

**UNIVERSITE DES SCIENCES ET TECHNOLOGIES DE LILLE (FRANCE)**  
Ecole doctorale Biologie-Santé

**UNIVERSITE DE SHERBROOKE (CANADA)**  
Département de Biochimie

## **THESE DE DOCTORAT**

En vue de l'obtention du grade de Docteur en Science  
de l'Université des Sciences et Technologies de Lille et de l'Université de Sherbrooke

Présentée par

**DAVID BONNEL**

**Développements et applications d'approches protéomiques pour la  
recherche de cibles du cancer de l'ovaire et de la prostate**

*Thèse sous la cotutelle internationale  
du Pr Michel SALZET et du Pr Robert DAY.*

Soutenue à Lille le 20 Octobre 2010

Devant le jury composé de :

**Pr Isabelle FOURNIER**

**Présidente**

**Dr Markus STOECKLI**

**Rapporteur**

**Pr Claude LAZURE**

**Rapporteur**

**Pr Denis VINATIER**

**Examineur**

**Pr Robert DAY**

**Co-directeur de Thèse**

**Pr Michel SALZET**

**Co-directeur de Thèse**



À ma femme, Céline.

*« La Science n'a pas de patrie »  
Louis Pasteur*



## Remerciements

---

J'ai eu le privilège d'effectuer cette thèse en cotutelle internationale entre les laboratoires du Pr Salzet de l'Université de Lille 1 en France et du Pr Day de l'Université de Sherbrooke au Québec.

C'est pourquoi, je souhaite tout d'abord exprimer ma plus profonde reconnaissance et mon plus profond respect envers mes deux directeurs de thèse, les professeurs Michel Salzet et Robert Day, et envers ma chef d'équipe en France la professeure Isabelle Fournier. Je les remercie pour toute la confiance qu'ils m'ont accordée et les enseignements d'une très grande qualité qu'ils m'ont apportés. Leur encadrement, leur passion de la Science et leurs connaissances ont nourri ma motivation tout au long de ces années de thèse.

Je tiens à remercier mes deux rapporteurs, le professeur Claude Lazure de l'institut de recherches cliniques de Montréal et le Docteur Markus Stoeckli du campus Novartis de Bâle pour avoir accepté d'évaluer ce travail de thèse.

Je remercie tous les membres de mon laboratoire d'accueil français, et plus particulièrement l'équipe d'imagerie MALDI : Maxence, Annie, Mohamed, Céline, Rémi, Céline et Laurent (les membres actuels) ; Rémi, Jonathan, Olivia, Julien, Karim, Amélie, Benjamin et Thomas (les anciens) avec lesquels j'ai partagé cette aventure.

De la même manière, je remercie les membres de mon laboratoire d'accueil québécois : Roxane, Sophie, Sandra, Hugo, Anna, Sarah, Kevin, Christine, Frédéric, Philippe et Xue Wen pour leur accueil chaleureux et leur sympathie. Je réserve un merci tout particulier au Dr François D'Anjou pour son amitié, son aide sans faille et ses conseils.

Je tiens à exprimer ma gratitude au Pr Morad Roudbaraki avec lequel j'ai eu le plaisir de collaborer. Je tiens également à remercier le Pr Vinatier et son service de l'hôpital Jeanne de Flandre à Lille et le Pr Gosset de l'hôpital saint Vincent de Paul à Lille.

Je remercie également les membres du jury de mon examen pré-doctoral à l'Université de Sherbrooke, soit les Professeurs Claude Lazure, Xavier Roucoux et François Corbin.

Je remercie la Région Nord-Pas de Calais pour les financements qui m'ont été attribués durant ma thèse, à savoir une allocation de recherche et une bourse régionale de mobilité à l'international. J'y associe l'école doctorale Biologie Santé de Lille et l'Université des Sciences et Technologies de Lille, ainsi que le département de Biochimie de Sherbrooke et l'Université de Sherbrooke.

Je remercie particulièrement mes parents Véronique et Bernard, ma sœur Séverine et toute ma famille pour leur disponibilité et leurs encouragements. Je pense également à mes amis les plus proches Rodrigue, Gaëlle, Julien, Aurélie, Mickaël, Toun, Vincent et Dominique pour leur soutien.

Enfin, je remercie de tout mon cœur ma femme, Céline, pour son courage, sa compréhension, son soutien au quotidien et son aide. Sans elle à mes côtés, rien n'aurait été possible.

Au cours de la thèse, nous avons dans un premier temps développé une approche de protéomique clinique basée sur une nouvelle méthode d'analyse directe des tissus par spectrométrie de masse *i.e.* l'imagerie MALDI (MSI). L'analyse directe de tissus par MALDI offre l'accès à des centaines de molécules tout en préservant l'intégrité du tissu et leurs régionalisations spatiales au sein de celui-ci. Par automatisation de cette méthode, des images moléculaires de la répartition des biomolécules peuvent être obtenues en une seule étape d'analyse. Cet outil est parfaitement adapté à la recherche de marqueurs au sein de biopsies. Par cette technologie, de nombreuses données moléculaires peuvent être obtenues sur une large gamme de biomolécules (peptides, protéines, lipides, sucres, etc.) à l'échelle cellulaire. Néanmoins ces nouveaux outils génèrent un grand nombre de données. C'est pourquoi nous avons développé et appliqué des outils d'analyses statistiques multivariées, comme les analyses en composantes principales (PCA) et la classification hiérarchique, qui offraient une combinaison intéressante avec l'imagerie MALDI et permettaient une simplification des données, une recherche fine des variations moléculaires et une classification sur la base des profils moléculaires. Cette nouvelle stratégie offre un outil de choix pour la recherche de séries de marqueurs spécifiques d'une pathologie ou d'une sous-pathologie. Il est ainsi possible d'obtenir des marqueurs de chaque type cellulaire et de mettre en évidence les régions d'un même tissu présentant des phénotypes particuliers. Ainsi nous avons pu appliquer cette méthodologie sur des biopsies de cancer d'ovaire, nous permettant d'identifier une trentaine de marqueurs potentiels de cette pathologie comme le fragment C-terminal reg-alpha (activateur de l'immunoprotéasome 11S), la mucine 9, tout comme des oncoprotéines virales de la famille des herpes viridae (HHV4 et HHV6). Ainsi, ces outils d'analyse multivariée permettent d'accéder à une classification des tissus sur une base moléculaire. Il était donc envisageable d'utiliser ces approches pour construire des modèles de classification des différents stades et sous-type de pathologies en recherchant non plus un ou deux marqueurs mais des jeux de marqueurs liés au type cellulaire considéré. Cette stratégie, rapide et non ciblée, a été appliquée et validée sur des échantillons de cancer de l'ovaire et de la prostate, et nous a permis de mettre en évidence certains biomarqueurs que nous avons identifiés, de générer de nouvelles cibles thérapeutiques potentielles et de corrélérer ces résultats avec une approche *in vitro*, tout en ouvrant la voie à une nouvelle méthode de diagnostic basée sur la signature spécifique protéique et/ou peptidique d'un type histologique.

Dans un second temps, nous nous sommes concentrés sur des études *in vitro* ce qui nous a permis de découvrir une nouvelle voie impliquant la proprotéine convertase PACE4 dans le cancer de la prostate. Nous avons donc entamé des expériences préliminaires avec la lignée cellulaire du cancer de la prostate DU145. Ces études ont conduit à démontrer que l'inhibition du gène PACE4 résulte en une prolifération et une progression réduite de la tumeur à la fois *ex vivo* et *in vivo*. Ainsi, la surexpression de certaines proprotéines convertases joue un rôle central dans la progression des tumeurs, au travers de l'activation et de la maturation des protéines associées aux cancers. La spectrométrie de masse est devenue un outil incontournable pour la caractérisation des biomolécules par l'analyse des mélanges complexes, et rapide avec une grande sensibilité renforcée par les possibilités d'analyse structurale. Après avoir utilisé une approche par MALDI-TOF/TOF qui nous a permis d'identifier TRPS1, un facteur de transcription impliqué dans l'apoptose et ainsi de mieux comprendre l'implication de cette enzyme de maturation, nous avons développé une approche de chromatographie sur puces, basée sur la technologie SELDI-TOF pour rechercher les substrats de PACE4 entre la lignée cellulaire DU145 et une lignée knock-down pour PACE4 (DU145/4-2). Ceci nous a permis de fractionner le protéome de chaque lignée et de simplifier son étude.

En conclusion, ces analyses protéomiques nous ont permis premièrement de caractériser de nouveaux biomarqueurs du cancer de la prostate et de l'ovaire, deuxièmement de valider une cible thérapeutique : PACE4 pour le cancer de la prostate et finalement de découvrir une possible étiologie de type viral pour le cancer de l'ovaire.



During the thesis, we initially developed a proteomic approach based on a new clinical method for direct tissue analysis by mass spectrometry imaging MALDI (MSI). Direct tissue analysis by MALDI offers access to hundreds of molecules while preserving tissue integrity and molecule localization. By automating this method, molecular images of the distribution of biomolecules can be obtained in one step of analysis. This tool is perfectly suited for markers discovery through biopsies, reflecting proteome changes associated with pathologies. Through this technology many of molecular data can be obtained over a wide range of biomolecules (peptides, proteins, lipids, sugars, etc.) with their localization at cellular and tissue level. However, these new tools generate a large number of data containing a wealth of information. For this reason, we developed and applied statistical analysis tools such as principal component analysis (PCA) and hierarchical clustering that offers an interesting combination with MALDI-imaging and allows data simplification, fine marker variations research and classification based on molecular profiles. This new strategy provides a tool for finding specific marker sets for a given disease. Thus, it is possible to obtain biomarkers of each cell type and highlight regions with the same phenotype. We have applied this methodology on ovarian cancer biopsies, which allowed us to identify approximately thirty markers of this pathology such as the fragment C-terminal reg-alpha (a fragment of the 11S immunoproteasome), the mucin 9, the viral oncoproteins of the herpes family of viridae (HHV4 and HHV6). Similarly, tools for categorizing, such as hierarchical clustering, allows histological type classification on molecular basis. It was therefore possible to use the method to construct classification models of different disease stages by looking for specific cell types marker sets. This strategy, fast and not-targeted, has been applied and validated in both ovarian and prostate cancer, and allowed us to validate biomarkers already known, to generate new targets and to correlate these results with *in vitro* approaches. This development paves the way for a new diagnostic method base on histologically specific signature. In a second step, we have discovered a new interesting pathway, showing that the proprotein convertase PACE4 has a role in prostate cancer and is selectively overexpressed in these tumors (i.e. other proprotein convertases are not altered). We have therefore begun preliminary experiments with cell line DU145 prostate cancer. These studies revealed that the inhibition of PACE4 gene results in reduced proliferation and progression of the tumor in both *ex vivo* and *in vivo*. Thus, the overexpression of proprotein convertases plays a central role in tumor

progression, through activation and maturation of proteins associated with cancer. Mass spectrometry has become an indispensable tool for biomolecules characterization by the analysis of complex mixtures rapidly with high sensitivity enhanced by the possibilities of structural analysis. Using a MALDI-TOF-TOF approach enabled us to identify TRPS1, a transcription factor involved in apoptosis and thus better understand the involvement of this maturation enzyme, we developed an approach on chromatography chips, based on SELDI-TOF technology to search PACE4 substrates between the DU145 cell line and a PACE4 knockdown line (DU145/4-2). This allowed us to fractionate the proteome of each line and simplify its study.

In conclusion, proteomic analysis allowed us first to characterize new biomarkers of prostate and ovarian cancer, secondly to validate PACE4 as potential target for prostate cancer and finally to explore a possible viral etiology of ovarian cancer.

# Table des matières générale

<b>Remerciements</b> .....	<b>I</b>
<b>Résumé</b> .....	<b>III</b>
<b>Table des matières générale</b> .....	<b>VII</b>
<b>Publications</b> .....	<b>IX</b>
<b>Liste des figures</b> .....	<b>XIII</b>
<b>Liste des tableaux</b> .....	<b>XVII</b>
<b>Liste des abréviations</b> .....	<b>XIX</b>
<b>Chapitre 1 : Introduction</b> .....	<b>1</b>
<b>La cellule : de l'émerveillement à l'inquiétude</b> .....	<b>1</b>
<b>Etudes menées au cours de cette thèse</b> .....	<b>2</b>
<b>Chapitre 2 Généralités</b> .....	<b>5</b>
<b>La Protéomique</b> .....	<b>5</b>
2.1.1 Les techniques de séparation des protéines et des peptides.....	6
2.1.1.1 Méthodes in-gel.....	7
2.1.1.2 Méthode off-gel.....	7
La chromatographie liquide .....	7
L'électrophorèse capillaire.....	8
Le SELDI.....	9
2.1.2 La spectrométrie de masse.....	11
2.1.2.1 Les sources d'ionisation .....	12
La source MALDI .....	13
La source ESI .....	16
2.1.2.2 Les analyseurs.....	20
L'analyseur TOF.....	22
La trappe ionique .....	23
2.1.2.3 L'identification des peptides et des protéines.....	25
Les empreintes peptidiques.....	25
La fragmentation.....	26
Le séquençage de novo.....	27
2.1.3 L'imagerie par spectrométrie de masse .....	28
2.1.3.1 L'imagerie MALDI.....	32
Etape 1 : Dissection et préparation de l'organe pour conservation et analyse .....	34
Etape 2 : Pré-traitement des tissus.....	36
Etape 3 : Matrice et Dépôt.....	42
Etapes 4 et 5 : Acquisition des données et reconstruction des images.....	48
2.1.3.2 Les différentes applications de l'imagerie MALDI. ....	54
Imagerie MALDI et protéomique .....	54
Imagerie MALDI et lipidomique .....	56
Imagerie MALDI et les études ADME .....	57
2.1.3.3 Les autres types d'imagerie par spectrométrie de masse .....	59
L'imagerie SIMS.....	59
L'imagerie DESI. ....	62
2.1.3.4 Gestion et traitement des données d'imagerie par spectrométrie de masse.....	64
L'Analyse en Composante Principale pour la MSI.....	65
La classification hiérarchique pour la MSI.....	67

<b>Les proprotéines convertases.....</b>	<b>70</b>
2.1.4 Origine et nomenclature des PCs. ....	71
2.1.5 Structure et mode d'action des PCs. ....	72
2.1.6 Proprotéines convertases et Cancers .....	74
2.1.6.1 Rôle des PCs dans la prolifération tumorale.....	75
2.1.6.2 PCs et adhésion cellulaire .....	78
2.1.6.3 Implications des PCs dans l'invasion tumorale .....	81
2.1.7 Inhibiteurs des PCs : vers une nouvelle voie thérapeutique ?.....	85
 <b>Chapitre 3 : Applications des analyses multivariées et de la classification aux données d'imagerie MALDI et application au cancer de la prostate et de l'ovaire. ....</b>	<b>87</b>
<b>Analyses Multivariées et classification des données .....</b>	<b>87</b>
<b>Résultats : Publication 1.....</b>	<b>88</b>
<b>Discussion et perspectives .....</b>	<b>113</b>
<b>Conclusion.....</b>	<b>116</b>
<b>Application des analyses multivariées au cancer de l'ovaire.....</b>	<b>117</b>
<b>Résultats : Publication 2.....</b>	<b>121</b>
<b>Discussion et perspectives. ....</b>	<b>134</b>
<b>Conclusion.....</b>	<b>140</b>
 <b>Chapitre 4 : Implications de la proprotéine convertase PACE4 dans le cancer de la prostate.....</b>	<b>141</b>
<b>Introduction .....</b>	<b>141</b>
<b>Résultats : Publication 3.....</b>	<b>144</b>
<b>Discussion et perspectives .....</b>	<b>173</b>
4.1.1 Choix du modèle cellulaire .....	173
4.1.2 Rôle de PACE4 dans le cancer de la prostate. ....	175
4.1.3 Approche protéomique différentielle. ....	179
<b>Conclusion.....</b>	<b>182</b>
 <b>Chapitre 5 : Conclusion générale et perspectives.....</b>	<b>183</b>
 <b>Références bibliographiques .....</b>	<b>191</b>
 <b>Annexes.....</b>	<b>207</b>
<b>Annexe I : MALDI imaging mass spectrometry: state of the art technology in clinical proteomics.....</b>	<b>208</b>
<b>Annexe II: MITICS (MALDI Imaging Team Imaging Computing System): a new open source mass spectrometry imaging software.....</b>	<b>220</b>
<b>Annexe III: MALDI imaging of formalin-fixed paraffin-embedded tissues: application to model animals of Parkinson disease for biomarker hunting. ....</b>	<b>235</b>
<b>Annexe IV: Contribution of MALDI tissue profiling and MALDI-Imaging for the study of saponins in Cuvierian tubules of Holothuria forskali .....</b>	<b>246</b>

### Publications acceptées :

- Mohamed Elayed\*, **David Bonnel\***, Remi Longuespée, Céline Castelier, Julien Franck, Daniele Vergara, Aurelie Tasiemski, Abderaouff Kenani, Denis Viniatier, Robert Day, Isabelle Fournier, Michel Salzet. MALDI Imaging Mass Spectrometry in ovarian cancer for tracking, identifying and validating biomarkers. Medical Science Monitor, 2010 16(8), BR233-245. \*:co-first authors

- Franck J\*, Arafah K\*, Elayed M, **Bonnel D**, Vergara D, Jacquet A, Viniatier D, Wisztorski M, Day R, Fournier I, Salzet M. MALDI imaging mass spectrometry: state of the art technology in clinical proteomics. Mol Cell Proteomics. 2009 Sep;8(9):2023-33. (annexe I)

- Jardin-Mathé O, **Bonnel D**, Franck J, Wisztorski M, Macagno E, Fournier I, Salzet M. MITICS (MALDI Imaging Team Imaging Computing System): a new open source mass spectrometry imaging software. J Proteomics. 2008 Aug 21;71(3):332-45. (annexe II)

- Stauber J, Lemaire R, Franck J, **Bonnel D**, Croix D, Day R, Wisztorski M, Fournier I, Salzet M. MALDI imaging of formalin-fixed paraffin-embedded tissues: application to model animals of Parkinson disease for biomarker hunting. J Proteome Res. 2008 Mar;7(3):969-78. (annexe III)

- Séverine Van Dyck, Céline Meriaux, **David Bonnel**, Michel Salzet, Isabelle Fournier, Patrick Flammang and Maxence Wisztorski. Contribution of MALDI tissue profiling and MALDI-Imaging for the study of saponins in Cuvierian tubules of *Holothuria forskali*. PLOS ONE. In Press (annexe IV)

### Publication soumise:

- David Bonnel, Rémi Longuespée, Julien Franck, Robert Day, Michel Salzet and Isabelle Fournier. Multivariate analyses for Biomarkers Hunting and Validation Through on Tissue Bottom-up or In Source Decay in MALDI-MSI. Soumise à MCP.

### Publications en cours :

- François D'Anjou, **David Bonnel**, Sophie Routhier, Isabelle Brochu, Alain Latil, Isabelle Fournier, Michel Salzet, Jean-Pierre Perreault and Robert Day. Role of PACE4 in prostate cancer: a cellular study using human DU145 cell line. A soumettre

- **Bonnel D.**, Wisztorski M., Day R., Salzet M., Fournier I. Prespotted Ionic Matrices for Standard Proteomic Studies, Direct Tissue MALDI analyses and Drug Quantification by MALDI Mass Spectrometry. A soumettre

- Arafah K., Longuespee R, El Ayed M, **Bonnell D.**, Desmons A., Buob D, Salzet M., Fournier. Developpements of new ionic matrices for the investigation of lipids directly on tissue by MALDI Imaging Mass Spectrometry: Application in rat brain and ovarian cancer hunting of lipid biomarkers. A soumettre.

### Communications orales :

- Journée Utilisateurs Brüker, Maison de la chimie, Paris, France, 25 Novembre 2008 :  
Conférence : "MALDI Imaging : Focus on sample preparation and biostatistical analysis."

- Département de Biochimie de L'Université de Sherbrooke, Québec, CANADA, 27 Février 2009 :

Séminaire: Imagerie MALDI et classification hiérarchique : un nouvel outil pour l'interprétation de la complexité des cancers humains.

- Département de Biochimie de L'Université de Sherbrooke, Québec, CANADA, 3 Juin 2010:

Séminaire : Développements et Applications d'approches protéomiques pour la recherche de cibles du cancer de la prostate.

### Posters :

-**Bonnell D.**, Franck J., Wisztorski M., Day R., Salzet M., Fournier I. On tissue fractionation: Unmasking difficult proteins for MALDI profiling, ASMS 2008, Denver, USA

-**Bonnell D.**, Franck J., Wisztorski M., Day R., Salzet M., Fournier I. On tissue fractionation: Unmasking difficult proteins for MALDI profiling, SFEAP 2008, Tours, France.

-**D. Bonnell**, J. Franck, M. El Ayed, R. Day, M. Salzet, I. Fournier. Imagerie MALDI et classification hiérarchique : un outil pour l'interprétation de la complexité des tissus humains et la recherche de biomarqueurs. SMAP 2009, Dijon, France.

-**D. Bonnell**, I. Fournier, M. Salzet, J. Stauber. Imagerie moléculaire MALDI pour l'analyse haut débit de médicaments et métabolites. SMAP 2009, Dijon, France.

-Maxence Wisztorski; Céline Meriaux; Jocelyne Bruand; Julien Franc, **David Bonnell**; Vineet Bafna ; Eduardo Macagno; Michel Salzet ; Isabelle Fournier. Exploring the Assembly of an Organism using MALDI Imaging: A Systems Approach to Understanding Embryonic Development. HUPO 2008, Amsterdam, Pays-Bas.

-Franck J., Wisztorski M., El Ayed M., **Bonnell D.**, Barnes A., Fournier I., Salet M. Automatic spotting solution for MALDI Imaging: process optimization and new developments, ASMS 2008, Denver, USA

- François d'Anjou, **David Bonnel**, Alain Latil, Robert Day. Implication of PACE4 in prostate cancer: genomic and proteomic studies with DU145 human cell line. Gordon research conferences 2010, New London, USA.

- Frédéric Couture, Christine Levesque, **David Bonnel**, Roxane Desjardins, Sophie Routhier, Michel Salzet, Isabelle Fournier, François D'Anjou and Robert Day. Role of PACE4 in the Proliferation of HT1080 Fibrosarcoma Cell Lines Validated with Proprotein Convertases shRNAs. Gordon research conferences 2010, New London, USA.





