. Kunze and H. H. Steiner, 'Clinical impact and functional aspects of tenascin-C expression during glioma progression', *Int J Cancer*, 98 (2002), pp. 362-9.
77. T. Ishiwata, K. Takahashi, Y. Shimanuki, R. Ohashi, R. Cui, F. Takahashi, K. Shimizu, K. Miura and Y. Fukuchi, 'Serum tenascin-C as a potential predictive marker of angiogenesis in non-small cell lung cancer', *Anticancer Res*, 25 (2005), pp. 489-95.
78. G. Watanabe, H. Nishimori, H. Irifune, Y. Sasaki, S. Ishida, H. Zembutsu, T. Tanaka, S. Kawaguchi, T. Wada, J. Hata, M. Kusakabe, K. Yoshida, Y. Nakamura and T. Tokino, 'Induction of tenascin-C by tumor-specific EWS-ETS fusion genes', *Genes Chromosomes Cancer*, 36 (2003), pp. 224-32.
79. J. Y. Ljubimova, M. Fujita, N. M. Khazenzon, A. V. Ljubimov and K. L. Black, 'Changes in laminin isoforms associated with brain tumor invasion and angiogenesis', *Front Biosci*, 11 (2006), pp. 81-8.
80. M. Fujita, N. M. Khazenzon, S. Bose, K. Sekiguchi, T. Sasaki, W. G. Carter, A. V. Ljubimov, K. L. Black and J. Y. Ljubimova, 'Overexpression of beta1-chain-containing laminins in capillary basement membranes of human breast cancer and its metastases', *Breast Cancer Res*, 7 (2005), pp. R411-21.
81. M. J. Moreno, M. Ball, M. F. Andrade, A. McDermid and D. B. Stanimirovic, 'Insulin-like growth factor binding protein-4 (IGFBP-4) is a novel anti-angiogenic and anti-tumorigenic mediator secreted by dibutyryl cyclic AMP (dB-cAMP)-differentiated glioblastoma cells', *Glia*, 53 (2006), pp. 845-57.
82. D. F. Peng, Y. Kanai, M. Sawada, S. Ushijima, N. Hiraoka, T. Kosuge and S. Hirohashi, 'Increased DNA methyltransferase 1 (DNMT1) protein expression in precancerous conditions and ductal carcinomas of the pancreas', *Cancer Sci*, 96 (2005), pp. 403-8.
83. G. Egger, S. Jeong, S. G. Escobar, C. C. Cortez, T. W. Li, Y. Saito, C. B. Yoo, P. A. Jones and G. Liang, 'Identification of DNMT1 (DNA methyltransferase 1) hypomorphs in somatic knockouts suggests an essential role for DNMT1 in cell survival', *Proc Natl Acad Sci U S A*, 103 (2006), pp. 14080-5.
84. L. McInroy and A. Maatta, 'Down-regulation of vimentin expression inhibits carcinoma cell migration and adhesion', *Biochem Biophys Res Commun*, 360 (2007), pp. 109-14.
85. C. Gilles, M. Polette, M. Mestdagt, B. Nawrocki-Raby, P. Ruggeri, P. Birembaut and J. M. Foidart, 'Transactivation of vimentin by beta-catenin in human breast cancer cells', *Cancer Res*, 63 (2003), pp. 2658-64.
86. T. Hoashi, T. Kadono, K. Kikuchi, T. Etoh and K. Tamaki, 'Differential growth regulation in human melanoma cell lines by TIMP-1 and TIMP-2', *Biochem Biophys Res Commun*, 288 (2001), pp. 371-9.
87. F. Lizarraga, V. Maldonado and J. Melendez-Zajgla, 'Tissue inhibitor of metalloproteinases-2 growth-stimulatory activity is mediated by nuclear factor-kappa B in A549 lung epithelial cells', *Int J Biochem Cell Biol*, 36 (2004), pp. 1655-63.
88. L. H. Shi, X. Q. Zhu, G. H. Zhao, Y. B. Xia and Y. S. Zhang, 'Expression of insulin-like growth factor binding protein-2 in gastric carcinoma and its relationship with cell proliferation', *World J Gastroenterol*, 12 (2006), pp. 6285-9.
89. E. J. Lee, C. Mircean, I. Shmulevich, H. Wang, J. Liu, A. Niemisto, J. J. Kavanagh, J. H. Lee and W. Zhang, 'Insulin-like growth factor binding protein 2 promotes ovarian cancer cell invasion', *Mol Cancer*, 4 (2005), p. 7.
90. C. Nagi, M. Guttman, S. Jaffer, R. Qiao, R. Keren, A. Triana, M. Li, J. Godbold, I. J. Bleiweiss and R. B. Hazan, 'N-cadherin expression in breast cancer: correlation with an aggressive histologic variant--invasive micropapillary carcinoma', *Breast Cancer Res Treat*, 94 (2005), pp. 225-35.
91. E. Rosivatz, I. Becker, M. Bamba, C. Schott, J. Diebold, D. Mayr, H. Hofler and K. F. Becker, 'Neoexpression of N-cadherin in E-cadherin positive colon cancers', *Int J Cancer*, 111 (2004), pp. 711-9.
92. A. Hurbin, J. L. Coll, L. Dubrez-Daloz, B. Mari, P. Auberger, C. Brambilla and M. C. Favrot, 'Cooperation of amphiregulin and insulin-like growth factor-1 inhibits Bax- and Bad-mediated apoptosis via a protein kinase C-dependent pathway in non-small cell lung cancer cells', *J Biol Chem*, 280 (2005), pp. 19757-67.
93. A. Hurbin, L. Dubrez, J. L. Coll and M. C. Favrot, 'Inhibition of apoptosis by amphiregulin via an insulin-like growth factor-1 receptor-dependent pathway in non-small cell lung cancer cell lines', *Ann N Y Acad Sci*, 1010 (2003), pp. 354-7.
94. N. E. Willmarth and S. P. Ethier, 'Autocrine and juxtacrine effects of amphiregulin on the proliferative, invasive, and migratory properties of normal and neoplastic human mammary epithelial cells', *J Biol Chem*, 281 (2006), pp. 37728-37.
95. C. Liu, C. Sun, H. Huang, K. Janda and T. Edgington, 'Overexpression of legumain in tumors is significant for invasion/metastasis and a candidate enzymatic target for prodrug therapy', *Cancer Res*, 63 (2003), pp. 2957-64.
96. V. Gocheva and J. A. Joyce, 'Cysteine cathepsins and the cutting edge of cancer invasion', *Cell Cycle*, 6 (2007), pp. 60-4.
97. S. Hagisawa, C. Ohyama, T. Takahashi, M. Endoh, T. Moriya, J. Nakayama, Y. Arai and M. Fukuda, 'Expression of core 2 beta1,6-N-acetylglucosaminyltransferase facilitates prostate cancer progression', *Glycobiology*, 15 (2005), pp. 1016-24.
98. E. Machida, J. Nakayama, J. Amano and M. Fukuda, 'Clinicopathological significance of core 2 beta1,6-N-acetylglucosaminyltransferase messenger RNA expressed in the pulmonary adenocarcinoma determined by in situ hybridization', *Cancer Res*, 61 (2001), pp. 2226-31.
99. A. Yuan, J. J. Chen, P. L. Yao and P. C. Yang, 'The role of interleukin-8 in cancer cells and microenvironment interaction', *Front Biosci*, 10 (2005), pp. 853-65.
100. J. J. Chen, P. L. Yao, A. Yuan, T. M. Hong, C. T. Shun, M. L. Kuo, Y. C. Lee and P. C. Yang, 'Up-regulation of tumor interleukin-8 expression by infiltrating macrophages: its correlation with tumor angiogenesis and patient survival in non-small cell lung cancer', *Clin Cancer Res*, 9 (2003), pp. 729-37.
101. H. Terada, T. Urano and H. Konno, 'Association of interleukin-8 and plasminogen activator system in the progression of colorectal cancer', *Eur Surg Res*, 37 (2005), pp. 166-72.
102. S. M. Pulukuri and J. S. Rao, 'Small interfering RNA directed reversal of urokinase plasminogen activator demethylation inhibits prostate tumor growth and metastasis', *Cancer Res*, 67 (2007), pp. 6637-46.
103. M. A. Gallicchio, C. Kaun, J. Wojta, B. Binder and L. A. Bach, 'Urokinase type plasminogen activator receptor is involved in insulin-like growth factor-induced migration of rhabdomyosarcoma cells in vitro', *J Cell Physiol*, 197 (2003), pp. 131-8.
104. S. Esseghir, A. Kennedy, P. Seedhar, A. Nerurkar, R. Poulsom, J. S. Reis-Filho and C. M. Isacke, 'Identification of NTN4, TRA1, and STC2 as prognostic markers in breast cancer in a screen for signal sequence encoding proteins', *Clin Cancer Res*, 13 (2007), pp. 3164-73.
105. I. Santamaria, G. Velasco, M. Cazorla, A. Fueyo, E. Campo and C. Lopez-Otin, 'Cathepsin L2, a novel human cysteine proteinase produced by breast and colorectal carcinomas', *Cancer Res*, 58 (1998), pp. 1624-30.
106. J. Wang, J. Shou and X. Chen, 'Dickkopf-1, an inhibitor of the Wnt signaling pathway, is induced by p53', *Oncogene*, 19 (2000), pp. 1843-8.
107. S. Peng, J. Li, C. Miao, L. Jia, Z. Hu, P. Zhao, J. Li, Y. Zhang, Q. Chen and E. Duan, 'Dickkopf-1 secreted by decidual cells promotes trophoblast cell invasion during murine placentation', *Reproduction*, 135 (2008), pp. 367-75.
108. H. Peinado, M. Del Carmen Iglesias-de la Cruz, D. Olmeda, K. Csiszar, K. S. Fong, S. Vega, M. A. Nieto, A. Cano and F. Portillo, 'A molecular role for lysyl oxidase-like 2 enzyme in snail regulation and tumor progression', *Embo J*, 24 (2005), pp. 3446-58.
109. K. Noma, K. S. Smalley, M. Lioni, Y. Naomoto, N. Tanaka, W. El-Deiry, A. J. King, H. Nakagawa and M. Herlyn, 'The essential role of fibroblasts in esophageal squamous cell carcinoma-induced angiogenesis', *Gastroenterology*, 134 (2008), pp. 1981-93.
110. W. Wick, M. Platten and M. Weller, 'Glioma cell invasion: regulation of metalloproteinase activity by TGF-beta', *J Neurooncol*, 53 (2001), pp. 177-85.
111. H. Lin, X. Chen, J. Wang and Z. Yu, 'Inhibition of apoptosis in rat mesangial cells by tissue inhibitor of metalloproteinase-1', *Kidney Int*, 62 (2002), pp. 60-9.
112. Y. Ikenaka, H. Yoshiji, S. Kuriyama, J. Yoshii, R. Noguchi, H. Tsujinoue, K. Yanase, T. Namisaki, H. Imazu, T. Masaki and H. Fukui, 'Tissue inhibitor of metalloproteinases-1 (TIMP-1) inhibits tumor growth and angiogenesis in the TIMP-1 transgenic mouse model', *Int J Cancer*, 105 (2003), pp. 340-6.
113. B. H. Lee, J. S. Bae, R. W. Park, J. E. Kim, J. Y. Park and I. S. Kim, 'betaig-h3 triggers signaling pathways mediating adhesion and migration of vascular smooth muscle cells through alphavbeta5 integrin', *Exp Mol Med*, 38 (2006), pp. 153-61.
114. J. Tang, H. W. Zhou, J. L. Jiang, X. M. Yang, Y. Li, H. X. Zhang, Z. N. Chen and W. P. Guo, 'BetaIg-h3 is involved in the HAb18G/CD147-mediated metastasis process in human hepatoma cells', *Exp Biol Med (Maywood)*, 232 (2007), pp. 344-52.

6. Résultats complémentaires

6.1. Dosage des protéines GDF-15, FGF-19 et Kallikrein-6 dans les échantillons de secretome de la lignée H358/TetOn/p53

Le dosage des protéines GDF-15, FGF-19 et Kallikrein-6 a été réalisé par ELISA pour différentes préparations d'échantillons « secretome » (Tableau 8). Ces dosages permettent d'établir des ratios d'expression qui confirment les variations déterminées en protéomique (Figure 64 A, B, C et Tableau 8).

Il est par ailleurs intéressant de constater que ces rapports sont très similaires de ceux précédemment obtenus en iTRAQ (Figure 64D). Ce résultat démontre l'efficacité de l'approche iTRAQ pour mettre en valeur des variations d'expression protéiques entre différentes conditions.

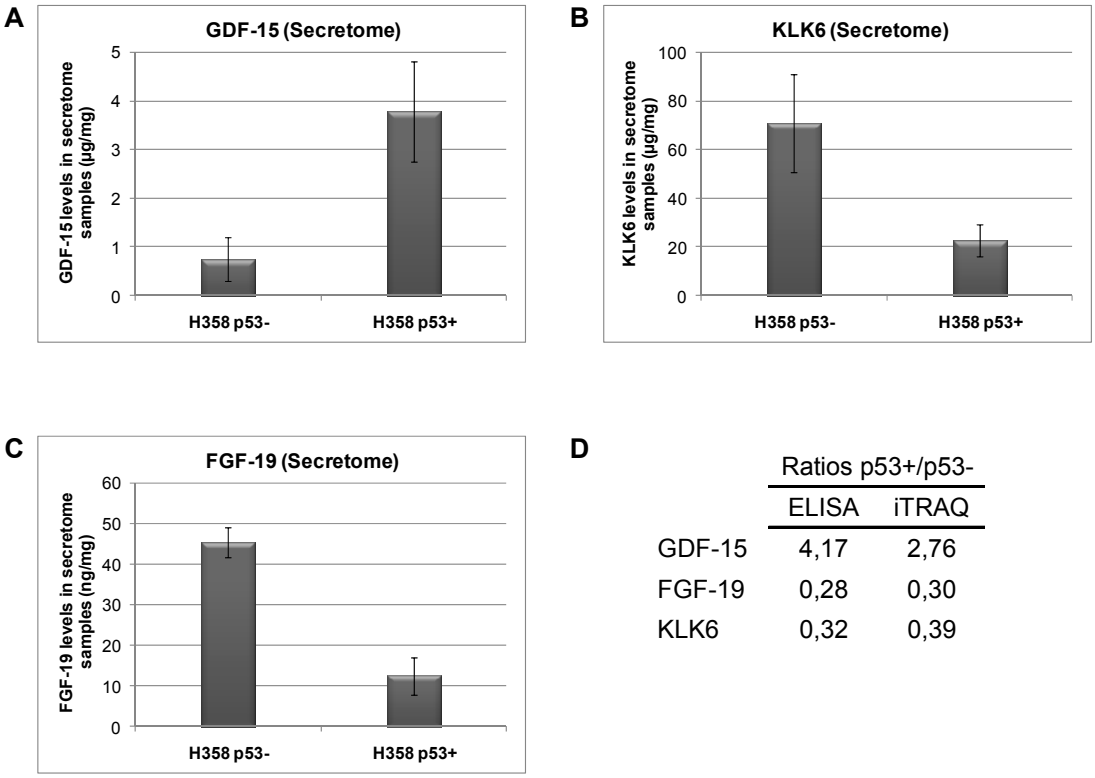
A	GDF-15 (Secretome)							
Echantillons	S1_p53	S1	S2_p53	S2	S3_p53	S3	S4_p53	S4
[GDF-15] (µg/mg prot. tot.)	4,94	0,28	2,57	0,47	3,13	1,27	4,46	0,97
Ratio p53+/p53-	17,78		5,44		2,46		4,61	

B	FGF-19 (Secretome)							
Echantillons	S1_p53	S1	S2_p53	S2	S3_p53	S3	S4_p53	S4
[FGF-19] (ng/mg prot. tot.)	15,2	46,6	9,4	46,8	12,6	42,7	-	-
Ratio p53+/p53-	0,33		0,20		0,30		-	

C	KLK6 (Secretome)							
Echantillons	S1_p53	S1	S2_p53	S2	S3_p53	S3	S4_p53	S4
[KLK6] (µg/mg prot. tot.)	30,1	84,4	16,0	45,1	21,5	83,3	-	-
Ratio p53+/p53-	0,36		0,35		0,26		-	

Tableau 8 : Dosages ELISA des protéines GDF-15 (A), FGF-19 (B) et Kallikrein-6 (C) dans des échantillons « secretome » après ou non ré-expression de p53 par les cellules H358/TetOn/p53.

Les concentrations sont rapportées pour 1 mg de protéines totales. Les ratios p53+/p53- représentent les rapports de concentration entre les échantillons exprimant ou non p53.



	Ratios p53+/p53-	
	ELISA	iTRAQ
GDF-15	4,17	2,76
FGF-19	0,28	0,30
KLK6	0,32	0,39

Figure 64 : Niveaux d'expression des protéines GDF-15, Kallikrein-6 et FGF-19 dans le secretome des cellules H358/TetOn/p53.

(A) Niveaux d'expression de GDF-15 dans le secretome (en µg/mg de protéines totales). Moyenne ± écart type (n=4). (B) Niveaux d'expression de Kallikrein-6 dans le secretome (en µg/mg de protéines totales). Moyenne ± écart type (n=3). (C) Niveaux d'expression de FGF-19 dans le secretome (en ng/mg de protéines totales). Moyenne ± écart type (n=3). (D) Tableau comparatif des ratios déterminés après ELISA et iTRAQ.
H358 p53+ : Secretome de la lignée H358 après induction de p53. H358 p53- : Secretome de la lignée H358 sans induction de p53.

6.2. Autres analyses bioinformatiques des données

6.2.1. PLIPS et le croisement des données protéomique

Les études protéomiques génèrent généralement des listes assez énormes de protéines. Parmi celles-ci, certaines apparaissent très fréquemment (les protéines de ménage) et ne sont donc pas du tout spécifique de l'étude réalisée. Des protéines comme la vimentine ou des protéines de l'inflammation sont, par exemple, souvent exprimées de façon différentielle dans les études et semblent donc varier quelque soit le stress appliqué. D'autres sont des contaminants souvent retrouvés (histones, kératines).

Dans le but d'évaluer la pertinence des protéines modulées par p53 identifiées dans le secretome des cellules H358, nous avons analysé l'antériorité de leur caractérisation dans les précédentes études protéomiques.

Le programme PLIPS (Protein Lists Identified in Proteomics Studies), récemment développé par Antonov et al., permet de confronter des listes protéiques aux données présentes dans près de 1000 études protéomiques précédentes et ainsi voir leur redondance [265].

Certaines protéines telles que la vimentine, l'histone H4 ou la clusterin ne sont pas du tout propre à notre étude (respectivement identifiées dans 119, 41 et 63 publications en protéomique). Au contraire, les protéines les plus influencées sont beaucoup plus spécifiques à notre étude. FGF-19, FGF-BP1, PIG3 ou la cadhérine-6 sont, par exemple, retrouvées dans aucune autre étude, et ZIP10, tPA, GDF-15 et VEGF dans un nombre très limité, respectivement 4, 3, 6 et 3 études.

6.2.2. Ingenuity systems

Le logiciel Ingenuity, commercialisé par Ingenuity systems, permet notamment de rechercher les relations entre les membres d'une liste protéique par l'interprétation des données de la littérature. Une telle analyse a été effectuée pour les 47 protéines surexprimées et les 44 sous-exprimées en présence de p53 dans le secretome (Figure 65). Les interactions ou modulations entre ces protéines sont représentées par une flèche. Un lien direct vers les publications ayant caractérisés ce lien est disponible et permet d'expliciter la relation entre les différentes protéines du réseau. Les liens connus entre les protéines et les procédés cellulaires ou pathologiques d'intérêts sont également accessibles grâce à ce logiciel. Cette possibilité n'a pas été approfondie dans notre étude.

Les voies générées établissent clairement p53 au centre de chacun des deux réseaux.

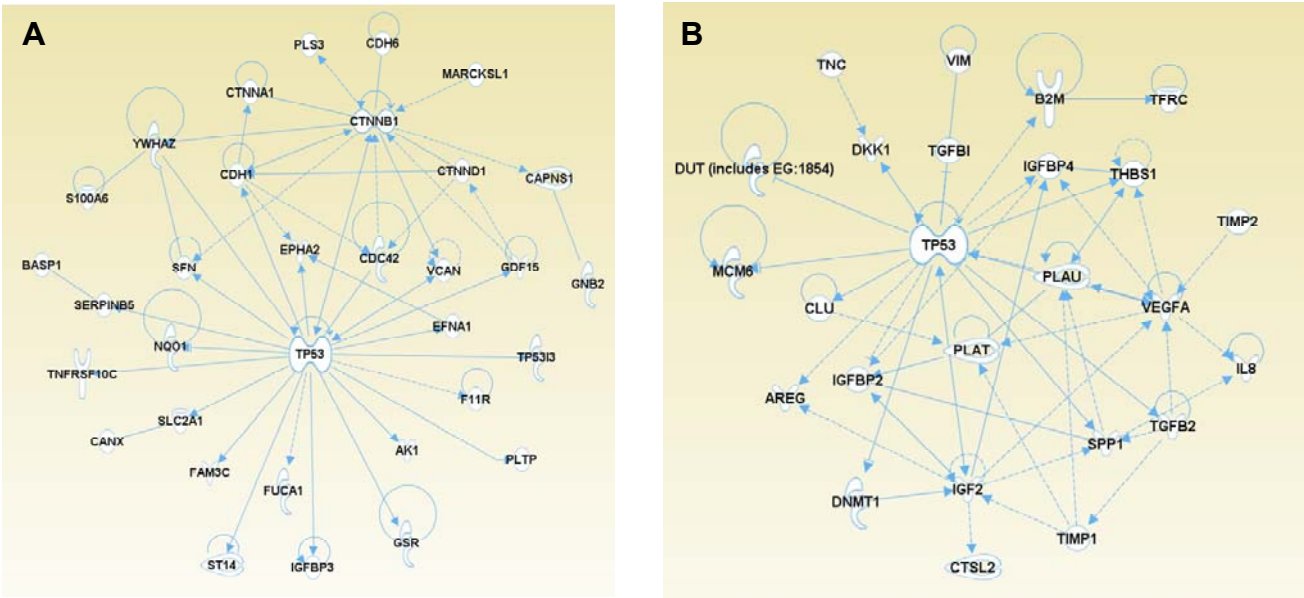


Figure 65 : Relation entre p53 et les protéines différentielles du secretome des cellules H358 établi par le logiciel Ingenuity.

A. Relation entre p53 et les protéines surexprimées dans le secretome.

B. Relation entre p53 et les protéines sous-exprimées dans le secretome.

6.3. Etude de l'influence de p53 sur la modulation d'autres sous-protéomes de la lignée H358/TetOn/p53

L'étude de l'influence de p53 sur la modulation du protéome cytosolique et du protéome des membranes & organelles (réticulum endoplasmique, appareil de Golgi, mitochondries,…) a été réalisée.

L'analyse du protéome cytosolique a pour but de s'assurer de la qualité de l'échantillon secretome. En effet, si ce dernier n'est pas contaminé par des protéines issues d'une lyse cellulaire anormale, chacun des deux sous-protéomes doit posséder une proportion significative de protéines qui lui sont propres.

L'analyse du protéome des organelles et membranes est intéressante car les protéines sécrétées par la voie classique transitent par le réticulum endoplasmique et l'appareil de Golgi avant leur sécrétion. De plus, l'analyse des protéines membranaires permettra d'identifier d'éventuels récepteurs des protéines sécrétées modulées par p53 et voir leur éventuelle régulation par p53.

Les cellules H358/TetOn/p53 sont cultivées selon le même protocole que pour l'obtention de l'échantillon secretome, à la seule différence que du sérum de veau fœtal (SVF) est supplémenté au milieu de culture. En effet, les fractions protéiques étudiées ne risquent pas d'être contaminées par des protéines sériques car les cellules sont lavées au PBS avant l'extraction.

L'extraction séquentielle des différents sous-protéomes est réalisée grâce au kit "ProteoExtract Subcellular Proteome Extraction" de Calbiochem.

L'analyse différentielle de l'influence de p53 sur ces protéomes est faite par la méthode iTRAQ. Un kit iTRAQ 8-plex est utilisé pour augmenter la robustesse de nos résultats. Cette analyse 8-plex est réalisée pour les échantillons cytosoliques et pour les échantillons des organites & membranes.

Les 8 conditions comparées sont les mêmes pour les deux expérimentations :

- Tag 113 : échantillon « dox+ » (cellules H358 non transfectées par le plasmide TetOn/p53 cultivées en présence de doxycycline).
- Tag 114 : échantillon « S2E1dox- » (Préparation n°2 de cellules H358 TetOn/p53 cultivées en absence de doxycycline, extraction n°1).
- Tag 115 : échantillon « S2E2dox+ » (Préparation n°2 de cellules H358 TetOn/p53 cultivées en présence de doxycycline, extraction n°2).
- Tag 116 : échantillon « S1dox+ » (Préparation n°1 de cellules H358 TetOn/p53 cultivées en présence de doxycycline).
- Tag 117 : échantillon « S2E2dox- » (Préparation n°2 de cellules H358 TetOn/p53 cultivées en absence de doxycycline, extraction n°2).
- Tag 118 : échantillon « S1dox- » (Préparation n°1 de cellules H358 TetOn/p53 cultivées en absence de doxycycline).
- Tag 119 : échantillon « S2E1dox+ » (Préparation n°2 de cellules H358 TetOn/p53 cultivées en présence de doxycycline, extraction n°1).
- Tag 121 : échantillon « dox- » (cellules H358 non transfectées par le plasmide TetOn/p53 cultivées en absence de doxycycline).

Cette analyse 8-plex permet de réaliser, en une seule expérimentation, une étude différentielle robuste en minimisant les éventuels biais expérimentaux. En effet, la reproductibilité de l'extraction des protéines est testée en réalisant, en parallèle, deux extractions indépendantes pour une même préparation de cellules (« S2P1 » *versus* « S2P2 » : pics à 114, 115, 117, 119). La reproductibilité inter-cultures est aussi considérée en comparant les échantillons issus de cultures cellulaires indépendantes (« S1 » *versus* « S2 » : pics à 116, 118 *versus* pics à 114, 115, 117, 119). Enfin, les éventuelles modifications d'expression dues à un effet doxycycline dépendant (et non p53-dépendant) sont éliminées en comparant les profils obtenus pour la lignée non transfectée soumise ou non à la doxycycline (« dox+ » *versus* « dox- » : pics à 113 *versus* 121). La variation d'expression d'une protéine est jugée indépendante de la doxycycline si le ratio 113/121 est compris entre 0,8 et 1,25. Au final, seuls les résultats concordants pour ces différentes analyses sont jugés fiables, l'objectif étant d'augmenter la confiance dans nos résultats en éliminant un maximum de faux positifs.

Les digestats marqués des 8 échantillons sont mélangés et séparés en 24 fractions par OFFGEL. Chacune des fractions est ensuite analysée en LC-MALDI-MS/MS. Ces études ont donné lieu à

l'identification de 820 protéines pour le sous-protéome « cytosol » et de 907 protéines pour la fraction « organelles & membranes ».

La réalisation d'un diagramme de Venn entre les échantillons « secretome », « cytosol » et « organelles&membranes » montre que chaque sous-protéome possède une part de son contenu qui lui est propre (Figure 66A). En effet, près de la moitié des protéines identifiées est spécifique du sous-protéome étudié (40,2 % pour le secretome, 43,3 % pour le cytosol et 46,2 % pour les organites & membranes) (Figure 66B). Il n'est, par ailleurs, pas surprenant que certaines protéines soient présentes dans plusieurs fractions. Par exemple, les protéines sécrétées par la voie classique cheminent par le réticulum endoplasmique et l'appareil de Golgi. Les protéines sécrétées par une voie non-classique ou exosomale sont généralement aussi présentes dans le cytosol. Enfin, certaines protéines peuvent avoir plusieurs localisations différentes.

A

Secretome

Présence	Nb prot	%
Unique	365	40.2
Cytosol	126	13.9
Organites & Membranes	149	16.4
Organites & Membranes + Cytosol	269	29.6
TOTAL	909	100.0

Cytosol

Présence	Nb prot	%
Unique	355	43.3
Secretome	126	15.4
Organites & Membranes	70	8.5
Organites & Membranes + Secretome	269	32.8
TOTAL	820	100.0

Organites & Membranes

Présence	Nb prot	%
Unique	419	46.2
Cytosol	70	7.7
Secretome	149	16.4
Cytosol + Secretome	269	29.7
TOTAL	907	100.0

B

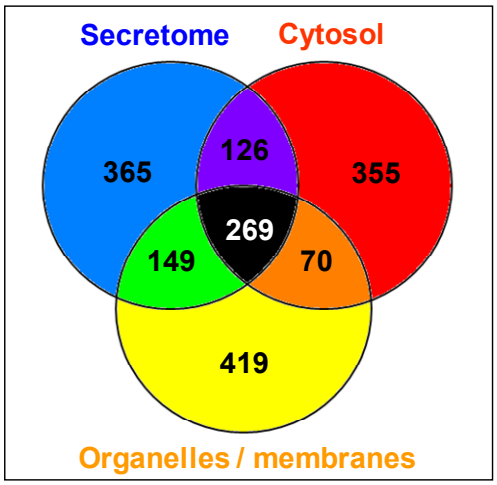


Figure 66 : Comparaison de trois sous-protéomes de la lignée cellulaire H358/TetOn/p53.

A. Tableaux d'analyse de la spécificité des protéines identifiées pour chacun des 3 sous-protéomes étudiés : Secretome ; Cytosol ; Organelles&Membranes.

Unique : protéines retrouvées uniquement dans le sous-protéome considéré. Pour chaque sous-protéome, la proportion de protéines retrouvées dans les autres sous-protéomes est renseignée. Les résultats sont indiqués sous forme de nombre de protéines (Nb prot) et de pourcentage (%).

B. Représentation sous forme de diagramme de Venn de la répartition de l'ensemble des protéines identifiées pour chacun des 3 sous-protéomes : Secretome ; Cytosol ; Organelles&Membranes.

6.3.1. Sous-protéome cytosolique

- 820 protéines ont été identifiées dans l'échantillon « Cytosol », parmi lesquelles 31 sont modulées par p53 : 25 sont surexprimées et 6 sous-exprimées en présence de p53 (Tableau 9).

- La purification de l'échantillon est correcte puisque 30 des 31 protéines différentielles ont été préalablement décrites comme localisées (au moins en partie) dans le cytosol (Tableau 9).

- Parmi les protéines surexprimées en présence de p53, 5 sur 25 sont des cibles transcriptionnelles de p53 préalablement décrites : Ribonucleoside-diphosphate reductase subunit M2 B, Putative quinone oxidoreductase, Serpin B5, Protein S100-A6 et Adenylate kinase 1 (Tableau 9). A l'inverse, 1 des 6 protéines sous-exprimées (Stathmin) est réprimée par p53 (Tableau 9). Comme pour l'étude secretome, cette information augmente la confiance dans les résultats observés.