## Liste des figures

---

Figure 1 : De la chenille au papillon : même génome mais protéomes différents. ....	5
Figure 2 : Nombre de publications scientifiques internationales par an dans les domaines de la génomique et de la protéomique de 2000 à Juillet 2010 (source : pubmed).....	6
Figure 3: Principe du SELDI-TOF .....	10
Figure 4 : Les différentes surfaces chromatographiques pour la technologie SELDI-TOF .....	10
Figure 5 : Représentation des trois différentes parties d'un spectromètre de masse. ...	11
Figure 6 : Historique des développements des principales sources et des principaux analyseurs en spectrométrie de masse. ....	12
Figure 7 : Principe de la source MALDI. ....	14
Figure 8 : Schéma du processus électrochimique de l'ESI. ....	17
Figure 9 : Théorie sur l'émission des ions en phase gazeuse avec la source ESI. ....	19
Figure 10 : Calcul de la résolution spectrale (A) et du pouvoir résolutif (B) d'un analyseur.....	21
Figure 11 : Schéma du mode linéaire d'un TOF.....	22
Figure 12 : Schéma du principe et de l'intérêt du mode réflecteur d'un TOF par rapport au mode linéaire .....	23
Figure 13 : Description de l'analyseur de type trappe ionique .....	24
Figure 14: Nomenclature de fragmentation des peptides selon Biemann. (R désigne les chaînes latérales) .....	27
Figure 15 : Représentation des deux approches en imagerie par spectrométrie de masse. (a) : le mode microsonde, (b) : le mode microscope.. ....	29
Figure 16 : Nombre de publications scientifiques internationales par an impliquant la technologie d'imagerie MALDI de 2000 à Juillet 2010 (source : PubMed).....	32
Figure 17 : L'imagerie MALDI : de la préparation de l'échantillon à la reconstruction des images. ....	33
Figure 18 : Préparation et conservation des tissus pour l'imagerie MALDI .....	34
Figure 19 : Effet de différents traitements chimiques sur le signal détecté lors de la préparation d'un échantillon pour l'imagerie MALDI.....	38
Figure 20 : Les différents modes de dépôt de la matrice pour l'imagerie MALDI : manuel (a), par spray (b) et par micro-dépôts (c).....	45

Figure 21 : Image MALDI de l'ion m/z 8404 (rouge) comme exemple de protéine surexprimée dans des tissus de cancer du sein HER2-positif (A) et absent dans les tissus HER2-négatif (B). Les ions m/z 6225 (jaune) et m/z 4969 (bleu, spécifique au stroma) sont cancer-spécifiques mais pas tissu-HER2 spécifiques. ....	55
Figure 22 : Images MALDI de différents peptides de digestions obtenus sur des coupes de tissu FFPE de cerveau de rat conservé depuis environ 1 an. ....	56
Figure 23 : Détection et localisation de cinq lipides différents propres à cinq zones histologiques sur une coupe de cancer d'ovaire. ....	57
Figure 24 : Apport de l'imagerie MALDI sur des coupes d'animaux entiers (WB : whole body) pour les études ADME : détection du composé d'intérêt et des métabolites (M1, M2). ....	58
Figure 25 : Comparaison entre l'autoradiographie et MALDI-MSI pour le suivi d'un composé et de son métabolite dans une coupe entière de rat. ....	58
Figure 26 : Détection de l'olanzapine et des métabolites dans une coupe sagittale de rat entier .....	59
Figure 27 : Principe de la source SIMS.....	60
Figure 28 : Images obtenues à partir d'une coupe de cerveau de rat des ions Na <sup>+</sup> (m/z 23), K <sup>+</sup> (m/z 39), choline (m/z 86), phosphocholine (m/z 184), cholestérol [M+H-H <sub>2</sub> O] <sup>+</sup> (m/z 369), vitamine E (m/z 430), phospholipides (m/z 769) et du tic. Images réalisées à 62,5µm et avec une source d'ions primaires Au <sup>3+</sup> . ....	61
Figure 29 : Images SIMS d'ions secondaires négatifs d'une coupe de thyroïde de rat. ....	61
Figure 30 : Principe de la source DESI . ....	63
Figure 31 : Images DESI d'ions négatifs de lipides et d'acides gras au niveau de la moelle épinière de rat. ....	63
Figure 32 : Représentation simplifiée d'un spectre de masse à l'aide d'un point dans un plan multidimensionnel. ....	65
Figure 33 : Simplification des données de MSI par PCA. ....	66
Figure 34 : Analyse des résultats PCA en mode supervisé, formation de deux groupes de spectres au niveau du loading plot (A) avec la possibilité de détecter les m/z spécifiques (représentés par des points) à ces groupes sur le loading plot (B). ....	66
Figure 35 : Principe de la classification hiérarchique aboutissant à la formation d'un dendrogramme.....	67
Figure 36 : Le répertoire des protéases humaines et murines .....	71

Figure 37 : Structure des PCs. ....	73
Figure 38 : Capacités acquises des cellules cancéreuses au cours du temps.....	74
Figure 39 : Schéma des événements impliquant les PCs dans la croissance tumorale et les métastases. Les croix vertes représentent les sites d'action potentiels des PCs.. ....	75
Figure 40 :Schéma représentant les voies par lesquelles les intégrines interagissent avec l'actine du cytosquelette . ....	79
Figure 41 : Potentiel et spécificité des inhibiteurs de PCs . ....	85
Figure 42 : Recherche et identification de biomarqueurs par imagerie MALDI en protéomique clinique.....	114
Figure 43 : Utilisation de la modélisation pour les données d'MSI : hypothèse de son application pour le diagnostic clinique.....	116
Figure 44 : Analyse par classification hiérarchique d'une coupe de cancer d'ovaire (A). (B) dendrogramme obtenu après classification des spectres. Les deux premiers embranchements sont indiqués en rouge et en vert. (C) images des localisations des deux groupes de spectres obtenus. ....	134
Figure 45 : Utilisation de la PCA pour la mise en évidence de m/z d'intérêt.....	135
Figure 46 : Composition et organisation du protéasome avec ses différents activateurs intracellulaires. ....	136
Figure 47 : Localisation des ARNm de différentes PCs par hybridation in situ dans des coupes sagittales de cerveau de rat adjacentes. ....	176
Figure 48 : Analyse SELDI-TOF des peptides de digestion différentiels entre la lignée DU145 et la lignée 4-2. ....	180
Figure 49 : Principe de la technique iTRAQ-TAILS pour la recherche des substrats de protéases . ....	181
Figure 50 : Représentation des Prs Damadian et Lauterbur avec leur première image IRM obtenue sur un corps humain en 1977. ....	188
Figure 51 : Evolution de l'imagerie MALDI depuis la première image en 1999 sur des cerveaux murins . ....	189



## Liste des tableaux

---

Tableau 1 : Exemples de laser utilisés en MALDI et leurs caractéristiques. ....	15
Tableau 2 : Listes des matrices utilisables en MALDI.....	16
Tableau 3 : Comparaison des analyseurs actuels en fonction de leur pouvoir résolutif, de leur précision de masse et de leur gamme de masse. ....	21
Tableau 4: Liste des différentes sources d'ionisation et de leurs caractéristiques utilisées en MSI.....	31
Tableau 5 : Listes et caractéristiques des différents instruments pour le dépôt de la matrice en MSI.....	48
Tableau 6: Liste des différents logiciels existants pour l'imagerie MALDI. ....	52
Tableau 7 : Liste et caractéristiques des différents spectromètres de type MALDI compatibles avec la MSI (www.maldi-msi.org).....	53
Tableau 8: Schématisation des méthodes de calcul des distances entre les points et des liaisons entre les clusters pour la classification hiérarchique.....	69
Tableau 9 : Nomenclature des PCs chez l'Homme d'après Fugère et al .....	72
Tableau 10 : Facteurs de croissance et sites de clivage pour les PCs. ....	77
Tableau 11 : Sites potentiels de clivage des récepteurs de facteur de croissance par les PCs. ....	78
Tableau 12 : Sites potentiels de clivage intégrines par les PCs.....	80
Tableau 13 : Sites de clivage potentiels des MMPs par les PCs. ....	83
Tableau 14 : Sites de clivage potentiels des ADAMs par les PCs. ....	83
Tableau 15 : Sites de clivage potentiels des ADAM-TSs par les PCs. ....	84
Tableau 16 : Liste des inhibiteurs des PCs.....	86
Tableau 17 : Comparaison des profils d'expression des mucines dans différents stades et types histologiques du cancer de l'ovaire. ....	139
Tableau 18 : Lignées cellulaires humaines de cancer de la prostate .....	173
Tableau 19 : Principaux modèles précliniques de cancer de la prostate actuellement utilisés.....	174
Tableau 20 : Liste des différents phénotypes de souris knock-out existants pour les PCs. ....	177



## Liste des abréviations

---

2D	bidimensionnel (électrophorèse)
AB	Applied Biosystem
ACE	Angiotensin-Converting Enzyme
ADAM	Adamalysin Metalloproteinases
ADAMTS	Adamalysine Metalloproteinases with Thrombospondin motifs
ADME	Absorption, Distribution, Métabolisme et Elimination
ADN (DNA)	Acide désoxyribonucléique
ADNc (cDNA)	Acide désoxyribonucléique complémentaire
ANI	Aniline
APCI	Atmospheric Pressure Chemical Ionization
APP	Amyloid Precursor Protein
ARN (RNA)	Acide Ribonucléique
ARNm (mRNA)	Acide Ribonucléique messenger
ARNt (tRNA)	Acide Ribonucléique de transfert
ATP	Adénosine TriPhosphate
BD	Bruker Daltonics
BMP	Bone Morphogenetic Protein
CA125	Cancer Antigen 125
CE	Capillary electrophoresis
CH	Classification Hiérarchique
CI	Chemical Ionisation
CID	Collision-Induced Dissociation
CMH	Complexe Majeur d'Histocompatibilité
CRIP1	Cysteine-Rich Intestinal Protein 1
CSF	Colony-Stimulating Factor
Da	Dalton
DAN	DiAminoNaphatalene
DESI	Desorption ElectroSpray Ionisation
DHB	Acide 2,5-Dihydroxybenzoïque
ECM	Extra Cellular Matrix
EGF	Epidermal Growth Factor
EI	Electronic Ionisation
EOC	Epithelial Ovarian Cancer
Er:YAG	Erbium-Doped Yttrium Aluminium Garnet
ESI	ElectroSpray Ionisation
EtOH	Ethanol
FAB	Fast Atoms Bombardement
FAK	Focal Adhesion Kinase
FD	Field Desorption
FFPE	Formalin-fixed, Paraffin-Embedded (tissue)
FGF	Fibroblast Growth Factor
FIGO	Fédération Internationale de Gynécologie Obstétrique

FT-ICR	Fourier Transform Ion Cyclotron Resonance
FWHM	Full Width at Half Maximum
GA	Genetic Algorithm
GHRH	Growth-Hormone-Releasing Hormone
HBEGF	Heparin-Binding Epidermal Growth Factor
HCCA	Acide $\alpha$ -cyano-4-hydroxycinnamique
HER2	Human Epidermal Growth Factor Receptor-2
HGF	Hepatocyte Growth Factor
ICAM	IntraCellular Adhesion Molecule
IEF	Iso-électrofocalisation
Ig	ImmunoGlobuline
IGF	Insulin-like Growth Factor
IL	Interleukin
INF	Interferon
IR	Infrarouge
ISD	In Source Decay
IT	Ion Trap
ITO	Indium Tin Oxyde
iTRAQ	isobaric Tag for Relative and Absolute Quantitation
KL	c-Kit-Ligand
LC	Liquid Chromatography
LD	Laser Desorption
LDI	Laser Desorption Ionisation
LDL	Low Density Lipoprotein
LPA	Lysophosphatidic Acid
LPL	LipoProtein-Lipase
m/z	rapport masse sur charge
MadCAM	Mucosal addressin Cell Adhesion Molecule
MALDI	Matrix assisted Laser Desorption Ionisation
MeOH	Methanol
MMPs	Matrix MetalloProteinases
MS	Mass Spectrometry
MSI	Mass Spectrometry Imaging
MT-MMPs	Membranes Type metalloprotéinases
NARC-1/PCSK9	NeuralApoptosis-Regulated Convertase-1 / Proprotein Convertase Subtilisin Kexin 9
Nd:YAG	Neodymium-Doped Yttrium Aluminium Garnet
NGF	Nerve Growth Factor
OCT	Optimal Cutting Temperature
ODNs	OligoDeoxyNucleotides
OSE	Ovarian Surface Epithelial
PACE4	Paired Basic amino Acid Cleaving Enzyme 4
PANTHER	Protein Analysis THrough Evolutionary Relationships
PC	Principal component



PCA	Principal Component Analysis
PCA-SDA	Principal Component Analysis-Symbolic Discriminant Analysis
PCR	Polymerase Chain Reaction
PCs	Proprotéines Convertases
PD	Plasma Desorption
PDF	Prostate Derived Factor
PDGF	Platelet-Derived Growth Factor
pH	potentiel Hydrogène
pI	point Isoelectrique
PMF	Peptide Mass Fingerprint
PSA	Prostate Specific Antigen
PSD	Post Source Decay
PTH	ParaThyroid Hormone
PTHRP	ParaThyroid-Hormone-Related Peptide ou Protein
PTMs	PostTranslational Modifications
Q	Quadripôle
QC	QuickClassifier
RE	Réticulum Endoplasmique
ROI	Region Of Interest
RT-PCR	Reverse Transcriptase-Polymerase Chain Reaction
SA	Acide sinapinique
SDS PAGE	Sodium Dodecyl Sulfate PolyAcrylamide Gel Electrophoresis
SELDI	Surface Enhanced Laser Desorption Ionisation
SIMS	Secondary Ion Mass Spectrometry
SKI-1/SP1	Subtilisin/Kexin-like Isoenzyme-1 - Site-1 Protease
SNN	Supervised Neural Network
SVM	Support Vector Machine
SZU	Shimatzu
TACE	Tumor necrosis factor- $\alpha$ converting enzyme
TAILS	Terminal Amine Isotopic Labeling of Substrates
TGF	Transforming Growth Factor
TGN	Trans Golgi Network
TIMPs	Tissues Inhibitors of MetalloProteinases
TNF	Tumor Necrosis Factor
TNFR	TNF receptor
TNM	Tumors, Nodes, Metastasis
TOF	Time Of Flight
TRPS1	TrichoRhinoPhalangeal Syndrome I
TS	Thermo Scientific
TSHR	Thyroid-Stimulating Hormone Receptor
u.m.a	Unité de masse atomique
UV	Ultraviolet
VCAM	Vascular Cell Adhesion Molecule
VEGF	Vascular Endothelial Growth Factor

WBA  
YAG

Whole Body Autoradiography  
Yttrium Aluminium Garnet

## La cellule : de l'émerveillement à l'inquiétude.

Comprendre la complexité des mécanismes biologiques mis en jeu au sein d'un être vivant, s'émerveiller de cette machinerie hors norme faisant intervenir une multitude de molécules ayant chacune un rôle précis et indispensable, n'est-ce pas là la chance qu'a tout chercheur dans le domaine des Sciences du vivant ? Depuis Aristote (-384, -322 av. J.-C.), qui fut considéré comme l'un des premiers « biologistes », l'homme n'a cessé d'aller de découvertes en découvertes pour ainsi mieux comprendre notre environnement et la vie qui y évolue. Cette vie qui repose entièrement sur la cellule (du latin *cellula* petite chambre), unité structurale, fonctionnelle et reproductrice constituant la base de tout être vivant. Néanmoins, nous savons de nos jours qu'un dérèglement ou une modification d'origine héréditaire, environnementale ou infectieuse dans une seule cellule peut menacer la survie d'un individu. Cette pathologie, appelée cancer, représente un réel fléau dans nos sociétés actuelles, devenant ainsi un sujet d'investigation prioritaire et primordial pour bon nombre de chercheurs dans le monde. Depuis le séquençage du génome humain en 2004 [1], de nombreux domaines d'étude se sont développés comme la génomique, la transcriptomique et la protéomique [2]. Cette dernière reflète aujourd'hui un réel engouement en raison du rôle « effecteur » que jouent les peptides et les protéines au sein de la cellule. La possibilité de comprendre, de caractériser, d'identifier les réseaux de protéines associées à des cibles thérapeutiques ouvre la possibilité d'améliorer le diagnostic et la thérapie des patients atteints de tumeurs. C'est pourquoi de nombreuses avancées technologiques ont eu lieu ces dernières années [3]. Parmi elles, on retrouve en première place la spectrométrie de masse. En effet, cette technique d'analyse permet de détecter, de quantifier et d'identifier des molécules d'intérêt par mesure de leur masse, de caractériser leur structure chimique et s'avère donc parfaitement appropriée pour la recherche de marqueurs pathologiques à la fois sur des fluides biologiques, des extractions cellulaires ou encore des tissus [4-6]. Cependant ce type d'étude ne permet pas la corrélation d'un marqueur donné avec une localisation au sein du tissu, de la zone histologique ou de la cellule et peut entraîner une perte d'informations en raison des nombreuses étapes nécessaires. L'imagerie par spectrométrie de masse peut palier à ces problèmes [7-11]. Cette technique, qui ne nécessite que très peu de traitements de l'échantillon tissulaire [12, 13], permet la détection, et la localisation de plusieurs centaines de

molécules en une seule analyse, ce qui en fait aujourd'hui un outil adapté à la recherche de marqueurs des pathologies, comme le cancer.

## **Etudes menées au cours de cette thèse**

Au travers des cinq chapitres de cette thèse, l'ensemble des travaux portera sur les développements et les applications d'approches en protéomiques pour la recherche de cibles dans le cadre de deux pathologies : le cancer de l'ovaire et le cancer de la prostate.

Le chapitre 2 introduit l'ensemble des généralités propres à cette thèse. Une première partie concerne les proprotéines convertases en se focalisant plus particulièrement sur leurs implications dans le cancer. Ces enzymes sont en effet impliquées dans la maturation d'un grand nombre de protéines décrites dans les phénomènes de prolifération, d'adhésion et d'invasion tumorales. C'est pourquoi la deuxième partie de ce chapitre traite des outils en protéomique permettant l'analyse des protéines d'un tissu à un temps donné. Nous nous focaliserons particulièrement sur l'imagerie par spectrométrie de masse dont l'approche a particulièrement été utilisée dans ces travaux de recherche.

Le chapitre 3 décrit le développement et l'application de méthodes statistiques multivariées et de la classification hiérarchique aux données d'imagerie MALDI. En effet, l'utilisation de ces méthodes d'analyse et de traitement des données permet de simplifier et d'aboutir à des identifications fines des protéines d'intérêt au sein d'une coupe histologique dépendamment des différents états physiologiques présents. A l'aide de ces méthodes, nous avons pu vérifier et valider les protocoles des différentes étapes de préparation de tissus nécessaires avant l'acquisition en imagerie MALDI à la fois pour les tissus frais et FFPE (Formalin Fixed and Paraffin Embedded). Nous avons également mis au point une nouvelle approche, appelée PCA-SDA (Principal Component Analysis – Symbolic Discriminant Analysis), qui permet de détecter des cibles potentielles directement à partir de coupe de tissu.

Ensuite, l'utilisation de la technique d'imagerie par spectrométrie de masse MALDI a été appliquée à la recherche de nouveaux marqueurs du cancer de l'ovaire. Ainsi à partir de biopsies de cancer d'ovaire, nous avons pu identifier des marqueurs potentiels de cette pathologie. Certains d'entre eux sont connus pour leur implication dans le système

immunitaire, ce qui nous a permis de mieux comprendre comment les cellules cancéreuses échappaient aux contrôles de l'organisme, en jouant sur l'adhésion et l'invasion de celles-ci via des modifications au niveau de la matrice extracellulaire et du cytosquelette, et en modulant la réponse immunitaire.

Enfin le chapitre 4 démontre l'implication de la proprotéine convertase PACE4 dans le cancer de la prostate. PACE4 s'est avérée être la seule PC surexprimée dans cette pathologie. C'est pourquoi son rôle et son impact ont été recherché à partir d'une lignée cellulaire de cancer de prostate (DU145) à l'aide d'un Knock-down pour cette enzyme. Après la mise en évidence préalable de son rôle majeur dans la prolifération des cellules, nous avons réussi à décrire et mieux comprendre les phénomènes d'évitement de l'apoptose mises en jeu à l'aide d'approches protéomiques. Cette étude ouvre des perspectives sur l'importance de l'implication des PCs dans les cancers, et plus particulièrement de PACE4 qui seraient une cible thérapeutique de choix pour le cancer de la prostate.

Le chapitre 5 conclut ce travail de thèse en faisant le bilan des résultats obtenus, de leur impact et des perspectives scientifiques.



### La Protéomique

Lorsqu'on explique ce qu'est la protéomique, très souvent l'image de la chenille et du papillon revient (**Figure 1**). En effet, cet organisme possède le même génome tout au long de sa vie, mais l'expression de ses gènes donnant les protéines sera totalement différente au cours de son évolution qui passe par deux stades morphologiques. Selon le cycle cellulaire, la différenciation, la réponse à différents signaux biologiques, l'état physiopathologique et l'environnement de cet organisme, son génome ou l'expression de son génome peuvent être modifiés.

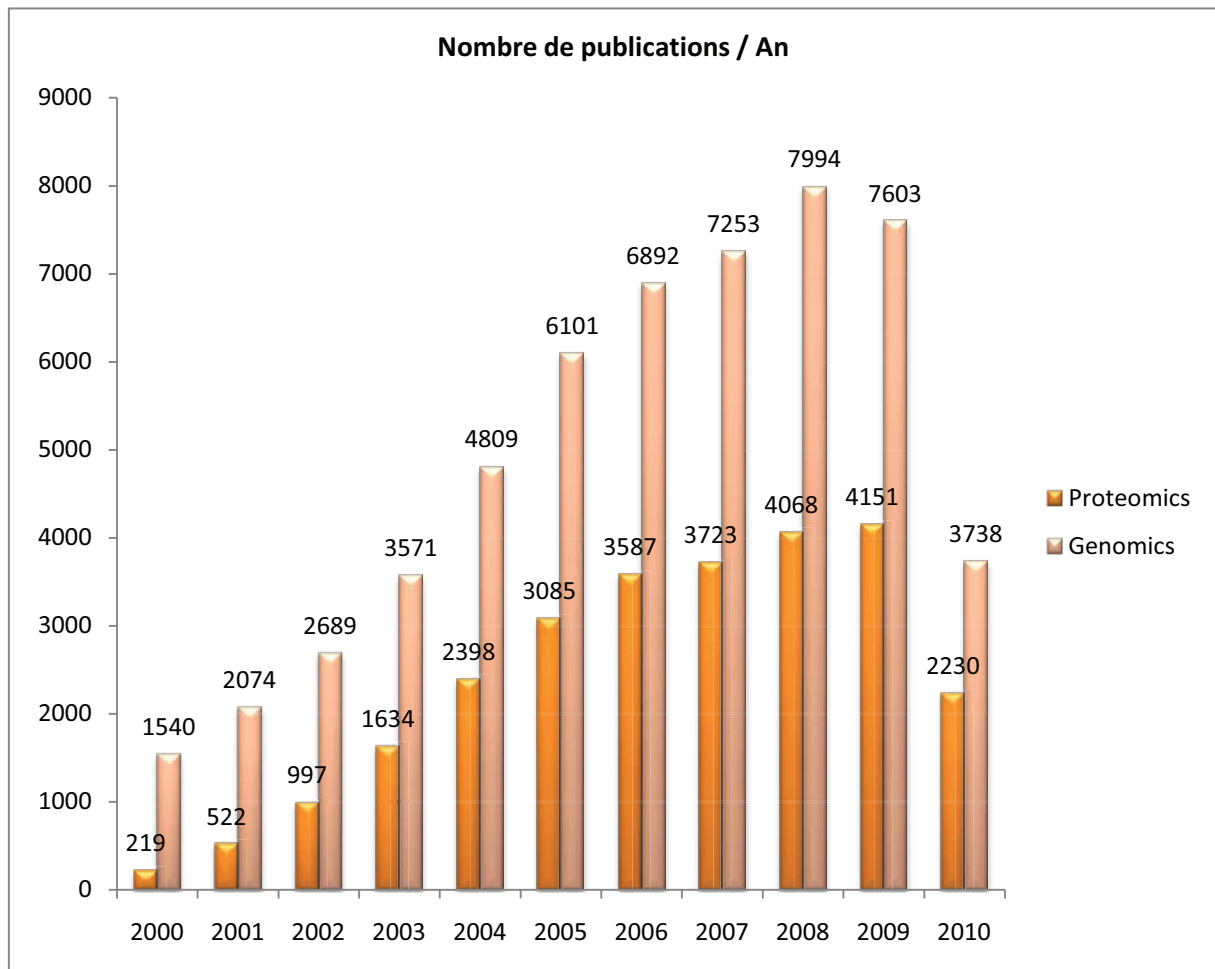


**Figure 1 : De la chenille au papillon : même génome mais protéomes différents.**

La protéomique désigne donc la science qui étudie les protéomes, c'est-à-dire l'ensemble des protéines d'un organite, d'une cellule, d'un tissu, d'un organe ou d'un organisme à un moment donné et sous des conditions données. Le terme protéomique fut utilisé pour la première fois en 1997 dans la publication scientifique de James P. [14].

En effet, depuis le séquençage de nombreux génomes dont celui de l'homme [1, 15] et avec l'accumulation de séquences d'ADN dans les banques génomiques, il est apparu qu'il n'était pas suffisant de disposer de séquences complètes de génomes pour comprendre les fonctions biologiques des protéines pour lesquelles ces séquences codent. L'identification des protéines, la caractérisation de leur séquence en acides aminés et de leurs modifications post-traductionnelles, la découverte de produits de l'épissage alternatif ne sont possibles qu'en

étudiant les protéines elles-mêmes. Les mécanismes de régulation et la dynamique de l'expression protéique font également partie des défis à relever pour traduire l'information issue des données génomiques en connaissances biologiques, faisant de ce domaine d'étude, un secteur en pleine expansion (**figure 2**).



**Figure 2 : Nombre de publications scientifiques internationales par an dans les domaines de la génomique et de la protéomique de 2000 à Juillet 2010 pour les termes de recherche « proteomics » et genomics » (source : pubmed)**

### 2.1.1 Les techniques de séparation des protéines et des peptides

La grande complexité des échantillons biologiques à partir desquels les analyses protéomiques sont effectuées, impose d'utiliser des techniques de séparation. Ces techniques en constant développement peuvent être in-gel ou off-gel.



### 2.1.1.1 Méthodes in-gel

L'outil séparatif le plus couramment utilisé en analyse protéomique est basé sur les techniques d'électrophorèse bi-dimensionnelle, possédant une résolution suffisante pour séparer des variants post-traductionnels et une dynamique d'analyse quantitative appréciable [16]. L'échantillon déposé sur gel est séparé suivant une première dimension en fonction du point isoélectrique (pI) des protéines (iso-électrofocalisation, IEF), puis suivant une deuxième dimension en fonction de leurs masses moléculaires (électrophorèse SDS PAGE) [17]. L'IEF possède une grande résolution qui permet de distinguer certains variants post-traductionnels contrairement à l'électrophorèse en milieu dénaturant SDS. Mais, le spectre reste limitant, c'est à dire qu'une certaine catégorie de protéines (hydrophobes, de pI extrêmes) ne seront pas visibles. L'introduction des gradients de pH immobilisés a permis de progresser pour l'analyse des protéines basiques et pour améliorer la résolution via les gradients de pH étroits. De la même manière, les progrès récents de la solubilisation des protéines ont déjà augmenté les possibilités d'analyse (spectres d'analyse). Cependant, malgré tous les progrès réalisés, il reste difficile d'aboutir à toute la complexité du protéome d'une cellule (ex : protéines de pH extrêmes, protéines membranaires). C'est pourquoi, d'autres approches ont été développées : approche par chromatographie liquide mono ou bidimensionnelle, techniques de marquage, couplage gel SDS/nano-LC.