- 8 protéines sur les 31 différentielles sont également retrouvées dans les 91 protéines du secretome modulées par p53. Dans les deux sous-protéomes, ces 8 protéines varient de façon similaire (Tableau 9 et Tableaux 1 et 2 de la publication).

A

Nom	No. Access.	Ratio +/-	SD	Ratio dox	Loc.	p53	Secr.
Stathmin	P16949	0.62	0.05	0.92	C	R	X
Novel protein similar to prothymosin alpha	Q5VYG3	0.71	0.09	0.98	U		
Pterin-4-alpha-carbinolamine dehydratase	P61457	0.71	0.10	1.25	C		
D-3-phosphoglycerate dehydrogenase	O43175	0.73	0.03	0.94	C+S		
Asparagine synthetase	P08243	0.77	0.07	0.92	C		
Myoglobin	B0QYF8	0.78	0.14	0.93	C		

B

Nom	No. Access.	Ratio +/-	SD	Ratio dox	Loc.	p53	Secr.
Ribonucleoside-diphosphate reductase subunit M2 B	Q7LG56	2.69	0.10	1.05	C	I	
Putative quinone oxidoreductase	Q53FA7	2.30	0.19	0.84	C+S (NC)	I	X
Serpin B5	P36952	1.60	0.17	0.96	S+exo	I	X
Putative uncharacterized protein DKFZp686G1675	Q7Z3A2	1.52	0.11	0.95	C		
Gelsolin	P06396	1.48	0.08	1.07	C	L	
Basic leucine zipper and W2 domain-containing protein 1	Q7L1Q6	1.43	0.28	0.76	C		
Protein S100-A6	P06703	1.43	0.03	0.88	C+S (NC)+exo	I	X
Cathepsin D	P07339	1.42	0.24	1.46	C+S		
UDP-glucose 4-epimerase	Q14376	1.42	0.08	0.98	C		
Tubulin beta-6 chain	Q9BUF5	1.42	0.17	1.16	C		
Annexin A4	P09525	1.42	0.05	1.09	C		
14 kDa phosphohistidine phosphatase	Q9NRX4	1.42	0.18	0.93	C		

Adenylate kinase 1	Q5T9B7	1.39	0.18	0.93	C+S (NC)	I	X
Vasodilator-stimulated phosphoprotein	P50552	1.37	0.12	0.83	C		
Plectin 1	Q6S383	1.37	0.14	0.99	C+exo		X
cAMP-dependent protein kinase type I-alpha regulatory subunit	P10644	1.36	0.17	1.08	C		
Echinoderm microtubule-associated protein-like 2	O95834	1.35	0.09	1.06	C		
Adseverin	Q9Y6U3	1.35	0.13	1.16	C		
Interferon-induced 17 kDa protein	P05161	1.32	0.10	0.98	C		
EH domain-containing protein 1	Q9H4M9	1.30	0.13	0.96	exo+mb (C side)		X
Protein S100-A11	P31949	1.27	0.06	0.85	C		
Myosin light polypeptide 6	P60660	1.26	0.12	0.69	C		
Actin-related protein 3	P61158	1.24	0.11	0.82	C		
Leukocyte elastase inhibitor	P30740	1.21	0.07	0.90	C+S (NC)		X
Puromycin-sensitive aminopeptidase	P55786	1.19	0.06	0.94	C+mb		X

Tableau 9 : Les protéines modulées par p53 dans la fraction cytosolique des cellules H358/TetOn/p53.

Les protéines sont listées avec leur numéro d'accession UniProt (No. Access.). Seules les protéines avec un score ProteinPilot > 4 (2 peptides distincts avec 99% de confiance) sont considérées. Les ratios donnés correspondent à la moyenne des ratios obtenus lors de l'expérience iTRAQ 8-plex. Le ratio +/- correspond à la quantification relative des protéines dans la condition avec p53 par rapport à celle sans p53. Les protéines sous-exprimées en présence de p53 ont un ratio < 0.8 (A) et les surexprimées > 1.25 (B). La déviation standard inter-ratio est notée pour chaque protéine (SD). L'effet dox est contrôlé sur les cellules H358 non transfectées et soumises ou non à la doxycycline (ratio dox+ / dox-). La localisation cellulaire des protéines est obtenue de la littérature (C : Cytosol ; mb : Membrane ; U : « Unknown » ; S : Sécrété ; NC : sécrété par une voie non-classique ; exo : sécrété par la voie des exosomes). Les relations entre p53 et les protéines modulées (colonne « p53 ») sont issues de la littérature (I : cible transcriptionnelle induite par p53; R : cible transcriptionnelle réprimée par p53; L: protéine ayant un lien avec p53).Les protéines modulées de façon similaire dans les échantillons du cytosol et du secretome sont cochées dans la colonne « Secr ».

- 17 des 91 protéines modulées par p53 dans le secretome sont également présentes dans le cytosol (Figure 67 et Tableau 10). Comme dit précédemment, 8 de ces protéines communes aux deux échantillons sont modulées de façon similaire. Neuf protéines modulées au niveau du secretome sont donc inchangées dans le cytosol. On peut émettre l'hypothèse que la modulation d'expression de ces protéines dans le secretome se fait donc par une action sur les processus de sécrétion ou sur la stabilité de ces protéines dans le milieu extracellulaire. Une étude transcriptomique permettrait de confirmer que la régulation de ces protéines n'a pas lieu au niveau transcriptionnel.

De façon surprenante, ces 17 protéines communes appartiennent aux 47 protéines surexprimées dans le secretome en présence de p53. Leur analyse démontre que 9 sont sécrétées par une voie non classique, 4 sont des protéines membranaires, 3 sont relarguées par les exosomes, et 1 n'est pas connue comme sécrétée (voie exosomale ?) (Figure 67). Cette observation va dans le sens d'une régulation par p53 des voies de sécrétions non classiques et notamment la voie de sécrétion par les exosomes comme démontré par Yu et al. [186].

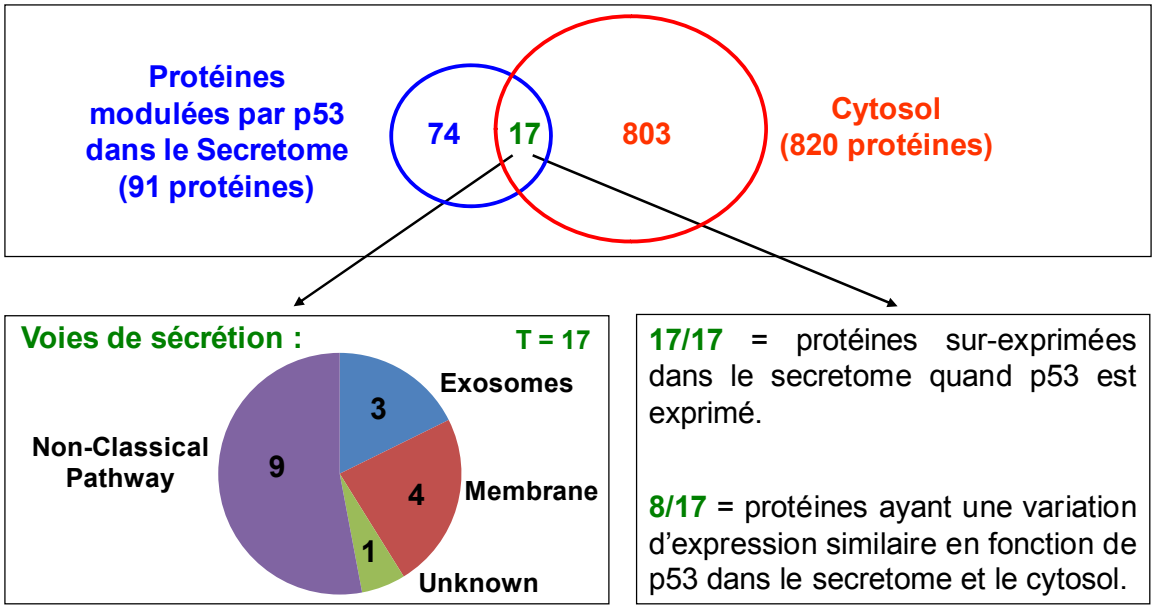


Figure 67 : Protéines modulées par p53 dans le secretome *versus* protéome cytosolique.

Haut : Diagramme de Venn représentant la proportion de protéines modulées par p53 dans le secretome identifiées également dans le sous protéome « Cytosol ».

Bas : Voies de sécrétion connues ou prédites pour les 17 protéines du secretome identifiées également dans le Cytosol.

No.	Nom	Ratio p53+ / p53- Secretome	Organelles & Membranes	Cytosol	Unique
Q14512	Fibroblast growth factor-binding protein 1	0,23	X	X	Yes
Q9ULF5	Zinc transporter ZIP10	0,30	X	X	Yes
O95750	Fibroblast growth factor 19	0,30	X	X	Yes
P00750	Tissue-type plasminogen activator	0,33	X	X	Yes
Q92820	Gamma-glutamyl hydrolase	0,34	X	X	Yes
P15692	Vascular endothelial growth factor A	0,34	X	X	Yes
P10909	Clusterin	0,38	X	X	Yes
Q92876	Kallikrein-6	0,38	X	X	Yes
P33316	Deoxyuridine 5'-triphosphate nucleotidohydrolase	0,44	0,61	X	
P10451	Osteopontin *	0,45	X	X	Yes
P22692	Insulin-like growth factor-binding protein 4	0,46	X	X	Yes
P61769	Beta-2-microglobulin	0,46	X	X	Yes
P98095	Fibulin-2	0,47	X	X	Yes
P19022	Cadherin-2	0,47	X	X	Yes
P23142	Fibulin-1	0,49	X	X	Yes
P33552	Cyclin-dependent kinases regulatory subunit 2	0,49	X	X	Yes
P07942	Laminin subunit beta-1	0,51	X	X	Yes
Q8NCL4	Polypeptide N-acetylgalactosaminyltransferase 6	0,51	X	X	Yes
P24821	Tenascin	0,51	X	X	Yes
P02786	Transferrin receptor protein 1	0,52	0,44	X	
P13284	Gamma-interferon-inducible lysosomal thiol reductase	0,53	X	X	Yes
P07996	Thrombospondin-1	0,53	X	X	Yes
P26358	DNA (cytosine-5)-methyltransferase 1	0,54	X	X	Yes

P01344	Insulin-like growth factor II *	0,56	X	X	Yes
Q12805	EGF-containing fibulin-like extracellular matrix protein 1	0,56	X	X	Yes
P61916	Epididymal secretory protein E1	0,56	1,10	X	
P08670	Vimentin	0,58	X	X	Yes
Q14566	DNA replication licensing factor MCM6	0,58	0,62	X	
P18065	Insulin-like growth factor-binding protein 2	0,58	X	X	Yes
P15514	Amphiregulin	0,60	X	X	Yes
P16035	Metalloproteinase inhibitor 2	0,61	X	X	Yes
Q99538	Legumain	0,62	X	X	Yes
P53634	Dipeptidyl-peptidase 1	0,62	0,94	X	
P10145	Interleukin-8	0,63	X	X	Yes
Q02742	Beta-1,3-galactosyl-O-glycosyl-glycoprotein beta-1,6-N-acetylglucosaminyltransferase	0,63	X	X	Yes
P00749	Urokinase-type plasminogen activator	0,63	X	X	Yes
O94907	Dickkopf-related protein 1	0,64	X	X	Yes
O76061	Stanniocalcin-2	0,64	X	X	Yes
P61812	Transforming growth factor beta-2	0,66	X	X	Yes
Q9Y4K0	Lysyl oxidase homolog 2	0,68	X	X	Yes
Q8NES3	Beta-1,3-N-acetylglucosaminyltransferase lunatic fringe	0,70	X	X	Yes
O60911	Cathepsin L2	0,70	X	X	Yes
P01033	Metalloproteinase inhibitor 1	0,74	X	X	Yes
Q15582	Transforming growth factor-beta-induced protein ig-h3	0,75	X	X	Yes
Q53FA7	Putative quinone oxidoreductase	5,57	X	2,30	
Q99988	Growth/differentiation factor 15	2,80	X	X	Yes
P55058	Phospholipid transfer protein	2,49	X	X	Yes
P55285	Cadherin-6	2,48	X	X	Yes
Q96QV6	Histone H2A type 1-A	2,47	X	X	Yes
O14798	Tumor necrosis factor receptor superfamily member 10C *	2,41	X	X	Yes
Q99880	Histone H2B type 1-L	2,23	X	X	Yes
P62805	Histone H4	2,20	X	X	Yes
P04066	Tissue alpha-L-fucosidase	2,07	X	X	Yes
P31431	Syndecan-4	1,99	X	X	Yes
P17936	Insulin-like growth factor-binding protein 3	1,80	X	X	Yes
P13611	Versican core protein	1,76	X	X	Yes
O00754	Lysosomal alpha-mannosidase	1,75	X	X	Yes
Q9H4M9	EH domain-containing protein 1	1,75	1,45	1,30	
P00568	Adenylate kinase isoenzyme 1	1,66	1,89	1,39	
P11166	Solute carrier family 2, facilitated glucose transporter member 1	1,65	1,27	X	
P29317	Ephrin type-A receptor 2	1,60	X	X	Yes
P62879	Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta 2	1,58	1,82	X	
P35221	Catenin alpha-1	1,57	1,04	0,91	
P35052	Glypican-1	1,57	X	X	Yes
P36952	Serpin B5	1,55	X	1,60	
Q9Y624	Junctional adhesion molecule A	1,55	X	X	Yes
Q13510	Acid ceramidase	1,54	1,52	X	
P35222	Catenin beta-1	1,54	0,97	1,01	
P27824	Calnexin	1,53	0,97	X	
P30740	Leukocyte elastase inhibitor	1,53	X	1,21	
P10619	Lysosomal protective protein	1,53	1,05	X	

P13797	Plastin-3	1,51	X	X	Yes
O60716	Catenin delta-1	1,50	0,98	0,98	
P00390	Glutathione reductase, mitochondrial	1,50	1,31	1,01	
P04632	Calpain small subunit 1	1,49	X	X	Yes
Q15149	Plectin-1	1,47	1,35	1,37	
Q92520	Protein FAM3C	1,46	1,39	X	
P07602	Proactivator polypeptide	1,45	1,31	X	
Q14517	Cadherin-related tumor suppressor homolog	1,45	X	X	Yes
P49006	MARCKS-related protein	1,44	X	X	Yes
P80723	Brain acid soluble protein 1	1,44	1,06	1,06	
Q13308	Tyrosine-protein kinase-like 7	1,44	X	X	Yes
P60953	Cell division control protein 42 homolog	1,44	1,32	0,97	
Q9UNN8	Endothelial protein C receptor	1,43	X	X	Yes
O76070	Gamma-synuclein	1,41	X	X	Yes
Q08431	Lactadherin	1,41	X	X	Yes
P06703	Protein S100-A6	1,38	1,31	1,43	
Q9Y5Y6	Suppressor of tumorigenicity protein 14	1,38	X	X	Yes
P55786	Puromycin-sensitive aminopeptidase	1,35	0,84	1,19	
P31947	14-3-3 protein sigma	1,34	1,48	1,17	
P12830	Epithelial-cadherin	1,33	1,41	X	
P15559	NAD(P)H dehydrogenase [quinone] 1	1,28	0,98	0,94	
P20827	Ephrin-A1	1,28	X	X	Yes
P63104	14-3-3 protein zeta/delta	1,27	1,05	1,20	

Tableau 10 : Représentation des protéines modulées par p53 dans le secretome de la lignée H358/TetOn/p53 avec leur ratio iTRAQ éventuel obtenu dans les échantillons « cytosol » et « organelles & membranes ».

Les protéines sont listées avec leur numéro d'accession UniProt (No. Access). Les protéines avec un « X » ne sont pas détectées dans l'échantillon correspondant. Les protéines identifiées uniquement dans l'échantillon « secretome » sont notées « Yes » dans la colonne « Unique ». Les ratios en rouge sont obtenus à partir d'un peptide unique.

6.3.2. Sous-protéome des organites et des membranes

- 907 protéines ont été identifiées dans l'échantillon « Organites & Membranes » dont 44 modulées par p53 : 28 sont surexprimées et 16 sous-exprimées en présence de p53 (Tableau 11).

- 25 des 44 protéines différentielles ont été préalablement décrites comme étant localisées (au moins en partie) dans les organites ou la membrane cellulaire (Tableau 11). Les autres protéines sont soit non référencées, soit cytosoliques avec une affinité pour des protéines membranaires (et donc co-purifiées) soit plus probablement présentes à cause d'une contamination résiduelle de la fraction cytosolique précédente ou de la fraction nucléaire suivante. Certaines protéines se retrouvent également dans le réticulum endoplasmique ou l'appareil de Golgi durant leur maturation pour acquérir des modifications post-traductionnelles.

- Parmi les protéines surexprimées en présence de p53, 8 sur 28 sont des cibles transcriptionnelles de p53 préalablement décrites (Tableau 11). A l'inverse, 2 des 16 protéines sous-exprimées sont réprimées par p53 (Tableau 11).

- 16 protéines sur les 44 différentielles sont également retrouvées dans les 91 protéines du secretome modulées par p53. Dans les deux sous-protéomes, ces 16 protéines varient de façon similaire (Tableau 11 et Tableaux 1 et 2 de la publication).

A

Nom	No. Access.	Ratio +/-	SD	Ratio dox	Loc.	p53	Secr.
DNA replication licensing factor MCM3	P25205	0.44	0.18	1.22	N		
Transferrin receptor protein 1	P02786	0.44	0.15	1.25	mb+S (NC)		X
DNA replication licensing factor MCM4	P33991	0.46	0.02	0.99	N		
6-phosphofructokinase, liver type	P17858	0.46	0.21	0.94	C		
4-hydroxyphenylpyruvate dioxygenase-like protein	Q96IR7	0.55	0.07	1.18	L		
DNA replication licensing factor MCM7	P33993	0.57	0.12	0.91	N	R	
DNA replication licensing factor MCM2	P49736	0.59	0.08	0.97	N		
Deoxyuridine 5'-triphosphate nucleotidohydrolase	P33316	0.61	0.17	0.79	M+S (NC)	R	X
DNA replication licensing factor MCM6	Q14566	0.62	0.22	1.02	N		X
NADH dehydrogenase iron-sulfur protein 2	O75306	0.70	0.16	0.91	M		
Nucleophosmin	P06748	0.70	0.04	0.96	C+N	I	
NADH dehydrogenase 1 alpha subcomplex subunit 9	Q16795	0.71	0.08	0.78	M		
ATPase family AAA domain-containing protein 3A	Q9NVI7	0.72	0.05	0.81	N		
ATP-dependent DNA helicase 2 subunit 1	P12956	0.74	0.05	0.79	C+N	L	
Translocon-associated protein subunit gamma	Q9UNL2	0.74	0.10	1.19	RE+mb		
FACT complex subunit SSRP1	Q08945	0.75	0.07	0.99	N		

B

Nom	No. Access.	Ratio +/-	SD	Ratio dox	Loc.	p53	Secr.
Carboxylesterase 2	O00748	4.39	1.09	1.23	RE	I	
NADPH:adrenodoxin oxidoreductase	P22570	3.64	0.25	1.19	M	I	
Apoptosis regulator BAX	Q07812	2.48	0.98	0.86	M+C	I	
Adenylate kinase isoenzyme 1	P00568	1.89	0.57	0.71	C+S (NC)	I	X
Guanine nucleotide-binding protein G(I)/G(S)/G(T) subunit beta-2	P62879	1.82	0.68	-	mb + exo	I	X
Acid ceramidase	Q13510	1.52	0.33	1.21	L+S	I	X
Pre-B-cell leukemia transcription factor-interacting protein 1	Q96AQ6	1.51	0.22	1.17	C		
14-3-3 protein sigma	P31947	1.48	0.16	0.74	C+N + S (NC)	I	X
NAD-dependent malic enzyme	P23368	1.47	0.20	1.05	M		
EH domain-containing protein 1	Q9H4M9	1.45	0.24	1.26	mb+V +exo		X
Basic leucine zipper and W2 domain-containing protein 1	Q7L1Q6	1.45	0.14	1.04	C		
Transmembrane emp24 domain-containing protein 4	Q7Z7H5	1.44	0.14	0.99	RE+mb		
Epithelial cadherin	P12830	1.41	0.15	1.21	mb		X
Abhydrolase domain-containing protein 11	Q8NFV4	1.41	0.32	0.93	M		

Protein FAM3C	Q92520	1.39	0.09	0.94	M+S		X
Coatomer subunit gamma	Q9Y678	1.36	0.12	1.14	G+C		
Plectin-1	Q15149	1.35	0.14	0.97	mb+C +exo		X
Protein S100-A11	P31949	1.34	0.07	0.93	C+N		
Myosin light polypeptide 6	P60660	1.34	0.14	1.12	C		
Sterol-4-alpha-carboxylate 3-dehydrogenase, decarboxylating	Q15738	1.34	0.23	0.96	mb		
Major vault protein	Q14764	1.33	0.13	0.93	C+N		
Cell division control protein 42 homolog	P60953	1.32	0.26	0.91	mb + exo	I	X
Protein S100-A6	P06703	1.31	0.11	0.78	C+N + exo	I	X
Glutathione reductase	P00390	1.31	0.22	1.03	M		X
Proactivator polypeptide	P07602	1.31	0.11	0.89	L+S		X
Protein transport protein Sec24D	O94855	1.28	0.07	0.93	RE+G +C		
Solute carrier family 2, facilitated glucose transporter member 1	P11166	1.27	0.07	0.92	mb + exo	L	X
Nicotinamide phosphoribosyltransferase	P43490	1.26	0.11	1.04	C		

Tableau 11 : Les protéines modulées par p53 dans la fraction « organelles & membranes » des cellules H358/TetOn/p53.

Les protéines sont listées avec leur numéro d'accession UniProt (No. Access.). Seules les protéines avec un score ProteinPilot > 4 (2 peptides distincts avec 99% de confiance) sont considérées. Les ratios donnés correspondent à la moyenne des ratios obtenus lors de l'expérience iTRAQ 8-plex. Le ratio +/- correspond à la quantification relative des protéines dans la condition avec p53 par rapport à celle sans p53. Les protéines sous-exprimées en présence de p53 ont un ratio < 0.8 (A) et les surexprimées > 1.25 (B). La déviation standard inter-ratio est notée pour chaque protéine (SD). L'effet dox est contrôlé sur les cellules H358 non transfectées et soumises ou non à la doxycycline (ratio dox+ / dox-). La localisation cellulaire des protéines est obtenue de la littérature (mb: Membrane; M: Mitochondrie; RE: Reticulum Endoplasmique; G: appareil de Golgi; L: Lyzosome; V: Vésicule; C: Cytosol; N: Noyau; U : « Unknown » ; S : Sécrété ; NC : sécrété par une voie non-classique ; exo : sécrété par la voie des exosomes). Les relations entre p53 et les protéines modulées (colonne « p53 ») sont issues de la littérature (I : cible transcriptionnelle induite par p53; R : cible transcriptionnelle réprimée par p53; L: protéine ayant un lien avec p53).Les protéines modulées de façon similaire dans les échantillons du cytosol et du secretome sont cochées dans la colonne « Secr ».

- 27 des 91 protéines modulées par p53 dans le secretome sont également présentes dans l'échantillon « Organites & Membranes » (Figure 68 et Tableau 10). Comme dit précédemment, 16 de ces protéines sont modulées de façon similaire. Ceci n'est pas surprenant car les protéines sécrétées par la voie classique cheminent via le reticulum endoplasmique et l'appareil de Golgi. De plus, les protéines membranaires peuvent être retrouvées dans le milieu extracellulaire via la voie exosomale ou par libération d'une partie extracellulaire de ces protéines par activité protéolytique. Les 11 autres protéines sont inchangées dans les organelles et membranes et doivent donc être modulées au niveau de leur sécrétion ou stabilité.

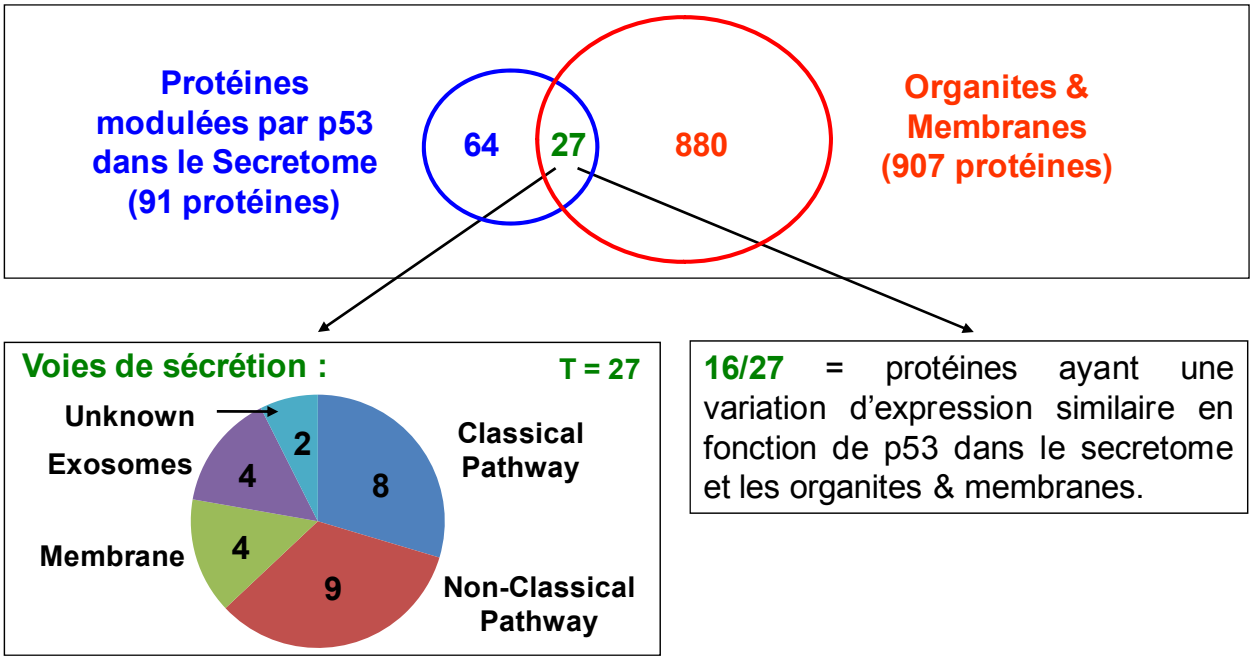


Figure 68 : Protéines modulées par p53 dans le secretome *versus* protéome des organelles et des membranes

Haut : Diagramme de Venn représentant la proportion de protéines modulées par p53 dans le secretome identifiées également dans le sous protéome « Organites & Membranes ».

Bas : Voies de sécrétion connues ou prédites pour les 27 protéines du secretome identifiées également dans les Organites et Membranes.

Sur ces 27 protéines, 8 sont sécrétées par la voie classique, 9 par une voie non classique, 4 sont membranaires, 4 relarguées par les exosomes, et 2 ne sont pas connues comme sécrétées (voie exosomale ?) (Figure 68). Ces résultats corroborent avec le fait de les retrouver dans les deux sous-protéomes.

6.4. Synthèse

Les 91 protéines modulées par p53 dans le secretome sont représentées dans le Tableau 10. Leur détection dans les deux autres sous-protéomes étudiés (« cytosol » et « organites&membranes ») ainsi que leur ratio d'expression sont également renseignés.

La Figure 69 récapitule la distribution des 91 protéines différentielles du secretome dans les autres protéomes. 50 % des protéines surexprimées et 88,6 % des sous-exprimées dans le secretome sont retrouvées uniquement dans ce sous-protéome. Cette observation corrobore avec une action de p53 sur le déclenchement de la sécrétion de protéines par des voies non classiques.

A

Protéines modulées par p53 dans le Secretome

Présence	Nb prot	%
Unique Secretome	62	68,1
Organites & Membranes	12	13,2
Cytosol	3	3,3
Organites & Membranes + Cytosol	14	15,4
TOTAL	91	100,0

Protéines surexprimées dans le Secretome en présence de p53

Présence	Nb prot	%
Unique Secretome	22	46,8
Organites & Membranes	8	17,0
Cytosol	3	6,4
Organites & Membranes + Cytosol	14	29,8
TOTAL	47	100,0

Protéines sous-exprimées dans le Secretome en présence de p53

Présence	Nb prot	%
Unique Secretome	39	88,6
Organites & Membranes	5	11,4
Cytosol	0	0,0
Organites & Membranes + Cytosol	0	0,0
TOTAL	44	100,0

B

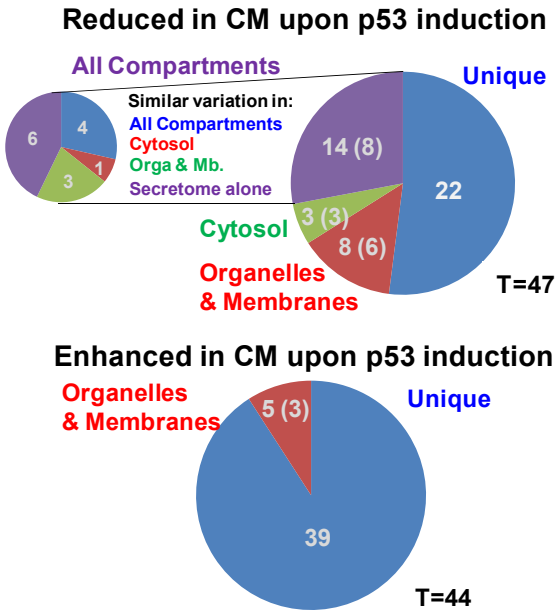


Figure 69 : Spécificité des protéines modulées par p53 dans le Secretome.

A. Tableaux d'analyse de la spécificité des protéines modulées par p53 identifiées dans le secretome.
Unique Secretome : protéines retrouvées uniquement dans le secretome ; Organites&Membranes : protéines du secretome aussi retrouvées dans l'échantillon « Organites&Membranes » ; Cytosol : protéines du secretome aussi retrouvées dans l'échantillon « Cytosol » ; « Organites&Membranes + Cytosol » : protéines retrouvées dans les 3 sous-protéomes.
Les résultats sont présentés sous forme de nombre de protéines (Nb prot) et de pourcentage (%).

B. Représentation sous forme de diagramme de Ven de la spécificité des protéines modulées par p53 identifiées dans l'échantillon « Secretome ». Les valeurs entre parenthèses représentent le nombre de protéines variant de façon similaire dans le secretome et le protéome concerné.

Troisième partie :

Le Secretome : une source de bio-marqueurs potentiels ?

CHAPITRE 3 : Le Secretome : une source de bio-marqueurs ?

1. Problématique et objectifs du travail

Les tumeurs libèrent de nombreux facteurs impliqués dans leur développement. Ceux-ci sont généralement retrouvés dans les fluides biologiques tels que le sang ou l'urine et peuvent donc théoriquement y être mesurés par des tests non invasifs [256]. La découverte de tels marqueurs est un véritable challenge de par leur utilisation potentielle en diagnostic, pronostic ou thérapeutique.

Le sang possède un contenu très riche et son prélèvement est un acte peu invasif pour le patient. Cela en fait donc une source de marqueurs idéale. Malheureusement, leur recherche par des approches protéomiques directes se heurte actuellement à la grande complexité du fluide sanguin, tant en terme de diversité de protéines et d'isoformes que de leur très large gamme dynamique de concentration (ordre de magnitude de 9 à 10) [259]. Les protéines bio-marqueurs potentiels sont généralement très faiblement concentrées dans le sang par rapport aux protéines abondantes. Des approches de déplétion visant à éliminer ces protéines majoritaires ont ainsi été développées [33, 266]. Néanmoins, leur efficacité reste relative car l'élimination des protéines ciblées n'est jamais totale et un certain nombre de protéines non concernées sont aussi enlevées soit par fixation non spécifique, soit par affinité avec les protéines devant être déplétées [267]. L'approche opposée vise à enrichir les protéines minoritaires. Celle-ci est en développement avec le système Protein Equalizer Technology (Biorad) mais son efficacité réelle est encore à démontrer [268].

En conséquence, le passage préalable par un modèle cellulaire constitue une approche pertinente afin d'établir une liste de marqueurs potentiels à rechercher et valider dans un second temps dans des cohortes de plasmas de patients. Le secretome est un modèle séduisant car les protéines libérées dans le milieu extracellulaire sont très souvent retrouvées dans le plasma. Notre modèle de xénogreffe chez la souris, développé dans la publication du chapitre précédent, appuie cette hypothèse. Chez l'homme, les variations de concentrations protéiques liées à une tumeur devront être suffisamment importantes pour être significatives et détectables dans les fluides biologiques.

Le gène *TP53* présente un fort taux de mutation dans les cancers du poumon (plus de 60 % des cas). En conséquence, les protéines modulées par p53 dans le secretome de la lignée H358 constituent des candidats bio-marqueurs potentiels (Chapitre 2).

Nous avons voulu ici vérifier si les protéines d'intérêt du secretome de la lignée H358 (protéines modulées par p53) étaient retrouvées dans le plasma humain sain. Dans cette optique une approche dot blot a été réalisée.

Dans un second temps, une étude préliminaire a été entreprise pour doser ces protéines dans des prélèvements biologiques. Ces dosages par ELISA, pour une faible cohorte de patients atteints d'un cancer du poumon avec caractérisation du statut p53, pourrait permettre de pointer certains marqueurs potentiels.

2. Matériels et méthodes

2.1. Echantillons

Pool de plasmas sains : plasma CQI (Contrôle Qualité Interne) fourni par la plate-forme protéomique de Dijon (IFR-santé-STIC, France). Ces plasmas sont aliquotés et conservés à -130°C jusqu'à utilisation. Ils seront utilisés pour les tests en dot-blot.

Plasmas de patients : 10 plasmas de patients atteints de cancers du poumon non à petites cellules de type adénocarcinome sont fournis par Dr Florence de Fraipont (protocole BIO-IFCT002). Les extraits protéiques totaux de la tumeur correspondant sont aussi disponibles. La caractérisation du statut p53 a été réalisée au sein du laboratoire de Pierre Hainaut (IFCT de Lyon) (Tableau 12). Cinq plasmas contrôles d'individus sains de sexe et âge appariés sont également utilisés.

Secretome des cellules H358/TetOn/p53 exprimant ou non la protéine p53 sauvage (voir Chapitre 2).

Extraits totaux des cellules H358/TetOn/p53 exprimant ou non la protéine p53 sauvage. Ces extraits protéiques sont obtenus après extraction dans un tampon RIPA.

					Statut p53		
N° IFCT	**Type**	**Naissance**	**Sexe**	**Prélèvement**	**codon**	**mutation**	**Changt. aa**
022102	ADK NSCLC	27/12/1949	masculin	25/05/2002	wt		
022106	ADK NSCLC	20/05/1946	masculin	20/03/2003	248	G->C	Arg->Pro
041113	ADK NSCLC	14/03/1934	masculin	26/12/2002	298	G->T	Glu->STOP
042105	ADK NSCLC	03/04/1946	masculin	20/05/2002	wt		
042125	ADK NSCLC	23/03/1956	masculin	03/07/2003	245	G->T	Gly->Cys
042126	ADK NSCLC	21/02/1934	masculin	29/09/2003	wt		
042129	ADK NSCLC	17/08/1956	masculin	26/04/2004	wt		
242227	ADK NSCLC	31/08/1949	masculin	17/02/2004	249	G->T	Arg->Ser
511105	ADK NSCLC	06/12/1945	masculin	23/04/2004	245	G->T	Gly->Cys
542103	ADK NSCLC	13/01/1944	masculin	28/11/2003	wt		

Tableau 12 : Liste des patients atteints d'un NSCLC de type ADK inclus dans l'étude ELISA.

2.2. Dot-blot

Les dot-blots sont réalisés sur le système DHM-96 (Scie-Plas) en utilisant des membranes de nitrocellulose (Hybond ECL membrane de GE Biosciences). Les anticorps primaires et secondaires utilisés sont les mêmes que lors de l'étude secretome (voir la publication du chapitre 2 des résultats). La révélation s'effectue avec le système de détection ECL (GE Biosciences).

2.3. ELISA

Les dosages ELISA sont réalisés pour les protéines FGF-19, GDF-15 (R&D Systems) et Kallikrein-6 (IBEX Pharmaceuticals) à partir de kits commerciaux en appliquant les protocoles fournisseurs.

3. Résultats

3.1. Analyses par dot blot

Nous avons voulu vérifier la présence des protéines modulées par p53 dans le plasma. Nous disposons d'un panel d'anticorps dirigés contre plusieurs protéines d'intérêt et d'un pool de plasma d'individus sains. Les anticorps sont utilisés pour rechercher, grâce à la technique dot blot, la présence de ces protéines dans le plasma.

L'approche dot blot a été validée au préalable sur les échantillons de secretome issus de cellules H358/TetOn/p53 exprimant ou non p53. La Figure 70A valide les différences d'expression attendues. La quasi-totalité de ces protéines est également détectable dans le plasma (Figure 70B). Elles sont donc libérées dans la circulation de façon physiologique par des cellules de l'organisme.

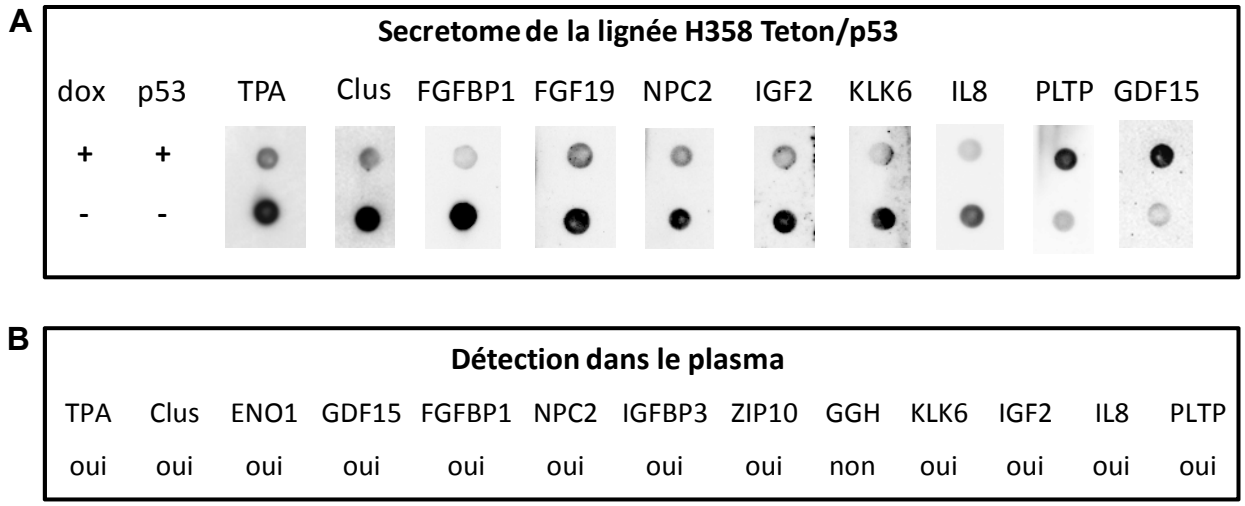


Figure 70 : Criblage des protéines par la technique dot-blot.

A. Validation des différences d'expression connues pour 10 protéines dans les échantillons de secretome issus de cellules H358/TetOn/p53 ré exprimant ou non p53.

B. Détection de la présence des protéines d'intérêt dans le plasma standardisé CQI.

A ce stade, nous pouvons émettre l'hypothèse que la libération conditionnelle de protéines par la tumeur entraîne une variation de leur concentration plasmatique suffisante pour y être détectée.

Afin de tester cette hypothèse, nous avons entrepris de doser en ELISA 3 protéines d'intérêt : GDF-15, FGF-19 et Kallikrein-6. Ces protéines sont détectées en dot blot dans le plasma (Figure 70B) et ont une importante modulation d'expression par p53 dans le secretome de la lignée H358/TetOn/p53 (Tableaux 1 et 2 de la publication du chapitre 2 des résultats).

3.2. Dosage par ELISA

Le dosage des protéines GDF-15, FGF-19 et Kallikrein-6 est réalisé dans différents buts.

Le premier est de détecter et doser ces protéines dans le plasma.

Le second est de réaliser une étude préliminaire sur des plasmas et des extraits protéiques totaux de biopsies de patients atteints de cancer du poumon non à petites cellules (avec caractérisation du statut p53) dans le but d'appréhender une éventuelle variation d'expression de ces protéines.

A	Extraits totaux (ng/mg prot. tot.)		Plasma CQI (pg/mL)
Echantillons	ET_p53	ET	459,6
[GDF-15]	12	1,6	
Ratio p53+/p53-	7,53		

B	Extraits totaux (pg/mg prot. tot.)		Plasma CQI (pg/mL)
Echantillons	ET_p53	ET	107,1
[FGF-19]	305,7	439,2	
Ratio p53+/p53-	0,70		

C	Extraits totaux (ng/mg prot. tot.)		Plasma CQI (ng/mL)
Echantillons	ET_p53	ET	2,76
[KLK6]	257,4	466,8	
Ratio p53+/p53-	0,55		

Tableau 13 : Dosages ELISA des protéines GDF-15 (A), FGF-19 (B) et Kallikrein-6 (C).

Les dosages sont réalisés sur des extraits totaux de protéines de la lignée H358/TetOn/p53 induite ou non pour l'expression de wt-p53 et sur du plasma contrôle CQI. Les ratios p53+/p53- représentent les rapports de concentration entre les échantillons exprimant ou non p53.

Le dosage des protéines dans les extraits protéique totaux de la lignée H358/TetOn/p53 est réalisé pour estimer la quantité d'extrait protéique total de biopsie nécessaire (Tableau 13). Les protéines sont dosables dans 100 µg d'extrait total. Nous partirons donc sur ces valeurs pour les dosages dans les extraits de biopsie.

Les trois protéines sont détectées et dosées dans le plasma sain (Tableau 13). La Kallikrein-6 est ainsi dosée à 2.76 ng/mL, GDF-15 à 460 pg/mL et FGF-19 à 107 pg/mL.

Ces deux contrôles servent à optimiser les conditions expérimentales dans le but de réaliser une étude préliminaire sur des échantillons cliniques. Les patients ont été diagnostiqués comme porteur d'un cancer du poumon non à petites cellules de type adénocarcinome et leur statut p53 a

été préalablement caractérisé. L'étude englobe 10 patients et 5 contrôles sains appariés (protocole BIO-IFCT002) (Tableau 12). Les dosages sont effectués sur les plasmas et les extraits totaux de tumeurs.

En clinique, la délétion totale de p53 est très rarement observée. Le gène *TP53* est généralement muté, ce qui aboutit à une protéine p53 plus ou moins fonctionnelle selon le lieu et l'incidence de la mutation sur la conformation de la protéine. Les 10 patients étudiés ont été sélectionnés de façon à constituer deux groupes homogènes en âge et en sexe. Un groupe est constitué de 5 patients possédant un gène *TP53* sauvage et l'autre d'un nombre équivalent avec un gène muté. Afin d'avoir une incidence sur la fonctionnalité de p53, quatre mutations sont localisées sur l'un des « hotspots » (codons 245 (x2), 248, 249), et une est responsable de la génération d'une protéine tronquée (E298Stop). Les dosages ELISA sont réalisés en aveugle.

Cette étude préliminaire n'a pas donné de résultats concluants. Chacune des trois protéines a été dosée dans le plasma (Tableau 14). Néanmoins, les concentrations déterminées ne révèlent aucune différence significative entre chaque groupe (« sain » *versus* « cancéreux p53+ » *versus* « cancéreux p53- »). Par ailleurs, aucune protéine n'a été détectée dans les extraits totaux de tumeur, probablement à cause d'une quantité de protéines analysées insuffisante.

Patients	Statut	[C] Protéines / p53	KLK6 (ng/mL)	FGF19 (pg/mL)	GDF15 (ng/mL)
3	NSCLC	E298Stop	3,04	69,91	1647,94
5	NSCLC	G245C	4,80	45,00	917,51
9	NSCLC	G245C	2,55	36,74	698,86
2	NSCLC	R248P	4,16	38,54	1215,33
8	NSCLC	R249S	3,12	483,80	731,81
1	NSCLC	wt	4,27	89,24	1144,20
4	NSCLC	wt	4,69	90,36	789,75
6	NSCLC	wt	4,01	220,99	1410,39
7	NSCLC	wt	3,90	121,18	967,21
10	NSCLC	wt	4,35	162,40	887,18
11	sain	wt	4,92	115,27	979,28
12	sain	wt	4,37	135,98	1176,68
13	sain	wt	5,16	328,01	835,80
14	sain	wt	4,70	137,08	877,73
15	sain	wt	4,73	80,79	971,23

Tableau 14 : Dosage plasmatique par ELISA des protéines KLK6, FGF19 et GDF15.

Les dosages ELISA sont réalisés avec des kits commerciaux sur des plasmas issus de 3 groupes de 5 individus. Groupe 1 (patients 2, 3, 5, 8, 9) : individus ayant un adénocarcinome du poumon avec mutation de *TP53* ; groupe 2 (patients 1, 4, 6, 7, 10) : individus ayant un adénocarcinome du poumon mais un *TP53* sauvage ; groupe 3 : individus sains appariés.

Au vu des résultats, aucune variation significative des concentrations plasmatiques de GDF-15, FGF-19 et Kallikrein-6 n'a été observée en présence de la tumeur et/ou d'un statut p53 muté. Les tissus normaux libèrent physiologiquement ces protéines dans le sang comme l'atteste les dosages effectués sur les plasmas de sujets sains. En conséquence, la modulation de la libération de ces protéines par la tumeur doit être suffisante pour induire un changement de concentration détectable dans le plasma, ce qui n'est sans doute pas le cas. Il est ainsi dommage que les dosages effectués sur les extraits totaux n'aient pas fonctionné car ils auraient permis d'observer directement l'expression des protéines au niveau de la tumeur.

Notre modèle cellulaire H358/TetOn/p53 (exprimant ou non p53 sauvage) ne correspond pas aux échantillons cliniques sélectionnés. Bien que les mutations choisies soient parmi les plus délétères pour sa fonction, le domaine de p53 modulant ces protéines n'est pas forcément affecté.

A

Patients	état	p53
1	CPNPC ADK	wt
2	CPNPC ADK	R248P
3	CPNPC ADK	E298Stop
4	CPNPC ADK	wt
5	CPNPC ADK	G245C
6	CPNPC ADK	wt
7	CPNPC ADK	wt
8	CPNPC ADK	R249S
9	CPNPC ADK	G245C
10	CPNPC ADK	wt
11	sain	wt
12	sain	wt
13	sain	wt
14	sain	wt
15	sain	wt

B

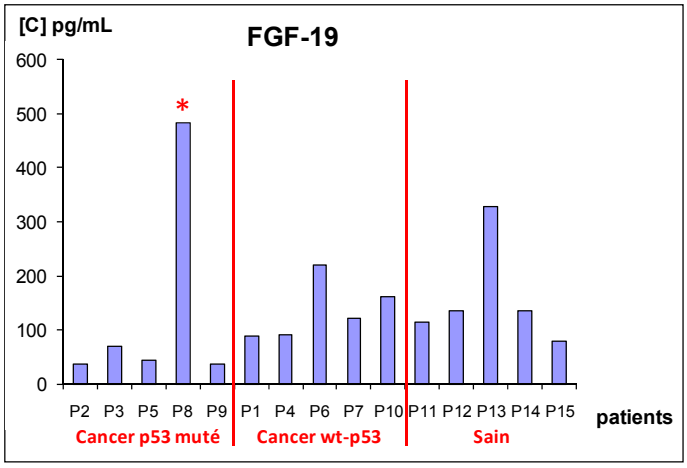


Figure 71 : Exemple de dosages ELISA de la protéine FGF-19 dans les plasmas.

A. Tableau des individus pour qui FGF-19 a été dosée dans le plasma. 5 individus ont un cancer du poumon avec mutation de p53, 5 un cancer du poumon sans mutation de p53 et 5 sont sains.

B. Représentation graphique des dosages de FGF19 pour chaque individu.

De plus, cette étude préliminaire a été réalisée sur une très faible cohorte (5 échantillons par groupe), qui n'est pas statistiquement fiable. Le groupe « p53 muté » n'est pas homogène car il regroupe différentes mutations de *TP53* qui n'ont probablement pas le même effet sur l'activité de p53. Le patient portant la mutation R249S, l'une des plus délétères décrites, a une concentration sérique de FGF-19 significativement augmentée (Figure 71). Ce résultat pourrait concorder avec nos données précédentes mais puisqu'aucun autre patient possédant cette mutation n'a été étudié, on ne peut donc rien conclure.

Au final, aucune information exploitable n'est ressortie de cette étude préliminaire. Des essais sur des cohortes plus grandes, avec des groupes de mutants mieux ciblés et plus homogènes seraient

intéressants à réaliser. Le dosage d'autres protéines d'intérêt pourrait également être pertinent. Enfin, l'analyse des protéines directement au niveau de la tumeur serait une étape intermédiaire très importante.

4. Conclusion

Les essais dot blot ont validé la présence dans le plasma des protéines du modèle secretome.

L'étude préliminaire de dosage en ELISA des protéines GDF-15, KLK6 et FGF-19 sur une faible cohorte de patients n'a pas donné de résultats concluants. Ces protéines sont détectées dans le plasma mais pas dans les extraits tumoraux. Les concentrations dans le plasma ne mettent pas en évidence de différences significatives.

DISCUSSION GENERALE

DISCUSSION GENERALE

Les travaux présentés ont visé à démontrer l'intérêt des études secretome pour la compréhension des changements de communication et d'interaction qui s'opèrent entre les cellules tumorales et leur environnement durant le développement de la tumeur. La méthodologie développée s'étend de l'analyse protéomique sur un modèle cellulaire *in vitro* à la validation de sa pertinence *in vivo*. Elle a été appliquée pour étudier l'influence d'un facteur clé de la régulation anti-tumorale, la protéine p53. La modulation du secretome par p53 a ainsi été investiguée dans un modèle de lignée cellulaire d'adénocarcinome du poumon non à petites cellules puis mise en valeur *in vivo* chez la souris. Outre son rôle primordial dans la régulation intracellulaire, p53 est donc aussi impliqué dans la modulation de l'environnement et de la communication cellulaire. De plus, nous avons démontré pour la première fois la relation entre les protéines du secretome et leur libération dans les fluides biologiques. Ces résultats font des études « secretome » des approches de choix pour la mise en valeur de cibles protéiques d'intérêt clinique.

1. Le modèle cellulaire et le procédé d'analyse

La première partie de ce travail a consisté à développer un mode opératoire adapté à l'analyse différentielle des secretomes. Il a nécessité la sélection d'un modèle biologique et d'approches protéomique adéquates.

Le choix de la lignée tumorale H358 comme modèle cellulaire s'explique par différentes raisons. Cette lignée a été bien étudiée et caractérisée au laboratoire [261, 263]. Elle est dérivée d'un CPNPC, type de cancer du poumon le plus couramment rencontré. Elle est délétée en *TP53* et l'expression conditionnelle de la protéine p53 sauvage, par transfection d'un plasmide TetOn/p53, provoque l'arrêt du cycle cellulaire sans augmentation d'apoptose. Cette lignée est aussi résistante à l'apoptose due à la privation de sérum de veau fœtal (SVF) dans le milieu de culture. Toutes ces caractéristiques en font un modèle de choix pour notre étude.

La réussite d'une étude protéomique dépendant grandement de la préparation de l'échantillon protéique, un protocole adapté au « secretome » a été développé. Les cellules sont passées en milieu sans sérum pour éviter toute contamination extérieure de l'échantillon (protéines du SVF). Après 72 heures de culture (temps adéquat pour que les cellules libèrent suffisamment de protéines sans augmentation d'apoptose significative), les milieux conditionnés contenant les protéines du secretome sont collectés, puis filtrés et centrifugés de façon à éliminer les cellules

flottantes et débris cellulaire. Les protéines, très diluées dans le milieu, sont alors concentrées et dessalées par ultrafiltration.

Le modèle biologique mis en place a montré son efficacité mais est-il suffisamment solide ? En effet, quelle proportion de protéines identifiées appartient réellement au sous-protéome « secretome » ?

Les cellules sont initialement cultivées en milieu complémenté en SVF. Lors de la préparation du secretome, elles sont lavées 6 fois pour éliminer les protéines sériques puis cultivées en milieu sans SVF. Malgré ces précautions, une légère contamination est toujours possible. Nous l'avons évalué en soumettant les données de spectrométrie de masse à une recherche dans une base de données mammifère. Seule l'albumine bovine est retrouvée, avec un score relativement faible. La contamination résiduelle par le SVF est donc très limitée et ne perturbe pas l'analyse des protéines d'intérêt.

Le secretome n'est également pas à l'abri d'une contamination par des protéines relarguées par nécrose ou éclatement des cellules. Nous avons contrôlé au maximum ce risque. Nous pouvons nous appuyer, tout d'abord, sur les travaux publiés au laboratoire, montrant que la culture en milieu sans SVF et l'expression de p53 sauvage n'induisaient pas d'augmentation d'apoptose [261, 263]. L'observation des cellules au microscope ne révèle pas d'aspect morphologique de souffrance anormal. De plus, seules 91 des 909 protéines identifiées dans le secretome sont modulées par p53. Si p53 induisait une augmentation importante de la mort cellulaire, les variations du secretome seraient plus drastiques. La plupart des protéines modulées par p53 sont connues comme sécrétées et donc leur présence dans le secretome est justifiée. Enfin, le suivi de l'activation de la pro-caspase-3 en cytométrie de flux n'a pas révélé d'augmentation d'apoptose liée à p53 (1-1,6 % de cellules en apoptose dans les deux conditions). La combinaison culture sans SVF et expression de p53 n'augmente donc pas l'apoptose des cellules. Les 1% d'apoptose correspondent aux taux habituellement observés lors de culture cellulaire *in vitro*. Celle-ci peut entrainer une légère contamination du secretome mais n'aura pas d'influence sur sa modulation entre les deux conditions. De plus, il faut resituer ce phénomène par rapport à la définition du terme « secretome », c'est-à-dire l'ensemble des protéines libérées dans le milieu conditionné par des cellules en culture. D'un certain côté, la mort cellulaire « naturelle » est un mode de libération de protéines ! Le secretome contient aussi certaines protéines cytosoliques libérées par des voies non-classiques (exosomes,…). L'étude des autres protéomes de la lignée H358 démontre une bonne spécificité de l'échantillon « secretome ».

Une autre question peut être posée : les cellules libèreraient elles les mêmes protéines si elles n'étaient pas cultivées en milieu sans SVF ? La réponse est non, l'absence de SVF perturbe forcément certaines voies cellulaires. Il faut garder en mémoire que le modèle H358 utilisé pour l'étude protéomique a des caractéristiques qui lui sont propres à cause des passages répétés de la lignée ou des voies particulières exprimées ou non (régulation de la voie amphiréguline/EGFR, dérégulation de la voie p53,…). Avant de pouvoir être généralisés, les résultats obtenus nécessitent de nombreuses validations *in vitro* sur d'autres modèles cellulaires, *in vivo* chez la souris ou sur des échantillons cliniques (biopsies, plasmas de patients…). Les résultats obtenus *in vivo* chez la souris confirment néanmoins les variations observées pour les trois protéines étudiées (GDF-15, FGF-19 et VEGF). De plus, les protéines modulées par p53 identifiées dans différents protéomes de la lignée H358 (protéine S100A6, plectin-1, PIG3, Serpin B5,…) varient de façon similaire. Or pour la préparation des échantillons autre que « secretome », les cellules ont été cultivées en milieu avec SVF. Le modèle utilisé reflète donc au moins en partie la réalité.

Le deuxième point important repose sur la qualité du fractionnement de l'échantillon et la pertinence du procédé d'analyse protéomique établi.

Le procédé mis en place intègre : marquage des peptides trypsiques de chaque échantillon par un tag iTRAQ spécifique puis mélange de l'ensemble, séparation bidimensionnelle couplant une approche OFFGEL et une chromatographie nano phase inverse (RP nano-LC), dépôt sur cible MALDI et analyse en mode MS/MS par MALDI-TOF/TOF (identification et quantification).

Nous avons validé l'utilisation de l'isoélectrofocalisation en solution OFFGEL comme nouvelle approche de fractionnement d'échantillons complexes. Celle-ci est utilisable pour focaliser au choix protéines ou peptides. Les séparations en fonction du pI (OFFGEL) et du caractère hydrophobe (RP-LC) s'avèrent très complémentaires. De plus, nous avons démontré que l'approche OFFGEL est compatible avec la méthode d'analyse différentielle par marquage iTRAQ. Les étapes de séparation sont ainsi effectuées sur le mix de peptides marqués issus des différents échantillons, ce qui diminue grandement les biais expérimentaux. Cette étude a permis d'expliquer les profils de séparation obtenus en OFFGEL et de valider cette approche pour l'analyse différentielle d'échantillons complexes dont le secretome. Ce travail s'est avéré primordial pour la suite des analyses, une séparation optimale des échantillons étant bien souvent l'étape clé de la réussite d'une étude protéomique. Cette technologie OFFGEL est d'ailleurs maintenant intégrée dans la plupart des projets LC-MS/MS développés sur la plate-forme.

Ce procédé analytique présente des avantages certains. Il repose sur l'utilisation d'un matériel dernière génération (OFFGEL, nanoHPLC, MALDI TOF/TOF). L'analyse différentielle utilise la

technologie de marquage isotopique iTRAQ, l'une des plus récentes et mieux adaptée. L'emploi de ProteinPilot comme moteur de recherche permet d'exploiter au maximum nos données sans augmenter le risque de faux positifs.

Néanmoins, certaines réserves peuvent être émises. Malgré ses nombreux avantages, le marquage iTRAQ se fait relativement tard dans le procédé de préparation des échantillons. En effet, il est réalisé sur les peptides obtenus après digestion à la trypsine. Les différents biais expérimentaux pouvant être introduits en amont du marquage ne sont pas pris en compte : les cellules de chaque condition sont cultivées indépendamment, et la préparation (concentration,…) et digestion des protéines sont également réalisées en parallèle. Il y a donc un risque de générer des faux positifs. Ce problème pourrait en partie être résolu par la réalisation d'une étude complémentaire par marquage de type SILAC qui s'effectue très en amont durant la culture cellulaire. Néanmoins, d'autres inconvénients apparaitraient : complexification du spectre MS, analyse 2-plex,… L'idéal serait, en fait, de pouvoir combiner les avantages des méthodes SILAC et iTRAQ, malheureusement une telle approche n'existe pas encore (!!).

Nous avons donc minimisé au mieux les biais dus à l'iTRAQ. Deux analyses iTRAQ ont ainsi été réalisées. Dans une première expérience, une analyse 4-plex a été effectuée sur deux paires (p53 ou non induit) d'échantillons indépendants pour contrôler la reproductibilité inter-échantillons. Une seconde analyse 2-plex a été entreprise sur l'une des 2 préparations d'échantillons en inversant les tags iTRAQ afin d'éliminer tout effet potentiel tag-dépendant. Seuls les peptides (et donc protéines) modulés de façon significative dans les 3 cas sont considérés comme variant de façon significative. ProteinPilot effectue également une correction automatique des biais expérimentaux prenant en compte les éventuels ajouts de quantité inégale de protéines lors du mélange des échantillons marqués. Enfin, l'analyse par électrophorèse 2D ainsi que les Western blots ont permis de valider les différences observées en iTRAQ.

Le procédé d'analyse protéomique développé est donc solide et pourrait être applicable à des études similaires sur d'autres modèles ou lignées cellulaires.

2. Les protéines modulées par p53 et la tumorogenèse

2.1. L'étude du secretome

L'analyse différentielle du secretome de la lignée H358 a permis d'identifier 909 protéines dont 91 significativement influencées par p53. Plusieurs facteurs impliqués dans des voies classiquement associées à la tumorogenèse sont retrouvés. D'autres protéines moins connues représentent des cibles potentielles d'intérêt. Ces résultats sont utiles pour mieux comprendre le rôle joué par p53 dans la tumorogenèse à travers sa régulation de facteurs sécrétés.

2.1.1. Influence sur les voies classiques de dégradation de la matrice extracellulaire

Le caractère invasif d'une cellule tumorale s'accompagne d'un remaniement de la matrice extracellulaire (MEC) impliquant sa dégradation par l'activation de différentes cascades protéolytiques telles que le système uPA-uPAR, les MMPs et les cathepsines (voir Chapitre 5 de la partie bibliographie).

Un certain nombre des protéines impliquées dans ces voies sont modulées par l'expression de p53 dans le secretome de la lignée H358. Les voies uPA-uPAR et MMPs sont vraisemblablement influencées mais pas les cathepsines B et D dont le niveau dans le secretome reste inchangé que p53 soit ou non exprimé (Figure 72).

Il est intéressant de constater que les protéines uPA (urokinase plasminogen activator) et tPA (tissue plasminogene activator) sont modulées sous tendant un caractère plus invasif en l'absence de p53. De plus, uPA et tPA sont réprimées au niveau transcriptionnel par p53, il est donc tout à fait normal que leur niveau dans le secretome soit influencé. La protéine tenascin formant un complexe avec notamment tPA varie de la même façon.