### 2.1.1.2 Méthode off-gel

#### **La chromatographie liquide**

En protéomique, la chromatographie liquide (LC) a pour objectif de concentrer et de séparer un mélange plus ou moins complexe de peptides. Plus le volume dans lequel est éluée la molécule est petit, plus sa concentration sera importante et meilleure sera l'intensité du signal. En effet, la réponse est concentration dépendante et non débit dépendant. Le développement de la nano-LC a 2 objectifs :

- Améliorer la limite de détection afin de pouvoir travailler avec des quantités de plus en plus faibles de matériel
- Rendre les nano-colonnes et la nano-LC compatibles avec les sources nano-ESI.

De ce fait, de grandes avancées techniques ont été réalisées au niveau des nano-colonnes (techniques de remplissage, nouvelles particules plus fines...) et au niveau de la chromatographie (nano-débits).

Ainsi, du fait de la bonne performance de ce couplage nanoLC-MS ou nanoLC-MS/MS, cette approche est devenue complémentaire à l'approche par gel 2D. Son avantage majeur reste sa sensibilité. En effet, la diminution du débit (200 nl/min) ne modifiant pas la concentration de l'échantillon, le temps d'analyse peut être augmenté, permettant l'obtention d'informations de séquences de bonnes qualités, sur de faibles quantités de matériels (environ 50 femtomoles). De plus, cette approche est assez aisément automatisable. Cependant, elle ne peut pas être utilisée seule, sans autre technique de séparation en amont, si l'échantillon est trop complexe. En effet, un extrait protéique contient plusieurs milliers de protéines d'où des centaines de milliers de peptides tryptiques qui ne pourront pas être séparés efficacement sur nano-colonnes. Cette technique sera donc couplée en amont à une autre technique séparative, généralement le gel 2D.

### **L'électrophorèse capillaire**

Introduite en 1979 [18], l'électrophorèse capillaire (CE) est basée sur la séparation d'espèces sous l'action d'un champ électrique, en fonction de leurs différences de mobilité électrophorétique, dans un capillaire en silice vierge de faible diamètre interne (< 100µm). Cette technique peut être utilisée pour la séparation des peptides, des protéines, des sucres, des métabolites, des molécules pharmaceutiques. Les travaux de Jorgenson et de ses collaborateurs [19] ont montré que l'électrophorèse capillaire, de par son aspect de miniaturisation, présente un grand nombre d'avantages. Le premier est le faible diamètre interne du capillaire (généralement inférieur à 100µm) générant un faible effet Joule grâce à une meilleure dissipation de la chaleur, permettant ainsi d'appliquer des champs électriques plus élevés que lors de séparations par électrophorèse conventionnelle, et ainsi d'accéder à des séparations plus rapides. La faible consommation d'échantillons et de tampon de séparation requis pour les séparations électrophorétiques, ainsi que la haute résolution et les temps d'analyse rapides obtenus, ont fait de la CE une méthode analytique de choix. De plus, l'automatisation des différentes étapes de l'analyse (conditionnement, injection, séparation) sont des avantages majeurs. Par ailleurs, les instruments commerciaux peuvent être couplés

aux spectromètres de masse (CE/MS) pour obtenir une meilleure détection et une meilleure sensibilité [20-22].

### **Le SELDI**

Le Surface Enhanced Laser Desorption Ionisation (SELDI) [23] est une source d'ionisation qui découle du MALDI. Elle utilise une cible possédant des affinités biochimiques avec certains composés de l'analyte. Un mélange complexe de protéines ou de peptides est déposé sur une surface fonctionnalisée, certaines molécules se lient à la surface et des étapes de lavages sont appliquées pour retirer les protéines et les peptides n'ayant pas d'affinité avec la phase. La matrice est ensuite déposée sur le spot et co-cristallise avec l'échantillon. Les molécules d'intérêt sont ensuite détectées suivant le même principe que le MALDI-TOF en mode linéaire (**Figure 3**). Le SELDI peut être représenté comme un couplage des technologies de chromatographie liquide avec un MALDI-TOF. Les différents types de surfaces chromatographiques sont listées dans la **figure 4**. Le principe du SELDI fait de cette technologie, une technique quantitative permettant la recherche de marqueurs dans différentes conditions, dans des extraits cellulaires, des extraits tissulaires et des fluides biologiques. Une suite de logiciels est proposé par la société Biorad qui commercialise aujourd'hui cet appareil, ces logiciels permettent des analyses statistiques poussées et robustes pour la mise en évidence de biomarqueurs [24-30] Néanmoins, il n'est pas possible d'identifier les molécules d'intérêt à l'aide de cet appareil. C'est pourquoi, depuis 2010, les sociétés Biorad et Bruker se sont associées pour mettre au point le système Lucid. Ce système permet l'identification des  $m/z$  d'intérêt, détectés par la technologie SELDI, en rendant possible l'introduction des barrettes dans les appareils Bruker de type UltraFlex qui permet de faire de la  $ms^2$  [31].

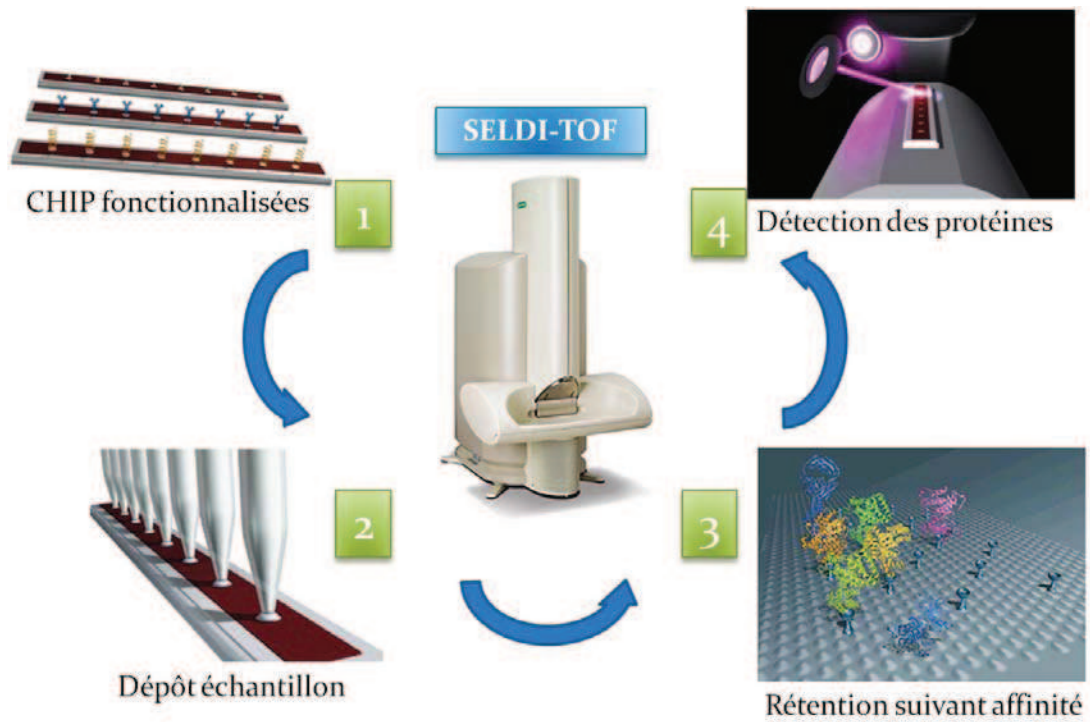


Figure 3: Principe du SELDI-TOF

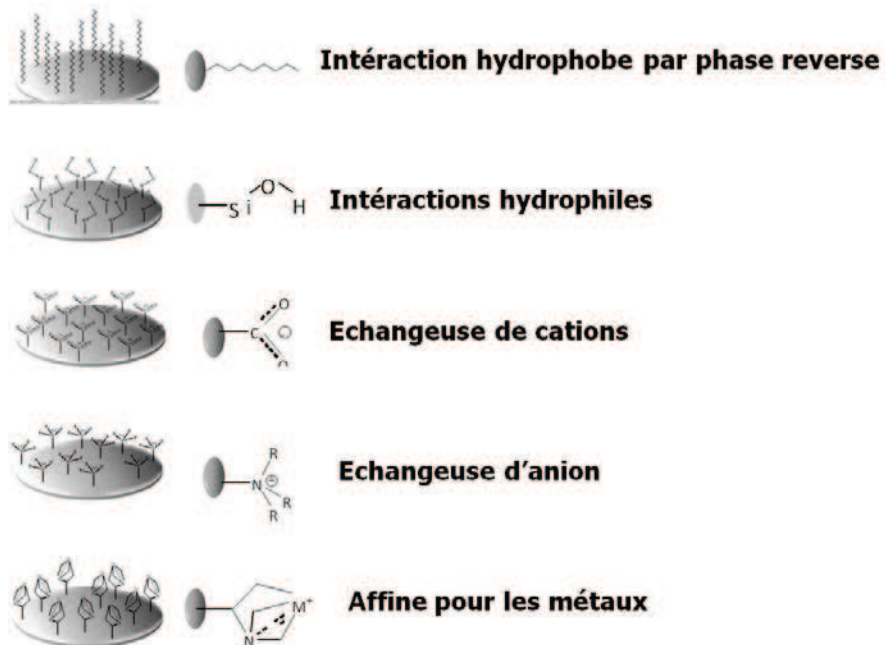
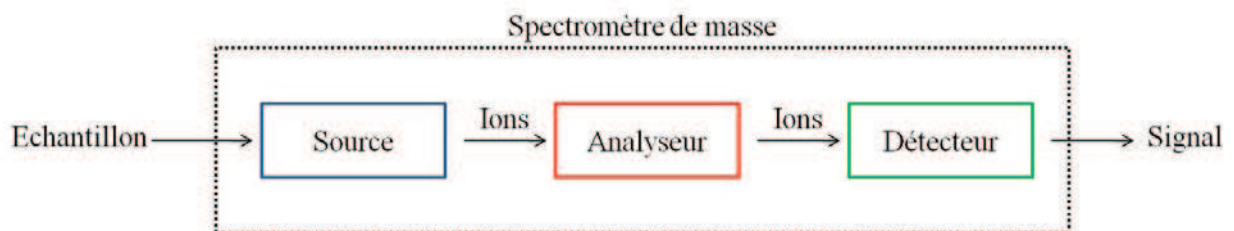


Figure 4 : les différentes surfaces chromatographiques pour la technologie SELDI-TOF

## 2.1.2 La spectrométrie de masse

La spectrométrie de masse s'est imposée comme l'outil de choix puisqu'aucune autre technologie ne peut aujourd'hui recueillir autant d'informations avec autant de sensibilité et de rapidité en protéomique. Les développements technologiques et méthodologiques pour l'analyse protéomique par spectrométrie de masse conduisent encore à la mise au point ou à l'amélioration de stratégies toujours plus poussées pour mieux comprendre les fondements moléculaires de processus cellulaires et physiologiques.

Un spectromètre de masse est classiquement composé de trois parties principales : une source d'ionisation, un analyseur de masse, et un détecteur de courant d'ions (**Figure 5**).



**Figure 5 : Représentation des trois différentes parties d'un spectromètre de masse.**

Les molécules, en phase liquide ou solide, sont ionisées et transférées en phase gazeuse au niveau de la source. Puis les ions formés sont séparés dans un analyseur en fonction de leur rapport masse-sur-charge ( $m/z$ ). Le mouvement des ions dans l'analyseur est assuré sous l'effet de champs électriques et/ou magnétiques : la vitesse et/ou la trajectoire des ions jusqu'au détecteur dépendent donc de leur  $m/z$  caractéristique. A la fin de leur course, les ions frappent le détecteur qui convertit alors ces éléments chimiques en signal électrique : le détecteur. Par la suite, ces informations sont exploitables par l'utilisateur au niveau d'une interface informatique où le rapport masse sur charge ( $m/z$ ) est donné en fonction de l'intensité du signal, sous la forme d'un spectre de masse. La **figure 6** retrace l'historique de découvertes des principales sources et des principaux analyseurs utilisés en spectrométrie de masse.

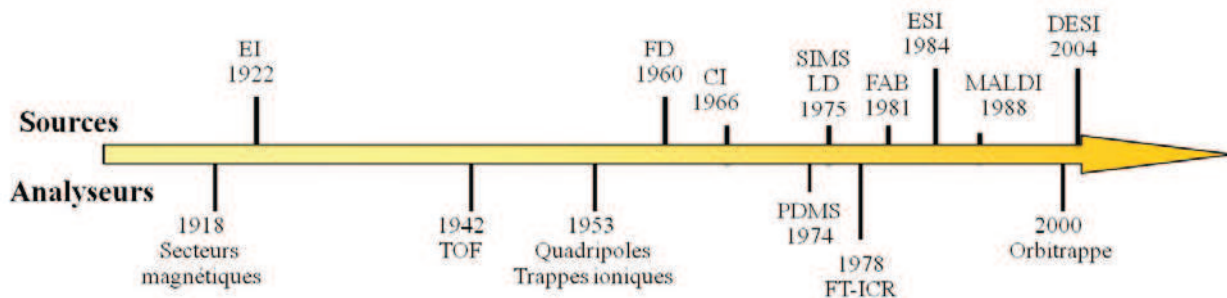


Figure 6 : Historique des développements des principales sources et des principaux analyseurs en spectrométrie de masse.

### 2.1.2.1 Les sources d'ionisation

Au sein du spectromètre, l'échantillon est ionisé dans la partie appelée source d'ionisation. On compte aujourd'hui de nombreux types de sources. Les notions les plus importantes sont l'énergie interne transférée durant l'ionisation et les propriétés physico-chimiques de l'analyte qui doit être ionisé. Certaines sources sont dites très énergétiques et entraînent la fragmentation, alors que d'autres seront plus douces et ne produiront que des ions stables.

Ces sources peuvent se présenter sous deux types : les sources d'ionisation en phase solide et les sources d'ionisation en phase liquide. En phase liquide, la solution est nébulisée sous forme de gouttelettes entraînant la formation d'ions sous pression atmosphérique. On retrouve parmi ces sources : l'électrospray (ESI), atmospheric pressure chemical ionization (APCI) et atmospheric pressure photoionization. En phase solide, l'échantillon peut être obtenu après plusieurs étapes de préparation, qui comprendront parfois l'introduction d'une matrice sous forme solide ou liquide. Le dépôt est ensuite irradié à l'aide de particules énergétiques ou de photons qui désorbent les ions à la surface du dépôt matrice-analyte. Ces ions peuvent être accélérés par un champ électrique avant d'être séparés au sein de l'analyseur. Les sources matrix-assisted laser desorption ionization (MALDI), secondary ion mass spectrometry (SIMS), plasma desorption (PD) permettent l'analyse des échantillons sous phase solide.

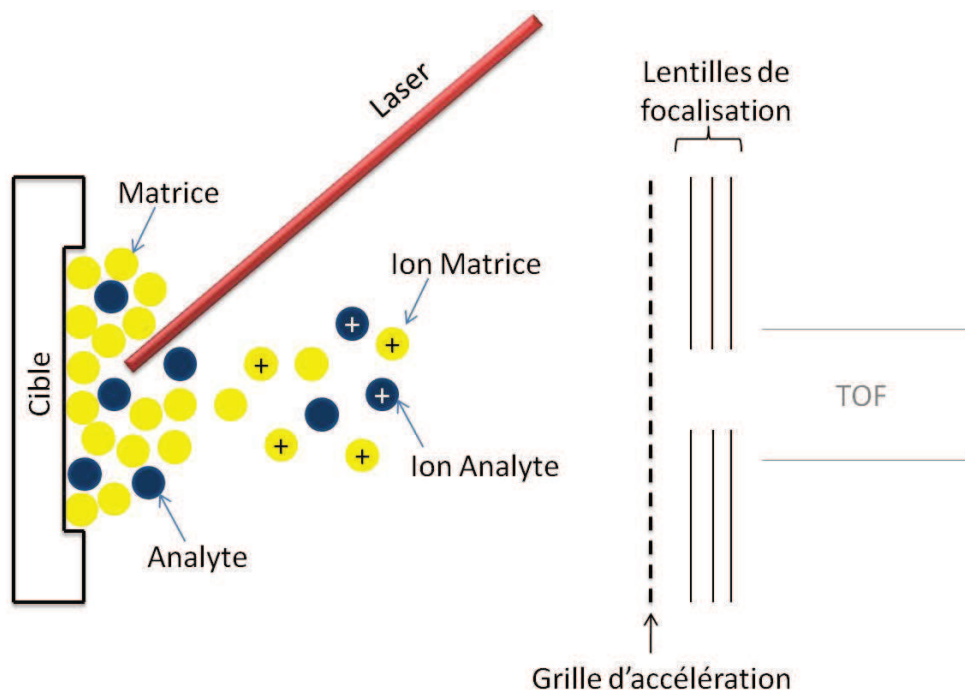
Les sources d'ionisation produisent généralement les ions à partir d'une molécule neutre en phase gazeuse par éjection ou capture d'électrons, protonation, déprotonation, formation d'adduits ou par transfert de charge.

Les sources MALDI et ESI feront l'objet d'une description plus détaillée à la suite de cette introduction en raison de leurs utilisations dans nos différents projets scientifiques.

### **La source MALDI**

Le MALDI constitue l'une des sources les plus utilisées en spectrométrie de masse. On doit son origine à deux équipes de recherche ayant travaillé, toutes deux, sur les processus de désorption et d'ionisation à l'aide d'une matrice. En effet, les méthodes de désorption laser (LDI) [32] permettent la production d'ions à partir de mélanges solides mais sont réellement limitées à la gamme de masse des composés détectés qui ne dépasse pas 300u. Pour pallier à cette limite, en 1988, l'équipe d'Hillenkamp et Karas a observé que l'utilisation d'un composé avec un groupement aromatique (acide nicotinique) permet d'obtenir un signal de meilleure qualité [33]. La matrice permet un transfert d'énergie efficace par absorption des photons laser tout en préservant l'analyte. Ceci leur permit de détecter pour la première fois l'albumine à une masse d'environ 67 KDa avec l'acide nicotinique comme matrice organique. En parallèle, l'équipe de Tanaka (prix Nobel de Chimie en 2002) a développé une approche reposant sur l'utilisation d'une poudre de métal en suspension dans du glycérol qui était irradié par un laser UV [34]. Néanmoins la sensibilité de la méthode utilisant une matrice organique est 500 à 1000 fois supérieure à celle utilisant le glycérol, et produit un signal de meilleure qualité, c'est pourquoi la méthodologie d'Hillenkamp et Karas est préférée de nos jours à celle de Tanaka.

Le principe du MALDI (**figure 7**) repose donc sur la co-cristallisation d'une matrice en large excès avec l'analyte sous la forme d'un réseau de microcristaux suite à l'évaporation du solvant.



**Figure 7 : Principe de la source MALDI.**

Un pulse laser, émettant le plus souvent dans l'UV ou dans l'IR (**Tableau 1**), vient percuter l'échantillon préalablement déposé sur une cible conductrice. La matrice, qui absorbe à la longueur d'onde du laser, va induire des transitions électroniques vers des états excités par absorption des photons [35, 36]. La relaxation de cette énergie vers un niveau plus stable au sein du solide matrice-analyte entrainera l'éjection de matière au niveau de la zone d'impact sous la forme d'un plasma. Celui-ci va former un nuage d'ions et de molécules neutres sous forme gazeuse (désorption), aussi appelé plume. L'énergie interne fournie reste suffisamment faible pour ne pas entrainer directement de fragmentation des analytes et pour que les ions formés soient suffisamment stables pour toute la durée de l'analyse. En mode d'ionisation positif, on retrouve le plus souvent des ions de matrice et des espèces protonées  $[M+H]^+$ . Elles peuvent être également cationisées si l'échantillon contient des traces de sels et il est alors possible d'observer des espèces cationisées par le sodium  $[M+Na]^+$  ou le potassium  $[M+K]^+$ . De la même façon, en mode d'ionisation négatif, les espèces les plus abondantes sont généralement sous la forme  $[M-H]^-$ . Ces ions seront par la suite accélérés par l'application d'une différence de potentiel entre la cible et la grille d'accélération qui sera fixée en fonction de l'analyse des ions positifs ou négatifs.



<b>Laser</b>	<b>Longueur d'onde</b>	<b>Energie (eV)</b>	<b>Temps de pulsation</b>
Nitrogen	337 nm	3.68	<1ns
Nd:YAG $\mu$ 3	355 nm	3.49	5 ns
Nd:YAG $\mu$ 4	266 nm	4.66	5 ns
Er:YAG	2.94 $\mu$ m	0.42	85 ns
CO2	10.6 $\mu$ m	0.12	100 ns

**Tableau 1 : Exemples de laser utilisés en MALDI et leurs caractéristiques.**

Depuis 1988, de nombreuses matrices ont été développées pour le MALDI. Leur utilisation va dépendre du type de molécules analysées. En effet, le MALDI permet la détection et l'identification de plusieurs classes de molécules parmi lesquelles on retrouve les peptides, les protéines, les lipides, les sucres et les oligonucléotides. Le **tableau 2** résume les matrices utilisables en MALDI en fonction du type de molécules souhaitées. Il est important de noter que la majorité des molécules détectées sera sous la forme mono-chargée, facilitant ainsi l'interprétation des spectres de masse. De même, ce type d'ionisation dite douce, permet l'étude de molécules intactes, c'est pourquoi la source MALDI sera le plus souvent couplée à l'analyseur TOF (time of flight) qui permet théoriquement une gamme de masse infinie, on parle alors de MALDI-TOF.

<b>Composé</b>	<b>Abréviation</b>	<b>Utilisation</b>
<b>Acide <math>\alpha</math>-cyano-4-hydroxycinnamique</b>	HCCA	Peptide (moins de 10kDa) Hydrate de carbone
<b>Acide sinapinique</b>	SA	Protéine (plus de 10kDa) Dendrimers Fullerenes
<b>Acide 2-(4-hydroxyphénylazo)benzoïque</b>	HABA	Protéine (plus de 10kDa)
<b>Acide succinique</b>	-	Peptide et protéine, Laser IR Polymère synthétique, Laser IR
<b>2,6-dihydroxyacétophénone</b>	2.6 DHAP	Peptide et protéine, Laser UV
<b>Acide férulique</b>	-	Peptide et protéine, Laser UV
<b>Acide cafféique</b>	-	Peptide et protéine, Laser UV
<b>Glycérol</b>	-	Peptide et protéine, Matrice liquide

<b>4-nitroaniline</b>	-	Peptide et protéine, Matrice liquide
<b>2,4,6-trihydroxyacétophénone</b>	THAP	Oligonucléotide de moins de 3,5kDa Hydrate de carbone acide
<b>Acide 3-hydroxypicolinique</b>	HPA	Oligonucléotide de plus de 3,5kDa
<b>Acide anthranilique</b>	-	Oligonucléotide de plus de 3,5kDa
<b>Acide nicotinique</b>	-	Oligonucléotide de plus de 3,5kDa
<b>Acide trans-3-indoleacrylique</b>	IAA	Polymère synthétique non polaire
<b>Dithranol</b>	DIT	Polymère synthétique non polaire Lipide Dendrimers
<b>Acide 2,5-Dihydroxybenzoïque</b>	DHB	Polymère synthétique non polaire Molécule organique Hydrate de carbone
<b>Isovanilline</b>	-	Molécule organique
<b>3-aminoquinoline</b>	3AQ	Hydrate de carbone
<b>T-2-(3-(4-t-butyl-phényl)-2-méthyl-2-propénylidène)malononitrile</b>	DCTB	Molécule inorganique
<b>1-isoquinolinol</b>	-	Oligosaccharide
<b>Acide picolinique</b>	PA	Oligonucléotide
<b>2,5-dihydroxyacétophénone</b>	DHAP	Protéine (plus de 10kDa)

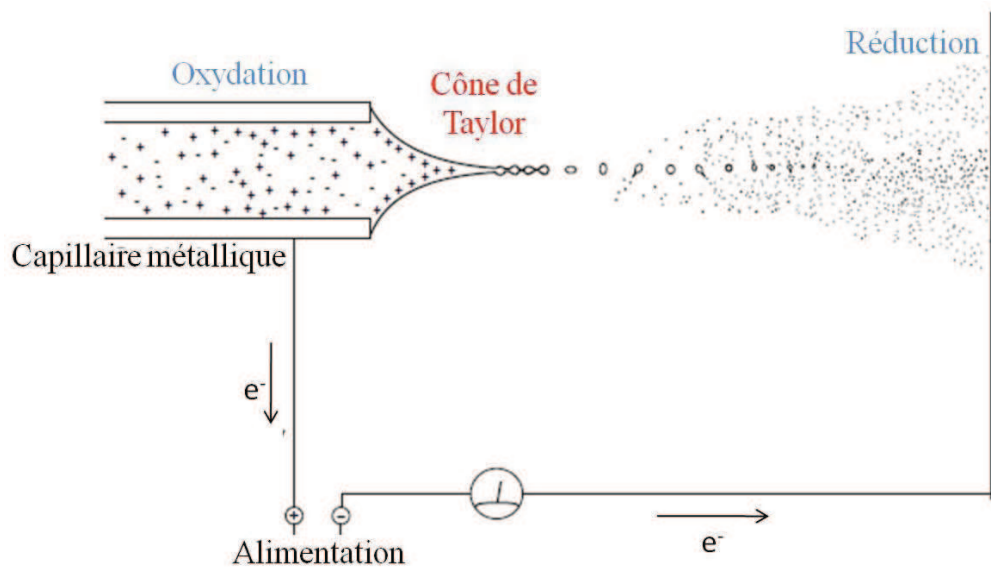
Tableau 2 : Listes des matrices utilisables en MALDI.

### La source ESI

L'ionisation ESI ou encore électro-nébulisation correspond à l'émission en phase gazeuse d'ions préformés en solution sous l'effet d'un champ électrique intense. Ce processus a lieu à pression atmosphérique contrairement au processus MALDI qui nécessite un vide poussé pour la formation des ions. Par ionisation ESI, l'échantillon est entraîné vers l'analyseur du spectromètre de masse à la fois par des différences de potentiels appliquées à des lentilles électrostatiques et par la différence de pression établie entre la source (pression atmosphérique) et l'analyseur (10<sup>-6</sup> – 10<sup>-7</sup> mbars). Les premiers travaux de Dole en 1968 sur

l'utilisation de l'ESI pour générer des ions à partir de polymères en solution n'ont pas été poursuivis à cette époque, en raison du manque de résultats pertinents obtenus [37]. L'intérêt de cette technique réapparaît avec les travaux de Fenn qui a pu, par un couplage entre cette source et un analyseur quadripolaire, effectuer les premières mesures de protéines de hauts poids moléculaires [38]. Ce couplage a également permis la mise en évidence de la nature multichargée des protéines durant ce processus, en présentant le premier spectre multichargé d'une protéine de 40kDa [39, 40]. A présent, le mécanisme macroscopique de l'électrospray est bien décrit. Même si quelques divergences persistent sur la désorption des ions de la solution vers la phase gazeuse, le processus d'ionisation-désorption ESI peut être divisé en 3 étapes majeures : la production des gouttelettes chargées à partir de l'électrolyte en solution, la fission des gouttelettes chargées en gouttelettes de plus petites tailles par explosions coulombiennes successives et l'émission des ions en phase gazeuse.

L'application d'un champ électrique intense (106 V/m) à la pointe du capillaire provoque une polarisation du liquide et une séparation des charges positives et négatives. L'accumulation de charges à la pointe du capillaire, déstabilise la surface du liquide qui prend alors la forme d'un cône appelé "cône de Taylor". Ce processus électrochimique entraîne la formation de gouttelettes chargées (**Figure 8**). Pour respecter l'équilibre des charges, la source ESI se comporte comme une cellule électrochimique [41, 42]. La nébulisation est par ailleurs assistée par un gaz de nébulisation (azote).



**Figure 8 :** Schéma du processus électrochimique de l'ESI. Tiré de [43].

Le solvant contenu dans les gouttelettes chargées va progressivement s'évaporer, entraînant une diminution de la taille de ces dernières. La densité de charge de la gouttelette va augmenter jusqu'à atteindre le rayon critique appelé le rayon de Rayleigh (**Equation 1**), moment où les forces de répulsion électrostatiques sont égales aux forces de tension de surface. Au moment où les forces de répulsion coulombiennes deviennent supérieures aux forces de cohésion du liquide, la gouttelette alors instable va exploser en gouttelettes filles, qui subiront à leur tour le même processus et ceci sur plusieurs générations de gouttelettes filles de plus en plus petites jusqu'à donner des ions complètement désolvatés.

$$Q = 8 \pi (\epsilon_o \gamma R_r^3)^{0.5}$$

**Équation 1** : Equation régissant le rayon de Rayleigh. (Q : charge de la gouttelette, R<sub>r</sub> : rayon critique de Rayleigh, ε<sub>o</sub> : perméabilité du vide et γ la tension de surface)

Jusqu'à aujourd'hui, la formation des ions à partir des gouttelettes chargées reste un sujet de controverse dans le domaine. Deux théories ont été proposées (**Figure 9**) :

- La théorie de Dole en 1968 [37] repose sur le modèle de la charge résiduelle. Une succession d'explosions coulombiennes aboutirait à des subdivisions et donc à la formation d'une dernière gouttelette ne contenant qu'un seul ion. Les dernières molécules de solvant comprises dans cette gouttelette vont s'évaporer afin d'aboutir à un ion complètement désolvaté.
- La théorie d'Iribarne et Thomson en 1976 [44] repose sur le principe de l'évaporation ionique. Les gouttelettes filles ont une densité de charge suffisante correspondant à un rayon précis ( $Q = 10^{-17}$  C et  $R < 10$  nm), elles ne subissent plus de fissions supplémentaires. Une fois cette densité de charges atteinte, les ions sont directement émis en phase gazeuse (désorption).

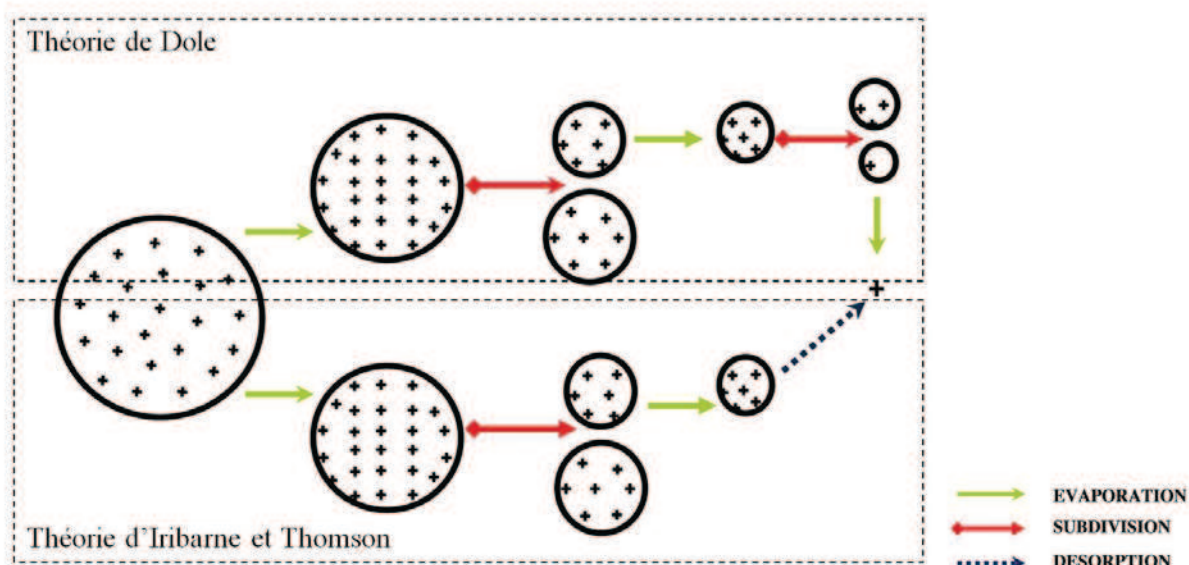


Figure 9 : Théorie sur l'émission des ions en phase gazeuse avec la source ESI. Tiré de [45].

Kebarle, en 2000, propose la théorie de Dole pour le processus d'ionisation des macromolécules et celle d'Iribarne et Thomson pour les plus petites molécules [46].

En mode d'ionisation electrospray, le spectromètre de masse détecte un signal (courant d'ions) dépendant de la concentration de l'analyte et non du débit avec lequel il est infusé [47, 48]. La diminution du débit permet aussi l'émission de plus petites gouttelettes à la pointe d'un cône de Taylor [49]. La plus grande densité de charge de ces gouttelettes contribue à l'augmentation du rendement d'ionisation/désorption des molécules, ce qui équivaut à une augmentation de la sensibilité. La diminution des débits (environ 50-100nL/min) prolonge le temps d'analyse, puisqu'avec 1 à 2 $\mu$ L d'analyte, l'analyse pourra durer environ 30 min.

Ainsi, pour répondre à ces nouvelles exigences, la source nano-électrospray ou nano-spray [50] a été développée. Cette source est constituée d'un capillaire en quartz étiré de 1 à 3  $\mu$ m de diamètre interne, recouvert d'une fine couche de métal qui assure sa conductivité. L'efficacité de l'ionisation nano-spray repose sur la faible taille des gouttelettes (environ 200 nm de diamètre), qui leur confère une densité de charge importante, un rapport surface/volume augmenté permettant une ionisation/désorption plus rapide [51]. Contrairement à l'ESI, ce montage ne nécessite plus, dans la majorité des cas, de gaz de nébulisation. De plus, la distance avec la contre-électrode étant plus courte pour éviter les décharges de Corona, les tensions appliquées sur le capillaire en quartz sont plus faibles (de l'ordre de 1000V). Ainsi, la miniaturisation des sources electrospray présente de nombreux

avantages pour l'analyse (tolérance aux sels non volatiles, augmentation de la sensibilité et, de ce fait, augmentation de la précision de la mesure de masse) [50]. Les diverses applications des sources nano-ESI, notamment dans le domaine de la protéomique, montrent aujourd'hui l'intérêt d'un tel dispositif [52], compatible avec des quantités de produit de l'ordre de la dizaine de femtomoles.

### 2.1.2.2 Les analyseurs

L'analyseur constitue la deuxième partie d'un spectromètre de masse (**Figure 5**). Suite à la formation d'ions en phase gazeuse dans la source, l'analyseur permet la séparation suivant le rapport masse-sur-charge ( $m/z$ ) de chaque molécule. Il existe plusieurs types d'analyseurs (**figure 6**) comme le quadripôle (Q), le temps de vol (TOF), la trappe ionique linéaire (IT) et l'orbitrappe qui en découle, à résonance cyclotronique (ICR) et enfin les secteurs magnétiques. Dans la grande majorité des cas, tous ces analyseurs peuvent être couplés avec n'importe quelle source d'ion. Il est également possible de coupler une source avec plusieurs analyseurs identiques ou différents.

Les analyseurs peuvent être comparés à l'aide de deux paramètres : la résolution spectrale et le pouvoir résolutif (**Figure 10**).

- La résolution spectrale est cruciale puisqu'elle permet de déterminer l'état de charge d'une molécule, pour caractériser les massifs isotopiques et pour discriminer les ions isobares (qui posséderaient le même nombre de nucléons). Celle-ci se calcule à partir de la différence en masse qui existe entre deux pics (**Figure 10A**).
- Le pouvoir résolutif est important puisqu'au plus sa valeur est élevée, plus le pic sera fin et donc plus la masse du pic sera précise. Il se calcule en divisant la masse par la largeur du pic à mi-hauteur (**Figure 10B**), on parle également dans ce cas de résolution FWHM (Full Width at Half Maximum).

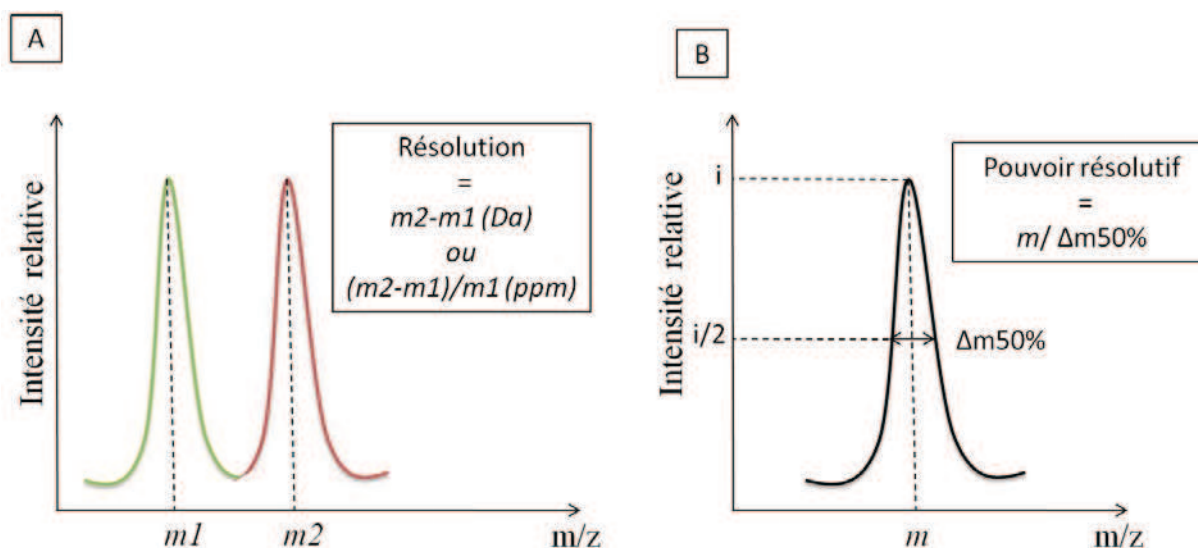


Figure 10 : Calcul de la résolution spectrale (A) et du pouvoir résolutif (B) d'un analyseur

L'utilisateur va donc choisir l'instrument en fonction du type de molécules qu'il souhaite analyser, des caractéristiques de la source et de(s) l'analyseur(s). En effet, chaque type d'analyseur aura un pouvoir résolutif, une précision de mesure et une gamme de masse qui lui est propre (**tableau 3**). Les constructeurs sont aujourd'hui dans une course concurrentielle effrénée pour l'amélioration de ces caractéristiques, permettant ainsi l'accès à des machines de plus en plus performantes pour la recherche. On peut considérer que les analyseurs les plus utilisés en protéomique sont le TOF, le quadripôle et la trappe ionique.

Analyseur	Pouvoir résolutif	Précision de mesure (ppm)	Gamme de masse (m/z)
ICR	1 000 000	1	< 6000
Orbitrappe	100 000	3	< 5000
TOF	25 000	50	∞
Trappe ionique	20 000	30	< 4000
Quadripôle	15 000	10	< 4000

Tableau 3 : Comparaison des analyseurs actuels en fonction de leur pouvoir résolutif, de leur précision de masse et de leur gamme de masse.

## L'analyseur TOF

Le TOF est sans doute le plus simple des analyseurs. En effet, suite à l'accélération des ions produits par la source par une différence de potentielle décroissante, ceux-ci rentrent dans le tube de vol qui est une zone libre de champ avec un vide poussé et se déplacent en fonction de la vitesse qu'ils auront acquise. Cette vitesse est directement liée à la masse des molécules. On distingue deux modes de séparation dans le TOF : le mode linéaire et le mode réflecteur.

En mode linéaire, les ions percutent le détecteur dans l'axe de la source. Plus la masse d'un ion sera élevée, moins sa vitesse sera importante et donc moins il percutera le détecteur rapidement (**Figure 11**). Les caractéristiques de ce mode sont : une gamme de masse de ce mode illimitée, une grande sensibilité, mais une mauvaise résolution spectrale.

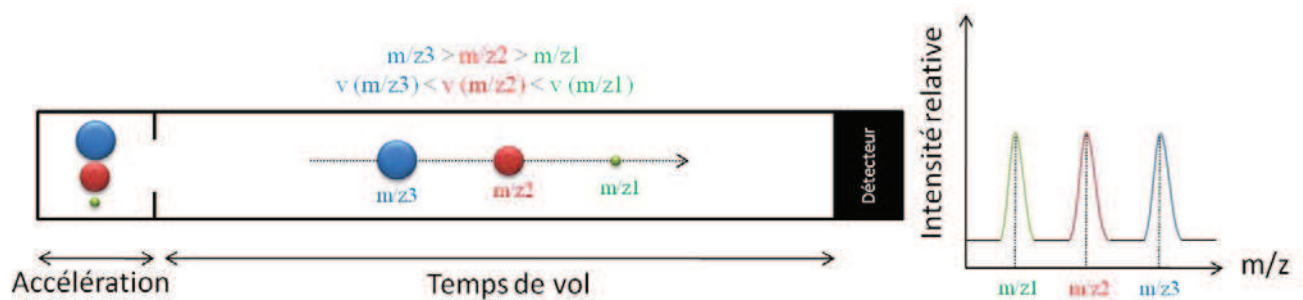


Figure 11 : Schéma du mode linéaire d'un TOF.

Pour pallier au problème de résolution spectrale, il est possible d'utiliser le TOF en mode réflecteur ou réflectron. En effet, certains ions ayant la même masse moléculaire peuvent posséder une variation de leurs énergies suite à un décalage au niveau de leur désorption et de leur ionisation. Leurs vitesses seront donc différentes. L'utilisation d'un réflecteur électrostatique composé de séries d'anneaux sur lesquelles des potentiels croissants sont appliqués permet de refocaliser les ions de même masse vers le deuxième détecteur. En fonction de l'énergie cinétique qu'ils possèdent, les ions pénétreront plus ou moins profondément dans ce miroir électrostatique et les ions de même rapport  $m/z$  vont donc atteindre le détecteur au même moment (**Figure 12**). Cet allongement du tube de vol permet donc une meilleure séparation des ions et une refocalisation des ions de même masse aboutissant à l'amélioration de la résolution spectrale par rapport au mode linéaire.



Néanmoins, ce mode présente des limites au niveau de sa gamme de détection (7000 m/z, d'où une moins bonne sensibilité.)

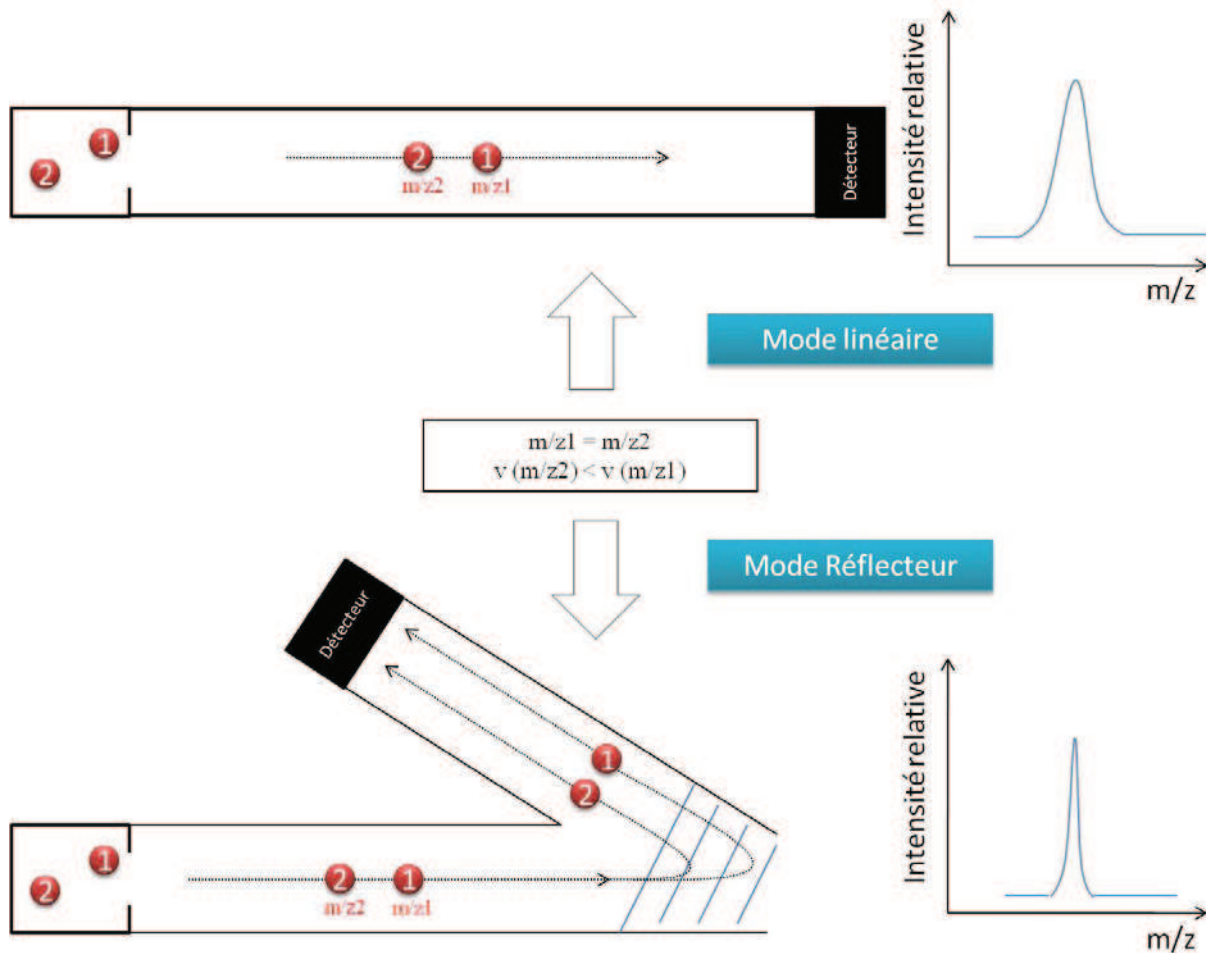


Figure 12 : Schéma du principe et de l'intérêt du mode réflecteur d'un TOF par rapport au mode linéaire

C'est pourquoi en protéomique, l'analyse des protéines jusqu'à 20 000 Da s'effectue en mode linéaire et l'analyse des peptides sera privilégiée avec le mode réflecteur d'un TOF.

### La trappe ionique

Contrairement au TOF et au Quadripôle, la trappe n'est pas traversée par les ions mais les piègent dans un espace d'environ 1cm<sup>2</sup>, on parle d'ailleurs de piège à ions. La trappe ionique est constituée de trois électrodes : deux électrodes chapeaux et une électrode annulaire (**Figure 13**). Les ions, provenant d'une source ESI ou nanoESI, vont pénétrer dans la trappe via des ouvertures au niveau des électrodes chapeaux. Une pression résiduelle d'hélium de

l'ordre de  $5 \cdot 10^{-5}$  bars est appliquée au sein de cette zone et permet d'augmenter l'efficacité du piégeage.

Une fois que les ions ont pénétré dans la trappe, une tension oscillante appliquée à l'électrode annulaire permet de générer un champ quadripolaire qui stabilise la trajectoire des ions en les piégeant dans un puits de potentiel dont la profondeur est proportionnelle à l'amplitude de la tension appliquée. Les ions décrivent alors une trajectoire appelée courbe de Lissajou, qui est régie par les équations de Mathieu. Les ions pourront être éjectés de deux manières différentes :

Une éjection simple : une diminution de l'amplitude de la tension réduit la profondeur du puits de potentiel et permet aux ions de  $m/z$  croissant de s'échapper du piège.

Une éjection résonante : les ions de  $m/z$  croissants, chacun oscillant avec une fréquence qui lui est propre, sont soumis à un champ électrique, oscillant, supplémentaire, appliqué entre les électrodes chapeaux et choisi spécifiquement pour faire entrer les ions de  $m/z$  croissant en résonance et donc les éjecter prématurément.

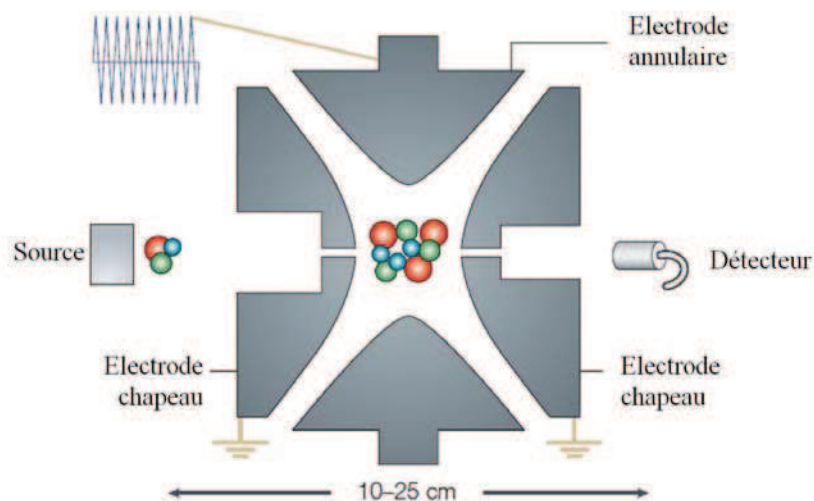


Figure 13 : Description de l'analyseur de type trappe ionique [53]

### 2.1.2.3 L'identification des peptides et des protéines

#### **Les empreintes peptidiques**

Cette stratégie fut décrite, dès 1993, par plusieurs équipes [54-56] comme alternative au séquençage peptidique d'Edman.

L'approche MALDI-MS permet l'identification rapide des protéines, sans séquençage, directement à partir de leurs cartes peptidiques massiques et par confrontation aux cartes théoriques générées *in silico*. Le crédit accordé à l'identification dépend du nombre de peptides reconnus, du pourcentage de recouvrement de séquence de la protéine, ainsi que de la précision de mesure. Le spectromètre de masse MALDI-TOF a été choisi, car il est adapté à ce type d'analyse :

- Pour sa précision de mesure.
- Pour sa tolérance aux sels et aux tampons.
- Pour sa sensibilité, détection de quantités femtomolaires, voire attomolaires.
- Pour sa possibilité d'analyse à haut-débit grâce à l'automatisation.
- Pour ses spectres directement exploitables. Seuls les ions monochargés sont détectés, donc aucune étape de déconvolution ne sera nécessaire contrairement à l'ESI.

Néanmoins il peut arriver que l'identification par cette approche échoue pour plusieurs raisons :

- Pourcentage de recouvrement insuffisant.
- Précision de mesure médiocre (supérieure à 50 ppm).
- Suppression de signal en fonction de la composition en acides aminés des peptides ou de la présence de contaminants.
- Peptides générés par la digestion trypsique (inférieurs à 800 Da) confondus avec les pics de matrice.
- Protéine absente des banques de données.