Aucune MMP n'est détectée dans le secretome. Néanmoins, on ne peut pas conclure qu'elles sont absentes, l'approche utilisée étant d'une sensibilité de détection limitée et donc non exhaustive. Plusieurs indices corroborent avec une présence et une modulation des MMPs. Un des rôles du système uPA est d'activer les MMPs. De plus, le TGF-β capable lui aussi d'activer les MMPs est identifié. Son expression dans le secretome est diminuée lorsque p53 est exprimé, corrélant bien avec l'action directe de p53 sur l'expression de TGF-β. Différents effecteurs de l'activation des MMPs sont donc présents et régulés. On observe également la présence de leur inhibiteur tissulaire spécifique, les TIMPs 1 et 2. Celles-ci sont surexprimées en absence de p53, sans doute pour moduler l'augmentation d'activité des MMPs. Enfin, le rôle des MMPs est de dégrader la MEC et un certain nombre de leurs cibles est retrouvé dans le secretome : la fibronectine, la laminine, la tenascine. La libération de ces protéines (ou de fragments) dans le milieu

extracellulaire est augmentée en l'absence de p53 (grâce à l'action des MMPs ??). Pour s'assurer de la présence et de la régulation des MMPs dans notre modèle, des Western-blots seront prochainement réalisés (notamment sur MMP-9 et MMP-2).

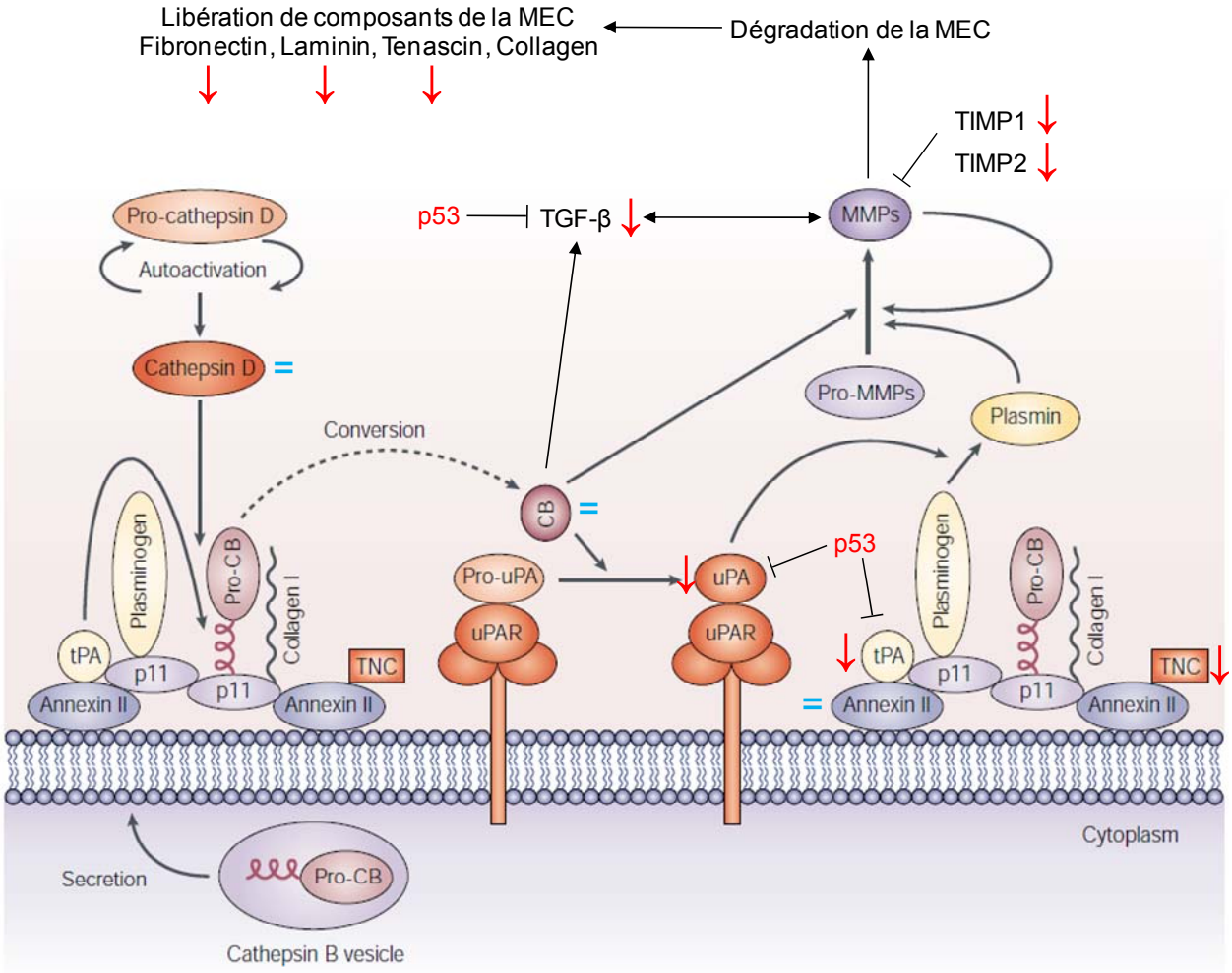


Figure 72 : Influence de p53 sur les cascades protéolytiques dégradant la MEC chez la lignée H358. Les protéines indiquées par une flèche rouge descendante ont leur expression diminuée dans le secretome lorsque p53 est exprimé. Les protéines avec un « = » bleu sont détectées mais non modulées. uPA : urokinase plasminogen activator, tPA : tissue plasminogene activator, CB : cathepsine B, TNC : tenascin.

2.1.2. L'ostéopontine et la tumorogenèse

L'augmentation d'ostéopontine est systématiquement corrélée avec le caractère métastatique des tumeurs [217]. Cette protéine se lie notamment au CD44 et aux intégrines, deux types de molécules impliquées dans le processus métastatique (Figure 73). L'ostéopontine a été identifiée dans notre étude et est surexprimée dans le secretome en absence de p53. Cela est logique puisque p53 est capable de la réprimer directement au niveau transcriptionnel. Cette augmentation en absence de p53 contribue au caractère tumorogénique observé en augmentant la

survie cellulaire par fixation sur CD44 et en promouvant la mobilité et la progression tumorale par fixation sur les intégrines αVβ3.

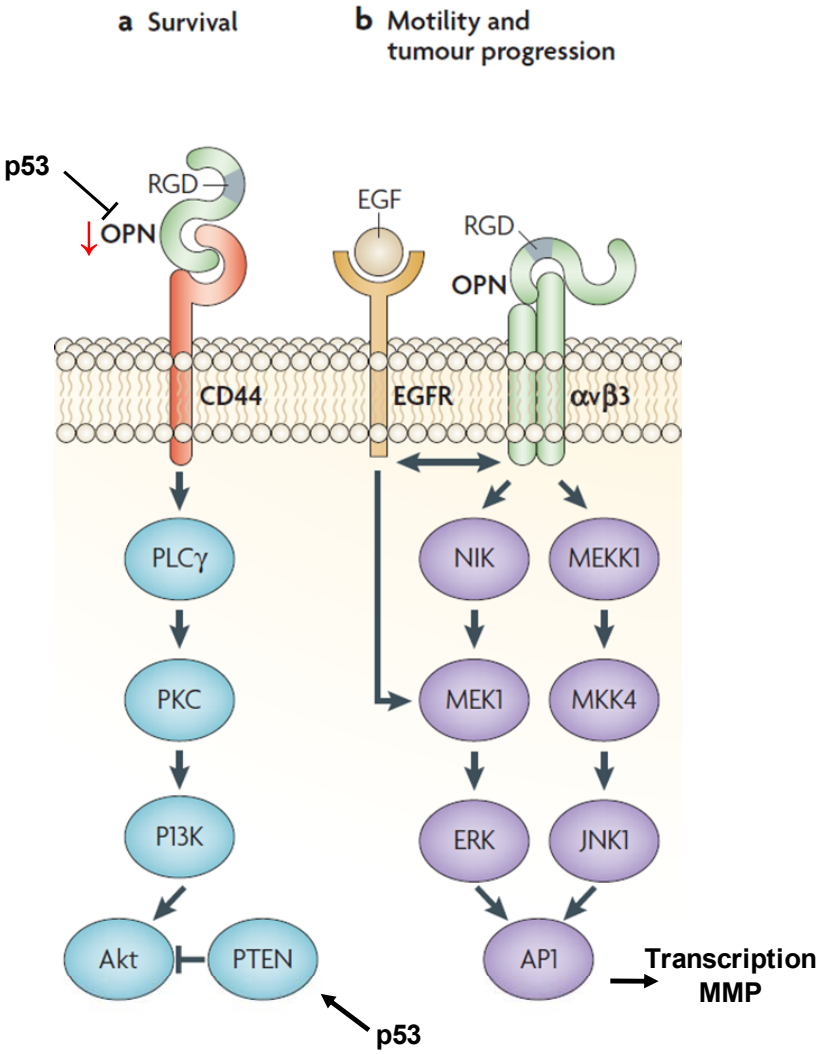


Figure 73 : Influence de la voie médiée par l'ostéopontine sur la tumorogenèse (d'après Bellahcène et al. 2008 [269]).
L'ostéopontine (OPN) est indiquée par une flèche rouge descendante car son expression est diminuée dans le secretome en présence de p53.

2.1.3. Influence sur la transition épithélio-mésenchymateuse (TEM)

L'acquisition des propriétés invasives et migratoires par les cellules tumorales s'accompagne de changements morphologiques caractéristiques appelés transition épithélio-mésenchymateuse (TEM).

Plusieurs facteurs propres au phénotype « TEM » sont observés dans le secretome de la lignée H358 en l'absence de p53 (Figure 74). La TEM se manifeste par un changement d'expression des molécules d'adhérence, incluant la perte de l'adhérence cellule-cellule due à l'E-cadhérine et l'augmentation d'expression de la N-cadhérine. Nos résultats montrent effectivement une expression opposée de ces deux protéines. Les variations d'expression de TGF-β et E-cadhérine

coïncident avec le modèle de la Figure 74 montrant un rôle négatif de TGF-β sur l'E-cadhérine pour favoriser la TEM. Deux protéines impliquées dans la TEM, la vimentine et la tenascine sont augmentées en absence de p53. Celles-ci pourraient être induites par une voie passant par le NF-κB, cible réprimée directement par p53. L'observation des niveaux cellulaires de NF-κB donnera des informations supplémentaires quant à ce phénomène.

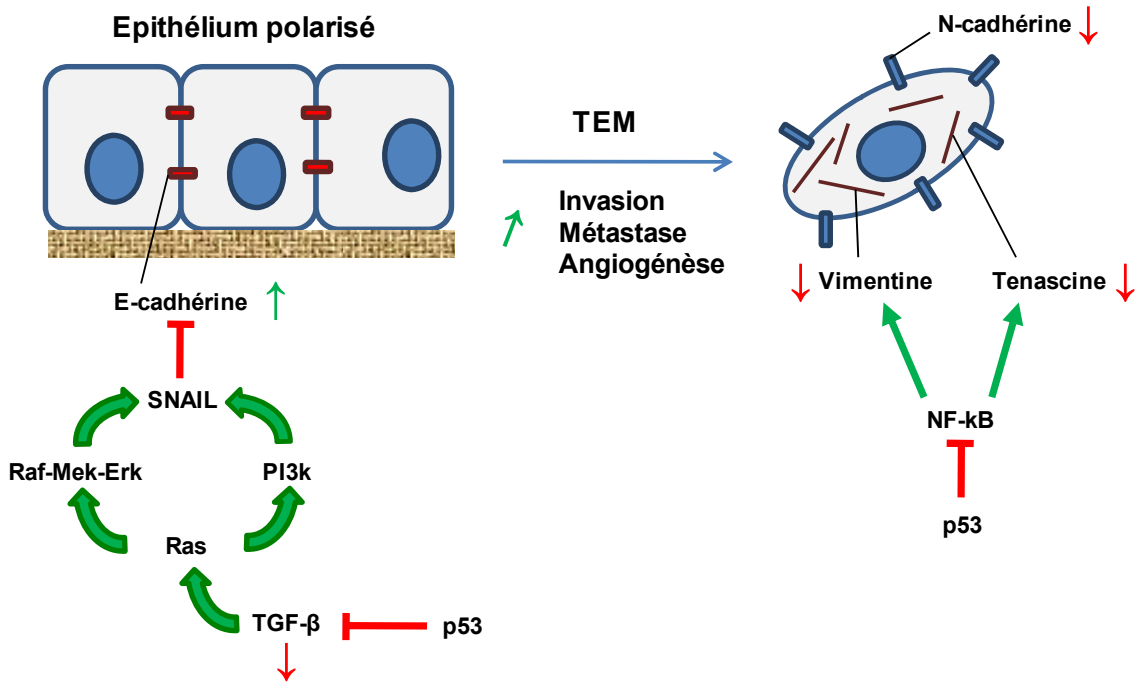


Figure 74 : Influence sur le phénotype TEM des cellules.
Les protéines indiquées par une flèche rouge descendante ont leur expression diminuée dans le secretome lorsque p53 est exprimé et celles avec une flèche verte montante augmentée.

Plusieurs protéines membranaires impliquées dans l'adhérence cellulaire sont également perturbées par l'expression de p53 : la cadhérine-6 (K-cadhérine), E-cadhérine, la cadherin-related tumor suppressor, l'ephrin type-A receptor 2 et l'ephrin A1 sont par exemple surexprimées en présence de p53. A l'inverse, la cadhérine-2 (N-cadhérine) est sous-exprimée.

En ce qui concerne la modulation de la « voie E-cadhérine », une hypothèse peut être formulée en rapport avec le modèle « β-caténine » de la Figure 75. La β-caténine joue deux rôles importants dans la cellule. Elle est tout d'abord impliquée dans l'adhérence inter-cellulaire par son action sur l'E-cadhérine. L'E-cadhérine forme un complexe multiprotéique avec notamment les protéines α-caténine, β-caténine et p120 (δ-caténine). Ces trois protéines sont, comme l'E-cadhérine, surexprimées en présence de p53 reflétant un phénotype moins invasif d'adhérence cellulaire. Lorsque p53 n'est pas exprimé la β-caténine se détache vraisemblablement de l'E-cadhérine pour jouer, dans la voie de signalisation Wnt, un rôle sur la transcription de nombreux gènes participant au caractère oncogénique de la cellule. Parmi les nombreux gènes cibles activés par la voie Wnt/β-caténine (voir la liste établie sur www.stanford.edu/~rnusse/pathways/targets.html),

on trouve uPAR [270], les MMP-2, 7, 9 [271, 272], les FGF-18, 20 [273, 274], le VEGF [275] et l'interleukin-8 [276]. A l'inverse, l'E-cadhérine est réprimée [277]. En l'absence de p53, nous trouvons une surexpression de VEGF et IL-8. Outre une régulation directe par p53, la β-caténine peut aussi participer à la modulation d'expression de ces protéines. On peut émettre l'hypothèse que la diminution de β-caténine observée dans le secretome en absence de p53 est due à son internalisation dans le noyau. L'étude plus approfondie de la voie Wnt/ β-caténine dans notre modèle sera intéressante.

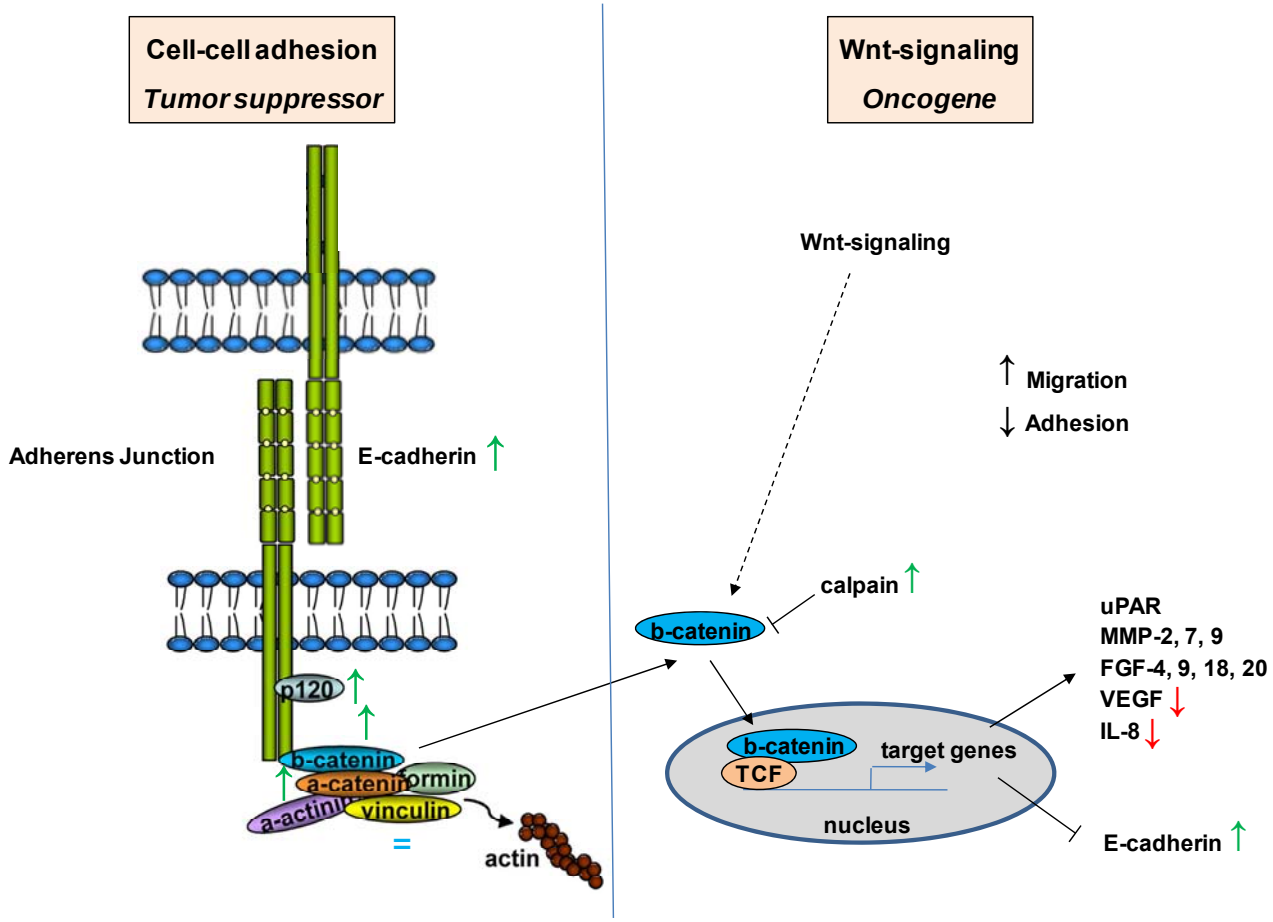


Figure 75 : Le double rôle de la β-caténine
La β-caténine a un rôle de suppresseur de tumeur en participant à l'adhérence cellule-cellule *via* l'E-cadhérine (gauche). Elle peut avoir un rôle oncogénique en agissant sur la transcription de gène par la voie Wnt (droite). Les protéines indiquées par une flèche rouge descendante ont leur expression diminuée dans le secretome lorsque p53 est exprimé et celles avec une flèche verte montante augmentée. Les protéines avec un « = » bleu sont détectées mais non modulées.

2.1.4. Influence sur les voies de néo-angiogenèse

La néo-angiogenèse observée dans les tumeurs nécessite notamment l'action de facteurs de croissance qui agissent sur la prolifération des cellules endothéliales (VEGF, bFGF, TGF-β, PDGF) (Figure 76). En absence de p53, VEGF, TGF-β et PDGF-B sont surexprimées dans le secretome de la lignée H358 (Rq : le PDGF-B a un ratio de 0.63. Il n'a pas été initialement renseigné dans le tableau car il n'est identifié qu'à partir d'un seul peptide). C'est aussi le cas

pour le facteur pro-angiogénique IL-8 [278], et l'IGF-II qui est capable de stimuler l'expression de VEGF [279]. Inversement, la serpin B5 (maspin), un facteur anti-angiogénique, est surexprimée en présence de p53. L'ensemble de ces protéines sont des cibles connues de p53. VEGF, TGF-β, IL-8, IGF-II sont réprimés et la serpin B5 induite par p53. Un cas contradictoire est observé avec la thrombospondine. C'est un facteur anti-angiogénique induit par p53. Pourtant celui-ci est surexprimé dans le secretome en absence de p53. D'autres voies de régulation sont sûrement en jeu dans ce cas précis.

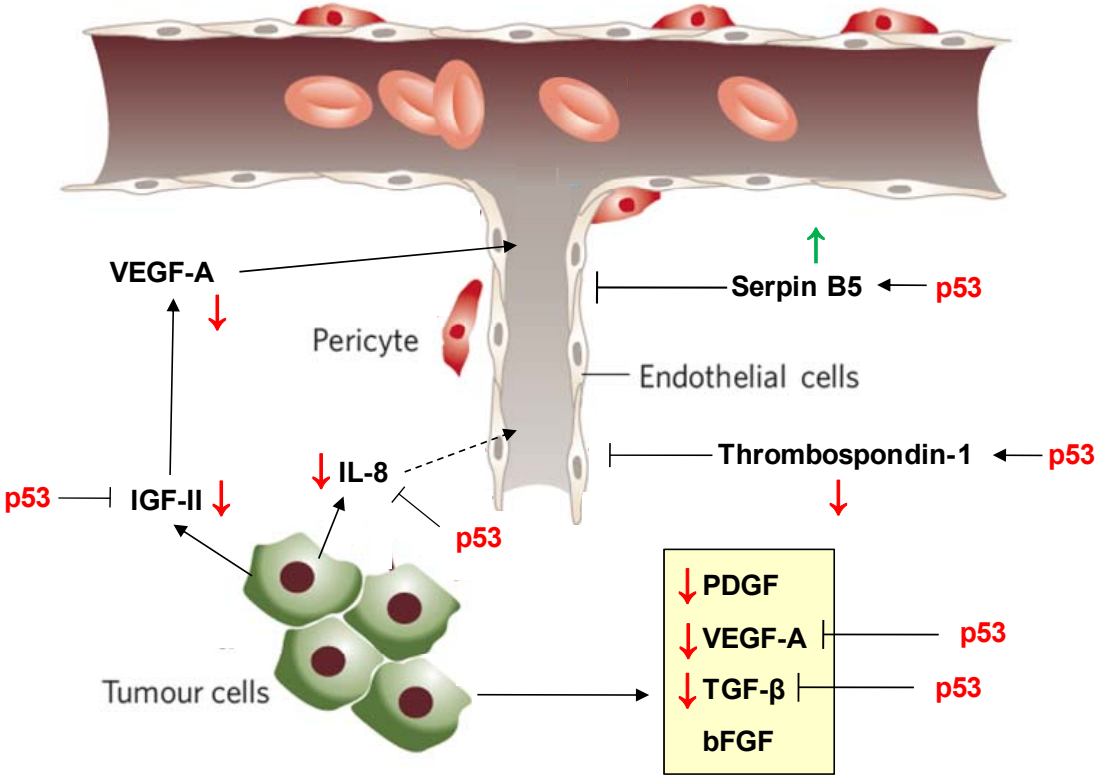


Figure 76 : Influence sur l'angiogenèse (d'après Ferrara et al. [221]).
Les protéines indiquées par une flèche rouge descendante ont leur expression diminuée dans le secretome lorsque p53 est exprimé et celles avec une flèche verte montante augmentée.

2.1.5. Influence sur la sénescence

La protéine p53 est capable, notamment par l'intermédiaire de l'expression de p21, d'induire la sénescence cellulaire. L'induction de l'expression de p21 en réponse à l'expression de p53 ainsi que l'arrêt de la prolifération des cellules ont été confirmées. D'autres voies de signalisation peuvent conduire à un état sénescent (Figure 77).

La voie IGF/IGFBP joue un rôle dans l'induction de la sénescence. Les IGFBPs sont capables de séquestrer les IGFs et de supprimer ainsi la fixation à leur récepteur et leur signalisation. Quatre membres de ce complexe sont régulés dans le secretome en fonction de p53 : IGF-II, IGFBP-2, 3, 4. IGFBP3 est une cible transcriptionnelle de p53 associée à la sénescence [234].

Les cellules sénescentes ont aussi une surexpression du régulateur PAI1 [235]. PAI1 est modulé par p53 et peut inhiber les deux activateurs du plasminogène tPA et uPA. Il empêche ainsi la dégradation de la MEC et la libération de facteurs de croissance nécessaires à la prolifération. En présence de p53, tPA et uPA sont tous les deux diminués dans le secretome. Il sera intéressant de voir si cette diminution passe bien par une induction de PAI1.

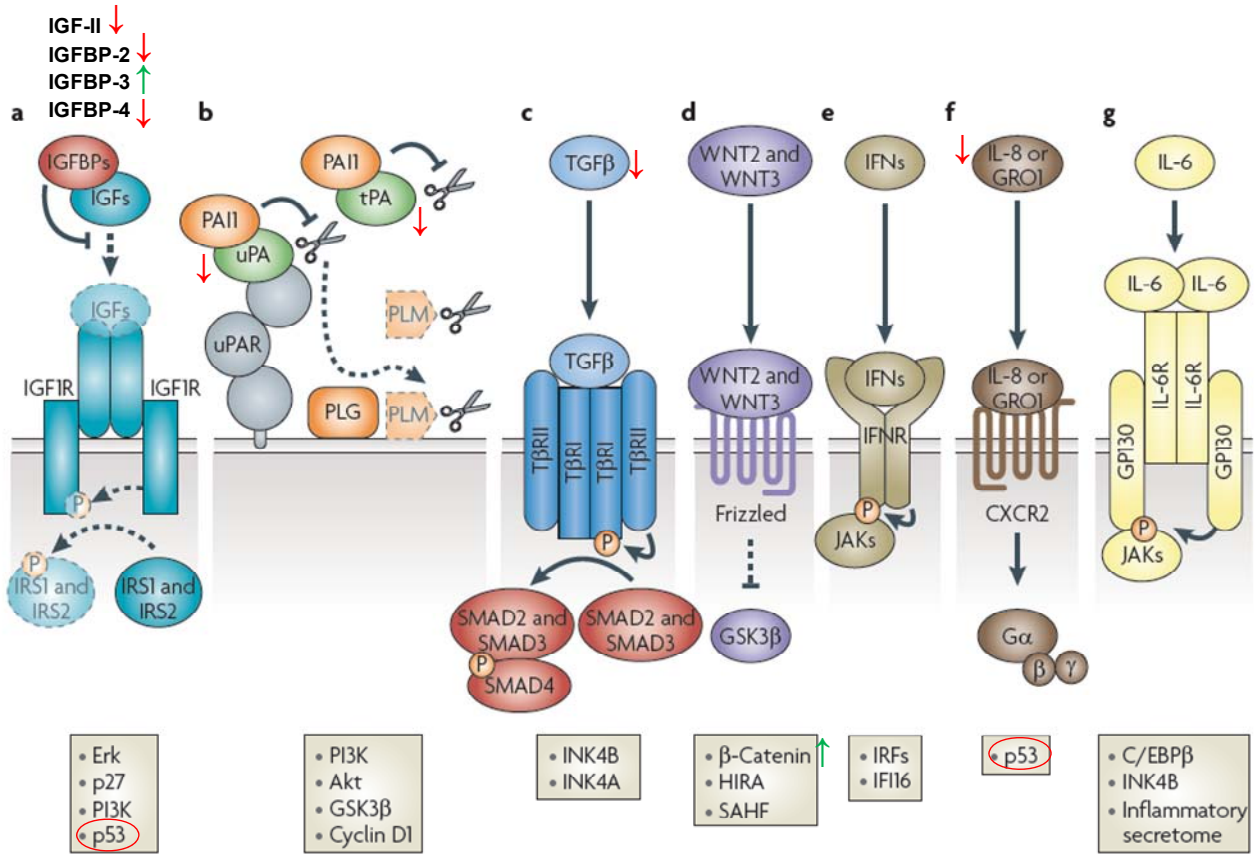


Figure 77 : Les différentes voies de transduction du signal de senescence (d'après Kuilman et Peeper [236]). Les protéines indiquées par une flèche rouge descendante ont leur expression diminuée dans le secretome lorsque p53 est exprimé et celles avec une flèche verte montante augmentée.

Un certain nombre des membres des voies « classiques » des procédés de métastase, invasion, angiogenèse, sénescence sont régulés dans notre étude. Cette régulation coïncide avec un caractère plus tumorogénique en absence de p53.

Certaines des protéines régulées du secretome n'entrent pas dans ces voies « classiques » mais peuvent néanmoins jouer un rôle important dans la tumorogenèse. Quelques unes sont discutées ci-dessous.

2.1.6. La kallikrein-6 et la modulation du micro-environnement tumoral

En parallèle du dogme central incluant uPA, les MMPs et les cathepsines, d'autres acteurs modulant le micro-environnement tumoral ont été mis en valeur récemment, notamment les kallikreins [280].

Les kallikrein-related peptidases appartiennent à la famille des sérine protéases et sont composées de 15 membres qui ont divers rôles physiologiques (voir les revues [281, 282]). Elles ont été mises en évidence comme bio-marqueurs potentiels pour le criblage, le diagnostic, le pronostic et le suivi de nombreux cancers (prostate, ovaires, sein, testicules et poumon). Une revue très complète sur la relation entre cette famille de protéines et les cancers a été publiée par Diamandis et al. [283]. La PSA (prostate-specific antigen), nommée aussi kallikrein-3, est le membre le plus connu de la famille et sert de bio-marqueur du cancer de la prostate.

La kallikrein-6 (KLK6) a été caractérisée comme marqueur de certains types de cancers ovariens et de l'utérus [284, 285]. Les patients atteints d'un carcinome ovarien ont une concentration de KLK6 dans leur sérum environ 2 fois plus importante que les patients normaux ou atteints de tumeurs bénignes [284]. Les cancers gastriques et du colon ont un niveau d'expression d'ARNm *KLK6* nettement plus important que les muqueuses normales [286, 287]. Dans chacune de ces études, la surexpression de *KLK6* est corrélée avec un pauvre pronostic. Cela est dû au rôle joué par *KLK6* dans la progression du cancer. Il est impliqué dans les procédés d'angiogenèse, migration et invasion par des mécanismes impliquant la dégradation de la matrice extra-cellulaire [280]. Les substrats connus de la KLK6 sont des protéines de la MEC telles que le collagène, le fibrinogène, la fibronectine et les laminines [280, 288].

Prezas et al. ont démontré que des cellules ovariennes transfectées de façon stable pour exprimer les KLK4, 5, 6 et 7sont plus invasives *in vitro* et forment de plus grosses tumeurs chez la souris [289]. Une autre étude démontre qu'un inhibiteur synthétique des kallikreins atténue l'invasion tumorale [290]. Plusieurs kallikreins jouent notamment ce rôle en activant les membres de la cascade protéolytique : uPA-uPAR-MMP [280].