Si la stratégie de l'empreinte peptidique massique ne permet pas d'identifier les protéines, le recours à l'approche MS/MS sera envisagée afin d'obtenir des informations de séquences.

## La fragmentation

Certains appareils permettent, suivant leur type d'analyseur ou leur couplage, d'effectuer des fragmentations en mode MS/MS ou MS<sup>n</sup> des molécules d'intérêt pour remonter à leurs séquences en acides aminés. En effet, les peptides se fragmentent principalement au niveau de la liaison peptidique (ions y et b). Mais d'autres fragments ont également été observés de part et d'autre de cette liaison (ions a, c et x, z). Une première nomenclature a été introduite par Roepstorff et Fohlmann en 1984 [57]. Elle a ensuite été complétée et simplifiée par Biemann quelques années plus tard et fait toujours référence dans ce domaine [58] (**Figure 14**).

Les 2 types d'ions issus de la fragmentation à basse énergie de la liaison peptidique sont les suivants :

- Ions a, b, c pour lesquels la charge positive est portée par l'acide aminé N terminal :
  - Ions b issus de la fragmentation de la liaison peptidique (CO---NH).
  - Ions a issus de la coupure en amont de la liaison peptidique (CH---CO).
  - Ions c issus de la coupure en aval de la liaison peptidique (NH---CH).
  
- Ions x, y, z pour lesquels la charge positive est portée par l'acide aminé C terminal :
  - Ions y issus de la fragmentation de la liaison peptidique (CO---NH).
  - Ions x issus de la coupure en amont de la liaison peptidique (CH---CO).
  - Ions z issus de la coupure en aval de la liaison peptidique (NH---CH).

D'autres fragments peuvent être observés dans le cas des fragmentations à haute énergie :

- Ions d pour lesquels la charge positive est portée par l'acide aminé N terminal suite à la coupure de la chaîne latérale après le carbone  $\beta$ .
- Ions w et v pour lesquels la charge positive est portée par l'acide aminé C terminal.
  - Ions v issus de la coupure de la chaîne latérale après le carbone  $\alpha$ .
  - Ions w issus de la coupure de la chaîne latérale après le carbone  $\beta$ .

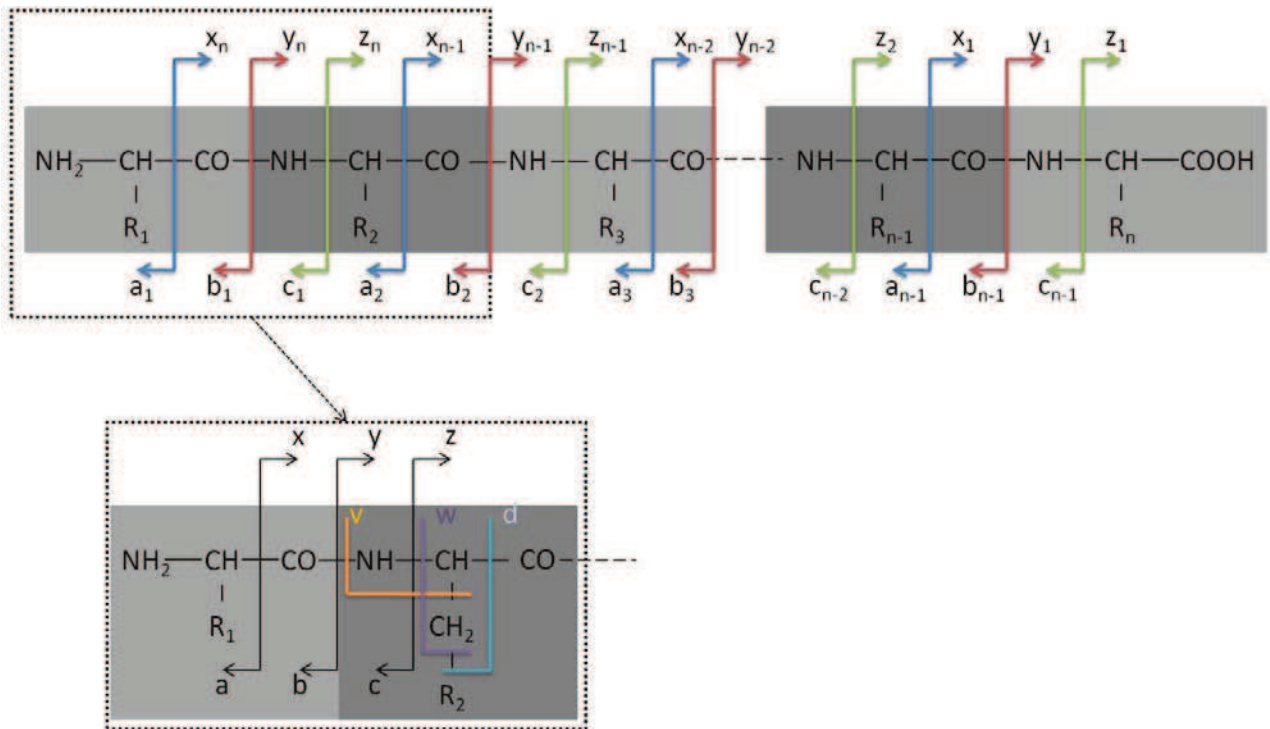


Figure 14: Nomenclature de fragmentation des peptides selon Biemann. (R désigne les chaînes latérales)

Le plus souvent, une liste de  $m/z$  associés à leur intensité relative est envoyée dans les banques de données via des logiciels d'interrogation (Mascot, Phenyx) pour identifier la protéine à partir de laquelle le peptide fragmenté provient. Notons qu'il est possible de combiner les spectres de fragmentation de peptides issus d'une digestion enzymatique au PMF de la protéine pour améliorer ou aboutir à une identification certaine.

### Le séquençage de novo

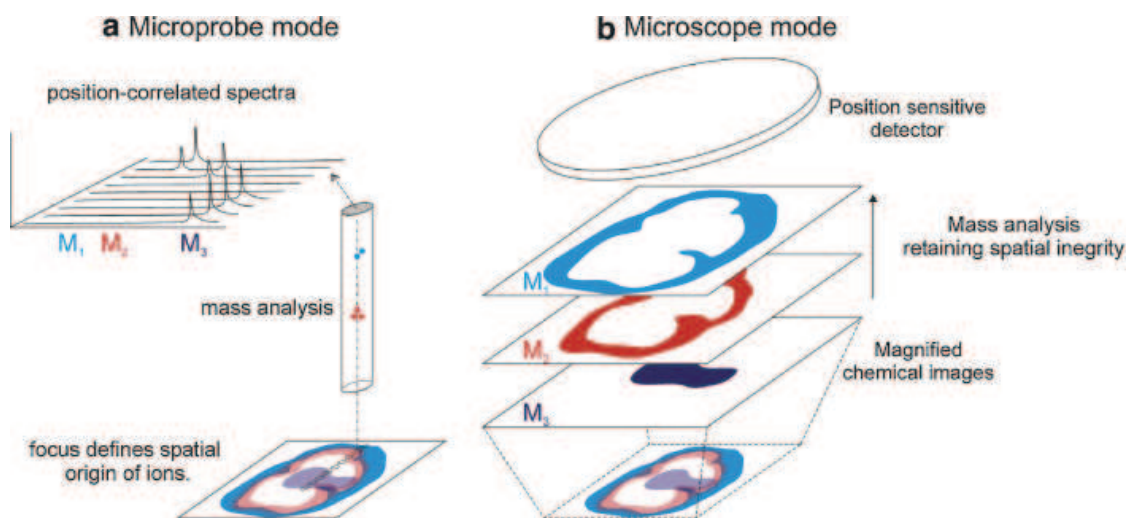
Si l'interrogation en banque de données des spectres de fragmentation via les listes de masse n'aboutit à aucune identification, il est possible de faire du séquençage de novo. Cette approche consiste en une analyse manuelle ou informatisée des spectres MS/MS pour déterminer des séquences hypothétiques. Certains logiciels intègrent maintenant cette fonctionnalité, comme Biotools de la société Bruker. Les séquences potentielles sont ensuite envoyées dans des logiciels de « blast » pour comparaison dans les banques de données de la séquence trouvée, mais ceci exclut toute utilisation de listes de  $m/z$  et donc de tolérance instrumentale.

### 2.1.3 L'imagerie par spectrométrie de masse

L'observation macroscopique puis microscopique des êtres, organismes, organes et cellules est définitivement l'un des moyens ayant conduit à une meilleure appréhension de leur fonctionnement. Des premiers microscopes d'Antoni van Leeuwenhoek (1632-1723) aux techniques d'imagerie cliniques in vivo les plus récentes comme l'IRM et le PET-SCAN, toutes ces techniques permettent une meilleure compréhension des problématiques au travers de la visualisation.

L'imagerie par spectrométrie de masse (MSI) est une technique d'imagerie moléculaire qui découle des analyses en profiling MALDI comme utilisées par pour l'étude des neuropeptides [59]. La puissance intrinsèque de la MSI repose sur la capacité de la MS à distinguer les différentes espèces moléculaires sur la base de leur masse. La MS repose sur la formation d'ions en phase gazeuse dans la source de l'instrument puis la séparation de ces ions en fonction de la quantité  $m/z$ . Une image se compose de pixels dont la taille et le nombre définissent la qualité. Ainsi, la réalisation d'images en MS requiert la capacité de générer des informations moléculaires à partir d'une surface décomposée en pixels adjacents qui serviront de base à la reconstruction de l'image finale. L'acquisition des données à partir de l'échantillon nécessite donc de pouvoir travailler à partir d'une surface. Pour cette simple raison, toutes les sources de production d'ions connues en MS ne pourront être utilisées en mode imagerie, les sources les plus adaptées se révélant être les sources fonctionnant par bombardement de l'échantillon par un faisceau primaire incident. Ainsi, les sources de production d'ions par bombardement par un faisceau d'ions primaires fortement accélérés (sources SIMS), un faisceau laser (sources LDI, MALDI, LA-ICP) ou un faisceau de gouttelettes chargées (sources DESI) ont conduit à l'émergence de différentes techniques de MSI. L'émergence de la MSI est relativement ancienne et remonte aux années 60, suite aux travaux de Castaing et Slodzian en 1962 [60] sur les sources SIMS. Depuis ces premiers travaux, d'autres techniques de MSI sont apparues. L'avènement des sources LDI a conduit au développement dans le milieu des années 70 de l'imagerie LDI avec l'arrivée du LAMMA (Laser Microprobe Mass Analyzer) [61]. Plus tardivement, les sources MALDI introduites en 1988 par Hillenkamp et Karas [33] ont conduit à l'imagerie MALDI. Il faut attendre 1997 pour voir apparaître les premières images MALDI suite aux travaux de Caprioli et coll. [62, 63]. Finalement, les sources DESI apparues en 2004 [64] ont rapidement conduit à des applications en imagerie [65].

A l'heure actuelle, deux modes d'acquisition des données ont été explorés en imagerie par spectrométrie de masse, le mode microsonde et le mode microscope. Le mode microsonde reste le mode le plus simple, il est utilisé dans la plupart des études. En effet, il est facilement mis en place sur les instruments les plus courants comme les appareils commerciaux. En mode microsonde (**Figure 15a**), l'échantillon est analysé de manière automatique pixel par pixel. Ainsi la réalisation du mode microsonde nécessite un contrôle de l'aire touchée par le faisceau incident. Pour chaque pixel balayé un spectre de masse est enregistré avec ses coordonnées correspondantes sur la surface. Par la suite l'image est reconstruite à partir de la translation sur une échelle de couleur des variations de l'intensité (ou de l'aire) du pic correspondant à cet ion dans la collection de spectre. En mode microscope (**Figure 15b**), la source d'ionisation est utilisée pour produire les images de la totalité de la coupe comme le fait la microscopie optique via les photons. Ainsi l'information spatiale est préservée à l'intérieur du spot puisque les ions conservent leur position relative pendant leur déplacement de la coupe jusqu'à un détecteur de position. Par conséquent, cette technique est indépendante des dimensions de l'aire irradiée par le faisceau incident. La résolution spatiale ne dépend alors que du grossissement du microscope et de la résolution du détecteur de position d'ions utilisé. Cependant les instruments sont extrêmement complexes de part la grande précision de l'optique mise en jeu pour guider les ions jusqu'au détecteur. Ce type d'approche n'est donc utilisé que dans des laboratoires ayant développé un instrument compatible et il n'existe à ce jour aucun appareil commercialisé pour l'imagerie par spectrométrie de masse en mode microscopique.



**Figure 15 : Représentation des deux approches en imagerie par spectrométrie de masse. (a) : le mode microsonde, (b) : le mode microscope. [66].**

Classiquement la MSI est réalisée à partir d'une surface mince pour des raisons d'encombrement dans la source de l'instrument, de perturbation des paramètres de cette dernière ou afin d'éviter de fausser la mesure des m/z. Les surfaces étudiées peuvent être de nature variée allant de surface de matériaux organiques ou inorganiques à des surfaces provenant d'organismes animaux ou végétaux. Différentes techniques peuvent être utilisées pour réaliser des sections à partir d'objets. Pour les applications biologiques, sont en particulier utilisées les coupes à l'aide de lames tranchantes à froid par cryostat ou à température ambiante par microtome.

Les techniques de MSI présentent différents atouts par rapport à d'autres technologies d'imagerie qui sont en particulier :

- i) La possibilité de reconstruire plusieurs images à partir d'une acquisition (précisément autant d'images que d'ions détectés),
- ii) L'aspect non ciblé de la méthode qui ne requiert pas de marquage et donc la connaissance de l'objet à observer,
- iii) La possibilité d'aller vers l'identification d'un composé d'intérêt,
- iv) et la quantification des composés regardés
- v) La capacité d'être applicables à un grand nombre de familles de composés différents (dans le cas d'un organisme, métabolites, lipides, peptides, protéines, composés exogènes),

Selon la source d'ionisation utilisée, la MSI présentera différentes caractéristiques en termes de résolution des images (résolution spatiale) ou de gamme de masse qui fixera le type de composé accessible. Elles ont chacune leurs avantages et inconvénients. Les caractéristiques des différentes sources sont regroupées dans le **Tableau 4**.



Source d'ionisations	Résolution spatiale	Gamme de masse	Référence
SIMS	100nm-10 $\mu$ m	m/z<2000	[60]
MALDI	25-200 $\mu$ m	0-1 000 000	[62]
DESI	500 $\mu$ m	m/z<2000	[67]
DIOS	Cellules isolées	m/z<5000	[68]
ICP	/	m/z<500	[69]
Microscope MALDI	<5 $\mu$ m	5- 100 000	[70]
SNOM	50 nm-1 $\mu$ m	/	[71]
SMALDI	500nm-10 $\mu$ m	m/z<2000	[72]
LA-ICP	40-100 $\mu$ m	m/z<1000	[73]
ESI	> 300 $\mu$ m	m/z<1000	[74]
MILDI	100 $\mu$ m	5-10 000	[75]
JEDI	1 $\mu$ m	m/z<5000	[76]
ELDI	150 $\mu$ m	5-100 000	[77]

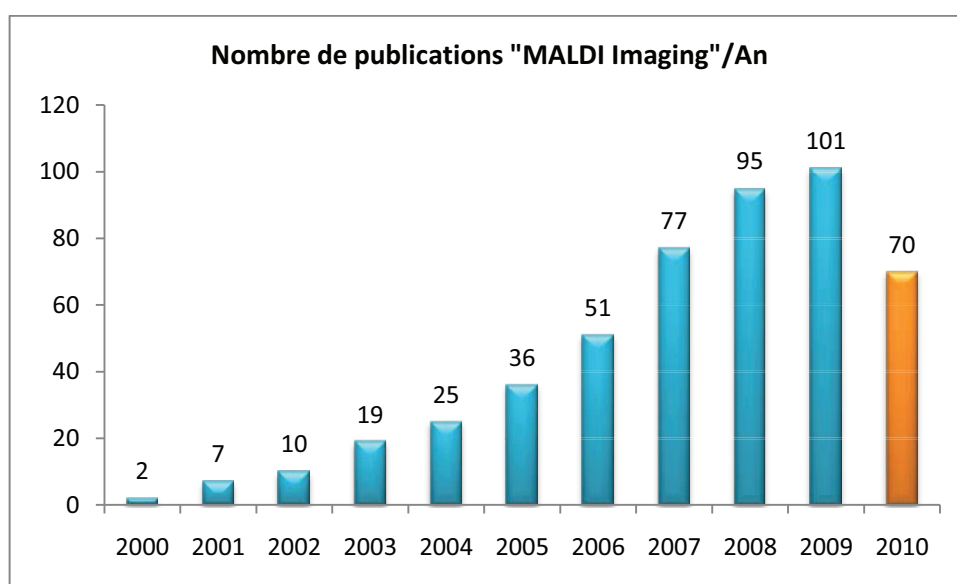
**Tableau 4: Liste des différentes sources d'ionisation et de leurs caractéristiques utilisées en MSI.**

Si toutefois la MSI présente moult avantages, il ne faut cependant pas oublier qu'elle présente pour le moment l'inconvénient d'être une technique destructrice et non applicable *in vivo* sur des organismes vivants, contrairement à l'IRM. Des travaux préliminaires ont permis de démontrer l'utilisation de la MSI directement sur des patients au bloc opératoire [78], mais ces premiers résultats ne constituent que les prémices de l'utilisation de la MSI *in vivo*.

### 2.1.3.1 L'imagerie MALDI

Les premières expériences d'analyse directe par MALDI ont été décrites en 1993 peu après l'introduction de ces sources. Li et coll. ont ainsi réalisé l'analyse directe de neurones géants de l'invertébré *Lymnaea stagnalis* [79]. En 1994, ces mêmes auteurs se sont par la suite intéressés à l'analyse de tissus disséqués à partir de ce même modèle animal. La même année, l'équipe du Pr. Richard Caprioli a présenté la première description de l'analyse directe d'une section de tissu mince d'un organe entier par imagerie MALDI. L'intérêt de pouvoir récolter des données MALDI tout en apportant une dimension de localisation sur les tissus a rapidement orienté les recherches vers la possibilité de passer du profilage des tissus à la réalisation des images. L'automatisation des méthodes d'acquisition en mode microsonde et du traitement des données a ainsi permis en 1997 la publication de la première image MALDI [62]. C'est cependant en 1999 que commence à être présentées les premières images réalisées à partir de sections minces de tissus animaux (sections de cerveau de rat) en routine. Parallèlement à ces développements, de nombreux travaux sur le profilage des tissus et l'analyse directe de cellules uniques ont été réalisés, confortant ainsi l'intérêt de la méthode pour les applications notamment biologiques [63, 80-83].

Depuis l'imagerie MALDI a connu de nombreux développements tant sur les plans technologique, instrumental, analytique que stratégique. Ces développements ont contribué à l'essor de la technologie et expliquent l'augmentation très nette du nombre de publication référençant le terme « MALDI Imaging » (**Figure 16**).



**Figure 16 : Nombre de publications scientifiques internationales par an impliquant la technologie d'imagerie MALDI de 2000 à Juillet 2010 (source : PubMed).**

La réalisation d'une image fait appel à différentes étapes classiquement rencontrées pour toutes les techniques de MSI. Une séquence d'imagerie comprend ainsi 5 phases qui correspondent (I) au prélèvement du tissu et son traitement pour la conservation, (II) à la réalisation de la section mince et le traitement du tissu, (III) au dépôt de la matrice MALDI, (IV) à l'acquisition de l'image et (V) au traitement des données permettant la reconstruction de l'image. La **Figure 17** présente une séquence d'imagerie MALDI classique avec ses 5 étapes en prenant en compte une acquisition des plus classiques en mode microsonde. L'étape spécifique à l'imagerie MALDI par rapport aux autres imageries telles que l'imagerie SIMS ou LDI est l'étape (III) correspondant au dépôt de la matrice MALDI. Nous reviendrons plus en détail dans la suite de ce chapitre sur ces différentes étapes et en particulier l'étape (III) qui joue un rôle fondamental sur la qualité des données générées.

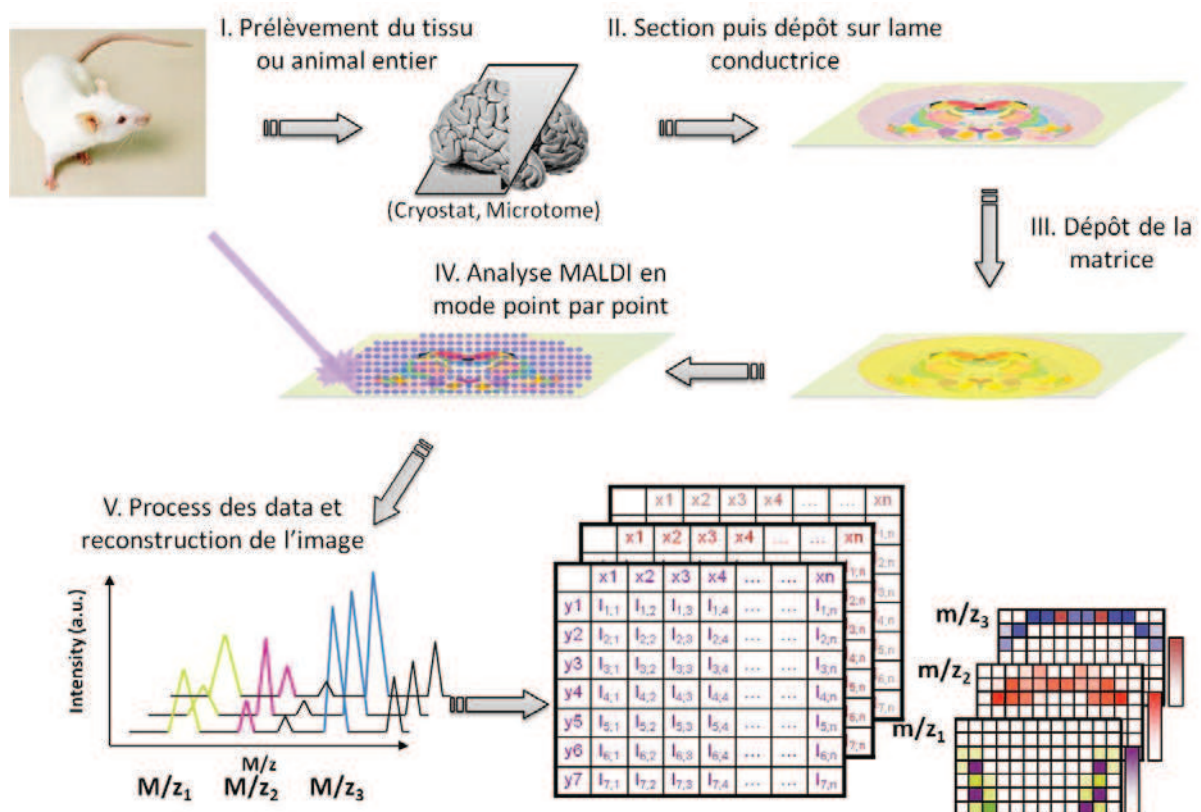
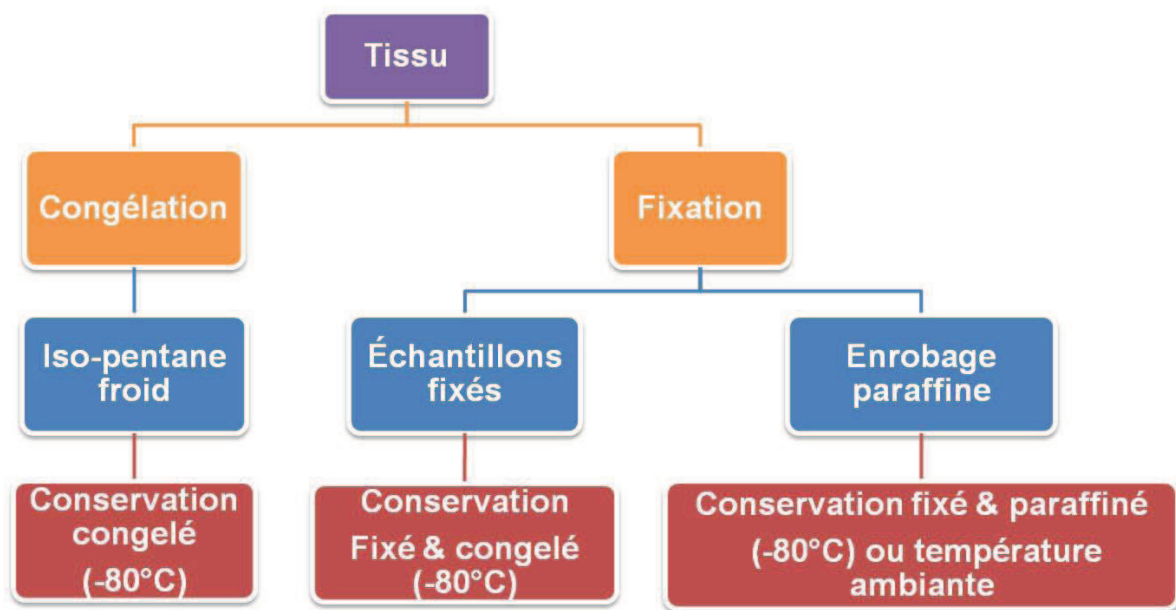


Figure 17 : L'imagerie MALDI : de la préparation de l'échantillon à la reconstruction des images.

## Etape 1 : Dissection et préparation de l'organe pour conservation et analyse

Différentes méthodes de préparation et de conservation des tissus sont couramment employées en biologie pour des domaines tels que l'histologie, l'immunohistochimie ou encore la microscopie électronique. Le principal objectif est la préservation de la morphologie du tissu. Pour les méthodes s'intéressant à un aspect moléculaire il est également important de s'intéresser à préserver les molécules par elles-mêmes des dégradations ou modifications (e.g. hydrolyses, oxydations,...). Globalement, les techniques utilisées classiquement pour l'histologie et l'IHC se sont révélées compatibles avec l'imagerie MALDI et se divisent en trois types de préparation pour conservation comme illustré par la **figure 18**. Ainsi, les échantillons après dissection peuvent être conservés par congélation à basse température (-80°C), par congélation après fixation du tissu ou à température ambiante après fixation des tissus et inclusions en paraffines (tissu FPE ou Fixed and Paraffin Embedded).



**Figure 18 : Préparation et conservation des tissus pour l'imagerie MALDI**

La conservation par simple congélation est de loin le protocole le plus simple et le plus rapide. La congélation est pratiquée par refroidissement rapide (max 1 min) de l'échantillon dans un bain d'isopentane refroidi à l'azote liquide. Ce protocole particulier permet la congélation de l'échantillon tout en offrant la meilleure préservation morphologique possible du tissu qui ne serait pas obtenue par un refroidissement trop soudain dans un bain d'azote liquide par exemple. L'échantillon est par la suite conservé à -80°C jusqu'à utilisation.

Alternativement, le tissu peut-être fixé en lieu et place de la congélation rapide. La fixation est pratiquée pour stabiliser l'échantillon et obtenue à l'aide de réactifs variés. Généralement la stabilisation est obtenue par utilisation d'un agent de pontage tel que le formalin ou le glutaraldéhyde. La fixation au formalin, par exemple, induit un pontage des molécules. Cependant, les mécanismes de fixation sont généralement mal décrits du point de vue des réactions chimiques qu'elles impliquent. Ainsi le formalin induit notamment un pontage des protéines par les groupements amines mais d'autres sous-réactions peuvent aussi intervenir. Le processus de fixation du formaldéhyde est décrit comme de cinétique rapide, mais à l'inverse la formation du pont par lui-même est plus longue et peut se poursuivre dans le bloc de tissus bien après que ce dernier ait été retiré de la solution de fixation. La fixation au formalin, qui est l'une des plus couramment utilisées, se pratique en plongeant le tissu dans une solution de formaline à quelques pourcents (généralement 4%) et à 60°C. Le temps de fixation doit être étudié en fonction du volume de la pièce à fixer et peut aller de quelques heures à plusieurs jours afin de permettre au fixateur de pénétrer l'intégralité du bloc. Il faut également prendre en compte que l'effet de telles fixations sur d'autres familles de molécules est assez mal connu. C'est le cas pour les lipides : en effet, il n'est pas clair si ceux-ci sont fixés, emprisonnés dans le maillage créé par les protéines, ou si certains peuvent être délocalisés ou retirés du tissu par le processus. Il est donc particulièrement important de bien adapter le choix de la préparation aux études qui seront menées par la suite. Les tissus fixés peuvent être alors conservés tels quels congelés, à basse température (-80°C), mais le plus souvent ils subissent un enrobage dans la paraffine après la fixation. Les tissus fixés au formalin et inclus en paraffine sont couramment nommés tissus FFPE (Formalin Fixed and Paraffin Embedded). Les tissus FFPE sont les tissus les plus fréquemment rencontrés dans les banques hospitalières. En effet, cette fixation rend les tissus particulièrement stables dans le temps. Les blocs peuvent être conservés plusieurs décennies sans qu'aucune dégradation des protéines ne soit observée, celles-ci étant pontées de manière irréversible. Les tissus FFPE qui ont l'avantage d'être stables dans le temps présentent également l'avantage de permettre la réalisation de coupes minces avec une excellente préservation des structures fines morphologiques. La préservation est alors supérieure à celle obtenue par coupe sur des tissus conservés par simple congélation. Cependant, ces tissus sont plus difficilement compatibles avec l'analyse par spectrométrie de masse et des stratégies particulières ont dû être développées pour permettre leur utilisation en imagerie MALDI [84, 85]. Les tissus congelés sont, de bien loin, beaucoup plus faciles pour l'imagerie MALDI. Cependant ces tissus ne sont que peu stables dans le temps et malgré une conservation à -80°C des dégradations de

certaines composés comme les protéines sont observées après plusieurs mois de conservation. Récemment, des méthodes permettant de stabiliser les tissus conservés par congélation ont été proposées. Ces méthodes font appel à un chauffage rapide de l'échantillon permettant entre autre d'inactiver les enzymes conduisant à la dégradation des protéines. Cette solution est proposée de manière commerciale par la société Denator [86]. Cependant, elle présente l'inconvénient de réduire la préservation de la morphologie des tissus.

Les tissus conservés par congélation sont ensuite sectionnés à froid ( $\sim -20^{\circ}\text{C}$ ) à l'aide d'un cryostat. Cette étape doit être réalisée avec précaution. En effet, les blocs sont maintenus sur leur support à l'aide d'une solution (OCT, Optimal Cutting Temperature) de polymères se solidifiant à basse température. Ainsi, si l'OCT ne gêne pas les expériences d'histologie ou d'IHC la contamination du tissu par l'OCT lors de la pose du bloc ou de la réalisation des coupes (traces d'OCT sur la lame du rasoir) peut engendrer de nombreux signaux parasites dans les spectres MALDI (en particulier distribution liée à la présence de PEG) qui masqueront en grande partie les signaux d'intérêt. Pour les tissus FFPE, les coupes sont classiquement réalisées à l'aide d'un microtome. Après réalisation de la section les coupes sont transférées sur une lame de verre. L'utilisation d'une lame de verre offre l'avantage de permettre les observations microscopiques avant l'analyse. Suivant l'instrument utilisé pour l'imagerie MALDI (i.e. dépendamment de l'analyseur) les lames de verres utilisées sont des lames classiques (e.g. orbitrappe) ou conductrices par coating avec une fine couche de matériau conducteur transparent (e.g. ITO Indium Teen Oxyde) (e.g. analyseurs TOF).

## **Etape 2 : Pré-traitement des tissus**

Préalablement au dépôt de la matrice MALDI, des traitements des tissus peuvent être réalisés. Ces traitements ont pour objectif de permettre d'accéder à une meilleure qualité d'analyse (nombre de composés observés, intensité des signaux, rapport signal/bruit) ou de développer des stratégies permettant de dépasser le cadre de la simple réalisation d'images. Ainsi dans une séquence classique d'imagerie les traitements permettent d'améliorer les spectres et d'accéder à l'analyse de certaines familles de composés plus difficiles d'accès dans des conditions standards de préparation. Il est également possible de dépasser ce cadre pour rechercher par exemple une identification des composés *in situ* qui nécessiteront la mise en place de stratégies particulières et recourant à de la chimie à la surface des échantillons.

*Traitement des échantillons dans le cadre de la réalisation d'images.* Les recherches menées par les différents groupes ont toujours fait état de l'importance de la préparation des échantillons en MALDI et de l'impact possible sur la qualité des résultats. Ici, le problème est complexifié par la nature de l'échantillon examiné, le tissu. Il apparaît donc clairement intéressant de pouvoir jouer sur les propriétés physico-chimiques du système (en particulier solvants et matrice) pour accéder plus spécifiquement à certaines familles de composés en favorisant leur extraction des tissus et leur solubilisation ou en supprimant d'autres composés plus majoritaires et pouvant gêner leur détection. A l'heure actuelle il n'apparaît pas possible d'analyser tous les composés constituant le tissu en une seule analyse, la complexité du milieu entraînant l'existence de nombreux phénomènes de suppression des ions. Ainsi, afin d'observer certains composés il faut trouver les conditions favorables à leur analyse tout en supprimant les composés facilement observables mais sans intérêt dans les études menées. La possibilité de jouer avec les propriétés physico-chimiques des composés offre un moyen unique de pouvoir, presque à volonté, contrôler les molécules qui seront observées. Ainsi, dès le début des années 2000, l'éthanol a intuitivement été utilisé pour augmenter la qualité des spectres en favorisant notamment la cristallisation de la matrice [87]. Un lavage du tissu à l'éthanol permet d'éliminer non seulement une grande partie des sels présents qui peuvent gêner la cristallisation de la matrice mais également certaines protéines abondantes comme l'hémoglobine. Ainsi, le lavage des tissus par des solvants est un procédé intéressant mais qu'il faut bien maîtriser et dont il faut être conscient de ce qui peut être éliminé ou délocalisé par le traitement choisi. En 2006, Lemaire et coll. [88] se sont intéressés à étudier l'impact de l'utilisation de différents traitements des tissus par des solvants organiques (chloroforme, acétone, hexane, toluène, benzène,...) sur l'analyse des peptides et des protéines de masse faible ( $M_w < 20\ 000$ ). Les expériences ont montré que le lavage des tissus permettait d'augmenter nettement le nombre de composés détectés et l'intensité des signaux pour les peptides et protéines. Il a également pu être montré que cette augmentation s'accompagnait d'une drastique diminution de l'intensité des signaux de faible  $m/z$  qui ont pu être identifiés comme appartenant à des lipides (**Figure 19**).

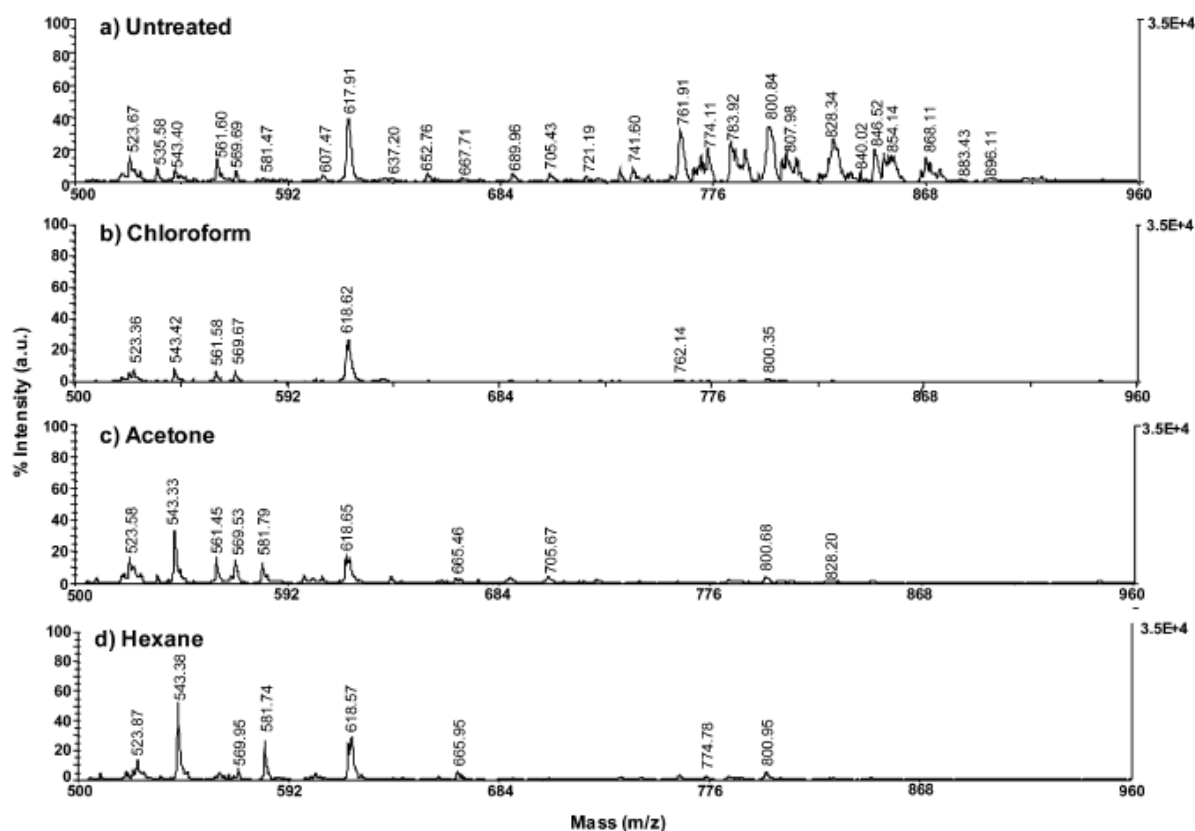


Figure 19 : Effet de différents traitements chimiques sur le signal détecté lors de la préparation d'un échantillon pour l'imagerie MALDI. [88]

Ainsi, il est possible d'améliorer l'analyse peptides/protéines en éliminant du tissu certains lipides très majoritaires, ce qui permet la détection d'autres signaux d'intérêt. Les expériences ont également permis de mettre en évidence la spécificité de chacun des solvants sur les lipides retirés et les protéines favorisées. Il faut noter que si le traitement est choisi de manière appropriée, les composés d'intérêt ne sont pas délocalisés par l'étape de traitement. Par la suite cette étude a été complétée par une étude de Seeley et coll. [89] qui ont confirmé les effets observés en testant notamment quelques nouveaux solvants. Ainsi suivant les protéines recherchées, certains traitements sont clairement plus adaptés que d'autres. Sur des protéines de propriétés intermédiaires l'isopropanol se révélant être avec le chloroforme l'un des meilleurs compromis. Il est également possible de cumuler les traitements pour optimiser l'analyse. Ainsi pour l'analyse de protéines le traitement couramment employé au laboratoire consiste en trois bains successifs de différents solvants respectivement acétone/chloroforme/éthanol [Thèse Julien Franck, [90]]. D'autres études ont par la suite montré l'intérêt des traitements pour accéder à des protéines particulières comme des protéines membranaires de type aquaporine à partir de cristallins de vertébrés par des traitements avec un solvant hydrophobe le HFIP [91]. Récemment, Leinweber et coll. [92] ont



développé des protocoles pour l'analyse des protéines membranaires en utilisant des traitements au Triton X-100 combiné au xylène et en suggérant que la non observation de ces protéines pouvait être liée à un manque d'extraction, de solubilisation et d'incorporation dans les cristaux de matrice de ces composés. Dernièrement, Franck et coll. [93] ont étudié différents traitements permettant d'accéder à l'analyse de protéines de plus haut poids moléculaire. En effet, les expériences de MSI classique conduisent à l'observation de signaux pour les protéines jusqu'à des Mw d'environ 20 000 et peu de protéines de Mw supérieur sont généralement observées. Les travaux montrent que l'utilisation de solvants hydrophobes permet de nettement augmenter la gamme de masse des protéines analysables. La combinaison des traitements des tissus avec l'utilisation d'un détecteur spécifique à la détection des hauts m/z permettant encore d'améliorer la gamme de masse accessible pour l'analyse des protéines [94]. Ainsi les traitements des tissus permettent clairement d'augmenter le nombre d'espèces détectables (et en particulier pour les peptides/protéines) et se révèlent être un aspect clef pour l'imagerie MALDI comme le démontrent les applications réalisées [93, 95].

**Traitement des échantillons et stratégies d'identification.** Le traitement des échantillons est également une étape incontournable pour dépasser la réalisation d'images MALDI et accéder à l'identification des composés. Cet aspect est particulièrement important pour les peptides de masse élevée et les protéines. L'identification des composés passe par l'obtention d'informations structurales (MS/MS ou MS<sup>n</sup>) directement *in situ* à partir des tissus. Pour les composés de masse molaire inférieure à quelques milliers d'u.m.a. cette étape peut-être réalisée directement à partir de la section de tissu utilisée pour réaliser l'image, l'instrumentation utilisée permettant une fragmentation efficace de ces composés au travers de l'utilisation de différentes méthodes classiques d'activation des ions (PSD et CID principalement). Cependant, les composés de Mw plus élevé sont difficilement fragmentables par ces méthodes classiques, l'énergie interne transférée se répartissant sur l'ensemble des degrés de liberté de l'ion considéré. Au cours de ces dix dernières années, différentes méthodes d'activation des ions se sont révélées efficaces pour la fragmentation de composés de haut poids moléculaire comme les protéines. Cependant, ces méthodes sont efficaces sur des ions présentant un état de charge minimum et s'adaptent donc tout naturellement aux ions formés par ESI et non par MALDI. Ainsi, les identifications par analyse en mode Top-Down ne sont pour le moment pas accessibles en MALDI hormis par fragmentation de type In Source Decay (ISD) [96-100]. Cependant, l'ISD pose d'autres problèmes dans la mesure où

les fragmentations ISD se produisent par augmentation de l'énergie laser dans les premiers moments de la désorption et n'offrent pas de possibilité de sélection de l'ion d'intérêt. L'ISD est donc une méthode puissante pour caractériser des parties importantes de la séquence primaire d'une protéine purifiée, son application à l'imagerie MALDI ne conduit à l'identification, pour le moment, que d'un nombre restreint de protéines (2 à 3). Ainsi si l'ISD est possible directement à partir des tissus [101], son application aux études biologiques nécessite encore recherches et développements. De plus, il faut noter qu'une sélection d'un ion fragment formé en ISD reste possible et devrait permettre de générer des spectres MS<sup>3</sup> (ou pseudo MS<sup>3</sup> en MALDI-TOF) permettant de remonter à l'identification. Cette approche s'est jusqu'alors montrée comme étant peu efficace par manque de sensibilité sur les instruments couramment accessibles. Ainsi, pour l'identification des protéines, les équipes se sont principalement orientées depuis près de 5 ans vers des stratégies de type Bottom-Up [85, 102].

Ces stratégies font appel à la digestion des protéines en fragments peptidiques facilement analysables par MS<sup>2</sup> pour des méthodes d'activation des ions classiques. Cependant, ces stratégies impliquent la capacité à réaliser les digestions enzymatiques in situ sur les tissus tout en contrôlant la délocalisation des peptides issus de la digestion. Ainsi, les différentes recherches ont conduit à développer une microchimie à la surface de l'échantillon permettant d'obtenir une digestion efficace à l'échelle de quelques dizaines à 100 µm [90]. Ces étapes sont réalisées notamment au travers de l'utilisation de robots de micro-dépôts et seront détaillées dans la suite de ce chapitre. L'utilisation des stratégies Bottom-Up par digestion à la trypsine s'est montrée être efficace pour l'identification des protéines. Cette stratégie s'est révélée présenter l'avantage d'être la seule à permettre l'étude des tissus FFPE en imagerie MALDI [85]. En effet, comme précédemment mentionné, le formaldéhyde induit un pontage irréversible de protéines par la formation de ponts méthylènes. Il devient alors impossible de détecter des peptides ou protéines à partir de ces échantillons, sauf pour des échantillons conservés sur de très courtes périodes (et présentant un pontage plus faible) et grâce à l'utilisation de matrices particulières comme le 2,4-DNPH [85]. Le seul moyen efficace permettant l'accès aux protéines consiste à employer une digestion par des enzymes. Cependant, l'efficacité de la digestion enzymatique est parfois réduite car le réseau de protéines rend l'accès difficile au site de clivage. De tels problèmes sont également rencontrés en IHC pour l'accès des anticorps aux épitopes des antigènes. Ainsi, en IHC se sont développés depuis de nombreuses années des protocoles dits de démasquage des antigènes (AR ou Antigen Retrieval). L'effet de l'AR sur les tissus FFPE n'est pas clair et les

hypothèses sont variées, voire parfois contradictoires [103]. Il semble cependant raisonnable de penser que l'AR ne permet pas de renverser la réaction de pontage mais conduit à un changement de conformation du réseau de protéines, ce qui rend l'accès plus facile aux sites d'intérêt. De tels protocoles ont été récemment utilisés en imagerie MALDI et se sont révélés efficaces [104, 105]. Il existe de nombreux protocoles d'AR différents qui vont du simple chauffage sous pression au chauffage par micro-ondes (Heat Induced Antigen Retrieval HIAR [106, 107]) ou bain marie, avec différents tampons tels que l'EDTA [105] ou les tampons citrates (voire même à l'acide citrique (Citric Acid Antigen Retrieval CAAR [104])). L'utilisation en imagerie MALDI de ces protocoles est aussi peu clarifiée qu'en HIC à l'heure actuelle. Il ressort que le protocole doit être optimisé pour chaque tissu en fonction de sa nature et de son temps de conservation. Pour les échantillons conservés peu de temps les processus AR améliorent rarement les stratégies Bottom-Up en imagerie. Il faut également noter que le processus AR s'effectue par bain sur le tissu avant l'étape de digestion mais après le déparaffinage du tissu, opération qui s'effectue classiquement par un bain de toluène (ou alternativement de n-octane) puis une réhydratation progressive du tissu dans des bains éthanol/eau. La stratégie Bottom-Up présente le désavantage d'augmenter la complexité des spectres par le nombre considérable de fragments peptidiques générés. Ceci pose un problème pour l'identification des composés car de nombreux ions se révèlent très proches en  $m/z$  et sont parfois non séparables sur des instruments classiques tels que les MALDI-TOF. De nombreuses améliorations sont à attendre dans le futur concernant l'utilisation de telles stratégies sur des instruments qui présentent une résolution spectrale supérieure et une meilleure précision sur la mesure des  $m/z$  (e.g. FT-ICR, orbitrappe). L'identification des protéines reste néanmoins parfois difficile dans la mesure où l'identification repose sur la simple interprétation du spectre MS/MS ou de son utilisation pour l'interrogation des banques de données. Ainsi, Franck et coll. [108, 109] ont développé l'utilisation de procédures de dérivation N-terminales qui permettent l'accès à une identification plus fiable au travers de la simplification des spectres MS/MS par une orientation de la fragmentation. Différentes dérivations N-terminales par additions de charge négative (SPITC, 3-SBA) ou positive (ion phosphonium de type TMPP) ont été mises au point in situ sur les tissus et ont permis de conduire à une identification fiable des protéines. La dérivation au TMPP a également pu être réalisée par microchimie, permettant ainsi la réalisation d'images après les étapes de digestion puis de dérivation. Le développement des stratégies d'identification renforce l'intérêt et les nombreux aspects que peuvent prendre le traitement du tissu avant son analyse pour l'imagerie MALDI.

### **Etape 3 : Matrice et Dépôt**

Le dépôt de la matrice fait partie de la préparation du tissu et tout comme cette dernière, il joue un rôle prépondérant sur la qualité des expériences, aussi bien sur la qualité des spectres que sur celle des images. Le rôle prépondérant de la matrice en MALDI n'est plus à démontrer. Le choix de cette dernière influence les familles de molécules qui seront observées selon leurs propriétés physico-chimiques (polarité, hydrophobicité, solubilité,...). La manière dont la matrice sera déposée peut également influencer la qualité des spectres (intensité, nombre de composés observés, rapport signal/bruit, abondance d'ions cationisés).

*Nature de la matrice.* De nombreuses matrices MALDI ont été proposées depuis l'introduction du MALDI en 1988. Elles ont évolué avec le progrès des lasers équipant les sources et des longueurs d'onde utilisées (266 nm, 355 nm, 337 nm). Ainsi, l'acide nicotinique utilisé dans les premières expériences MALDI à 266 nm a été remplacé par d'autres matrices plus performantes à des longueurs d'onde de 335 ou 337 nm. Quelques matrices se sont montrées être particulièrement bien adaptées à l'étude de certaines familles de composés et sont encore utilisées de nos jours, telles que l'acide 2,5-dihydroxybenzoïque (2,5-DHB) introduit par l'équipe de F. Hillenkamp en 1991 pour l'analyse des protéines [35] ou encore l'acide sinapinique (SA) et l'acide alpha-cyano-4 hydroxycinnamique (HCCA), deux matrices introduites en 1996 par Chait et Beavis pour l'analyse des peptides/protéines [110]. En imagerie MALDI, les premières expériences ont été réalisées pour la recherche de peptides/protéines à partir de sections de tissus et ainsi dès les premières expériences les matrices SA, 2,5-DHB et HCCA furent utilisées. Ces matrices sont encore utilisées en imagerie en routine. SA est principalement utilisée pour l'analyse des protéines et HCCA pour l'analyse des peptides mais le 2,5-DHB est en revanche une matrice plus usitée pour les lipides et métabolites que pour les peptides/protéines. Cependant, hormis les différentes matrices ayant déjà montré leur intérêt en MALDI classique et transplantées à l'imagerie, de nouvelles matrices dédiées à l'analyse des tissus ont été recherchées. Ainsi, Lemaire et coll. [111] ont démontré l'intérêt de l'utilisation de matrices de type matrice ionique pour l'analyse de tissu et l'imagerie.