Les kallikreins sont aussi capables d'interférer avec la voie IGF/IGFBP. Ainsi les KLK1, 2 et 3 sont des protéases anti-IGFBP capables de dégrader les IGFBP2, 3, 4 et 5 supprimant leur effet suppresseur de tumeur et diminuant ainsi leur affinité pour les IGFs. L'IGF peut ainsi moduler la croissance de cellules du cancer de la prostate *in vitro* [291-293]. Il peut ainsi y avoir une relation entre la surexpression de l'IGF-II, la diminution de IGFBP3 et la surexpression de KLK6 observés.

La kallikrein-6 n'a jamais été caractérisée dans des tissus tumoraux du poumon. Par contre, au moins une kallikrein (kallikrein-11) surexprimée dans les carcinomes du poumon a été identifiée

lors d'une étude microarray (particulièrement dans les tumeurs neuroendocrines au devenir peu favorable) [294].

En nous basant sur le modèle de Henkhaus et al. [295], nous pouvons proposer le suivant (Figure 78). Henkhaus a montré que la protéine kinase AKT permettait l'expression du gène *KLK6*. Or quand p53 est exprimé, celle-ci induit l'expression du gène *PTEN* dont la protéine est capable d'inhiber AKT et donc potentiellement de diminuer l'expression de *KLK6*. Pour confirmer cette hypothèse, il sera intéressant d'analyser les niveaux de AKT et PTEN dans la cellule.

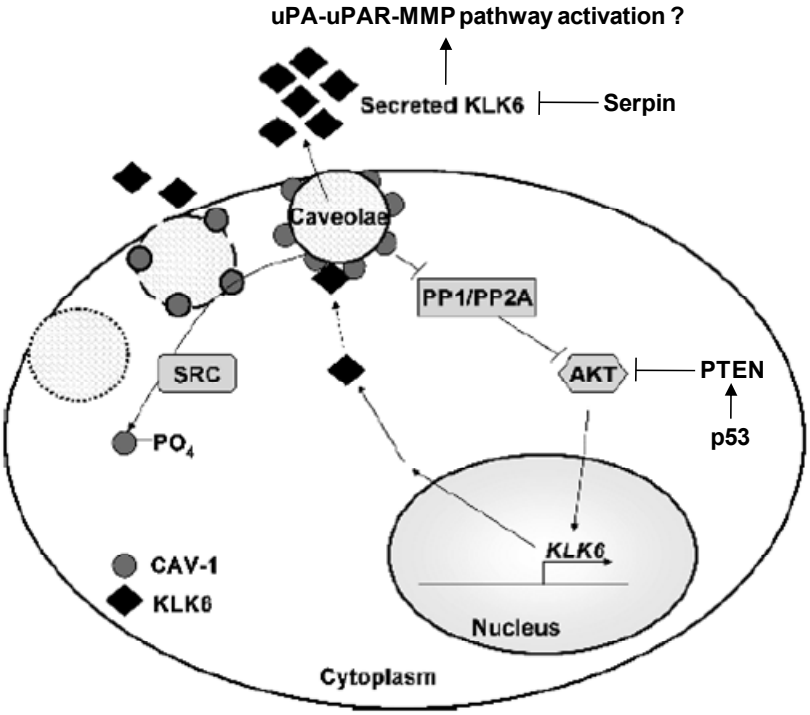


Figure 78 : Modèle proposé pour expliquer l'expression différentielle de KLK6 dans le secretome de H358 en fonction de p53 (D'après Henkhaus et al. [295])

L'identification de KLK6 dans notre étude, ainsi que sa surexpression dans le secretome en l'absence de p53 corrèle avec ces données.

Les kallikreins ont leur activité bloquée par un groupe d'inhibiteurs de sérine protéase, les serpins [283]. De façon intéressante, notre étude identifie deux membres de cette famille, la serpin B5 (maspin) et la serpin B1 (leucocyte elastase inhibitor) dont la surexpression en présence de p53 pourrait diminuer la quantité de KLK6.

S'il parait clair que KLK6 joue un rôle dans la tumorogenèse, le nombre de cancers dans lequel il a déjà été caractérisé exclu vraisemblablement une spécificité de ce marqueur pour les cancers du poumon. Enfin, aucune variation d'expression significative n'a été mise en évidence lors de notre étude préliminaire faite sur une faible cohorte de patients.

2.1.7. FGF-19

FGF-19 appartient à la famille des fibroblast growth factor (FGF) composée de 22 membres se fixant sur 4 récepteurs tyrosine kinases et (FGFR1-4) [296]. L'interaction de FGF avec son récepteur conduit à une dimérisation et une autophosphorylation du récepteur, au recrutement d'adaptateurs cytosoliques tels que FRS2 (fibroblast growth factor receptor substrate 2) et à l'initiation de multiples voies de signalisation [297]. FGF-19 se fixe de façon spécifique sur le FGFR4 [298]. L'expression systémique de FGF-19 chez des souris transgéniques provoque une carcinogénèse hépatique [299]. Desnoyers et al. ont démontré récemment que FGF-19 est exprimé dans plusieurs tumeurs humaines dont les carcinomes à cellules squameuses du poumon et que ce facteur est impliqué dans la progression tumorale [300]. Cette action passe par FGFR4, FRS2, ERK et l'activation de la β-caténine (Figure 79). Pai et al. ont montré que FGF-19 augmente la phosphorylation sur tyrosine de la β-caténine et cause une perte de l'interaction E-cadhérine/β-caténine [301]. Cela active la voie Wnt avec une translocation nucléaire de la β-caténine et l'expression de gènes favorisant le développement tumoral (uPAR,…). La protéine FGF-19 est surexprimée dans le secretome de H358 en absence de p53. Cette variation est retrouvée *in vivo* dans notre modèle de souris. Cette protéine peut être une cible intéressante pour des investigations futures.

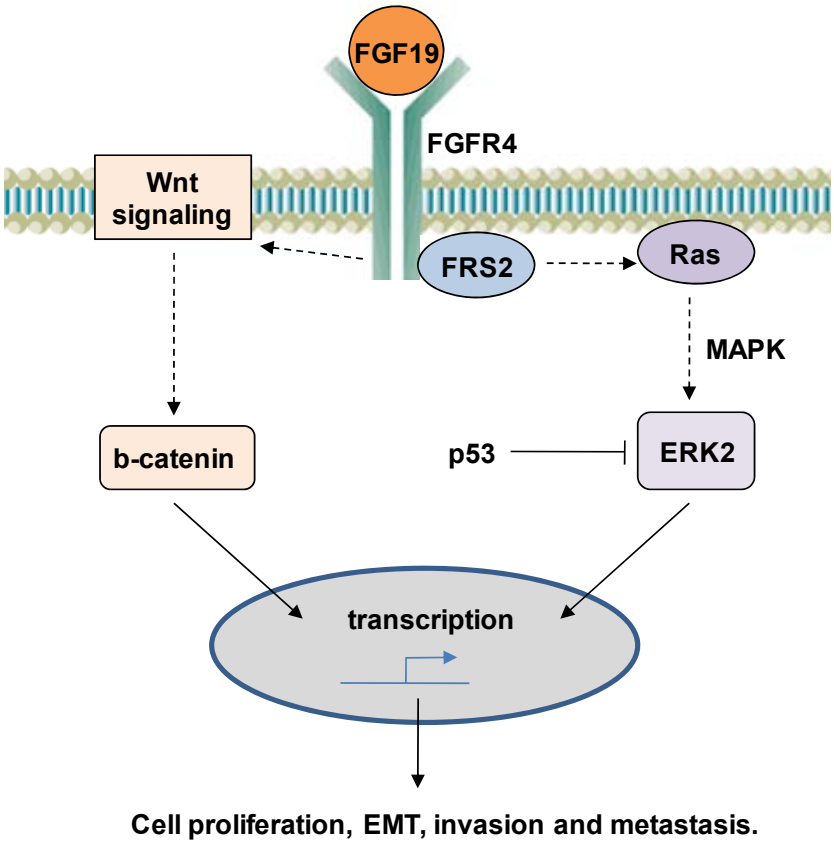


Figure 79 : Modèle de l'action de FGF-19 sur le développement tumoral.

2.1.8. FGF-BP1

Les FGFs sont impliquées dans des conditions pathologiques telles que la croissance et l'angiogenèse tumorale. Les FGFs sont étroitement liés aux protéoglycans et donc enchâssés dans la MEC. Leur libération dans le milieu est une étape critique de leur activation qui, outre l'action de protéases sur la dégradation de la MEC, nécessite l'action d'une protéine chaperone FGF-binding protein (FGF-BP) [302]. Ces FGF-BP sont surexprimés dans de nombreuses tumeurs et sont particulièrement associés au stade précoce de la formation tumorale, quand la néo-angiogenèse joue un rôle critique [303, 304]. De façon intéressante, Sherif et al. ont montré une relation entre FGF-BP et p53 [305]. Lorsque p53 est exprimé dans des cellules de carcinome de la tête et du cou, ils observent une diminution de FGF-BP associée avec une inhibition de la croissance cellulaire *in vitro* et *in vivo*. De plus, l'action nucléaire de la β-caténine lors de l'activation de la voie Wnt (Figure 75) induit une augmentation de l'expression de FGF-BP durant la carcinogénèse du colon [306]

FGF-BP1 est un des membres de la famille FGF-BP peu documenté dans la littérature. Tassi et al. ont suggéré que FGF-BP1 représente un régulateur important qui module de façon positive les réponses cellulaires médiées par les FGFs telles que la signalisation, la prolifération et l'angiogenèse [307]. De plus, les différentes voies décrites pour FGF-BP pourraient correspondre aux variations observées dans notre étude du secretome. FGF-BP1 est sous-exprimée en présence de p53. Il pourrait avoir une relation avec la β-caténine. Enfin, son action couplée à celle des protéases (uPA, tPA, KLK6) pourrait participer à la libération de FGF-19 de la MEC et potentialiser son action passant par ERK2 et la β-caténine. En tenant compte de ces paramètres, on peut proposer le modèle suivant :

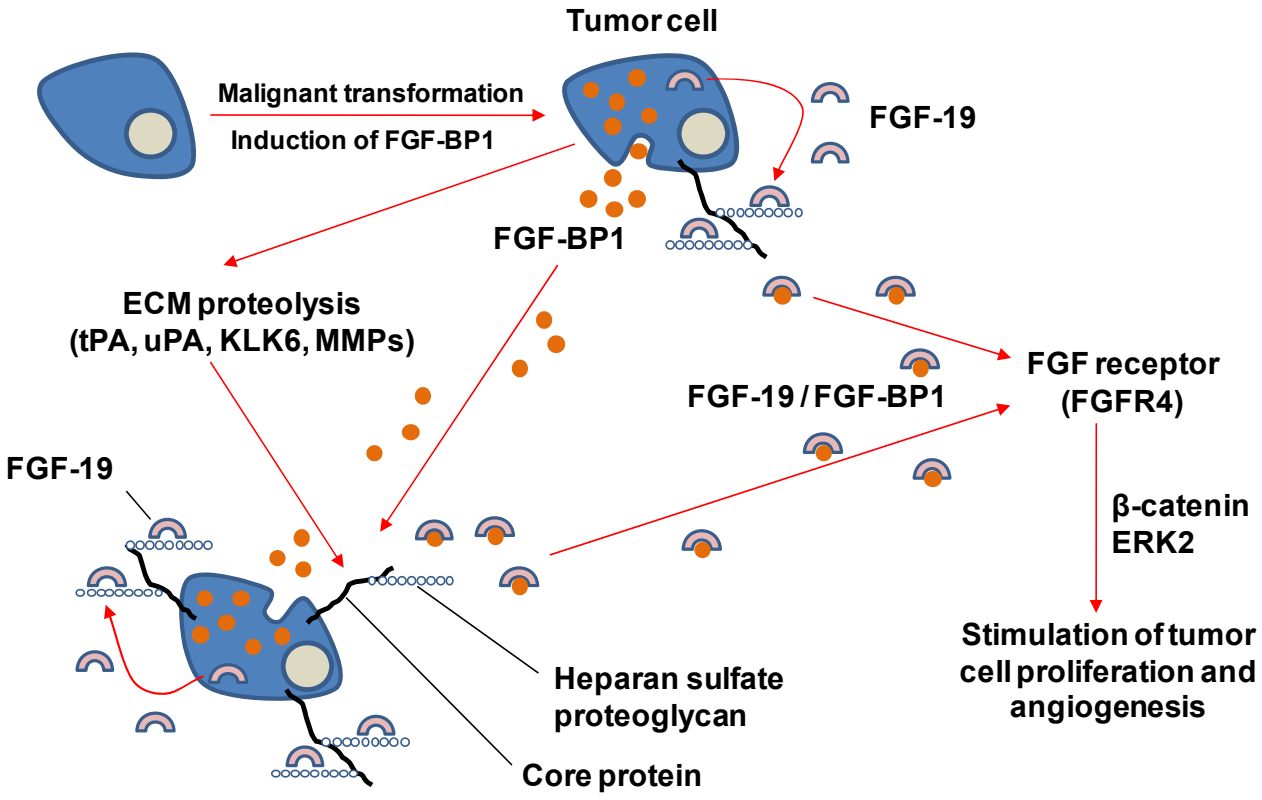


Figure 80 : Modèle du mécanisme d'action de la protéine FGF-BP1 (d'après Abuharbeid et al. [302]).

2.1.9. Le transporteur de zinc ZIP10

ZIP10 appartient à la famille des transporteurs de zinc ZIP (ZRT IRT-like protein), protéines impliquées dans l'influx de zinc dans le cytosol [308]. Des études ont démontré que le zinc est impliqué dans la progression des cancers [309], comme par exemple dans les carcinomes de la prostate [310] et du foie [311]. Certains ZIP sont impliqués dans la progression des cancers, comme ZIP6/LIV1 qui joue un rôle dans la métastase du cancer du sein vers les ganglions lymphatiques [312].

Très peu d'études sur ZIP10 sont référencées dans la littérature, néanmoins Kagara et al. ont montré son rôle dans la progression du cancer du sein [313]. Cette activité de migration est inhibée par le knockdown de ZIP10 ou la chélation de zinc. L'analyse d'échantillons cliniques montre que les cancers du sein métastatiques ont une expression plus importante de ZIP10 que ceux sans métastase. Cet effet pourrait résulter d'une action de ZIP10 dans la TEM observée lors de la progression des cancers. En effet, il pourrait augmenter l'invasion en ciblant la voie ERK1/2-Snail/Slug [314]. Nous trouvons dans notre étude que ZIP10 est grandement diminuée dans le secretome lors de la réexpression de p53. Cette protéine pourrait donc être une cible intéressante.

2.1.10. GDF-15

La protéine GDF-15 (growth differentiation factor 15) appartient à la superfamille des TGF-β [315]. Sa découverte ayant été faite simultanément par différentes équipes, elle est aussi connue sous d'autres noms tels que NAG-1 (NSAID activated gene), MIC-1 (Macrophage Inhibitory Cytokine-1) ou PTGFB (Placental Transformation Growth Factor-β). Son récepteur et sa voie de signalisation n'ont pas encore été décrits. Les effets de GDF-15 sont contradictoires puisqu'il peut, selon les conditions, aussi bien jouer un rôle tumorogène qu'anti-tumorogène [315]. La surexpression de GDF-15 dans des cellules cancéreuses déclenche un arrêt du cycle cellulaire et/ou une augmentation d'apoptose [316]. *In vivo*, la surexpression de GDF-15 par des cellules de glioblastome ou du cancer du colon inhibe la formation de tumeur chez des souris nude [316]. L'expression de p53 [317] et EGR-1 [318] et l'inhibition de la voie AKT/GSK-3β augmentent l'expression de GDF-15 [319] (Figure 81).

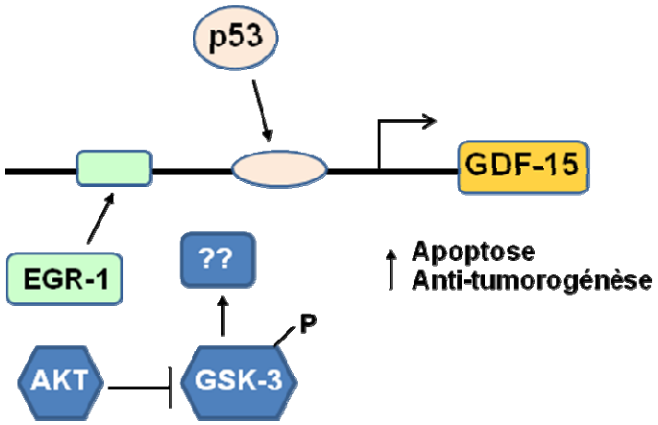


Figure 81 : Régulation de l'expression de GDF-15 (d'après Eling [315]).

Lors de notre étude du secretome, GDF-15 est surexprimé en présence de p53 en accord avec sa régulation transcriptionnelle par ce dernier. L'arrêt du cycle cellulaire observé chez les cellules H358 correspond aussi à la fonction de GDF-15. De façon intéressante, sa variation est retrouvée *in vivo* chez notre modèle de souris et est aussi observée dans leur plasma.

2.2. L'étude des protéomes cytosoliques et des organelles&membranes

L'influence de p53 sur la modulation du protéome cytosolique et du protéome des organelles (réticulum endoplasmique, appareil de Golgi, mitochondries,…) et membranes des cellules H358 a été étudiée. Ces protéines régulées par p53 présentent également un intérêt pour la compréhension des mécansimes cellulaires impliqués.

2.2.1. Le complexe MCM

La réplication de l'ADN est un mécanisme très conservé qui nécessite la présence d'un complexe hétérohexamérique de MCM 2-7 [320]. Les MCMs peuvent constituer des bio-marqueurs de « l'état » du cycle cellulaire, c'est-à-dire refléter une cellule capable de proliférer plutôt qu'une cellule quiescente ou senescente [320]. L'expression des 6 MCMs est observée à toutes les phases du cycle cellulaire et est diminuée lorsque la cellule quitte le cycle pour la quiescence, la différentiation ou la senescence [321]. Les cellules tumorales étant caractérisées par une forte prolifération, les MCMs peuvent représenter des bio-marqueurs de cancers. Le rôle potentiel des MCMs comme marqueur prédictif de survie des patients a été montré pour différents cancers dont les carcinomes du poumon non à petites cellules [322, 323].

Cinq de ces 6 protéines sont sous-exprimées en présence de p53 dans le protéome des organelles & membranes de la lignée H358 (protéines MCM 2, 3, 4, 6, 7). Ce résultat coïncide bien avec une augmentation de la prolifération cellulaire en absence de p53. Il a d'ailleurs été montré que p53 est capable de diminuer la transcription de MCM7 [324]. Par contre, ces protéines sont normalement exclusivement nucléaires et on peut donc être surpris de les retrouver dans le protéome des organelles & membranes. L'hypothèse la plus probable réside en l'éclatement d'un certain nombre de noyaux durant la préparation des échantillons. La même variation est observée pour chacun des réplicas effectués. L'éclatement des noyaux est donc « homogène » entre les différentes préparations et cette différence d'expression donc fiable.

2.2.2. Carboxylestérase 2

La carboxylestérase 2 est impliquée dans la détoxification de la cellule par l'hydrolyse des liaisons esters de certaines drogues (cocaine, heroine,…). Son expression est régulée au niveau transcriptionnel par la protéine p53 [325]. Très peu d'information sur cette protéine est disponible dans la littérature. Néanmoins, une récente étude a montré que son expression était diminuée dans les tissus de cancer colorectal [326]. Notre étude met en valeur une forte augmentation (>4 fois) de cette protéine dans le protéome organelles lorsque p53 est exprimé. Une étude plus approfondie de ce phénomène pourrait donc être intéressante.

2.2.3. Ferredoxin reductase (NADPH:adrenodoxin oxidoreductase)

La ferredoxin reductase est une protéine mitochondriale responsable du transfert d'électrons du NADPH au cytochrome p450 via la ferredoxine [327]. C'est une cible transcriptionnelle de p53 [327]. Sa surexpression par la cellule inhibe la croissance de cellules du cancer du colon [328]. Une étude clinique a montré que l'expression de la ferredoxin reductase n'affectait pas le devenir de patients atteints d'un cancer colorectal sans mutation de p53. Par contre, parmi les patients

porteurs d'une mutation de p53, ceux possédant une faible expression de la ferredoxin reductase ont un plus mauvais devenir que ceux qui en ont une forte expression [329]. Notre étude met en valeur une forte augmentation de cette protéine dans le protéome organelles lorsque p53 est exprimé en accord avec les résultats précédents.

2.2.4. Ribonucleotide reductase M2B

La ribonucleotide reductase M2B (p53R2) est une protéine induite par p53 qui joue un rôle crucial dans la réparation des dommages de l'ADN [330]. La corrélation entre l'expression de p53R2 et l'invasion / métastase des cancers reste controversée. La protéine p53R2 a été rapportée comme associée avec la résistance aux agents anti-cancéreux pour les cancers de la cavité orale et de l'œsophage [331, 332]. Liu et al. ont au contraire montré que l'augmentation de l'expression de p53R2 entraine une réduction de l'invasion des cellules cancéreuses et qu'à l'inverse sa diminution augmente significativement l'invasion [333].

Une autre étude a établi que la sous-expression de p53R2 augmente le potentiel d'invasion dans des cellules tumorales mutées en p53. Cela entraine une augmentation de la translocation nucléaire de la β-caténine et une perte de l'expression cellulaire totale d'E-cadhérine [334]. Cette observation est à mettre en parallèle avec la voie présentée Figure 75 et donc une activation de la voie Wnt associée. Lorsque p53 est exprimé par H358, nous observons une augmentation du p53R2 cytosolique.

3. Les cibles transcriptionnelles de p53 régulées dans la lignée H358

La Figure 82 résume l'ensemble des cibles transcriptionnelles de p53 identifiées dans les différents protéomes de la lignée H358/TetOn/p53. Chacune des protéines est modulée par p53 dans le sens attendu. Seule la thrombospondin-1 (tsp1), normalement induite par p53, est surexprimée dans notre étude en absence de p53. Ceci s'explique sans doute par l'intervention d'autres voies de régulation. Le facteur de transcription aryl hydrocarbon receptor (AhR) est, par exemple, capable d'induire l'expression de la thrombospondin-1. De plus, une régulation autre que transcriptionnelle peut aussi intervenir.

Ces résultats valident notre modèle et notamment l'expression d'une protéine p53 sauvage et fonctionnelle.

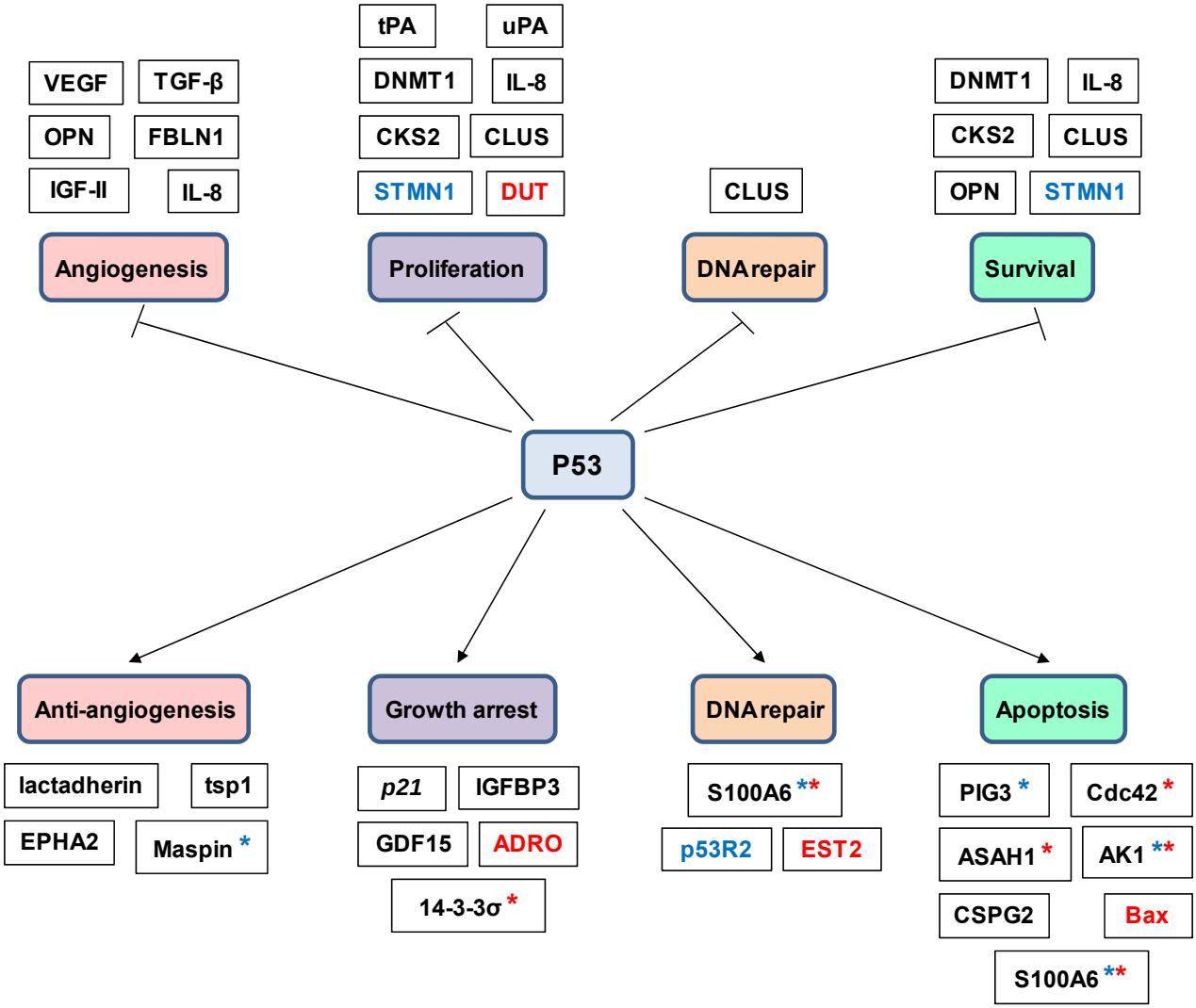


Figure 82 : Les cibles transcriptionnelles de p53 régulées dans la lignée H358.

En noir : les protéines régulées dans le secretome, en bleu celles du cytosol et en rouge celles des organites&membranes. Les astérisques (*) représentent les protéines identifiées dans différents protéomes avec le même code couleur.

VEGF: Vascular Endothelial Growth Factor A, TGF-β: Transforming Growth factor β2, OPN:Osteopontin, FBLN1:Fibulin-1, IGF-II: Insulin Growth Factor II, IL-8: Interleukin-8, tPA: Tissue Plasminogen activator, uPA: Urokinase Plasminogen Activator, DNMT1: DNA (cytosine-5)-methyltransferase 1, CKS2: Cyclin-dependant kinase regulatory subunit 2, CLUS: Clusterin, STMN1: Stathmin, DUT: Deoxyuridine 5'-triphosphate nucleotidohydrolase, tsp1: Thrombospondin-1, EPHA2: Ephrin Type-A receptor 2, IGFBP3: IGF binding protein 3, GDF15: Growth Differentiation Factor 15, ADRO: NADPH:adrenodoxin oxidoreductase, p53R2: Ribonucleoside-diphosphate reductase subunit M2 B, EST2: Carboxylesterase 2, PIG3: Putative Quinone Oxidoreductase, ASAH1: Acid ceramidase, AK1: Adenylate kinase 1, CSP2: Versican.

4. Protéomique et oncologie

La forte mortalité associée aux cancers est due en grande partie à l'incapacité actuelle d'établir un diagnostic suffisament précoce. En effet, les outils utilisés en clinique (imagerie,…) ne détectent la tumeur qu'à un stade déjà avancé. La découverte de bio-marqueurs spécifiques est donc primordiale. La protéomique est un outil de choix pour la recherche de cibles protéiques d'intérêt clinique. Néanmoins, elle se heurte à de nombreuses difficultés freinant actuellement son emploi pour une telle utilisation.

Aujourd'hui il est impossible de penser protéomique clinique sans penser standardisation. En effet, il est très difficile de valider la découverte de bio-marqueurs, par exemple, au sein de laboratoires indépendants si ceux-ci n'ont pas fixé préalablement les conditions de prélèvement ou de stockage de l'échantillon ainsi que les différents facteurs d'expérimentation permettant de faire reconnaître et donc de valider leurs découvertes. L'élaboration de standards et la standardisation posent un vrai problème de crédibilité à long terme pour l'analyse protéomique, entraînant des problèmes de confiance envers les résultats obtenus. En effet, il existe un très large panel de techniques aussi bien pour la séparation des analytes que pour leur identification, caractérisation ou quantification en spectrométrie de masse (voir Chapitres 1 et 2 de la partie Bibliographie), aboutissant ainsi à un volume de données énorme où la nature et le format des données sont variables. C'est pourquoi l'organisation HUPO (HUman Proteome Organisation) a mis en place PSI (Proteomics Standards Initiative www.psidev.info) dont le but est de définir des standards afin d'aboutir à une meilleure analyse des données protéomiques au niveau de leur comparaison, de leur échange et de leur vérification.

Les échantillons utilisés en protéomique clinique peuvent être de trois sources : le sang, les autres fluides biologiques accessibles (salive, sécrétions…) et les tissus.
Les échantillons tissulaires constituent la source originelle des bio-marqueurs tumoraux et donnent accès à de nombreuses protéines non observables dans les fluides biologiques. Par contre, leur collecte représente un acte invasif pour le patient.
Le sang est une source de protéines très avantageuse pour le criblage et la détection précoce des cancers car son prélèvement est un acte très peu invasif. Comme suggéré par Liotta : « the blood contains a treasure trove of previously unstudied biomarkers that could reflect the ongoing physiologic state of all tissues » [335]. Néanmoins, son analyse nécessite de relever un certain nombre de challenges. Les bio-marqueurs de cancers sont beaucoup plus dilués dans le fluide sanguin que dans leur tissu d'origine. De plus, c'est un fluide très complexe contenant une grande diversité de protéines ou d'isoformes protéiques avec une gamme dynamique de concentration

d'une magnitude 9 à 10 [254]. Les protéines abondantes du sang telles que l'albumine ou les immunoglobulines peuvent masquer les protéines bio-marqueurs potentiels [259]. Plusieurs procédures d'élimination de ces protéines ont été décrites dont les colonnes de déplétion mais celles-ci entrainent une perte de protéines par fixations non spécifiques et donc une diminution de l'efficacité du criblage [267]. Enfin, de nombreuses protéines se trouvent sous forme de fragments peptidiques dans le sang par l'action de protéases [336]. De nombreux développements technologiques sont donc encore nécessaires.

Début 2009, Lin et al. ont listé 57 études protéomiques faites directement sur le plasma ou le sérum parmi lesquelles 54 sont basées sur une approche MS et 3 sur une approche bio-puces à anticorps [337]. Dans la majorité des cas l'approche MS utilisée est de type SELDI-TOF (38 études sur 54).
Le SELDI-TOF (Surface Enhanced Laser Desorption Ionization - Time Of Flight) repose sur un système de criblage sur barrette d'affinité en amont de l'analyse MS. Elle est utilisée pour l'analyse différentielle de type patient / témoin afin de mettre en évidence des profils d'expression spécifiques. C'est une approche assez simple sur le plan technique possédant une capacité d'analyse haut débit facilitant son application clinique. Néanmoins, elle se heurte à de nombreuses limites [337]. La plupart des études mettent en valeur des pics m/z discriminants mais non rattachés à l'identification de la protéine correspondante. De plus, très peu d'études indépendantes caractérisent les mêmes pics discriminants reflétant ainsi les différences techniques dans la collecte des échantillons, le traitement, le type de barrette SELDI... Cette variabilité inter-études limite son transfert vers la clinique. Enfin, les quelques études, identifiant les protéines correspondantes, ont mis en valeur des protéines abondantes ou non spécifiques telles que des protéines du stress, de coagulation ou de l'inflammation.

Il est donc actuellement ardu de travailler directement sur le plasma. En conséquence, l'approche « secretome » représente une classe de protéines potentiellement détectables dans le sang et donc un modèle d'étude indirecte très prometteur. La pertinence clinique des marqueurs caractérisés pourra être appréhendée, dans un second temps, en les dosant dans des prélèvements de patients. Le secretome reflète la sécrétion protéique *in vitro* obtenue pour une lignée cellulaire, et donc des divergences avec ce qui se passe cliniquement sont inévitables.

Notre étude, réalisée sur des souris xéno-greffées par la lignée H358, démontre que les variations ciblées sont observables dans un environnement *in vivo*. La modulation des protéines GDF-15, FGF-19 et VEGF a été confirmée dans la tumeur *in vivo.*

Plusieurs études ont montré précédemment que des protéines humaines libérées par des greffons tumoraux pouvaient être dosées dans le plasma des souris [338]. Une récente étude a également démontré que des protéines libérées par un greffon tumoral pouvaient être observées dans le sang par des approches protéomiques [339]. Nous avons montré que GDF-15 est également modulée par p53 dans le plasma des souris. L'expression différentielle d'une protéine issue d'une étude secretome peut donc être retrouvée et quantifiée dans le plasma de souris xénogreffées. C'est la première fois qu'une démonstration de la pertinence du secretome au niveau plasmatique est établie. Le secretome représente donc un modèle fiable pour la caractérisation de marqueurs tumoraux potentiels.

Le but premier de notre étude était de mieux comprendre le rôle joué par la perte de p53 dans la tumorogenèse et le changement d'interaction des cellules tumorales avec leur environnement. Cette perte de fonction étant un élément très précoce et communément observé dans les cancers du poumon, notre étude pourrait aussi ouvrir des pistes intéressantes pour le diagnostic où le suivi de certains sous-types de cancer. Trois protéines (GDF-15, FGF-19 et Kallikrein-6) ont été dosées dans quelques échantillons de patients mais aucune variation significative n'a été observée. Les dosages effectués sur des plasmas de sujets sains attestent que les tissus normaux libèrent physiologiquement ces protéines dans le sang. La modulation due à la tumeur n'est sans doute pas suffisante pour être détectable. L'analyse des protéines directement au niveau des tissus tumoraux serait intéressante. De plus, des essais sur de plus grandes cohortes, avec des mutants p53 mieux ciblés et plus homogènes seraient pertinents. Enfin, il serait intéressant d'orienter ces essais ultérieurs en tenant compte des critères de classification des cancers du poumon établis par les anatomopathologistes [340].

A ce jour, aucun bio-marqueur potentiel identifié par protéomique n'est utilisé cliniquement et aucune étude de validation indépendante des marqueurs potentiels décrits n'a été publiée [341]. On peut donc logiquement se poser la question suivante : la protéomique est elle adaptée à la découverte de bio-marqueur utilisable en clinique ? Même si aucune étude ne valide ce concept, beaucoup d'espoirs persistent. Il ne faut pas oublier que la protéomique est une science récente. De nombreux développements technologiques et méthodologiques ne cessent de repousser les limites. Les études en cours dans les différents laboratoires apporteront sans doute prochainement des résultats plus concrets.

CONCLUSIONS ET PERSPECTIVES

CONCLUSIONS ET PERSPECTIVES

Ce travail a permis de développer un procédé d'analyse protéomique des secretomes efficace, couplant notamment un marquage iTRAQ et une séparation par OFFGEL. Celui-ci a été appliqué à l'étude du rôle joué par p53 sur la modulation des interactions des cellules tumorales entre elles et avec leur environnement. Nous avons ainsi caractérisé 91 protéines d'intérêt dont le niveau d'expression dans le secretome est modifié en présence de p53. Parmi elles, plusieurs ont une relation intéressante avec la tumorogenèse et constituent donc des cibles d'intérêt potentiel. Nos résultats *in vivo* sur un modèle de souris xénogreffées valident la pertinence du secretome et constituent la première validation *in vivo* du concept « secretome ».

De nombreuses perspectives à ce travail sont envisageables.

Ce travail a mis en valeur plusieurs cibles protéiques d'intérêt modulées par p53. Afin de mieux comprendre leurs mécanismes d'action dans le milieu extracellulaire, plusieurs études sont envisageables. Une partie des protéines est régulée indirectement par p53 et la caractérisation des voies de régulation impliquées serait donc intéressante.
De plus, certaines protéines nécessitent la présence d'un récepteur cellulaire spécifique. C'est le cas, par exemple, des facteurs de croissance FGF-19, VEGF ou TGF-β mais aussi de l'uPA, l'osteopontine, l'amphiréguline ou l'interleukine-8,… Ces récepteurs ne sont pas forcément tous caractérisés. Chacun déclenche des voies de signalisation intracellulaire particulières aboutissant à des effets biologiques ou pathologiques (TEM, modulation du cycle cellulaire, des interactions, métastases, invasions, angiogenèses,…).
La voie uPA-uPAR-MMPs semble clairement activée en absence de p53, participant ainsi à la dégradation de la MEC associée au caractère invasif. PAI, une cible de p53 régulant uPA et tPA, participe à la modulation de cette voie et il serait donc intéressant de caractériser son expression dans notre modèle. De même, de nombreuses MMPs (MMP-9, MMP-3,…) sont impliquées dans la dégradation de la MEC et il serait donc utile d'évaluer leur présence et leur effet.
Enfin, un certain nombre de protéines semble avoir une action passant par les voies Ras/ERK (FGF-19, FGF-BP1, ZIP10) ou AKT (KLK6, GDF-15). Il serait donc pertinent d'étudier ces facteurs et leurs voies, ainsi que leur régulation éventuelle par p53 (PTEN, une cible de p53, a par exemple un effet inhibiteur sur AKT).

La caractérisation des interactions protéiques (interactome) serait également intéressante. Cela permettrait, par exemple, d'investiguer une possible interaction entre FGF-19 et FGF-BP1, d'étudier la relation entre IGF-II et IGF-BP 3, 4 et 5 ou encore celle entre l'E-cadhérine, les

caténines β, α et δ et la voie Wnt. Différentes analyses pourraient être ainsi entreprises comme les co-immunoprécipitations, le double hybride ou le TAP-tag [342].

L'étude plus approfondie d'autres sous-protéomes de la lignée H358 pourrait aussi être intéressante, notamment le protéome nucléaire. Certaines cibles de p53 sont elles-mêmes des facteurs de transcription et peuvent donc contribuer à une régulation indirecte de p53 sur la transcription de gènes codant pour des protéines sécrétées.

Afin d'établir plus précisément quelles sont les protéines du secretome régulées par p53 au niveau transcriptionnel, il serait intéressant d'étudier le niveau d'expression des transcrits par bio-puces ou PCR quantitative. Aucune étude de ce type n'a été réalisée à ce jour pour les gènes cibles de p53 dans la lignée H358. Une telle étude transcriptomique permettrait éventuellement la mise en évidence de nouvelles cibles transcriptionnelles de p53 et de mieux comprendre la régulation du secretome par p53.

Une étude plus approfondie de l'action de p53 sur les voies de sécrétion serait également très utile. Certaines protéines du secretome sont libérées dans le milieu extracellulaire par des voies non classiques. Yu et al. ont notamment montré le rôle de p53 sur l'activation de la sécrétion par la voie exosomale [186]. L'action de p53 passe par l'induction de la transcription du gène *TSAP6* dont le produit est suffisant pour déclencher la libération exosomale. L'étude du niveau d'expression de TSAP6 permettrait de conforter l'emploi ou non de cette voie dans notre modèle. De plus, la purification et l'analyse du protéome exosomal pourrait également apporter des données intéressantes.

Notre étude porte essentiellement sur la lignée cellulaire H358. L'utilisation d'autres modèles cellulaires serait pertinente pour conforter et généraliser nos résultats. L'étude inverse pourrait aussi être élaborée en analysant une lignée cellulaire d'adénocarcinome de type NSCLC non délétée en *TP53 versus* la même lignée traitée avec des siRNA anti-*TP53*. Ces autres modèles perettraient de confirmer nos résultats et d'éliminer définitivement le risque d'avoir des variations d'expression doxycycline-dépendantes.

Les résultats obtenus lors de notre étude *in vivo* sont très encourageants. Plusieurs autres candidats potentiellement intéressants n'ont pas pu être testés faute de temps ou de kits commerciaux de dosage disponibles. De plus, des coupes de tumeur des souris ont été conservées en vue de leur analyse en immunohistochimie. Cette étude sera très intéressante pour valider les différences observées et les corréler avec les niveaux d'expression de p53, de caspase 3 active (apoptose), de CD31 (angiogenèse) ou de Ki67 (prolifération). De plus, celle-ci donnera des

informations sur la localisation cellulaire des protéines d'intérêts et notamment la confirmation de leur sécrétion dans le milieu extracellulaire.

L'objectif à plus long terme sera de regarder le niveau d'expression des protéines d'intérêt dans des échantillons cliniques issus de patients atteins de NSCLC avec caractérisation du statut p53. La délétion totale de p53 étant rare cliniquement, il sera important de cibler de façon pertinente les mutations de p53 considérées en se concentrant notamment sur les « hotspots » de mutation. Une première étude pourrait être réalisée en anatomopathologie sur des coupes de tissus tumoraux de patients. Les cibles les plus intéressantes pourront ensuite être dosées dans le plasma de patients en vue d'une possible utilisation comme marqueur pronostic, diagnostic ou thérapeutique.

La validation de bio-marqueurs potentiels dans des échantillons cliniques repose actuellement sur l'utilisation d'anticorps spécifiques dirigés contre ce marqueur. Typiquement, une cohorte de patients est incluse dans une étude et les échantillons sont criblés par ELISA (dans le cas des plasmas) ou immunohistochimie (dans le cas des biopsies de tumeur). La nécessité d'avoir l'anticorps adéquat est généralement consommateur en temps et en coûts.
La protéomique pourrait prochainement apporter une alternative intéressante. Les spectromètres de masse de type triple quadrupôle utilisables en mode MRM (Multiple Reaction Monitoring) permettent la quantification absolue de bio-marqueurs dans le sang ou les extraits tissulaires [343]. La MRM, initialement développée sur le spectromètre de masse Q-TRAP 4000 (Applied Biosystems), permet la détection multiplex de nombreuses caractéristiques des peptides et leur quantification absolue par l'utilisation de peptides marqués isotopiquement. Ce type de spectrométrie peut être couplé à des approches de quantification absolue telles que le QCAT ou le PSAQ. Puisqu'elle n'utilise pas d'anticorps et que les bio-marqueurs découverts par spectrométrie de masse produisent généralement des peptides détectables en MRM, la méthode MRM représente une approche de choix pour la validation de bio-marqueurs sur de grandes cohortes de patients. Un certain nombre d'études référencées dans la revue de Kitteringham et al. [343] vont dans ce sens, dont par exemple le récent dosage de la PSA par MRM dans le plasma et le sérum [344]. Cette approche est encore jeune mais si les premiers résultats obtenus se confirment, elle pourrait représenter une méthode incontournable dans les prochaines années.

REFERENCES BIBLIOGRAPHIQUES

REFERENCES BIBLIOGRAPHIQUES

[1] Wilkins, M. R., Sanchez, J. C., Gooley, A. A., Appel, R. D., *et al.*, Progress with proteome projects: why all proteins expressed by a genome should be identified and how to do it. *Biotechnology & genetic engineering reviews* 1996, *13*, 19-50.
[2] Wilkins, M. R., Sanchez, J. C., Williams, K. L., Hochstrasser, D. F., Current challenges and future applications for protein maps and post-translational vector maps in proteome projects. *Electrophoresis* 1996, *17*, 830-838.
[3] Kenyon, G. L., DeMarini, D. M., Fuchs, E., Galas, D. J., *et al.*, Defining the mandate of proteomics in the post-genomics era: workshop report. *Mol Cell Proteomics* 2002, *1*, 763-780.
[4] O'Farrell, P. H., High resolution two-dimensional electrophoresis of proteins. *The Journal of biological chemistry* 1975, *250*, 4007-4021.
[5] Klose, J., Protein mapping by combined isoelectric focusing and electrophoresis of mouse tissues. A novel approach to testing for induced point mutations in mammals. *Humangenetik* 1975, *26*, 231-243.
[6] Scheele, G. A., Two-dimensional gel analysis of soluble proteins. Charaterization of guinea pig exocrine pancreatic proteins. *The Journal of biological chemistry* 1975, *250*, 5375-5385.
[7] Pandey, A., Mann, M., Proteomics to study genes and genomes. *Nature* 2000, *405*, 837-846.
[8] Fleischmann, R. D., Adams, M. D., White, O., Clayton, R. A., *et al.*, Whole-genome random sequencing and assembly of Haemophilus influenzae Rd. *Science (New York, N.Y* 1995, *269*, 496-512.
[9] Schmutz, J., Wheeler, J., Grimwood, J., Dickson, M., *et al.*, Quality assessment of the human genome sequence. *Nature* 2004, *429*, 365-368.
[10] O'Donovan, C., Apweiler, R., Bairoch, A., The human proteomics initiative (HPI). *Trends in biotechnology* 2001, *19*, 178-181.
[11] Venter, J. C., Adams, M. D., Myers, E. W., Li, P. W., *et al.*, The sequence of the human genome. *Science (New York, N.Y* 2001, *291*, 1304-1351.
[12] Lottspeich, F., Proteome Analysis: A Pathway to the Functional Analysis of Proteins. *Angewandte Chemie (International ed* 1999, *38*, 2476-2492.
[13] Gygi, S. P., Rochon, Y., Franza, B. R., Aebersold, R., Correlation between protein and mRNA abundance in yeast. *Molecular and cellular biology* 1999, *19*, 1720-1730.
[14] Patterson, S. D., Aebersold, R. H., Proteomics: the first decade and beyond. *Nature genetics* 2003, *33 Suppl*, 311-323.
[15] Aebersold, R., Mann, M., Mass spectrometry-based proteomics. *Nature* 2003, *422*, 198-207.
[16] Roepstorff, P., Fohlman, J., Proposal for a common nomenclature for sequence ions in mass spectra of peptides. *Biomedical mass spectrometry* 1984, *11*, 601.
[17] Biemann, K., Mass spectrometry of peptides and proteins. *Annual review of biochemistry* 1992, *61*, 977-1010.
[18] Karas, M., Hillenkamp, F., Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons. *Analytical chemistry* 1988, *60*, 2299-2301.
[19] Tanaka, K., Waki, H., Ido, Y., Akita, S., *et al.*, Protein and Polymer Analyses up to m/z 100 000 by Laser Ionization Time-of flight Mass Spectrometry *Rapid Commun Mass Spectrom* 1988, *2*, 151-153.
[20] Kussmann, M., Lassing, U., Sturmer, C. A., Przybylski, M., Roepstorff, P., Matrix-assisted laser desorption/ionization mass spectrometric peptide mapping of the neural cell adhesion protein neurolin purified by sodium dodecyl sulfate polyacrylamide gel electrophoresis or acidic precipitation. *J Mass Spectrom* 1997, *32*, 483-493.
[21] Rusconi, F., Schmitter, J. M., Préparation des échantillons en vue d'une analyse par spectrométrie de masse MALDI-TOF. *Analusis Magazine* 1998, *10*.

[22] Salehpour, M., Perera, I., Kjellberg, J., Hedin, A., *et al.*, Laser-induced desorption of proteins. *Rapid Commun Mass Spectrom* 1989, *3*, 259-263.
[23] Niu, S., Zhang, W., Chait, B. T., Direct comparison of infrared and ultraviolet wavelength matrix-assisted laser desorption/ionization mass spectrometry of proteins. *Journal of the American Society for Mass Spectrometry* 1998, *9*, 1-7.
[24] Cotter, R. J., Time-of-flight mass spectrometry: an increasing role in the life sciences. *Biomedical & environmental mass spectrometry* 1989, *18*, 513-532.
[25] Vestal, M. L., Juhasz, P., Martin, A., Delayed extraction matrix-assisted laser desorption time-of-flight mass spectrometry. *Rapid Commun Mass Spectrom* 2005, *9*, 1044-1050.
[26] Mamyrin, B. A., Karataev, V. I., Shmikk, D. V., Zagulin, V. A., The mass-reflectron, a new nonmagnetic time-of-flight mass spectrometer with high resolution. *Soviet Physics JETP* 1973, *37*, 45.
[27] Dole, M., Mack, L. L., Hines, R. L., Mobley, R. C., *et al.*, Molecular Beams of Macroions. *J Chem Phys* 1968, *49*, 2240-2249.
[28] Yamashita, M., Fenn, J. B., Electrospray Ion Source. Another Variation on the Free-Jet Theme. *J. Phys. Chem.* 1984, *88*, 4451-4460.
[29] Fenn, J. B., Mann, M., Meng, C. K., Wong, S. F., Whitehouse, C. M., Electrospray ionization for mass spectrometry of large biomolecules. *Science (New York, N.Y* 1989, *246*, 64-71.
[30] Paul, W., Raether, M., Das eletrische Massenfilter. *Z Phys* 1955, *140*, 262-273.
[31] Stafford, J. G. C., Kelley, P. E., Syka, J. E. P., Reynolds, W. E., Todd, J. F. J., Recent improvements in and analytical applications of advanced ion trap technology. *International Journal of Mass Spectrometry and Ion Processes* 1984, *60*, 85-98.
[32] Mirzaei, H., Regnier, F., Structure specific chromatographic selection in targeted proteomics. *Journal of chromatography* 2005, *817*, 23-34.
[33] Ahmed, F. E., Sample preparation and fractionation for proteome analysis and cancer biomarker discovery by mass spectrometry. *Journal of separation science* 2009, *32*, 771-798.
[34] Guilak, F., Alexopoulos, L. G., Haider, M. A., Ting-Beall, H. P., Setton, L. A., Zonal uniformity in mechanical properties of the chondrocyte pericellular matrix: micropipette aspiration of canine chondrons isolated by cartilage homogenization. *Annals of biomedical engineering* 2005, *33*, 1312-1318.
[35] D'Souza, R., Brown, L. R., Newland, J. R., Levy, B. M., Lachman, L. B., Detection and characterization of interleukin-1 in human dental pulps. *Archives of oral biology* 1989, *34*, 307-313.
[36] Bodzon-Kulakowska, A., Bierczynska-Krzysik, A., Dylag, T., Drabik, A., *et al.*, Methods for samples preparation in proteomic research. *Journal of chromatography* 2007, *849*, 1-31.
[37] Rabilloud, T., Adessi, C., Giraudel, A., Lunardi, J., Improvement of the solubilization of proteins in two-dimensional electrophoresis with immobilized pH gradients. *Electrophoresis* 1997, *18*, 307-316.
[38] Visser, N. F., Lingeman, H., Irth, H., Sample preparation for peptides and proteins in biological matrices prior to liquid chromatography and capillary zone electrophoresis. *Analytical and bioanalytical chemistry* 2005, *382*, 535-558.
[39] Fountoulakis, M., Application of proteomics technologies in the investigation of the brain. *Mass spectrometry reviews* 2004, *23*, 231-258.
[40] Kenrick, K. G., Margolis, J., Isoelectric focusing and gradient gel electrophoresis: a two-dimensional technique. *Analytical biochemistry* 1970, *33*, 204-207.
[41] Harris, L. R., Churchward, M. A., Butt, R. H., Coorssen, J. R., Assessing detection methods for gel-based proteomic analyses. *Journal of proteome research* 2007, *6*, 1418-1425.
[42] Chevallet, M., Luche, S., Rabilloud, T., Silver staining of proteins in polyacrylamide gels. *Nature protocols* 2006, *1*, 1852-1858.
[43] Michel, P. E., Reymond, F., Arnaud, I. L., Josserand, J., *et al.*, Protein fractionation in a multicompartment device using Off-Gel isoelectric focusing. *Electrophoresis* 2003, *24*, 3-11.

[44] Arnaud, I. L., Josserand, J., Jensen, H., Lion, N., *et al.*, Salt removal during Off-Gel electrophoresis of protein samples. *Electrophoresis* 2005, *26*, 1650-1658.
[45] Mitulovic, G., Mechtler, K., HPLC techniques for proteomics analysis--a short overview of latest developments. *Briefings in functional genomics & proteomics* 2006, *5*, 249-260.
[46] Washburn, M. P., Wolters, D., Yates, J. R., 3rd, Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nature biotechnology* 2001, *19*, 242-247.
[47] Paoletti, A. C., Zybailov, B., Washburn, M. P., Principles and applications of multidimensional protein identification technology. *Expert review of proteomics* 2004, *1*, 275-282.
[48] Nagele, E., Vollmer, M., Horth, P., Vad, C., 2D-LC/MS techniques for the identification of proteins in highly complex mixtures. *Expert review of proteomics* 2004, *1*, 37-46.
[49] Henzel, W. J., Billeci, T. M., Stults, J. T., Wong, S. C., *et al.*, Identifying proteins from two-dimensional gels by molecular mass searching of peptide fragments in protein sequence databases. *Proceedings of the National Academy of Sciences of the United States of America* 1993, *90*, 5011-5015.
[50] James, P., Quadroni, M., Carafoli, E., Gonnet, G., Protein identification by mass profile fingerprinting. *Biochemical and biophysical research communications* 1993, *195*, 58-64.
[51] Mann, M., Hojrup, P., Roepstorff, P., Use of mass spectrometric molecular weight information to identify proteins in sequence databases. *Biological mass spectrometry* 1993, *22*, 338-345.
[52] Yates, J. R., 3rd, Speicher, S., Griffin, P. R., Hunkapiller, T., Peptide mass maps: a highly informative approach to protein identification. *Analytical biochemistry* 1993, *214*, 397-408.
[53] Nesvizhskii, A. I., Vitek, O., Aebersold, R., Analysis and validation of proteomic data generated by tandem mass spectrometry. *Nature methods* 2007, *4*, 787-797.
[54] McGregor, E., Dunn, M. J., Proteomics of the heart: unraveling disease. *Circulation research* 2006, *98*, 309-321.
[55] Charbonnier, S., Gallego, O., Gavin, A. C., The social network of a cell: Recent advances in interactome mapping. *Biotechnology annual review* 2008, *14*, 1-28.
[56] Mann, M., Jensen, O. N., Proteomic analysis of post-translational modifications. *Nature biotechnology* 2003, *21*, 255-261.
[57] Kim, Y. M., Song, E. J., Seo, J., Kim, H. J., Lee, K. J., Proteomic analysis of tyrosine phosphorylations in vascular endothelial growth factor- and reactive oxygen species-mediated signaling pathway. *Journal of proteome research* 2007, *6*, 593-601.
[58] Martin, K., Steinberg, T. H., Goodman, T., Schulenberg, B., *et al.*, Strategies and solid-phase formats for the analysis of protein and peptide phosphorylation employing a novel fluorescent phosphorylation sensor dye. *Combinatorial chemistry & high throughput screening* 2003, *6*, 331-339.
[59] Garcia, A., Senis, Y. A., Antrobus, R., Hughes, C. E., *et al.*, A global proteomics approach identifies novel phosphorylated signaling proteins in GPVI-activated platelets: involvement of G6f, a novel platelet Grb2-binding membrane adapter. *Proteomics* 2006, *6*, 5332-5343.
[60] Neubauer, G., Gottschalk, A., Fabrizio, P., Seraphin, B., *et al.*, Identification of the proteins of the yeast U1 small nuclear ribonucleoprotein complex by mass spectrometry. *Proceedings of the National Academy of Sciences of the United States of America* 1997, *94*, 385-390.
[61] Panchaud, A., Affolter, M., Moreillon, P., Kussmann, M., Experimental and computational approaches to quantitative proteomics: status quo and outlook. *Journal of proteomics* 2008, *71*, 19-33.
[62] Unlu, M., Morgan, M. E., Minden, J. S., Difference gel electrophoresis: a single gel method for detecting changes in protein extracts. *Electrophoresis* 1997, *18*, 2071-2077.
[63] Viswanathan, S., Unlu, M., Minden, J. S., Two-dimensional difference gel electrophoresis. *Nature protocols* 2006, *1*, 1351-1358.
[64] Kellermann, J., ICPL--isotope-coded protein label. *Methods in molecular biology (Clifton, N.J* 2008, *424*, 113-123.

[65] Gygi, S. P., Rist, B., Gerber, S. A., Turecek, F., *et al.*, Quantitative analysis of complex protein mixtures using isotope-coded affinity tags. *Nature biotechnology* 1999, *17*, 994-999.
[66] Ross, P. L., Huang, Y. N., Marchese, J. N., Williamson, B., *et al.*, Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents. *Mol Cell Proteomics* 2004, *3*, 1154-1169.
[67] Stewart, II, Thomson, T., Figeys, D., 18O labeling: a tool for proteomics. *Rapid Commun Mass Spectrom* 2001, *15*, 2456-2465.
[68] Ong, S. E., Blagoev, B., Kratchmarova, I., Kristensen, D. B., *et al.*, Stable isotope labeling by amino acids in cell culture, SILAC, as a simple and accurate approach to expression proteomics. *Mol Cell Proteomics* 2002, *1*, 376-386.
[69] Yang, W., Steen, H., Freeman, M. R., Proteomic approaches to the analysis of multiprotein signaling complexes. *Proteomics* 2008, *8*, 832-851.
[70] Fenselau, C., A review of quantitative methods for proteomic studies. *Journal of chromatography* 2007, *855*, 14-20.
[71] Oda, Y., Huang, K., Cross, F. R., Cowburn, D., Chait, B. T., Accurate quantitation of protein expression and site-specific phosphorylation. *Proceedings of the National Academy of Sciences of the United States of America* 1999, *96*, 6591-6596.
[72] Gu, S., Pan, S., Bradbury, E. M., Chen, X., Use of deuterium-labeled lysine for efficient protein identification and peptide de novo sequencing. *Analytical chemistry* 2002, *74*, 5774-5785.
[73] Yao, X., Freas, A., Ramirez, J., Demirev, P. A., Fenselau, C., Proteolytic 18O labeling for comparative proteomics: model studies with two serotypes of adenovirus. *Analytical chemistry* 2001, *73*, 2836-2842.
[74] Choe, L., D'Ascenzo, M., Relkin, N. R., Pappin, D., *et al.*, 8-plex quantitation of changes in cerebrospinal fluid protein expression in subjects undergoing intravenous immunoglobulin treatment for Alzheimer's disease. *Proteomics* 2007, *7*, 3651-3660.
[75] Pierce, A., Unwin, R. D., Evans, C. A., Griffiths, S., *et al.*, Eight-channel iTRAQ enables comparison of the activity of six leukemogenic tyrosine kinases. *Mol Cell Proteomics* 2008, *7*, 853-863.
[76] Ow, S. Y., Cardona, T., Taton, A., Magnuson, A., *et al.*, Quantitative shotgun proteomics of enriched heterocysts from Nostoc sp. PCC 7120 using 8-plex isobaric peptide tags. *Journal of proteome research* 2008, *7*, 1615-1628.
[77] Thompson, A., Schafer, J., Kuhn, K., Kienle, S., *et al.*, Tandem mass tags: a novel quantification strategy for comparative analysis of complex protein mixtures by MS/MS. *Analytical chemistry* 2003, *75*, 1895-1904.
[78] Dayon, L., Hainard, A., Licker, V., Turck, N., *et al.*, Relative quantification of proteins in human cerebrospinal fluids by MS/MS using 6-plex isobaric tags. *Analytical chemistry* 2008, *80*, 2921-2931.
[79] Xie, B., Parman, C., Patton, W. F., *Poster PerkinElmer* 2008.
[80] Allet, N., Barrillat, N., Baussant, T., Boiteau, C., *et al.*, In vitro and in silico processes to identify differentially expressed proteins. *Proteomics* 2004, *4*, 2333-2351.
[81] Liu, H., Sadygov, R. G., Yates, J. R., 3rd, A model for random sampling and estimation of relative protein abundance in shotgun proteomics. *Analytical chemistry* 2004, *76*, 4193-4201.
[82] Rappsilber, J., Ryder, U., Lamond, A. I., Mann, M., Large-scale proteomic analysis of the human spliceosome. *Genome research* 2002, *12*, 1231-1245.
[83] Sanders, S. L., Jennings, J., Canutescu, A., Link, A. J., Weil, P. A., Proteomics of the eukaryotic transcription machinery: identification of proteins associated with components of yeast TFIID by multidimensional mass spectrometry. *Molecular and cellular biology* 2002, *22*, 4723-4738.
[84] Ishihama, Y., Oda, Y., Tabata, T., Sato, T., *et al.*, Exponentially modified protein abundance index (emPAI) for estimation of absolute protein amount in proteomics by the number of sequenced peptides per protein. *Mol Cell Proteomics* 2005, *4*, 1265-1272.