Les matrices ioniques sont des sels organiques comprenant deux partenaires et synthétisées à partir d'une réaction acido-basique tout comme les liquides ioniques en synthèse organique [111]. Ainsi, il est possible d'obtenir un grand panel de matrices différentes en utilisant les matrices MALDI conventionnelles (généralement de type acide) et

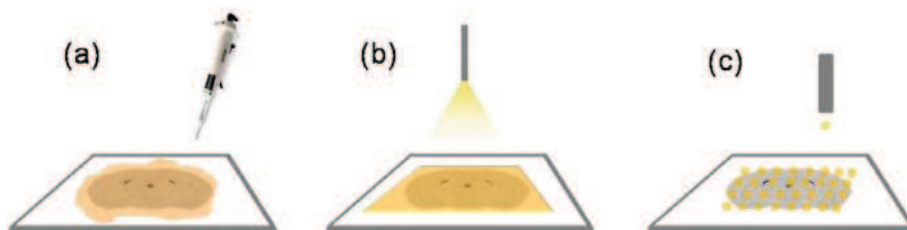
différentes bases organiques. Les matrices ioniques ainsi formées sont en phase liquide ou solide dépendamment du couple acide/base choisi et de l'équivalence de la réaction. Habituellement les matrices ioniques sont synthétisées à partir de mélanges équimolaire acide/base. Les premières matrices ioniques synthétisées et utilisées pour l'analyse de tissu ont été des matrices ioniques solides telles que HCCA/ANI, HCCA/DIANI et HCCA/A4M5NP [111]. L'utilisation de ces matrices en imagerie MALDI s'est révélée particulièrement intéressante. En effet, leur utilisation a conduit à observer une augmentation de l'intensité et du nombre de signaux détectés ainsi que du rapport signal/bruit pour l'analyse principalement des peptides. Ces matrices se sont également révélées beaucoup plus stables sous vide poussé et ont conduit à une meilleure reproductibilité des analyses intra et inter expériences. Par la suite d'autres matrices ioniques ont été proposées par ce groupe : matrices pour différentes gammes de masse de protéines et de lipides [90], matrices ioniques en phase liquide pour l'analyse des lipides [112].

Il faut noter l'intérêt de l'utilisation de matrices liquides pour des études de quantification, le dépôt liquide étant parfaitement homogène contrairement au dépôt solide. Comme nous le verrons dans le paragraphe suivant, outre les améliorations spectrales intrinsèques, ces matrices se sont également révélées intéressantes pour l'utilisation de modes de dépôts utilisés en imagerie MALDI. Chan et coll. se sont également intéressés aux matrices ioniques et ont proposé en 2009 une matrice ionique liquide à base d'HCCA et de 1-méthylimidazole pour l'analyse de lipides de type gangliosides dans des tissus de cerveau de rat [113]. La matrice ionique a ainsi permis d'accéder à différentes classes de gangliosides dont certains généralement difficiles d'observation. D'autres matrices non conventionnelles se sont également révélées intéressantes pour l'analyse de tissu. Il faut en particulier mentionner la paranitroaniline (PNA) proposée comme alternative au 2,5-DHB pour l'analyse des lipides neutres par Rujoi en coll. en 2004 [114]. En 2005, les matrices 6-aza-2-thiothymine (ATT) et la matrice 2,6-dihydroxyacetophenone (DHAP) sont utilisées pour l'analyse de différents lipides et en particulier les phospholipides à partir de tissus animaux [115-117]. Malheureusement si le DHAP est une matrice donnant de bons résultats pour l'analyse directe de lipides à partir de tissus, elle est peu stable sous vide (sublimation des cristaux de matrice au cours de l'expérience) et conduit à des images de moindre qualité. Pour l'analyse de petits composés tels que les lipides et métabolites, des matrices de type inorganique ont également été utilisées. Ainsi le graphite s'est révélé une bonne matrice pour l'analyse des lipides à partir de cerveaux de rats et en particulier pour l'analyse de cérébrosides [118] ou encore pour l'analyse de métabolites à partir de fruits [119]. Le même groupe a également proposé

l'utilisation du graphite colloïdal pour l'analyse de métabolites à partir de plantes [120]. En 2008, Astigarraga et coll. ont proposé le 2-Mercaptobenzothiazole (MBT) pour l'analyse des lipides [121]. Cette matrice s'est révélée plus efficace pour l'analyse des phospholipides et glycérophospholipides de haut poids moléculaire avec des signaux nombreux et intenses à la fois en mode positif et négatif. En 2009, Benabdellah et coll. se sont intéressés à la matrice 9-aminoacridine initialement proposée pour l'analyse des lipides en mode négatif. Utilisée sur tissu, cette matrice, outre son intérêt pour les lipides, s'est montrée intéressante pour l'analyse de petits métabolites primaires tels que AMP, ADP, ATP, UDP-GlcNAc [122]. Récemment, Lorkiewicz et coll. ont utilisé des microparticules de Titane pour l'analyse de lipides [123]. Celles-ci ont permis l'analyse de lipides (PC, PE, PI) complémentaires aux matrices 2,5-DHB et PNA mais conduisent en revanche à la fragmentation des espèces. Ces matrices type microparticules de Titane sont néanmoins intéressantes pour l'analyse de certains tissus tels que des tissus de plantes pour lesquelles le dépôt de matrices conventionnelles est difficile et ne conduit parfois à aucun résultat. Ainsi, en mode négatif sur des feuilles de plantes, les microparticules de Titane permettent l'observation de métabolites variés non observés par utilisation des matrices MALDI type 2,5-DHB et PNA.

De nombreuses matrices sont à l'heure actuelle à disposition pour l'imagerie MALDI. Celles-ci sont généralement employées en priorité pour l'analyse des peptides, protéines, médicaments et lipides. Cependant, d'autres familles de composés tels que les différents métabolites sont encore mal connues et suscitent de ce fait la recherche de nouvelles matrices (organiques ou inorganiques) mieux adaptées. Progressivement, différents progrès sont rencontrés au fil du temps et le panel de matrices utilisables s'élargit. Il faut également noter l'importance du choix des solvants pour dissoudre la matrice. Le solvant doit être à la fois adapté à la matrice, au composé recherché et au mode de dépôt utilisé. Ainsi, le mélange de solvant utilisé peut influencer les profils des spectres qui seront enregistrés. Il convient bien entendu d'optimiser la composition de ce mélange pour chaque expérience, en fonction du composé recherché et du tissu étudié.

***Dépôt de la matrice.*** Le mode de dépôt de la matrice revêt également une forte importance dans la préparation des échantillons puisqu'il conditionne non seulement la résolution spatiale des images, au travers de la taille des particules ou des gouttes de solution déposées, mais aussi la qualité des analyses au travers de la capacité de solubilisation, d'extraction voire d'incorporation dans les cristaux de matrice (**Figure 20**).



**Figure 20 : Les différents modes de dépôt de la matrice pour l'imagerie MALDI : manuel (a), par spray (b) et par micro-dépôts (c).**

Le plus couramment, dans une expérience MALDI traditionnelle, la matrice est en solution. La méthode la plus simple consiste donc à déposer à la micropipette un volume suffisant de la solution qui permet de recouvrir l'intégralité du tissu (**Figure 20a**). Pour un cerveau de rat, soit un échantillon d'environ 1.5-2 cm<sup>2</sup> avec une solution de matrice classique, il faut environ 15 à 20 µL de solution de matrice pour couvrir intégralement le tissu. Cette méthode, avec un peu de savoir faire, permet de générer des spectres présentant de nombreux signaux avec un bon rapport signal/bruit. Cependant, dès les débuts de l'imagerie MALDI les chercheurs se sont interrogés quant à la validité de cette méthode du point de vue du maintien de l'intégrité moléculaire du tissu. En effet, l'addition d'un volume non négligeable de solvant peut conduire à la délocalisation de certains des composés d'intérêt dans le tissu. Rapidement les recherches se sont donc orientées vers des méthodes de dépôt permettant de réduire et de contrôler la délocalisation liée à l'ajout de la solution de matrice.

L'une des méthodes les plus simples de mises en œuvre qui a été utilisée pour déposer la matrice est le dépôt par nébulisation en utilisant un simple nébulisateur pneumatique (en particulier nébulisateur en verre pour Chromatographie sur couche mince CCM) ou un système d'aérographe (systèmes utilisés en graphisme). Ces méthodes conduisent au dépôt de gouttelettes de solution de matrice de taille assez faible (500 µm -1mm). Cependant, elles présentent le très net désavantage d'être peu reproductibles en terme de résultats sur les spectres (grandes variations du nombre de signaux et de leur intensité d'une expérience à l'autre). Au cours du temps, différents systèmes automatisés se sont développés et ont été utilisés pour l'imagerie MALDI. Ces robots fonctionnent suivant deux modes de dépôt distincts qui sont le micro-dépôt d'une gouttelette de solution de très faible volume (nL) ou la micro-nébulisation par vaporisation d'un nuage de fines gouttelettes sur le tissu (**Figure 20b**). Pour les deux méthodes, différents systèmes sont à disposition et se basent sur des principes différents.

Pour le mode micro-dépôt, on peut utiliser par exemple le micro-dépôt par capillarité à partir d'un système robotisé muni d'un fin capillaire de silice fondue (**Figure 20c**). Ce système peut être adapté à partir d'un robot de collecte pour la nanoLC. Dans ce système, la taille du dépôt est uniquement conditionnée par le diamètre interne du capillaire utilisé et permet d'atteindre des dépôts de l'ordre de 200-300  $\mu\text{m}$ . Les expériences réalisées avec ce mode de dépôt conduisent à une qualité de spectres suffisante et une délocalisation parfaitement contrôlée [85].

D'autres systèmes automatisés fonctionnent par éjection de microgouttelettes par onde acoustique. Ces systèmes permettent de déposer des gouttes de 170 pL à une vitesse de 200 gouttes/s pour des gouttes d'une taille d'environ 100-200  $\mu\text{m}$ . Ce système de dépôt est à l'heure actuelle couramment rencontré dans les laboratoires pour l'imagerie [124]. Il est également possible de réaliser le micro-dépôt à l'aide d'un système de type piézo-électrique similaire aux systèmes employés pour la réalisation des puces à protéines ou ADN/ARN. Les caractéristiques obtenues sont relativement similaires à celles du dépôt par onde acoustique avec un dépôt d'une dimension d'environ 100-200  $\mu\text{m}$  et un volume déposé de 100 pL [90]. Il est également fait mention dans la littérature de l'utilisation d'une imprimante jet d'encre pour le micro-dépôt de la solution de matrice qui est injectée en remplissant une cartouche d'encre et qui fournit un dépôt de dimensions particulièrement faibles (taille du spot 10-50  $\mu\text{m}$ ) avec des gouttes de faible volume (3 pL) pour une vitesse de dépôt particulièrement rapide (30000 gouttes/s) et un moindre coût d'acquisition [125]. Il est également possible d'utiliser des dépôts par nébulisation. Deux systèmes sont actuellement couramment rencontrés. Le premier repose sur la formation d'un nébulisât de très fines gouttelettes via l'utilisation d'une grille vibrant à très haute fréquence. Ce système automatisé permet de générer de très petites gouttes (en moyenne 25  $\mu\text{m}$  et au maximum 50  $\mu\text{m}$ ) qui correspondent à des volumes à l'échelle du pL. Celui-ci présente l'avantage d'être muni d'un diffractomètre permettant de contrôler l'évolution du dépôt de matrice (temps de séchage, épaisseur du dépôt). Des systèmes de nébulisateurs pneumatiques automatisés sont également rencontrés. Un système de nébulisation par utilisation d'un capillaire oscillant a également été proposé [126].

Bien évidemment il est également possible d'utiliser un système de type électronébulisation (ES) pour le dépôt de la matrice [125]. Des études comparant différentes méthodes de dépôt telles que l'étude de Baluya et coll. [125] montrent que toutes ces méthodes de dépôt ne sont pas équivalentes en termes de signaux observés et de qualité spectrale. Ce phénomène est également dépendant de la famille de composé étudiée. Ainsi, les études réalisées en MEB (microscopie électronique à balayage) sur la cristallisation de la



matrice montrent que le mode de dépôt qui permet de générer les cristaux de matrice les plus fins et de la taille la plus homogène (et donc par là le dépôt qui devrait permettre d'atteindre la meilleure résolution spatiale) est le dépôt par électronébulisation. Cependant, à l'inverse, ce dépôt est celui qui conduit aux spectres de plus faible qualité, en particulier sur le plan du nombre de signaux observés et de leur intensité. Ce phénomène est particulièrement important pour les peptides et protéines. Il semble que pour ces composés, il existe une relation directe entre le volume de solvant ajouté et la qualité spectrale. L'hypothèse la plus raisonnable pour expliquer ce phénomène repose sur i) l'extraction des composés et ii) la nécessité d'une incorporation réelle dans le réseau cristallin de la matrice pour ces composés. En particulier, il semble que l'incorporation dans les cristaux de la matrice ne soit pas une condition nécessaire pour les autres familles de composés de plus faible poids moléculaire tels que les lipides, les médicaments ou les métabolites. Ainsi, le mode d'électronébulisation n'est pas spécialement performant en imagerie. Pour les peptides/protéines il est également clair que le dépôt micropipette est le dépôt qui permet de conduire à l'observation du plus grand nombre de composés, par opposition aux dépôts de type micro-dépôt ou micro-nébulisation. Le mode dépôt est donc un équilibre précaire à trouver en bonne qualité spectrale et en bonne résolution spatiale. Il faut noter que la micro-nébulisation présente l'avantage d'être plus rapide car l'intégralité du tissu est traitée (~45 min de préparation) alors que le micro-dépôt en mode goutte par goutte sur le tissu est plus long (plusieurs heures). Cependant, ce dernier permet l'observation d'un nombre de signaux un peu supérieur. Il faut également remarquer que dans tous les cas il est nécessaire de réaliser plusieurs cycles de dépôts avant d'atteindre une épaisseur de couche minimum pour l'observation de signaux. Ainsi, le nombre de cycles de dépôt doit être optimisé pour obtenir la meilleure capacité d'analyse et ce processus doit être réalisé pour chaque matrice. Il est important de noter que pour les composés de faible masse des méthodes de dépôt alternatives peuvent être employées. En effet, les composés de faible poids moléculaire sont, de manière générale, plus faciles à analyser à partir des tissus. Différents facteurs contribuent à cet effet comme leur plus grande solubilité, leur diffusion plus facile dans les tissus (ce qui favorise leur extraction), la non nécessité de leur incorporation dans les cristaux de matrice (comme tendent à le démontrer les différentes études) mais aussi l'instrumentation par elle-même qui favorise souvent l'analyse et la détection des composés plus petits. Ainsi, pour ces petits composés il est possible par exemple de minimiser la taille des cristaux de matrice et d'augmenter l'homogénéité de cristallisation en déposant la matrice par sublimation [127]. Les expériences comparées entre le dépôt par sublimation et le dépôt par électronébulisation montrent de bien meilleurs résultats pour le

dépôt par sublimation pour l'analyse de phospholipides à partir d'une coupe de cerveau de rat [127]. Il faut noter également la reproductibilité du signal. Il est également possible d'utiliser une méthode dépôt de la matrice sans solvant. Cette méthode évite toute délocalisation des composés à étudier et permet d'obtenir un dépôt particulièrement fin et composé de cristaux de taille homogène. La comparaison de ce mode de dépôt avec un dépôt par nébulisation à partir d'une coupe de cerveau de rat montre que le mode de dépôt sans matrice est aussi voire plus performant que le mode classique de nébulisation de la solution de matrice pour l'analyse des lipides [128]. Dans une publication récente [129] il est même montré que ce mode de dépôt permet la détection et la réalisation d'images de médicaments à des doses pour lesquels celui-ci n'est pas détecté par préparation classique de l'échantillon avec un dépôt par nébulisation de la solution. Les différentes méthodes de dépôt décrites sont regroupées avec leurs caractéristiques dans le **Tableau 5** présenté ci-dessous.

<b>Instrument</b>	<b>Sociétés</b>	<b>Mode de dépôt</b>	<b>Taille de dépôt</b>
ImagePrep	Bruker	Vaporisation vibrationnelle	25µm
Chemical inkjet printing (CHIP 1000)	Shimadzu	Piézo-électrique	150 µm
TMiD	LEAP Technologies	Ejection par Nano-Pompe	200 µm
MALDI-spotter Suncollect	SunChrom	Dépôt par un Capillaire	150 µm
Portrait™ 630 Reagent Multi-Spotter	Labcyte Inc	Ejection acoustique	200 µm
NC	Siliflow	Piézo-électrique	300-600µm

**Tableau 5 : Liste et caractéristiques des différents instruments pour le dépôt de la matrice en MSI.**

#### **Etapes 4 et 5 : Acquisition des données et reconstruction des images.**

Comme décrit dans le paragraphe précédent, l'acquisition des données se fait dans la plupart des cas, et sur tous les appareils commerciaux en mode microsonde, par déplacement de l'échantillon sous le faisceau laser. Au cours de ces dernières années, l'imagerie MALDI a considérablement progressé sur le plan de l'instrumentation ce qui a permis de nettement diminuer le temps d'acquisition et la résolution spatiale ainsi que d'augmenter la résolution, le pouvoir séparatif et la sensibilité.

Du point de vue du temps d'acquisition, les améliorations sont de deux sortes. D'une part les instruments sont équipés de laser de répétition des tirs de plus en plus grande, d'autre part les systèmes de digitalisation et conversion des données sont de plus en plus rapides.

Au début des années 2000, les acquisitions étaient principalement réalisées avec des lasers N<sub>2</sub> ( $\lambda=337$  nm) de 3 à 20 Hz suivant les instruments, ce qui conduisait à plus d'une journée d'acquisition pour une image d'une résolution spatiale satisfaisante. Par la suite vers 2004 les appareils ont rapidement été équipés de laser N<sub>2</sub> pouvant atteindre des répétitions de 50 Hz. Ceux-ci ont eux-mêmes été remplacés vers 2006 par des lasers Nd:YAG à 355 nm (3<sup>ème</sup> harmonique) de fréquence de répétition jusqu'à 200 Hz. Le changement de lasers N<sub>2</sub> vers des lasers de type Nd:YAG a également été motivé par la durée de vie de ces lasers. L'acquisition d'une image correspond à quelques centaines de tirs laser par point pour plusieurs dizaines de milliers de points en général. Sur un laser N<sub>2</sub> le remplacement de la cartouche est nécessaire au bout de quelques millions de tirs et la durée de vie du laser est donc trop faible pour une utilisation en imagerie alors que les lasers Nd:YAG peuvent atteindre plus d'1 milliard de tirs. Cependant, le laser N<sub>2</sub> semble conduire à des rendements de production d'ions supérieurs aux lasers Nd:YAG et à une plus faible ablation de matière. Ceci s'attribue au profil d'énergie très hétérogène de ces lasers par comparaison aux lasers Nd:YAG de profil presque parfaitement gaussien. Un très bon compromis a été décrit par Holle et coll. [130] par utilisation d'un laser Nd:YAG dont le profil est volontairement déstructuré et qui permet d'atteindre des rendements de production d'ions supérieurs aux lasers N<sub>2</sub> classiques. Il semble donc que des effets d'absorption locale d'énergie (existence de pointes d'énergie dans le faisceau) contribuent à augmenter le rendement de production des ions comme des recherches précédentes avaient déjà pu le mettre en évidence [131]. Ce système est au cœur de la technologie SmartBeam équipant certains instruments. Depuis 2009 sont commercialisés des instruments avec des lasers Nd:YAG pouvant atteindre 1 kHz et très récemment une publication a fait état de l'utilisation d'un laser à 20 kHz [132]. L'augmentation de la fréquence de répétition des tirs a ainsi permis de passer de plusieurs jours d'analyse à quelques heures pour une image équivalente en taille et en nombre de pixels.

La résolution spatiale a quant à elle également considérablement évoluée. Celle-ci a été principalement améliorée par l'utilisation de systèmes de focalisation de lasers plus performants. A ses débuts, l'imagerie présentait une résolution spatiale d'environ 100-150  $\mu\text{m}$ . L'utilisation de systèmes de focalisation plus travaillés a permis d'atteindre dans la seconde moitié des années 2000 une résolution en routine de 75  $\mu\text{m}$  et pouvant aller jusqu'à

50  $\mu\text{m}$ . Actuellement, les instruments commerciaux permettent d'atteindre une focalisation jusqu'à 10  $\mu\text{m}$ . Cependant, il faut noter que la réduction du diamètre du faisceau laser s'accompagne d'une diminution relativement drastique de l'intensité des signaux. Cette diminution provient d'une part de la plus faible quantité de molécules analysées lorsque le diamètre du spot laser est réduit mais également à des phénomènes plus fondamentaux encore mal élucidés. Récemment Spengler et coll [133] ont mis au point un système expérimental permettant d'atteindre une résolution de 5-10  $\mu\text{m}$  et d'obtenir des images de qualité pour ces résolutions. Cependant, si ces expériences permettent l'analyse de petits composés tels que les lipides, il reste encore difficile de travailler sur des composés de plus haut poids moléculaire tels que les peptides et les protéines. Il faut tout de même noter le remarquable bon en avant effectué par l'imagerie MALDI sur le plan de la résolution spatiale puisque des images à l'échelle sub-cellulaire ont pu être présentées pour ces petits composés organiques [133]. A terme, la diffraction sera une limite pour la résolution spatiale en MALDI, mais de nombreux travaux de recherches restent encore à mener pour atteindre une imagerie subcellulaire performante pour toutes les familles de composés.

L'instrumentation a également été un point clef de développement. Du point de vue de l'acquisition, un facteur limitant reste le temps d'analyse. Ce temps est tributaire de la fréquence de répétition des tirs laser, de la rapidité de déplacement du support pendant l'analyse et du temps d'acquisition de l'électronique. L'utilisation de lasers pouvant atteindre de grandes fréquences de répétition est particulièrement souhaitable. Par exemple, à quelques Hz l'acquisition peut durer plus d'une journée alors qu'à 1000 Hz le temps d'acquisition sera réduit à quelques minutes selon la superficie du tissu à analyser (notons que très récemment, un laser à 20 KHz a été utilisé pour la MSI [134]). Ce point permet également de prévenir toute dégradation de l'échantillon (sublimation de la matrice progressive sous vide par exemple) pouvant intervenir avec une durée d'analyse trop importante. La fréquence de tirs des lasers représente donc un enjeu majeur pour les constructeurs aujourd'hui, et ce qui fut une limite hier, ne l'est plus de nos jours. En effet, de réelles avancées ont été faites et ce n'est plus la fréquence qui limite le temps d'analyse mais bien le système de déplacement de la cible au sein de l'appareil. De la même façon, les progrès dans le domaine de l'électronique permettent d'améliorer sans cesse la capacité des instruments. De plus, la reproductibilité du déplacement est un paramètre important pour la qualité des images, tout comme l'aire analysée en chaque point est également primordiale. Le nombre de points pouvant être faits va

définir la résolution spatiale de l'image reconstruite. La distance entre 2 points sera, dans la plupart des cas, au minimum égale au diamètre du faisceau laser.

Lors de l'acquisition, un autre paramètre important est l'accumulation progressive de matrice sur les lentilles de la source suite à la répétition des tirs lasers. Ce processus, visible sur des tissus de grande taille, entraîne au fil du temps une perte de signal en raison de la diminution des capacités d'accélération des ions. C'est pourquoi, la société Bruker a intégré dans sa dernière version de l'UltrafleXtreme un laser IR permettant le nettoyage de ces lentilles entre les différentes acquisitions sans avoir à les nettoyer manuellement, ce qui représente un gain de temps non négligeable.

Du point de vue de la reconstruction de l'image, un grand nombre de paramètres vont en affecter le rendu. Par exemple, les traitements des spectres doivent-ils être effectués ou non avant extraction des données (lissage, retrait de bruit, filtrage du bruit,...) et dans quelle mesure ces traitements affecteront-ils la qualité des images ? La question de savoir si l'intensité ou l'aire d'un pic décrit le mieux la réalité peut aussi se poser. De la même manière, sachant que plusieurs ions distincts peuvent correspondre à un même composé ( $MH^+$ ,  $MNa^+$ ,  $MK^+$ ,  $MH_2^{2+}$ ,  $[M+m]^+$  etc.), l'image de chaque ion doit-elle être reconstruite distinctement ou l'intensité des différents ions correspondant à ce même composé doit-elle être prise en compte simultanément ? Des critères de sélection peuvent être aussi ajoutés : utiliser un seuil d'intensité minimum au-dessous duquel le signal sera considéré comme nul ou un seuil d'intensité maximum au-delà duquel il sera estimé que les valeurs du signal sont automatiquement attribuées à 100%. L'utilisation d'échelles multi chromatiques peut également jouer sur le rendu visuel de l'image ainsi que la superposition avec des images de microscopie classique. Les logiciels développés pour l'acquisition et/ou la reconstruction des images sont listés dans le **tableau 6**.