[85] Chelius, D., Bondarenko, P. V., Quantitative profiling of proteins in complex mixtures using liquid chromatography and mass spectrometry. *Journal of proteome research* 2002, *1*, 317-323.
[86] Lasonder, E., Ishihama, Y., Andersen, J. S., Vermunt, A. M., *et al.*, Analysis of the Plasmodium falciparum proteome by high-accuracy mass spectrometry. *Nature* 2002, *419*, 537-542.
[87] Palagi, P. M., Walther, D., Quadroni, M., Catherinet, S., *et al.*, MSight: an image analysis software for liquid chromatography-mass spectrometry. *Proteomics* 2005, *5*, 2381-2384.
[88] Mueller, L. N., Rinner, O., Schmidt, A., Letarte, S., *et al.*, SuperHirn - a novel tool for high resolution LC-MS-based peptide/protein profiling. *Proteomics* 2007, *7*, 3470-3480.
[89] Leptos, K. C., Sarracino, D. A., Jaffe, J. D., Krastins, B., Church, G. M., MapQuant: open-source software for large-scale protein quantification. *Proteomics* 2006, *6*, 1770-1782.
[90] Li, X. J., Yi, E. C., Kemp, C. J., Zhang, H., Aebersold, R., A software suite for the generation and comparison of peptide arrays from sets of data collected by liquid chromatography-mass spectrometry. *Mol Cell Proteomics* 2005, *4*, 1328-1340.
[91] Matthiesen, R., Bunkenborg, J., Stensballe, A., Jensen, O. N., *et al.*, Database-independent, database-dependent, and extended interpretation of peptide mass spectra in VEMS V2.0. *Proteomics* 2004, *4*, 2583-2593.
[92] Matthiesen, R., Lundsgaard, M., Welinder, K. G., Bauw, G., Interpreting peptide mass spectra by VEMS. *Bioinformatics (Oxford, England)* 2003, *19*, 792-793.
[93] Matthiesen, R., Trelle, M. B., Hojrup, P., Bunkenborg, J., Jensen, O. N., VEMS 3.0: algorithms and computational tools for tandem mass spectrometry based identification of post-translational modifications in proteins. *Journal of proteome research* 2005, *4*, 2338-2347.
[94] Gerber, S. A., Rush, J., Stemman, O., Kirschner, M. W., Gygi, S. P., Absolute quantification of proteins and phosphoproteins from cell lysates by tandem MS. *Proceedings of the National Academy of Sciences of the United States of America* 2003, *100*, 6940-6945.
[95] Beynon, R. J., Doherty, M. K., Pratt, J. M., Gaskell, S. J., Multiplexed absolute quantification in proteomics using artificial QCAT proteins of concatenated signature peptides. *Nature methods* 2005, *2*, 587-589.
[96] Pratt, J. M., Simpson, D. M., Doherty, M. K., Rivers, J., *et al.*, Multiplexed absolute quantification for proteomics using concatenated signature peptides encoded by QconCAT genes. *Nature protocols* 2006, *1*, 1029-1043.
[97] Brun, V., Dupuis, A., Adrait, A., Marcellin, M., *et al.*, Isotope-labeled protein standards: toward absolute quantitative proteomics. *Mol Cell Proteomics* 2007, *6*, 2139-2149.
[98] Dupuis, A., Hennekinne, J. A., Garin, J., Brun, V., Protein Standard Absolute Quantification (PSAQ) for improved investigation of staphylococcal food poisoning outbreaks. *Proteomics* 2008, *8*, 4633-4636.
[99] Parkin, D. M., Bray, F., Ferlay, J., Pisani, P., Global cancer statistics, 2002. *CA: a cancer journal for clinicians* 2005, *55*, 74-108.
[100] Greenlee, R. T., Murray, T., Bolden, S., Wingo, P. A., Cancer statistics, 2000. *CA: a cancer journal for clinicians* 2000, *50*, 7-33.
[101] Parkin, D. M., International variation. *Oncogene* 2004, *23*, 6329-6340.
[102] Travis, W. D., Travis, L. B., Devesa, S. S., Lung cancer. *Cancer* 1995, *75*, 191-202.
[103] Hamilton, W., Peters, T. J., Round, A., Sharp, D., What are the clinical features of lung cancer before the diagnosis is made? A population based case-control study. *Thorax* 2005, *60*, 1059-1065.
[104] Henschke, C. I., McCauley, D. I., Yankelevitz, D. F., Naidich, D. P., *et al.*, Early Lung Cancer Action Project: overall design and findings from baseline screening. *Lancet* 1999, *354*, 99-105.
[105] Kaneko, M., Kusumoto, M., Kobayashi, T., Moriyama, N., *et al.*, Computed tomography screening for lung carcinoma in Japan. *Cancer* 2000, *89*, 2485-2488.
[106] Spira, A., Ettinger, D. S., Multidisciplinary management of lung cancer. *The New England journal of medicine* 2004, *350*, 379-392.

[107] Schiller, J. H., Current standards of care in small-cell and non-small-cell lung cancer. *Oncology* 2001, *61 Suppl 1*, 3-13.
[108] Stupp, R., Monnerat, C., Turrisi, A. T., 3rd, Perry, M. C., Leyvraz, S., Small cell lung cancer: state of the art and future perspectives. *Lung cancer (Amsterdam, Netherlands)* 2004, *45*, 105-117.
[109] Dalton, W. S., Friend, S. H., Cancer biomarkers--an invitation to the table. *Science (New York, N.Y* 2006, *312*, 1165-1168.
[110] Fung, E. T., Wright, G. L., Jr., Dalmasso, E. A., Proteomic strategies for biomarker identification: progress and challenges. *Current opinion in molecular therapeutics* 2000, *2*, 643-650.
[111] Arteaga, C. L., Moulder, S. L., Yakes, F. M., HER (erbB) tyrosine kinase inhibitors in the treatment of breast cancer. *Seminars in oncology* 2002, *29*, 4-10.
[112] Baylin, S. B., DNA methylation and gene silencing in cancer. *Nature clinical practice* 2005, *2 Suppl 1*, S4-11.
[113] Belinsky, S. A., Gene-promoter hypermethylation as a biomarker in lung cancer. *Nature reviews* 2004, *4*, 707-717.
[114] Oh, P., Li, Y., Yu, J., Durr, E., *et al.*, Subtractive proteomic mapping of the endothelial surface in lung and solid tumours for tissue-specific therapy. *Nature* 2004, *429*, 629-635.
[115] Sung, H. J., Cho, J. Y., Biomarkers for the lung cancer diagnosis and their advances in proteomics. *BMB reports* 2008, *41*, 615-625.
[116] Stieber, P., R.Hatz, Holdenrieder, S., Molina, R., *et al.*, National Academy of Clinical Biochemistry Guidelines for the Use of Tumor Markers in Lung Cancer.
[117] Simpson, R. J., Bernhard, O. K., Greening, D. W., Moritz, R. L., Proteomics-driven cancer biomarker discovery: looking to the future. *Current opinion in chemical biology* 2008, *12*, 72-77.
[118] Lane, D. P., Crawford, L. V., T antigen is bound to a host protein in SV40-transformed cells. *Nature* 1979, *278*, 261-263.
[119] Linzer, D. I., Levine, A. J., Characterization of a 54K dalton cellular SV40 tumor antigen present in SV40-transformed cells and uninfected embryonal carcinoma cells. *Cell* 1979, *17*, 43-52.
[120] Finlay, C. A., Hinds, P. W., Levine, A. J., The p53 proto-oncogene can act as a suppressor of transformation. *Cell* 1989, *57*, 1083-1093.
[121] Eliyahu, D., Michalovitz, D., Eliyahu, S., Pinhasi-Kimhi, O., Oren, M., Wild-type p53 can inhibit oncogene-mediated focus formation. *Proceedings of the National Academy of Sciences of the United States of America* 1989, *86*, 8763-8767.
[122] Levine, A. J., p53, the cellular gatekeeper for growth and division. *Cell* 1997, *88*, 323-331.
[123] Cheah, P. L., Looi, L. M., p53: an overview of over two decades of study. *The Malaysian journal of pathology* 2001, *23*, 9-16.
[124] Soussi, T., May, P., Structural aspects of the p53 protein in relation to gene evolution: a second look. *Journal of molecular biology* 1996, *260*, 623-637.
[125] Prives, C., Hall, P. A., The p53 pathway. *The Journal of pathology* 1999, *187*, 112-126.
[126] Kubbutat, M. H., Jones, S. N., Vousden, K. H., Regulation of p53 stability by Mdm2. *Nature* 1997, *387*, 299-303.
[127] el-Deiry, W. S., Kern, S. E., Pietenpol, J. A., Kinzler, K. W., Vogelstein, B., Definition of a consensus binding site for p53. *Nature genetics* 1992, *1*, 45-49.
[128] Bourdon, J. C., Deguin-Chambon, V., Lelong, J. C., Dessen, P., *et al.*, Further characterisation of the p53 responsive element--identification of new candidate genes for trans-activation by p53. *Oncogene* 1997, *14*, 85-94.
[129] Jeffrey, P. D., Gorina, S., Pavletich, N. P., Crystal structure of the tetramerization domain of the p53 tumor suppressor at 1.7 angstroms. *Science (New York, N.Y* 1995, *267*, 1498-1502.
[130] Hupp, T. R., Meek, D. W., Midgley, C. A., Lane, D. P., Regulation of the specific DNA binding function of p53. *Cell* 1992, *71*, 875-886.

[131] Jayaraman, J., Prives, C., Activation of p53 sequence-specific DNA binding by short single strands of DNA requires the p53 C-terminus. *Cell* 1995, *81*, 1021-1029.
[132] Pluquet, O., Hainaut, P., Genotoxic and non-genotoxic pathways of p53 induction. *Cancer letters* 2001, *174*, 1-15.
[133] Appella, E., Anderson, C. W., Signaling to p53: breaking the posttranslational modification code. *Pathologie-biologie* 2000, *48*, 227-245.
[134] Haupt, Y., Maya, R., Kazaz, A., Oren, M., Mdm2 promotes the rapid degradation of p53. *Nature* 1997, *387*, 296-299.
[135] Honda, R., Tanaka, H., Yasuda, H., Oncoprotein MDM2 is a ubiquitin ligase E3 for tumor suppressor p53. *FEBS letters* 1997, *420*, 25-27.
[136] Sherr, C. J., The Pezcoller lecture: cancer cell cycles revisited. *Cancer research* 2000, *60*, 3689-3695.
[137] Wu, X., Bayle, J. H., Olson, D., Levine, A. J., The p53-mdm-2 autoregulatory feedback loop. *Genes & development* 1993, *7*, 1126-1132.
[138] Hainaut, P., Hollstein, M., p53 and human cancer: the first ten thousand mutations. *Advances in cancer research* 2000, *77*, 81-137.
[139] Oren, M., Rotter, V., Introduction: p53--the first twenty years. *Cell Mol Life Sci* 1999, *55*, 9-11.
[140] Jin, S., Levine, A. J., The p53 functional circuit. *Journal of cell science* 2001, *114*, 4139-4140.
[141] Wang, W., Rastinejad, F., El-Deiry, W. S., Restoring p53-dependent tumor suppression. *Cancer biology & therapy* 2003, *2*, S55-63.
[142] Hanahan, D., Weinberg, R. A., The hallmarks of cancer. *Cell* 2000, *100*, 57-70.
[143] Guimaraes, D. P., Hainaut, P., TP53: a key gene in human cancer. *Biochimie* 2002, *84*, 83-93.
[144] Szymanska, K., Hainaut, P., TP53 and mutations in human cancer. *Acta biochimica Polonica* 2003, *50*, 231-238.
[145] Hernandez-Boussard, T., Rodriguez-Tome, P., Montesano, R., Hainaut, P., IARC p53 mutation database: a relational database to compile and analyze p53 mutations in human tumors and cell lines. International Agency for Research on Cancer. *Human mutation* 1999, *14*, 1-8.
[146] Donehower, L. A., Harvey, M., Slagle, B. L., McArthur, M. J., *et al.*, Mice deficient for p53 are developmentally normal but susceptible to spontaneous tumours. *Nature* 1992, *356*, 215-221.
[147] Lin, J., Teresky, A. K., Levine, A. J., Two critical hydrophobic amino acids in the N-terminal domain of the p53 protein are required for the gain of function phenotypes of human p53 mutants. *Oncogene* 1995, *10*, 2387-2390.
[148] Matas, D., Sigal, A., Stambolsky, P., Milyavsky, M., *et al.*, Integrity of the N-terminal transcription domain of p53 is required for mutant p53 interference with drug-induced apoptosis. *The EMBO journal* 2001, *20*, 4163-4172.
[149] May, P., May, E., Twenty years of p53 research: structural and functional aspects of the p53 protein. *Oncogene* 1999, *18*, 7621-7636.
[150] Cho, Y., Gorina, S., Jeffrey, P. D., Pavletich, N. P., Crystal structure of a p53 tumor suppressor-DNA complex: understanding tumorigenic mutations. *Science (New York, N.Y* 1994, *265*, 346-355.
[151] Meplan, C., Richard, M. J., Hainaut, P., Redox signalling and transition metals in the control of the p53 pathway. *Biochemical pharmacology* 2000, *59*, 25-33.
[152] Bouchet, B. P., de Fromentel, C. C., Puisieux, A., Galmarini, C. M., p53 as a target for anti-cancer drug development. *Critical reviews in oncology/hematology* 2006, *58*, 190-207.
[153] Vousden, K. H., Lu, X., Live or let die: the cell's response to p53. *Nature reviews* 2002, *2*, 594-604.
[154] Olivier, M., Petitjean, A., Marcel, V., Petre, A., *et al.*, Recent advances in p53 research: an interdisciplinary perspective. *Cancer gene therapy* 2009, *16*, 1-12.

[155] Gannon, J. V., Greaves, R., Iggo, R., Lane, D. P., Activating mutations in p53 produce a common conformational effect. A monoclonal antibody specific for the mutant form. *The EMBO journal* 1990, *9*, 1595-1602.
[156] Lane, D. P., Benchimol, S., p53: oncogene or anti-oncogene? *Genes & development* 1990, *4*, 1-8.
[157] Soussi, T., The p53 tumour suppressor gene: a model for molecular epidemiology of human cancer. *Molecular medicine today* 1996, *2*, 32-37.
[158] Brambilla, C., Fievet, F., Jeanmart, M., de Fraipont, F., *et al.*, Early detection of lung cancer: role of biomarkers. *The European respiratory journal* 2003, *39*, 36s-44s.
[159] Tammemagi, M. C., McLaughlin, J. R., Bull, S. B., Meta-analyses of p53 tumor suppressor gene alterations and clinicopathological features in resected lung cancers. *Cancer Epidemiol Biomarkers Prev* 1999, *8*, 625-634.
[160] Hecht, S. S., Adams, J. D., Numoto, S., Hoffmann, D., Induction of respiratory tract tumors in Syrian golden hamsters by a single dose of 4-(methylnitrosamino)-1-(3-pyridyl)-1-butanone (NNK) and the effect of smoke inhalation. *Carcinogenesis* 1983, *4*, 1287-1290.
[161] Rodin, S. N., Rodin, A. S., On the excess of G --> T transversions in the p53 gene in lung cancer cell lines. Reply to Pfeifer and Hainaut. *Mutation research* 2004, *545*, 141-144; author reply 145-146; discussion 147.
[162] Hainaut, P., Pfeifer, G. P., Patterns of p53 G-->T transversions in lung cancers reflect the primary mutagenic signature of DNA-damage by tobacco smoke. *Carcinogenesis* 2001, *22*, 367-374.
[163] Gouas, D., Shi, H., Hainaut, P., The aflatoxin-induced TP53 mutation at codon 249 (R249S): Biomarker of exposure, early detection and target for therapy. *Cancer letters* 2009.
[164] Hosny, G., Farahat, N., Tayel, H., Hainaut, P., Ser-249 TP53 and CTNNB1 mutations in circulating free DNA of Egyptian patients with hepatocellular carcinoma versus chronic liver diseases. *Cancer letters* 2008, *264*, 201-208.
[165] Kirk, G. D., Lesi, O. A., Mendy, M., Szymanska, K., *et al.*, 249(ser) TP53 mutation in plasma DNA, hepatitis B viral infection, and risk of hepatocellular carcinoma. *Oncogene* 2005, *24*, 5858-5867.
[166] Huang, X. H., Sun, L. H., Lu, D. D., Sun, Y., *et al.*, Codon 249 mutation in exon 7 of p53 gene in plasma DNA: maybe a new early diagnostic marker of hepatocellular carcinoma in Qidong risk area, China. *World J Gastroenterol* 2003, *9*, 692-695.
[167] Liu, Z., Muehlbauer, K. R., Schmeiser, H. H., Hergenhahn, M., *et al.*, p53 mutations in benzo(a)pyrene-exposed human p53 knock-in murine fibroblasts correlate with p53 mutations in human lung tumors. *Cancer research* 2005, *65*, 2583-2587.
[168] Ludovini, V., Pistola, L., Gregorc, V., Floriani, I., *et al.*, Plasma DNA, microsatellite alterations, and p53 tumor mutations are associated with disease-free survival in radically resected non-small cell lung cancer patients: a study of the perugia multidisciplinary team for thoracic oncology. *J Thorac Oncol* 2008, *3*, 365-373.
[169] Cai, D. W., Mukhopadhyay, T., Liu, Y., Fujiwara, T., Roth, J. A., Stable expression of the wild-type p53 gene in human lung cancer cells after retrovirus-mediated gene transfer. *Human gene therapy* 1993, *4*, 617-624.
[170] Fujiwara, T., Grimm, E. A., Mukhopadhyay, T., Cai, D. W., *et al.*, A retroviral wild-type p53 expression vector penetrates human lung cancer spheroids and inhibits growth by inducing apoptosis. *Cancer research* 1993, *53*, 4129-4133.
[171] Runnebaum, I. B., Wang, S., Kreienberg, R., Retrovirally mediated wild-type p53 restores S-phase modulation without inducing WAF1 mRNA in breast carcinoma cells containing mutant p53. *Journal of cellular biochemistry* 1995, *59*, 537-544.
[172] Roth, J. A., Nguyen, D., Lawrence, D. D., Kemp, B. L., *et al.*, Retrovirus-mediated wild-type p53 gene transfer to tumors of patients with lung cancer. *Nature medicine* 1996, *2*, 985-991.
[173] Brandt, C. D., Kim, H. W., Vargosko, A. J., Jeffries, B. C., *et al.*, Infections in 18,000 infants and children in a controlled study of respiratory tract disease. I. Adenovirus pathogenicity

in relation to serologic type and illness syndrome. *American journal of epidemiology* 1969, *90*, 484-500.
[174] Swisher, S. G., Roth, J. A., Nemunaitis, J., Lawrence, D. D., *et al.*, Adenovirus-mediated p53 gene transfer in advanced non-small-cell lung cancer. *Journal of the National Cancer Institute* 1999, *91*, 763-771.
[175] Nemunaitis, J., Swisher, S. G., Timmons, T., Connors, D., *et al.*, Adenovirus-mediated p53 gene transfer in sequence with cisplatin to tumors of patients with non-small-cell lung cancer. *J Clin Oncol* 2000, *18*, 609-622.
[176] Keedy, V., Wang, W., Schiller, J., Chada, S., *et al.*, Phase I study of adenovirus p53 administered by bronchoalveolar lavage in patients with bronchioloalveolar cell lung carcinoma: ECOG 6597. *J Clin Oncol* 2008, *26*, 4166-4171.
[177] Schuler, M., Herrmann, R., De Greve, J. L., Stewart, A. K., *et al.*, Adenovirus-mediated wild-type p53 gene transfer in patients receiving chemotherapy for advanced non-small-cell lung cancer: results of a multicenter phase II study. *J Clin Oncol* 2001, *19*, 1750-1758.
[178] Swisher, S. G., Roth, J. A., Komaki, R., Gu, J., *et al.*, Induction of p53-regulated genes and tumor regression in lung cancer patients after intratumoral delivery of adenoviral p53 (INGN 201) and radiation therapy. *Clin Cancer Res* 2003, *9*, 93-101.
[179] Pearson, S., Jia, H., Kandachi, K., China approves first gene therapy. *Nature biotechnology* 2004, *22*, 3-4.
[180] Barnard, D. L., Technology evaluation: Sch-58500, Canji. *Current opinion in molecular therapeutics* 2000, *2*, 586-592.
[181] Shangary, S., Wang, S., Small-molecule inhibitors of the MDM2-p53 protein-protein interaction to reactivate p53 function: a novel approach for cancer therapy. *Annual review of pharmacology and toxicology* 2009, *49*, 223-241.
[182] Vassilev, L. T., Vu, B. T., Graves, B., Carvajal, D., *et al.*, In vivo activation of the p53 pathway by small-molecule antagonists of MDM2. *Science (New York, N.Y* 2004, *303*, 844-848.
[183] Bykov, V. J., Issaeva, N., Shilov, A., Hultcrantz, M., *et al.*, Restoration of the tumor suppressor function to mutant p53 by a low-molecular-weight compound. *Nature medicine* 2002, *8*, 282-288.
[184] Doggrell, S. A., RITA--a small-molecule anticancer drug that targets p53. *Expert opinion on investigational drugs* 2005, *14*, 739-742.
[185] Zhao, R., Gish, K., Murphy, M., Yin, Y., *et al.*, Analysis of p53-regulated gene expression patterns using oligonucleotide arrays. *Genes & development* 2000, *14*, 981-993.
[186] Yu, X., Harris, S. L., Levine, A. J., The regulation of exosome secretion: a novel function of the p53 protein. *Cancer research* 2006, *66*, 4795-4801.
[187] Lespagnol, A., Duflaut, D., Beekman, C., Blanc, L., *et al.*, Exosome secretion, including the DNA damage-induced p53-dependent secretory pathway, is severely compromised in TSAP6/Steap3-null mice. *Cell death and differentiation* 2008, *15*, 1723-1733.
[188] Tarasov, V., Jung, P., Verdoodt, B., Lodygin, D., *et al.*, Differential regulation of microRNAs by p53 revealed by massively parallel sequencing: miR-34a is a p53 target that induces apoptosis and G1-arrest. *Cell cycle (Georgetown, Tex* 2007, *6*, 1586-1593.
[189] Bommer, G. T., Gerin, I., Feng, Y., Kaczorowski, A. J., *et al.*, p53-mediated activation of miRNA34 candidate tumor-suppressor genes. *Curr Biol* 2007, *17*, 1298-1307.
[190] Zigrino, P., Loffek, S., Mauch, C., Tumor-stroma interactions: their role in the control of tumor cell invasion. *Biochimie* 2005, *87*, 321-328.
[191] Khwaja, F. W., Svoboda, P., Reed, M., Pohl, J., *et al.*, Proteomic identification of the wt-p53-regulated tumor cell secretome. *Oncogene* 2006.
[192] Sieber, O. M., Heinimann, K., Tomlinson, I. P., Genomic instability--the engine of tumorigenesis? *Nature reviews* 2003, *3*, 701-708.
[193] Lengauer, C., Kinzler, K. W., Vogelstein, B., Genetic instabilities in human cancers. *Nature* 1998, *396*, 643-649.

[194] Jones, P. A., Baylin, S. B., The fundamental role of epigenetic events in cancer. *Nat Rev Genet* 2002, *3*, 415-428.
[195] Rao, J. S., Molecular mechanisms of glioma invasiveness: the role of proteases. *Nature reviews* 2003, *3*, 489-501.
[196] Puente, X. S., Sanchez, L. M., Overall, C. M., Lopez-Otin, C., Human and mouse proteases: a comparative genomic approach. *Nat Rev Genet* 2003, *4*, 544-558.
[197] Werb, Z., ECM and cell surface proteolysis: regulating cellular ecology. *Cell* 1997, *91*, 439-442.
[198] Gomez, D. E., Alonso, D. F., Yoshiji, H., Thorgeirsson, U. P., Tissue inhibitors of metalloproteinases: structure, regulation and biological functions. *European journal of cell biology* 1997, *74*, 111-122.
[199] Wilson, C. L., Heppner, K. J., Labosky, P. A., Hogan, B. L., Matrisian, L. M., Intestinal tumorigenesis is suppressed in mice lacking the metalloproteinase matrilysin. *Proceedings of the National Academy of Sciences of the United States of America* 1997, *94*, 1402-1407.
[200] Masson, R., Lefebvre, O., Noel, A., Fahime, M. E., *et al.*, In vivo evidence that the stromelysin-3 metalloproteinase contributes in a paracrine manner to epithelial cell malignancy. *The Journal of cell biology* 1998, *140*, 1535-1541.
[201] Itoh, T., Tanioka, M., Matsuda, H., Nishimoto, H., *et al.*, Experimental metastasis is suppressed in MMP-9-deficient mice. *Clinical & experimental metastasis* 1999, *17*, 177-181.
[202] Itoh, T., Tanioka, M., Yoshida, H., Yoshioka, T., *et al.*, Reduced angiogenesis and tumor progression in gelatinase A-deficient mice. *Cancer research* 1998, *58*, 1048-1051.
[203] Bergers, G., Brekken, R., McMahon, G., Vu, T. H., *et al.*, Matrix metalloproteinase-9 triggers the angiogenic switch during carcinogenesis. *Nature cell biology* 2000, *2*, 737-744.
[204] Chandrasekar, N., Jasti, S., Alfred-Yung, W. K., Ali-Osman, F., *et al.*, Modulation of endothelial cell morphogenesis in vitro by MMP-9 during glial-endothelial cell interactions. *Clinical & experimental metastasis* 2000, *18*, 337-342.
[205] Sun, Y., Cheung, J. M., Martel-Pelletier, J., Pelletier, J. P., *et al.*, Wild type and mutant p53 differentially regulate the gene expression of human collagenase-3 (hMMP-13). *The Journal of biological chemistry* 2000, *275*, 11327-11332.
[206] Spiess, E., Bruning, A., Gack, S., Ulbricht, B., *et al.*, Cathepsin B activity in human lung tumor cell lines: ultrastructural localization, pH sensitivity, and inhibitor status at the cellular level. *J Histochem Cytochem* 1994, *42*, 917-929.
[207] Kobayashi, H., Ohi, H., Sugimura, M., Shinohara, H., *et al.*, Inhibition of in vitro ovarian cancer cell invasion by modulation of urokinase-type plasminogen activator and cathepsin B. *Cancer research* 1992, *52*, 3610-3614.
[208] Koblinski, J. E., Dosescu, J., Sameni, M., Moin, K., *et al.*, Interaction of human breast fibroblasts with collagen I increases secretion of procathepsin B. *The Journal of biological chemistry* 2002, *277*, 32220-32227.
[209] Somanna, A., Mundodi, V., Gedamu, L., Functional analysis of cathepsin B-like cysteine proteases from Leishmania donovani complex. Evidence for the activation of latent transforming growth factor beta. *The Journal of biological chemistry* 2002, *277*, 25305-25312.
[210] Eeckhout, Y., Vaes, G., Further studies on the activation of procollagenase, the latent precursor of bone collagenase. Effects of lysosomal cathepsin B, plasmin and kallikrein, and spontaneous activation. *The Biochemical journal* 1977, *166*, 21-31.
[211] Kostoulas, G., Lang, A., Nagase, H., Baici, A., Stimulation of angiogenesis through cathepsin B inactivation of the tissue inhibitors of matrix metalloproteinases. *FEBS letters* 1999, *455*, 286-290.
[212] Lin, A. H., Eliceiri, B. P., Levin, E. G., FAK mediates the inhibition of glioma cell migration by truncated 24kDa FGF-2. *Biochemical and biophysical research communications* 2009.
[213] Guarino, M., Rubino, B., Ballabio, G., The role of epithelial-mesenchymal transition in cancer pathology. *Pathology* 2007, *39*, 305-318.

[214] Xie, L., Law, B. K., Chytil, A. M., Brown, K. A., *et al.*, Activation of the Erk pathway is required for TGF-beta1-induced EMT in vitro. *Neoplasia (New York, N.Y* 2004, *6*, 603-610.
[215] Kuphal, S., Bosserhoff, A. K., Influence of the cytoplasmic domain of E-cadherin on endogenous N-cadherin expression in malignant melanoma. *Oncogene* 2006, *25*, 248-259.
[216] Dandachi, N., Hauser-Kronberger, C., More, E., Wiesener, B., *et al.*, Co-expression of tenascin-C and vimentin in human breast cancer cells indicates phenotypic transdifferentiation during tumour progression: correlation with histopathological parameters, hormone receptors, and oncoproteins. *The Journal of pathology* 2001, *193*, 181-189.
[217] El-Tanani, M. K., Role of osteopontin in cellular signaling and metastatic phenotype. *Front Biosci* 2008, *13*, 4276-4284.
[218] Hsu, E. L., Chen, N., Westbrook, A., Wang, F., *et al.*, Modulation of CXCR4, CXCL12, and Tumor Cell Invasion Potential In Vitro by Phytochemicals. *Journal of oncology* 2009, *2009*, 491985.
[219] Nyberg, P., Salo, T., Kalluri, R., Tumor microenvironment and angiogenesis. *Front Biosci* 2008, *13*, 6537-6553.
[220] Gupta, M. K., Qin, R. Y., Mechanism and its regulation of tumor-induced angiogenesis. *World J Gastroenterol* 2003, *9*, 1144-1155.
[221] Ferrara, N., Kerbel, R. S., Angiogenesis as a therapeutic target. *Nature* 2005, *438*, 967-974.
[222] Zhang, H., Molecular signaling and genetic pathways of senescence: Its role in tumorigenesis and aging. *Journal of cellular physiology* 2007, *210*, 567-574.
[223] Sherwood, S. W., Rush, D., Ellsworth, J. L., Schimke, R. T., Defining cellular senescence in IMR-90 cells: a flow cytometric analysis. *Proceedings of the National Academy of Sciences of the United States of America* 1988, *85*, 9086-9090.
[224] Matsumura, T., Zerrudo, Z., Hayflick, L., Senescent human diploid cells in culture: survival, DNA synthesis and morphology. *Journal of gerontology* 1979, *34*, 328-334.
[225] Hampel, B., Wagner, M., Teis, D., Zwerschke, W., *et al.*, Apoptosis resistance of senescent human fibroblasts is correlated with the absence of nuclear IGFBP-3. *Aging cell* 2005, *4*, 325-330.
[226] Wei, W., Herbig, U., Wei, S., Dutriaux, A., Sedivy, J. M., Loss of retinoblastoma but not p16 function allows bypass of replicative senescence in human fibroblasts. *EMBO reports* 2003, *4*, 1061-1066.
[227] Beausejour, C. M., Krtolica, A., Galimi, F., Narita, M., *et al.*, Reversal of human cellular senescence: roles of the p53 and p16 pathways. *The EMBO journal* 2003, *22*, 4212-4222.
[228] Sherr, C. J., Roberts, J. M., CDK inhibitors: positive and negative regulators of G1-phase progression. *Genes & development* 1999, *13*, 1501-1512.
[229] McConnell, B. B., Starborg, M., Brookes, S., Peters, G., Inhibitors of cyclin-dependent kinases induce features of replicative senescence in early passage human diploid fibroblasts. *Curr Biol* 1998, *8*, 351-354.
[230] Chang, B. D., Watanabe, K., Broude, E. V., Fang, J., *et al.*, Effects of p21Waf1/Cip1/Sdi1 on cellular gene expression: implications for carcinogenesis, senescence, and age-related diseases. *Proceedings of the National Academy of Sciences of the United States of America* 2000, *97*, 4291-4296.
[231] Brown, J. P., Wei, W., Sedivy, J. M., Bypass of senescence after disruption of p21CIP1/WAF1 gene in normal diploid human fibroblasts. *Science (New York, N.Y* 1997, *277*, 831-834.
[232] Cristofalo, V. J., Pignolo, R. J., Molecular markers of senescence in fibroblast-like cultures. *Experimental gerontology* 1996, *31*, 111-123.
[233] Ferber, A., Chang, C., Sell, C., Ptasznik, A., *et al.*, Failure of senescent human fibroblasts to express the insulin-like growth factor-1 gene. *The Journal of biological chemistry* 1993, *268*, 17883-17888.
[234] Goldstein, S., Moerman, E. J., Jones, R. A., Baxter, R. C., Insulin-like growth factor binding protein 3 accumulates to high levels in culture medium of senescent and quiescent human

fibroblasts. *Proceedings of the National Academy of Sciences of the United States of America* 1991, *88*, 9680-9684.
[235] Goldstein, S., Moerman, E. J., Fujii, S., Sobel, B. E., Overexpression of plasminogen activator inhibitor type-1 in senescent fibroblasts from normal subjects and those with Werner syndrome. *Journal of cellular physiology* 1994, *161*, 571-579.
[236] Kuilman, T., Peeper, D. S., Senescence-messaging secretome: SMS-ing cellular stress. *Nature reviews* 2009, *9*, 81-94.
[237] Firth, S. M., Baxter, R. C., Cellular actions of the insulin-like growth factor binding proteins. *Endocrine reviews* 2002, *23*, 824-854.
[238] Wajapeyee, N., Serra, R. W., Zhu, X., Mahalingam, M., Green, M. R., Oncogenic BRAF induces senescence and apoptosis through pathways mediated by the secreted protein IGFBP7. *Cell* 2008, *132*, 363-374.
[239] Chenau, J., Michelland, S., Seve, M., [Secretome: Definitions and biomedical interest.]. *La Revue de medecine interne / fondee* .. 2008, *29*, 606-608.
[240] Volmer, M. W., Stuhler, K., Zapatka, M., Schoneck, A., *et al.*, Differential proteome analysis of conditioned media to detect Smad4 regulated secreted biomarkers in colon cancer. *Proteomics* 2005, *5*, 2587-2601.
[241] Tjalsma, H., Bolhuis, A., Jongbloed, J. D., Bron, S., van Dijl, J. M., Signal peptide-dependent protein transport in Bacillus subtilis: a genome-based survey of the secretome. *Microbiol Mol Biol Rev* 2000, *64*, 515-547.
[242] Dupont, A., Corseaux, D., Dekeyzer, O., Drobecq, H., *et al.*, The proteome and secretome of human arterial smooth muscle cells. *Proteomics* 2005, *5*, 585-596.
[243] Askonas, B. A., Immunoglobulin formation in B lymphoid cells. *Journal of clinical pathology* 1975, *6*, 8-12.
[244] Nickel, W., The mystery of nonclassical protein secretion. A current view on cargo proteins and potential export routes. *European journal of biochemistry / FEBS* 2003, *270*, 2109-2119.
[245] Pisitkun, T., Shen, R. F., Knepper, M. A., Identification and proteomic profiling of exosomes in human urine. *Proceedings of the National Academy of Sciences of the United States of America* 2004, *101*, 13368-13373.
[246] Bendtsen, J. D., Nielsen, H., von Heijne, G., Brunak, S., Improved prediction of signal peptides: SignalP 3.0. *Journal of molecular biology* 2004, *340*, 783-795.
[247] Bendtsen, J. D., Jensen, L. J., Blom, N., Von Heijne, G., Brunak, S., Feature-based prediction of non-classical and leaderless protein secretion. *Protein Eng Des Sel* 2004, *17*, 349-356.
[248] Bendtsen, J. D., Kiemer, L., Fausboll, A., Brunak, S., Non-classical protein secretion in bacteria. *BMC microbiology* 2005, *5*, 58.
[249] Altschul, S. F., Gish, W., Miller, W., Myers, E. W., Lipman, D. J., Basic local alignment search tool. *Journal of molecular biology* 1990, *215*, 403-410.
[250] Klee, E. W., Carlson, D. F., Fahrenkrug, S. C., Ekker, S. C., Ellis, L. B., Identifying secretomes in people, pufferfish and pigs. *Nucleic acids research* 2004, *32*, 1414-1421.
[251] Gronborg, M., Kristiansen, T. Z., Iwahori, A., Chang, R., *et al.*, Biomarker discovery from pancreatic cancer secretome using a differential proteomic approach. *Mol Cell Proteomics* 2006, *5*, 157-171.
[252] Martin, D. B., Gifford, D. R., Wright, M. E., Keller, A., *et al.*, Quantitative proteomic analysis of proteins released by neoplastic prostate epithelium. *Cancer research* 2004, *64*, 347-355.
[253] Wu, W. W., Wang, G., Baek, S. J., Shen, R. F., Comparative study of three proteomic quantitative methods, DIGE, cICAT, and iTRAQ, using 2D gel- or LC-MALDI TOF/TOF. *Journal of proteome research* 2006, *5*, 651-658.
[254] Anderson, N. L., Polanski, M., Pieper, R., Gatlin, T., *et al.*, The human plasma proteome: a nonredundant list developed by combination of four separate sources. *Mol Cell Proteomics* 2004, *3*, 311-326.

[255] Oesterling, J. E., Prostate-specific antigen: a valuable clinical tool. *Oncology (Williston Park, N.Y* 1991, *5*, 107-122; discussion 122, 125-106, 128.
[256] Xue, H., Lu, B., Lai, M., The cancer secretome: a reservoir of biomarkers. *Journal of translational medicine* 2008, *6*, 52.
[257] Huang, C. M., Ananthaswamy, H. N., Barnes, S., Ma, Y., *et al.*, Mass spectrometric proteomics profiles of in vivo tumor secretomes: capillary ultrafiltration sampling of regressive tumor masses. *Proteomics* 2006, *6*, 6107-6116.
[258] Wrotnoski, C., The Future of Plasma Proteins. *Genetic engineering news* 1998, *18*, 4.
[259] Omenn, G. S., States, D. J., Adamski, M., Blackwell, T. W., *et al.*, Overview of the HUPO Plasma Proteome Project: results from the pilot phase with 35 collaborating laboratories and multiple analytical groups, generating a core dataset of 3020 proteins and a publicly-available database. *Proteomics* 2005, *5*, 3226-3245.
[260] Schenk, S., Schoenhals, G. J., de Souza, G., Mann, M., A high confidence, manually validated human blood plasma protein reference set. *BMC medical genomics* 2008, *1*, 41.
[261] Hurbin, A., Dubrez, L., Coll, J. L., Favrot, M. C., Inhibition of apoptosis by amphiregulin via an insulin-like growth factor-1 receptor-dependent pathway in non-small cell lung cancer cell lines. *Annals of the New York Academy of Sciences* 2003, *1010*, 354-357.
[262] Hurbin, A., Coll, J. L., Dubrez-Daloz, L., Mari, B., *et al.*, Cooperation of amphiregulin and insulin-like growth factor-1 inhibits Bax- and Bad-mediated apoptosis via a protein kinase C-dependent pathway in non-small cell lung cancer cells. *The Journal of biological chemistry* 2005, *280*, 19757-19767.
[263] Dubrez, L., Coll, J. L., Hurbin, A., Solary, E., Favrot, M. C., Caffeine sensitizes human H358 cell line to p53-mediated apoptosis by inducing mitochondrial translocation and conformational change of BAX protein. *The Journal of biological chemistry* 2001, *276*, 38980-38987.
[264] Dubrez, L., Coll, J. L., Hurbin, A., de Fraipont, F., *et al.*, Cell cycle arrest is sufficient for p53-mediated tumor regression. *Gene therapy* 2001, *8*, 1705-1712.
[265] Antonov, A. V., Dietmann, S., Wong, P., Igor, R., Mewes, H. W., PLIPS, an Automatically Collected Database of Protein Lists Reported by Proteomics Studies. *Journal of proteome research* 2009.
[266] Bjorhall, K., Miliotis, T., Davidsson, P., Comparison of different depletion strategies for improved resolution in proteomic analysis of human serum samples. *Proteomics* 2005, *5*, 307-317.
[267] Yocum, A. K., Yu, K., Oe, T., Blair, I. A., Effect of immunoaffinity depletion of human serum during proteomic investigations. *Journal of proteome research* 2005, *4*, 1722-1731.
[268] Righetti, P. G., Boschetti, E., Lomas, L., Citterio, A., Protein Equalizer Technology : the quest for a "democratic proteome". *Proteomics* 2006, *6*, 3980-3992.
[269] Bellahcene, A., Castronovo, V., Ogbureke, K. U., Fisher, L. W., Fedarko, N. S., Small integrin-binding ligand N-linked glycoproteins (SIBLINGs): multifunctional proteins in cancer. *Nature reviews* 2008, *8*, 212-226.
[270] Mann, B., Gelos, M., Siedow, A., Hanski, M. L., *et al.*, Target genes of beta-catenin-T cell-factor/lymphoid-enhancer-factor signaling in human colorectal carcinomas. *Proceedings of the National Academy of Sciences of the United States of America* 1999, *96*, 1603-1608.
[271] Crawford, H. C., Fingleton, B. M., Rudolph-Owen, L. A., Goss, K. J., *et al.*, The metalloproteinase matrilysin is a target of beta-catenin transactivation in intestinal tumors. *Oncogene* 1999, *18*, 2883-2891.
[272] Wu, B., Crampton, S. P., Hughes, C. C., Wnt signaling induces matrix metalloproteinase expression and regulates T cell transmigration. *Immunity* 2007, *26*, 227-239.
[273] Shimokawa, T., Furukawa, Y., Sakai, M., Li, M., *et al.*, Involvement of the FGF18 gene in colorectal carcinogenesis, as a novel downstream target of the beta-catenin/T-cell factor complex. *Cancer research* 2003, *63*, 6116-6120.

[274] Chamorro, M. N., Schwartz, D. R., Vonica, A., Brivanlou, A. H., *et al.*, FGF-20 and DKK1 are transcriptional targets of beta-catenin and FGF-20 is implicated in cancer and development. *The EMBO journal* 2005, *24*, 73-84.
[275] Zhang, X., Gaspard, J. P., Chung, D. C., Regulation of vascular endothelial growth factor by the Wnt and K-ras pathways in colonic neoplasia. *Cancer research* 2001, *61*, 6050-6054.
[276] Masckauchan, T. N., Shawber, C. J., Funahashi, Y., Li, C. M., Kitajewski, J., Wnt/beta-catenin signaling induces proliferation, survival and interleukin-8 in human endothelial cells. *Angiogenesis* 2005, *8*, 43-51.
[277] Jamora, C., DasGupta, R., Kocieniewski, P., Fuchs, E., Links between signal transduction, transcription and adhesion in epithelial bud development. *Nature* 2003, *422*, 317-322.
[278] Hu, D. E., Hori, Y., Fan, T. P., Interleukin-8 stimulates angiogenesis in rats. *Inflammation* 1993, *17*, 135-143.
[279] Kwon, Y. W., Kwon, K. S., Moon, H. E., Park, J. A., *et al.*, Insulin-like growth factor-II regulates the expression of vascular endothelial growth factor by the human keratinocyte cell line HaCaT. *The Journal of investigative dermatology* 2004, *123*, 152-158.
[280] Borgono, C. A., Diamandis, E. P., The emerging roles of human tissue kallikreins in cancer. *Nature reviews* 2004, *4*, 876-890.
[281] Diamandis, E. P., Yousef, G. M., Olsson, A. Y., An update on human and mouse glandular kallikreins. *Clinical biochemistry* 2004, *37*, 258-260.
[282] Yousef, G. M., Diamandis, E. P., An overview of the kallikrein gene families in humans and other species: emerging candidate tumour markers. *Clinical biochemistry* 2003, *36*, 443-452.
[283] Paliouras, M., Borgono, C., Diamandis, E. P., Human tissue kallikreins: the cancer biomarker family. *Cancer letters* 2007, *249*, 61-79.
[284] Diamandis, E. P., Scorilas, A., Fracchioli, S., Van Gramberen, M., *et al.*, Human kallikrein 6 (hK6): a new potential serum biomarker for diagnosis and prognosis of ovarian carcinoma. *J Clin Oncol* 2003, *21*, 1035-1043.
[285] Santin, A. D., Diamandis, E. P., Bellone, S., Soosaipillai, A., *et al.*, Human kallikrein 6: a new potential serum biomarker for uterine serous papillary cancer. *Clin Cancer Res* 2005, *11*, 3320-3325.
[286] Nagahara, H., Mimori, K., Utsunomiya, T., Barnard, G. F., *et al.*, Clinicopathologic and biological significance of kallikrein 6 overexpression in human gastric cancer. *Clin Cancer Res* 2005, *11*, 6800-6806.
[287] Ogawa, K., Utsunomiya, T., Mimori, K., Tanaka, F., *et al.*, Clinical significance of human kallikrein gene 6 messenger RNA expression in colorectal cancer. *Clin Cancer Res* 2005, *11*, 2889-2893.
[288] Ghosh, M. C., Grass, L., Soosaipillai, A., Sotiropoulou, G., Diamandis, E. P., Human kallikrein 6 degrades extracellular matrix proteins and may enhance the metastatic potential of tumour cells. *Tumour Biol* 2004, *25*, 193-199.
[289] Prezas, P., Arlt, M. J., Viktorov, P., Soosaipillai, A., *et al.*, Overexpression of the human tissue kallikrein genes KLK4, 5, 6, and 7 increases the malignant phenotype of ovarian cancer cells. *Biological chemistry* 2006, *387*, 807-811.
[290] Wolf, W. C., Evans, D. M., Chao, L., Chao, J., A synthetic tissue kallikrein inhibitor suppresses cancer cell invasiveness. *The American journal of pathology* 2001, *159*, 1797-1805.
[291] Cohen, P., Graves, H. C., Peehl, D. M., Kamarei, M., *et al.*, Prostate-specific antigen (PSA) is an insulin-like growth factor binding protein-3 protease found in seminal plasma. *The Journal of clinical endocrinology and metabolism* 1992, *75*, 1046-1053.
[292] Rehault, S., Monget, P., Mazerbourg, S., Tremblay, R., *et al.*, Insulin-like growth factor binding proteins (IGFBPs) as potential physiological substrates for human kallikreins hK2 and hK3. *European journal of biochemistry / FEBS* 2001, *268*, 2960-2968.
[293] Sutkowski, D. M., Goode, R. L., Baniel, J., Teater, C., *et al.*, Growth regulation of prostatic stromal cells by prostate-specific antigen. *Journal of the National Cancer Institute* 1999, *91*, 1663-1669.

[294] Bhattacharjee, A., Richards, W. G., Staunton, J., Li, C., *et al.*, Classification of human lung carcinomas by mRNA expression profiling reveals distinct adenocarcinoma subclasses. *Proceedings of the National Academy of Sciences of the United States of America* 2001, *98*, 13790-13795.
[295] Henkhaus, R. S., Roy, U. K., Cavallo-Medved, D., Sloane, B. F., *et al.*, Caveolin-1-mediated expression and secretion of kallikrein 6 in colon cancer cells. *Neoplasia (New York, N.Y* 2008, *10*, 140-148.
[296] Ornitz, D. M., Itoh, N., Fibroblast growth factors. *Genome biology* 2001, *2*, REVIEWS3005.
[297] Powers, C. J., McLeskey, S. W., Wellstein, A., Fibroblast growth factors, their receptors and signaling. *Endocrine-related cancer* 2000, *7*, 165-197.
[298] Xie, M. H., Holcomb, I., Deuel, B., Dowd, P., *et al.*, FGF-19, a novel fibroblast growth factor with unique specificity for FGFR4. *Cytokine* 1999, *11*, 729-735.
[299] Nicholes, K., Guillet, S., Tomlinson, E., Hillan, K., *et al.*, A mouse model of hepatocellular carcinoma: ectopic expression of fibroblast growth factor 19 in skeletal muscle of transgenic mice. *The American journal of pathology* 2002, *160*, 2295-2307.
[300] Desnoyers, L. R., Pai, R., Ferrando, R. E., Hotzel, K., *et al.*, Targeting FGF19 inhibits tumor growth in colon cancer xenograft and FGF19 transgenic hepatocellular carcinoma models. *Oncogene* 2008, *27*, 85-97.
[301] Pai, R., Dunlap, D., Qing, J., Mohtashemi, I., *et al.*, Inhibition of fibroblast growth factor 19 reduces tumor growth by modulating beta-catenin signaling. *Cancer research* 2008, *68*, 5086-5095.
[302] Abuharbeid, S., Czubayko, F., Aigner, A., The fibroblast growth factor-binding protein FGF-BP. *The international journal of biochemistry & cell biology* 2006, *38*, 1463-1468.
[303] Liaudet-Coopman, E. D., Berchem, G. J., Wellstein, A., In vivo inhibition of angiogenesis and induction of apoptosis by retinoic acid in squamous cell carcinoma. *Clin Cancer Res* 1997, *3*, 179-184.
[304] Czubayko, F., Liaudet-Coopman, E. D., Aigner, A., Tuveson, A. T., *et al.*, A secreted FGF-binding protein can serve as the angiogenic switch in human cancer. *Nature medicine* 1997, *3*, 1137-1140.
[305] Sherif, Z. A., Nakai, S., Pirollo, K. F., Rait, A., Chang, E. H., Downmodulation of bFGF-binding protein expression following restoration of p53 function. *Cancer gene therapy* 2001, *8*, 771-782.
[306] Ray, R., Cabal-Manzano, R., Moser, A. R., Waldman, T., *et al.*, Up-regulation of fibroblast growth factor-binding protein, by beta-catenin during colon carcinogenesis. *Cancer research* 2003, *63*, 8085-8089.
[307] Tassi, E., Al-Attar, A., Aigner, A., Swift, M. R., *et al.*, Enhancement of fibroblast growth factor (FGF) activity by an FGF-binding protein. *The Journal of biological chemistry* 2001, *276*, 40247-40253.
[308] Eide, D. J., The SLC39 family of metal ion transporters. *Pflugers Arch* 2004, *447*, 796-800.
[309] Prasad, A. S., Kucuk, O., Zinc in cancer prevention. *Cancer metastasis reviews* 2002, *21*, 291-295.
[310] Costello, L. C., Franklin, R. B., The clinical relevance of the metabolism of prostate cancer; zinc and tumor suppression: connecting the dots. *Molecular cancer* 2006, *5*, 17.
[311] Tashiro, H., Kawamoto, T., Okubo, T., Koide, O., Variation in the distribution of trace elements in hepatoma. *Biological trace element research* 2003, *95*, 49-63.
[312] Manning, D. L., Robertson, J. F., Ellis, I. O., Elston, C. W., *et al.*, Oestrogen-regulated genes in breast cancer: association of pLIV1 with lymph node involvement. *Eur J Cancer* 1994, *30A*, 675-678.
[313] Kagara, N., Tanaka, N., Noguchi, S., Hirano, T., Zinc and its transporter ZIP10 are involved in invasive behavior of breast cancer cells. *Cancer science* 2007, *98*, 692-697.

[314] Zhao, L., Chen, W., Taylor, K. M., Cai, B., Li, X., LIV-1 suppression inhibits HeLa cell invasion by targeting ERK1/2-Snail/Slug pathway. *Biochemical and biophysical research communications* 2007, *363*, 82-88.
[315] Eling, T. E., Baek, S. J., Shim, M., Lee, C. H., NSAID activated gene (NAG-1), a modulator of tumorigenesis. *Journal of biochemistry and molecular biology* 2006, *39*, 649-655.
[316] Baek, S. J., Kim, K. S., Nixon, J. B., Wilson, L. C., Eling, T. E., Cyclooxygenase inhibitors regulate the expression of a TGF-beta superfamily member that has proapoptotic and antitumorigenic activities. *Molecular pharmacology* 2001, *59*, 901-908.
[317] Baek, S. J., Wilson, L. C., Eling, T. E., Resveratrol enhances the expression of non-steroidal anti-inflammatory drug-activated gene (NAG-1) by increasing the expression of p53. *Carcinogenesis* 2002, *23*, 425-434.
[318] Baek, S. J., Kim, J. S., Moore, S. M., Lee, S. H., *et al.*, Cyclooxygenase inhibitors induce the expression of the tumor suppressor gene EGR-1, which results in the up-regulation of NAG-1, an antitumorigenic protein. *Molecular pharmacology* 2005, *67*, 356-364.
[319] Yamaguchi, K., Lee, S. H., Eling, T. E., Baek, S. J., Identification of nonsteroidal anti-inflammatory drug-activated gene (NAG-1) as a novel downstream target of phosphatidylinositol 3-kinase/AKT/GSK-3beta pathway. *The Journal of biological chemistry* 2004, *279*, 49617-49623.
[320] Gonzalez, M. A., Tachibana, K. E., Laskey, R. A., Coleman, N., Control of DNA replication and its potential clinical exploitation. *Nature reviews* 2005, *5*, 135-141.
[321] Stoeber, K., Tlsty, T. D., Happerfield, L., Thomas, G. A., *et al.*, DNA replication licensing and human cell proliferation. *Journal of cell science* 2001, *114*, 2027-2041.
[322] Ramnath, N., Hernandez, F. J., Tan, D. F., Huberman, J. A., *et al.*, MCM2 is an independent predictor of survival in patients with non-small-cell lung cancer. *J Clin Oncol* 2001, *19*, 4259-4266.
[323] Hashimoto, K., Araki, K., Osaki, M., Nakamura, H., *et al.*, MCM2 and Ki-67 expression in human lung adenocarcinoma: prognostic implications. *Pathobiology* 2004, *71*, 193-200.
[324] Guida, T., Salvatore, G., Faviana, P., Giannini, R., *et al.*, Mitogenic effects of the up-regulation of minichromosome maintenance proteins in anaplastic thyroid carcinoma. *The Journal of clinical endocrinology and metabolism* 2005, *90*, 4703-4709.
[325] Choi, W., Cogdell, D., Feng, Y., Hamilton, S. R., Zhang, W., Transcriptional activation of the carboxylesterase 2 gene by the p53 pathway. *Cancer biology & therapy* 2006, *5*, 1450-1456.
[326] Tang, X., Wu, H., Wu, Z., Wang, G., *et al.*, Carboxylesterase 2 is downregulated in colorectal cancer following progression of the disease. *Cancer investigation* 2008, *26*, 178-181.
[327] Liu, G., Chen, X., The ferredoxin reductase gene is regulated by the p53 family and sensitizes cells to oxidative stress-induced apoptosis. *Oncogene* 2002, *21*, 7195-7204.
[328] Hwang, P. M., Bunz, F., Yu, J., Rago, C., *et al.*, Ferredoxin reductase affects p53-dependent, 5-fluorouracil-induced apoptosis in colorectal cancer cells. *Nature medicine* 2001, *7*, 1111-1117.
[329] Ichikawa, W., Ooyama, A., Toda, E., Sugimoto, Y., *et al.*, Gene expression of ferredoxin reductase predicts outcome in patients with metastatic colorectal cancer treated by 5-fluorouracil plus leucovorin. *Cancer chemotherapy and pharmacology* 2006, *58*, 794-801.
[330] Tanaka, H., Arakawa, H., Yamaguchi, T., Shiraishi, K., *et al.*, A ribonucleotide reductase gene involved in a p53-dependent cell-cycle checkpoint for DNA damage. *Nature* 2000, *404*, 42-49.
[331] Yanamoto, S., Kawasaki, G., Yoshitomi, I., Mizuno, A., Expression of p53R2, newly p53 target in oral normal epithelium, epithelial dysplasia and squamous cell carcinoma. *Cancer letters* 2003, *190*, 233-243.
[332] Okumura, H., Natsugoe, S., Matsumoto, M., Mataki, Y., *et al.*, The predictive value of p53, p53R2, and p21 for the effect of chemoradiation therapy on oesophageal squamous cell carcinoma. *British journal of cancer* 2005, *92*, 284-289.

[333] Liu, X., Zhou, B., Xue, L., Shih, J., *et al.*, Metastasis-suppressing potential of ribonucleotide reductase small subunit p53R2 in human cancer cells. *Clin Cancer Res* 2006, *12*, 6337-6344.
[334] Yanamoto, S., Kawasaki, G., Yamada, S. I., Yoshitomi, I., *et al.*, Ribonucleotide reductase small subunit p53R2 promotes oral cancer invasion via the E-cadherin/beta-catenin pathway. *Oral oncology* 2008.
[335] Liotta, L. A., Ferrari, M., Petricoin, E., Clinical proteomics: written in blood. *Nature* 2003, *425*, 905.
[336] Zheng, X., Baker, H., Hancock, W. S., Analysis of the low molecular weight serum peptidome using ultrafiltration and a hybrid ion trap-Fourier transform mass spectrometer. *Journal of chromatography* 2006, *1120*, 173-184.
[337] Lin, Y., Dynan, W. S., Lee, J. R., Zhu, Z. H., Schade, R. R., The current state of proteomics in GI oncology. *Digestive diseases and sciences* 2009, *54*, 431-457.
[338] Akhtar, S., Meeran, S. M., Katiyar, N., Katiyar, S. K., Grape seed proanthocyanidins inhibit the growth of human non-small cell lung cancer xenografts by targeting insulin-like growth factor binding protein-3, tumor cell proliferation, and angiogenic factors. *Clin Cancer Res* 2009, *15*, 821-831.
[339] Orazine, C. I., Hincapie, M., Hancock, W. S., Hattersley, M., Hanke, J. H., A proteomic analysis of the plasma glycoproteins of a MCF-7 mouse xenograft: a model system for the detection of tumor markers. *Journal of proteome research* 2008, *7*, 1542-1554.
[340] Beasley, M. B., Brambilla, E., Travis, W. D., The 2004 World Health Organization classification of lung tumors. *Seminars in roentgenology* 2005, *40*, 90-97.
[341] Kulasingam, V., Diamandis, E. P., Tissue culture-based breast cancer biomarker discovery platform. *International journal of cancer* 2008, *123*, 2007-2012.
[342] Puig, O., Caspary, F., Rigaut, G., Rutz, B., *et al.*, The tandem affinity purification (TAP) method: a general procedure of protein complex purification. *Methods (San Diego, Calif* 2001, *24*, 218-229.
[343] Kitteringham, N. R., Jenkins, R. E., Lane, C. S., Elliott, V. L., Park, B. K., Multiple reaction monitoring for quantitative biomarker analysis in proteomics and metabolomics. *Journal of chromatography* 2009, *877*, 1229-1239.
[344] Fortin, T., Salvador, A., Charrier, J. P., Lenz, C., *et al.*, Clinical quantitation of prostate specific antigen biomarker in the low nanogram/mL range by conventional bore liquid chromatography-tandem mass spectrometry (MRM) coupling and correlation with ELISA tests. *Mol Cell Proteomics* 2008.

RESUME

RESUME

Le cancer est un véritable problème de santé publique et représentera en 2010 la première cause de mortalité dans le monde. La lutte contre cette maladie est donc, plus que jamais, un enjeu capital. Les processus tumoraux associés à la carcinogénèse, tels que la métastase, l'invasion ou l'angiogenèse, dépendent étroitement de la composition du milieu extracellulaire, lui-même affecté par la sécrétion de protéines par les cellules tumorales. La mutation ou la variation d'expression de facteurs pro- ou anti-oncogénique joue un rôle important dans le développement et la progression de la tumeur. Durant ce travail de thèse, nous avons voulu apporter des éléments de réponse à la problématique suivante : l'étude par protéomique des protéines sécrétées par la tumeur apporte t'elle des informations contribuant à la compréhension des mécanismes de la tumorogenèse et peut elle permettre de mettre en valeur de nouvelles cibles d'intérêt clinique ? Nous avons ciblé notre étude sur le cancer du poumon, le plus fréquent au niveau mondial tant en terme d'incidence que de mortalité. Ce cancer présente le plus fort taux de mutation du gène p53, un anti-oncogène clé de la cellule. Cette perte de fonction module la sécrétion de nombreuses protéines dont l'investigation est primordiale, bien que paradoxalement, peu étudiée. Notre travail s'est ainsi focalisé sur l'étude de l'influence de p53 sur la modulation de la sécrétion de protéines par les tumeurs du poumon. Nous avons ainsi développé un procédé d'analyse protéomique adapté, basé notamment sur un marquage iTRAQ, une séparation par isoélectrofocalisation de type OFFGEL et une analyse LC-MALDI-MS/MS. Ce procédé a été appliqué à l'étude du rôle joué par l'expression conditionnelle de p53 sur la modulation du secretome d'une lignée cellulaire d'adénocarcinome du cancer du poumon non à petites cellules. Plusieurs protéines d'intérêt ont ainsi été caractérisées et confirmées *in vivo* chez la souris au niveau tumoral et plasmatique. Ces résultats apportent une meilleure compréhension du rôle primordial joué par l'altération de p53 dans la modulation de l'environnement tumoral et font des secretomes cellulaires un modèle de choix pour l'identification de marqueurs tumoraux.

Mots-clés : protéomique, secretome, tumorogenèse, cancer du poumon, p53, xénogreffe.

ABSTRACT

Cancer is a real public health problem and will represent in 2010 the first cause of mortality. So the knowledge of this disease and its mechanisms remains a great challenge. Malignant processes such as metastasis, invasion, or angiogenesis are tightly dependent on the composition of the extracellular medium, which is itself affected by the release of proteins by the tumor cells. The mutation of pro- and anti-oncogenic factors plays an important role in tumor progression. In this work, we want to ask the following major question: is it interesting to study by proteomics the secreted proteins by tumor to better understand tumorigenesis mechanisms and is it useful for clinical biomarker discovery? We focused on lung cancer which is the cancer with the greater mortality and incidency frequences in the world. This cancer has the greater p53 mutation rates. This protein is a major tumor suppressor known to modulate the release of proteins by the tumor cells; however, while p53-modulated intracellular proteins have been extensively studied, little is known concerning their extracellular counterparts. Here, we developed a proteomics process using iTRAQ labeling, OFFGEL isoelectric focusing and LC-MALDI-MS/MS analysis efficient to the study of this class of proteins. We used it to characterize the p53-dependent secretome of a lung tumor model *in vitro*. We identified 909 proteins released by these cells, among which 91 are p53-modulated. This differential secretome analysis promises more detailed of how the loss of wt-p53 contributes to tumorigenic processes through the modulation of secreted proteins. We also demonstrated that the modulation of exported proteins can be detected *in vivo* in total protein extracts and plasma of tumor-bearing mice. This is the first report establishing that the *in vitro* cell line secretome is reliable and reflects the extracellular release of proteins from tumor cells *in vivo* and could be used to identify putative tumor markers.

Keywords: proteomics, secretome, tumorigenesis, lung cancer, p53 xenograft mouse model.