Logiciels	Type	Compatibilité MS	Gratuité
MMSIT	Acquisition	AB MALDI/TOF	Oui
4700/4800 Imaging	Acquisition	AB MALDI-TOF/ TOF	Oui
Create Target Analyze This!	Acquisition/ conversion	BD MALDI- TOF/TOF	Oui
Axima2 Analyze	Acquisition/ conversion	SZU MALDI- TOF/TOF	Oui
oMALDI Server5.1	Acquisition/ reconstruction image	AB HybridLC/MS/MS	Non
FlexImaging	Acquisition/ reconstruction image	BD MALDI-TOF/TOF	Non
Tissue view1.0	Acquisition/ reconstruction image	AB MALDI-TOF/TOF HybridLC/MS/MS	Non
Image Quest	Acquisition/ conversion	TS MALDI-LTQ	Non
BIOMAP	reconstruction image	AB MALDI-TOF/ TOF	Oui
MITICS	reconstruction image	AB MALDI/TOF BD MALDI-TOF/TOF	Oui

**Tableau 6: Liste des différents logiciels existants pour l'imagerie MALDI. AB: Applied Biosystems; BD: Bruker Daltonics; SZU: Shimadzu; TS: Thermo Scientific. Tiré de [135].**

Comme nous l'avons vu, la source d'ionisation MALDI est donc parfaitement adaptée à l'imagerie par spectrométrie de masse. L'analyseur le plus souvent couplé à cette source est le TOF en raison de sa gamme de masse théorique illimitée et de la simplicité d'analyse des spectres, mais celui-ci reste limité par sa précision de mesure et son pouvoir résolutif. Plusieurs types d'analyseurs peuvent être utilisés avec cette source. Historiquement, les plus utilisés sont les temps de vol (TOF). Avec les progrès technologiques, sont apparus ces dernières années des couplages avec des analyseurs présentant de hautes performances et de plus grandes capacités pour les études structurales (MS/MS). A titre d'exemple, peuvent être cités les couplages avec des pièges ioniques et des pièges linéaires (LTQ), avec des instruments de Résonance Cyclotronique d'ions à transformer de Fourier, (FT-ICR) ou encore les instruments hybrides quadripôles (Q) et TOF (MALDI-Q/TOF), ainsi que les MALDI-TOF/TOF. L'ensemble des instruments utilisables en imagerie MALDI est résumé dans le **tableau 7**.

<b>Instrument</b>	<b>Sociétés</b>	<b>Type</b>	<b>Laser</b>	<b>Gamme de masse</b>	<b>Résolution max</b>
QSTAR Elite system	Applied Biosystems	QqTOF	Nd-YAG 1KHz	40-40 Kda	150 µm
QSTAR Pulsar i system	Applied Biosystems	QqTOF	Nd-YAG 1KHz	40-40 Kda	150 µm
QSTAR XL System	Applied Biosystems	QqTOF	Nd-YAG 1KHz	40-40 Kda	150 µm
4800 MALDI TOF/TOF Analyser	Applied Biosystems	TOF-TOF	Nd-YAG 200Hz	full	50 µm
4700 Proteomics Analyser	Applied Biosystems	TOF-TOF	Nd-YAG 200Hz	full	50 µm
Voyager DE STR Workstation	Applied Biosystems	TOF	N2 20 Hz	full	100 µm
Voyager DE PRO workstation	Applied Biosystems	TOF	N2 20 Hz	full	100 µm
LTQ Orbitrap XL	Thermo Scientific	LTQ-Orbitrap	N2 60 Hz	0-4000 Da	100µm
UltrafleXtreme	Bruker	TOF-TOF	Smartbeam-II 1KHz	full	10µm
Ultraflex III	Bruker	TOF-TOF	Smartbeam 200Hz	full	30µm
Autoflex IV	Bruker	TOF	Smartbeam 1 KHz	full	10µm
Solarix MI	Bruker	Hybrid FTMS	Smartbeam-II 1KHz	100-10 Kda	20µm
AXIMA QIT	Shimadzu	ion trap TOF	N2	0-20 Kda	100µm
AXIMA confidence	Shimadzu	TOF	N2 50Hz	full	100µm
AXIMA performance	Shimadzu	TOF-TOF	N2 50Hz	full	100µm
MALDI-Q-TOF premier	Waters	Qq-oaTOF	NdYAG-200Hz	20-40KDa	100µm
MALDI microMX	Waters	TOF	N2 20Hz	full	200µm
MALDI SYNAPT HDMS	Waters	Q-IMS-oaTOF	NdYAG-200Hz	20-40 Kda	100µm
MALDI SYNAPT MS	Waters	Q-Twave- oaTOF	NdYAG-200Hz	20-40 Kda	100µm

**Tableau 7 : liste et caractéristiques des différents spectromètres de type MALDI compatibles avec la MSI (www.maldi-msi.org)**

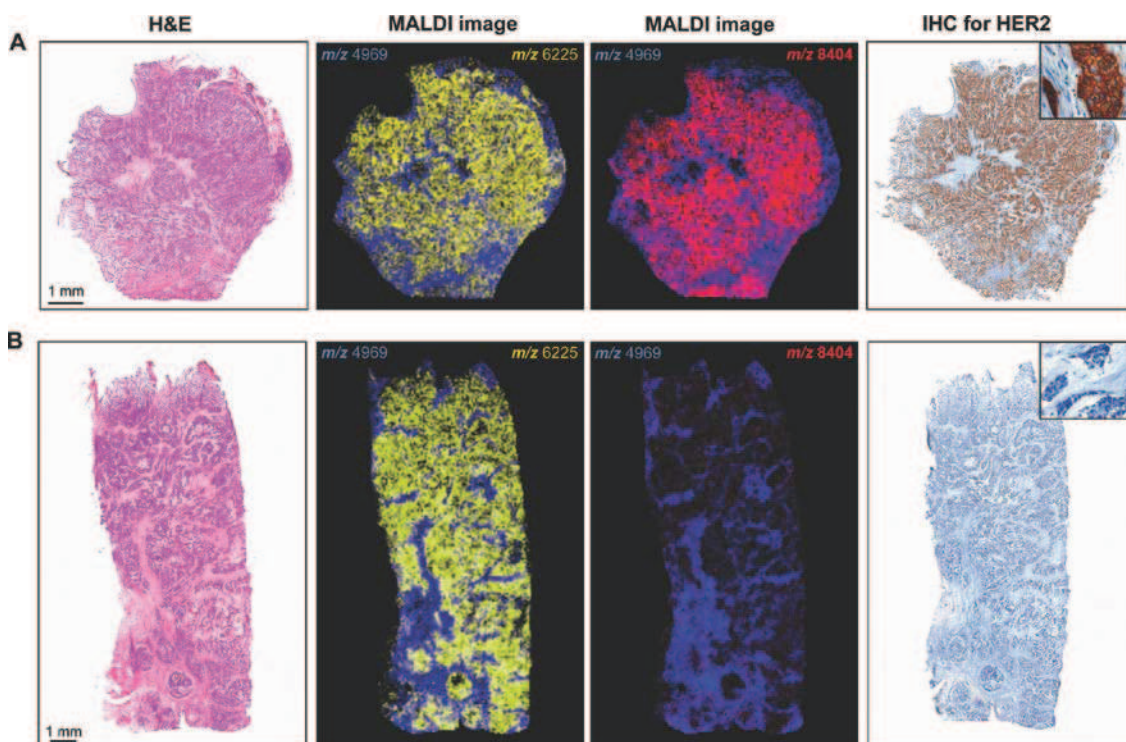
### 2.1.3.2 Les différentes applications de l'imagerie MALDI.

#### **Imagerie MALDI et protéomique**

La source d'ionisation MALDI est parfaitement adaptée aux études protéomiques en raison de ses caractéristiques que nous avons vues précédemment, c'est pourquoi l'imagerie MALDI est particulièrement utilisée dans le cadre d'étude des protéines dans divers tissus pouvant aller du végétal à l'animal (invertébrés et vertébrés), en passant par les bactéries. En effet, en raison du nombre de peptides et/ou protéines détectés sur un tissu couplé à l'information de localisation, l'imagerie MALDI est devenue un outil puissant pour la recherche de marqueurs de pathologies ou d'états physiologiques d'un organisme. Le développement de cette approche dans l'étude de pathologies comme le cancer a connu un fort intérêt ces dernières années. Parmi les types de tissus cancéreux étudiés, on retrouve les cancers de la prostate [136, 137], de l'ovaire [138], du sein [139], du colon [140], des poumons [141], et différents gliomes [142, 143].

La **figure 21** illustre l'intérêt de l'imagerie MALDI pour la recherche de biomarqueurs dans le cadre du cancer du sein sur tissu frais. Des chercheurs ont pu remarquer dans certains cas de cancer du sein que l'expression du gène codant la protéine HER2 (Human Epidermal Growth Factor Receptor-2) est amplifiée chez 20 à 30% des patientes. Cette protéine liée à des pronostiques pessimistes est aujourd'hui la cible de nouvelles thérapies [144]. C'est pourquoi la recherche des protéines associées à HER2 et impliquées dans les processus tumoraux a fait l'objet d'une étude en imagerie MALDI (**Figure 21**) aboutissant à l'identification et l'implication de cysteine-rich intestinal protein 1 (CRIP1) dans le cancer du sein HER2-positif [139].





**Figure 21** : Image MALDI de l'ion  $m/z$  8404 (rouge) comme exemple de protéine surexprimée dans des tissus de cancer du sein HER2-positif (A) et absent dans les tissus HER2-négatif (B). Les ions  $m/z$  6225 (jaune) et  $m/z$  4969 (bleu, spécifique au stroma) sont cancer-spécifiques mais pas tissu-HER2 spécifiques. [139]

L'intérêt pour les tissus FFPE n'a cessé d'augmenter depuis leur première étude en imagerie MALDI en 2007 [85]. En effet, ce mode de conservation des tissus est très largement utilisé dans les hôpitaux et d'importantes banques tissulaires existent. En raison de la fixation, une digestion trypsique est le plus souvent utilisée pour accéder aux protéines du tissu. Les peptides de digestion sont ensuite analysés et les ions d'intérêt sont sélectionnés pour identification. La séquence du peptide de digestion permet alors de remonter à l'identification de la protéine dont il est issu. La même approche protéomique de recherche de biomarqueurs peut donc être appliquée aux tissus FFPE. **La figure 22** montre la qualité des images MALDI d'ions obtenues sur des cerveaux de rat FFPE et frais.

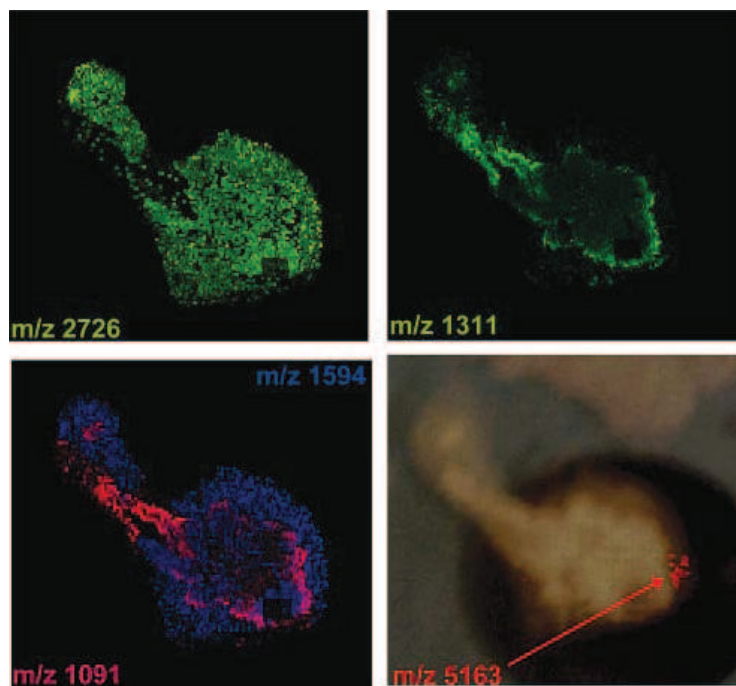


Figure 22 : Images MALDI de différents peptides de digestion obtenus sur des coupes de tissu FFPE de cerveau de rat conservé depuis environ 1 an. [85]

### Imagerie MALDI et lipidomique

Constituant majeur de la membrane cellulaire, les lipides jouent un rôle important dans les processus biologiques tel que la transduction des signaux. Des modifications au niveau lipidique sont observées dans de nombreuses pathologies comme la maladie d'Alzheimer [145], le syndrome de Down (trisomie 21)[146], le diabète [147], dans les cancers [148, 149] ou encore lors de processus d'apoptose [150, 151]. Les lipides peuvent être organisés en 8 catégories : les acides gras, les glycérolipides, les glycérophospholipides, les sphingolipides, les stérols, les prényls, les saccharolipides et les polykétides [152]. Longtemps l'apanage de l'imagerie SIMS, l'imagerie des lipides est devenue une application très intéressante de l'imagerie MALDI [148, 153-156] en raison de l'amélioration des instruments (les lipides requièrent une grande résolution spectrale car de nombreuses espèces sont très proches en masse, certaines étant souvent même isobariques) et des développements de nouvelles matrices permettant leur détection comme le montre la **figure 23** dans le cadre du cancer de l'ovaire [157].

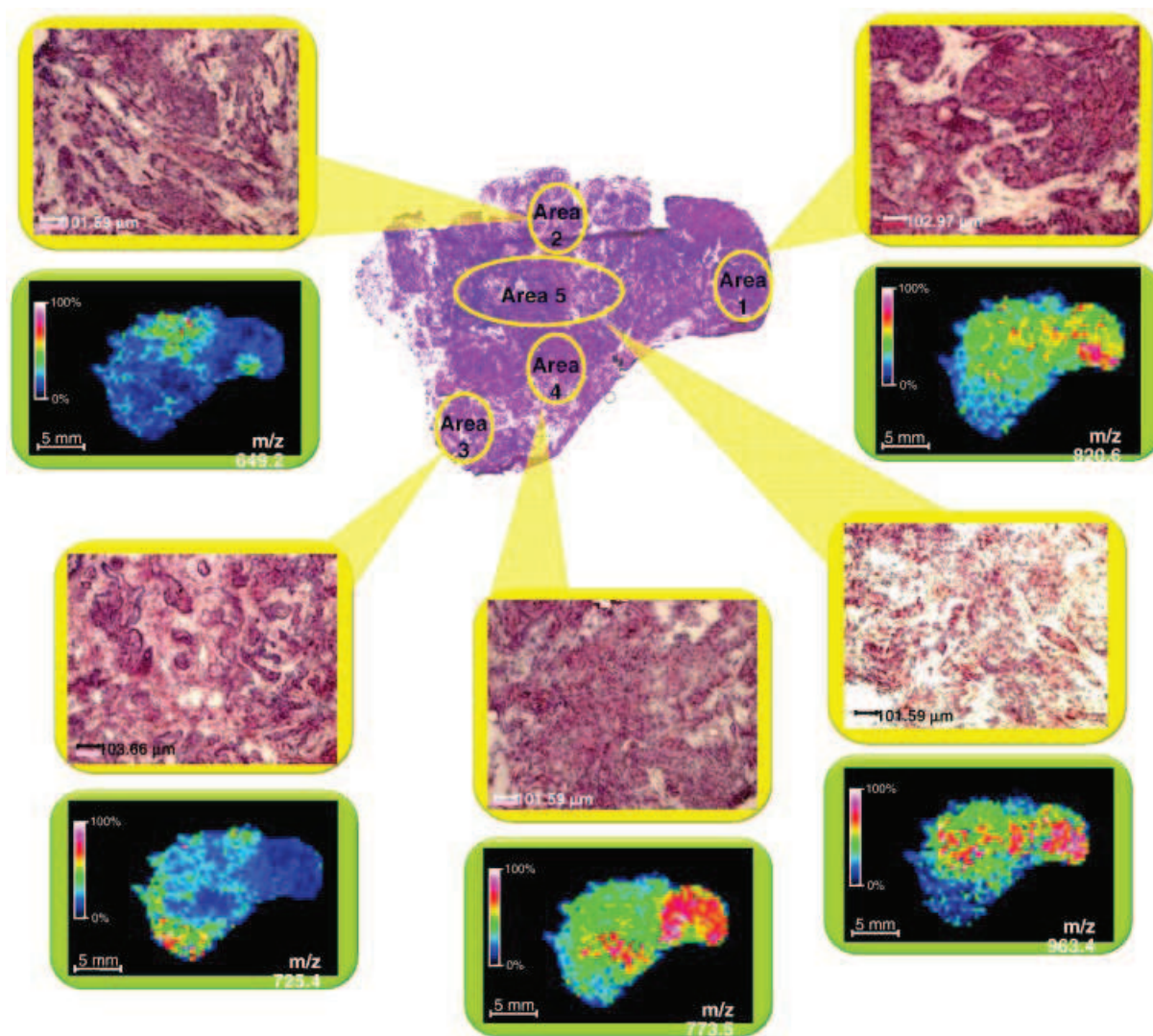


Figure 23 : Détection et localisation de cinq lipides différents propres à cinq zones histologiques sur une coupe de cancer d'ovaire. [157]

### Imagerie MALDI et les études ADME

L'imagerie MALDI a aujourd'hui trouvé sa place dans de nombreuses études ADME (Absorption, Distribution, Métabolisme et Elimination) et dans les suivis pharmacocinétiques. Habituellement, ces études sont réalisées par des méthodes, comme l'autoradiographie sur animal entier (whole body autoradiography, WBA), nécessitant un marquage préalable des composés, les rendant longues et coûteuses. Le suivi de molécules exogènes par imagerie MALDI, comme les médicaments et leurs métabolites (**figure 24**), a donc grandement simplifié ces études en y ajoutant la possibilité d'identifier les molécules d'intérêt directement sur le tissu.

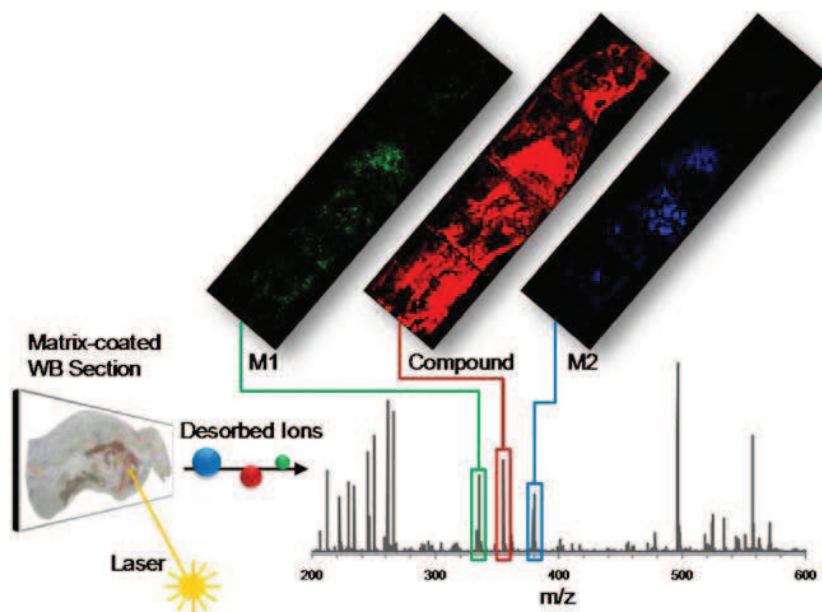


Figure 24 : Apport de l'imagerie MALDI sur des coupes d'animaux entiers (WB : whole body) pour les études ADME : détection du composé d'intérêt et des métabolites (M1, M2). [158]

Par exemple, après l'injection d'une dose donnée d'un composé radioactif d'intérêt dans un rat, l'animal est sacrifié à un temps donné et deux coupes adjacentes d'animal entier sont effectuées (notons que MALDI-MSI ne nécessite pas de radioactivité). La **figure 25** permet de comparer le résultat obtenu. On remarque que la même localisation est déterminée par les deux approches mais qu'il est possible de discriminer le composé de son métabolite par MALDI-MSI. L'autre intérêt majeur est l'étude cinétique d'un composé et de ses métabolites au sein d'un modèle animal (**figure 26**) [159].

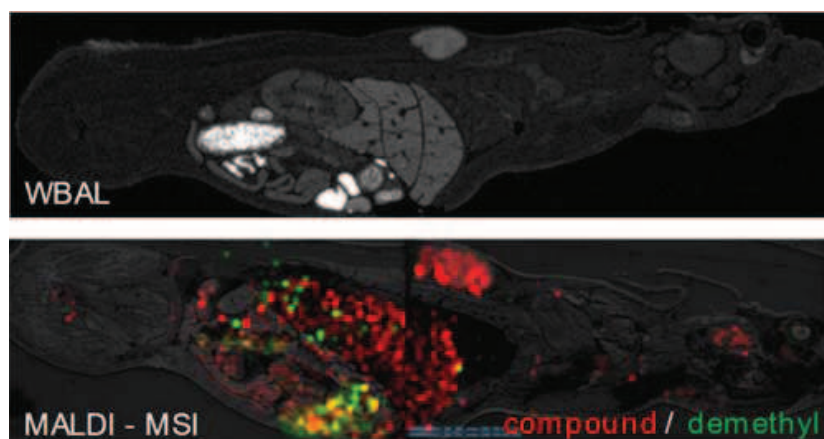


Figure 25 : comparaison entre l'autoradiographie et MALDI-MSI pour le suivi d'un composé et de son métabolite dans une coupe entière de rat. (Composé : rouge, métabolites : vert ou jaune si localisé avec le composé) [158]

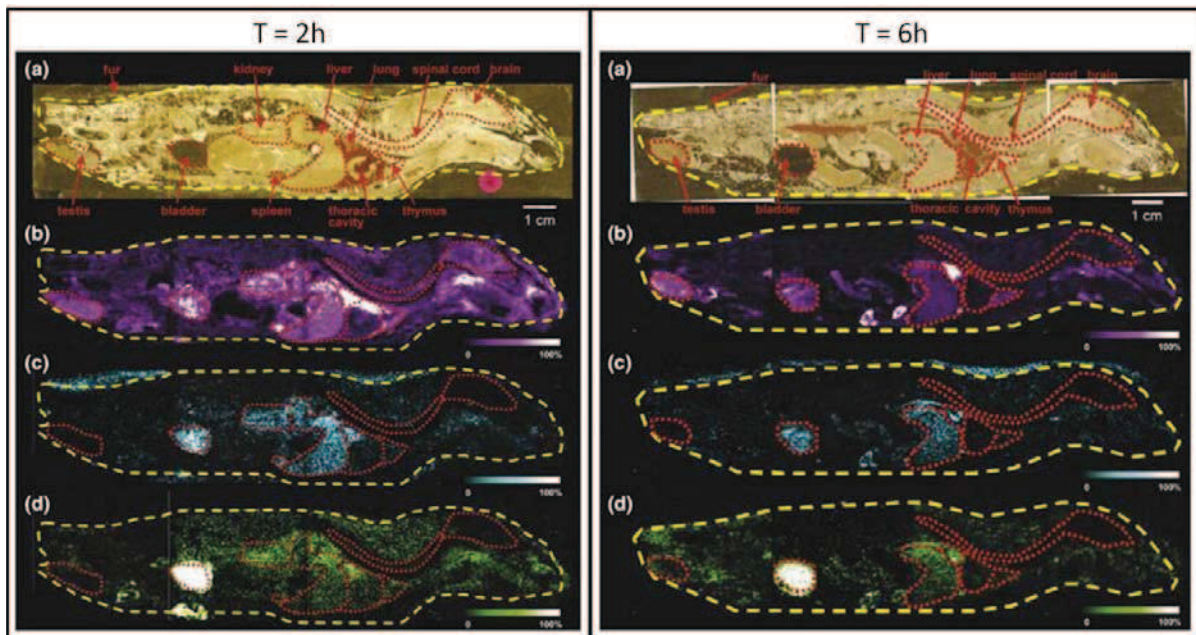


Figure 26 : Détection de l'olanzapine et des métabolites dans une coupe sagittale de rat entier à 2h (à gauche) et 6h (à droite) après injection. (a) Image optique de la coupe de rat entier avec les organes entourés en rouge. (b) Image MS/MS de l'olanzapine ( $m/z$  313 $\rightarrow$  $m/z$  256), (c) image MS/MS du N-desmethyl olanzapine ( $m/z$  299 $\rightarrow$  $m/z$  256) et (d) image MS/MS du 2-hydroxymethyl olanzapine ( $m/z$  329 $\rightarrow$  $m/z$  272). [159]

### 2.1.3.3 Les autres types d'imagerie par spectrométrie de masse

#### L'imagerie SIMS.

Le Secondary Ion Mass Spectrometry ou SIMS est une technique de spectrométrie de masse ancienne puisque les premières recherches de Thomson datent de 1900. Il faut attendre les années 1960 pour voir apparaître les premiers résultats d'imagerie SIMS par Castaing [60]. En SIMS, les ions de l'analyte (ou ions secondaires) sont désorbés à l'aide d'un bombardement de la surface par un faisceau d'ions primaire (**figure 27**). Ce faisceau ne peut fonctionner que dans un vide poussé ( $<10^{-6}$  torr) et est généralement composé d'ions monoatomiques, polyatomiques ou de clusters d'ions ayant une haute énergie (keV). Une large variété d'ions peut être utilisée comme source en analyse SIMS tels que  $Ar^+$ ,  $Xe^+$ ,  $Ga^+$ ,  $In^+$ ,  $Cs^+$ ,  $Au^+$ ,  $CO_2^+$  ou  $C_{60}^+$ .

Différents modes d'analyse sont possibles en SIMS ; le mode dynamique est un mode où la densité d'ions du faisceau primaire est importante, et où une érosion rapide et profonde

de la surface est observée. Les ions détectés sont principalement des ions monoatomiques ou des petits agrégats de molécules ayant une grande énergie interne. Le mode statique, où la densité d'ions est plus faible, n'endommage que les premières couches de l'analyte ou du tissu et permet de détecter des composés de plus hautes masses, voire des molécules entières. Ce mode est le plus adapté pour la réalisation d'images [160, 161].

Les ions secondaires obtenus peuvent être neutres (99,9%), de charges positives ou négatives (0,1%) ; différents analyseurs peuvent être utilisés avec ce type de source (quadripôle, secteurs magnétiques) et notamment le temps de vol, qui permettent une bonne transmission des ions et une bonne résolution des signaux.

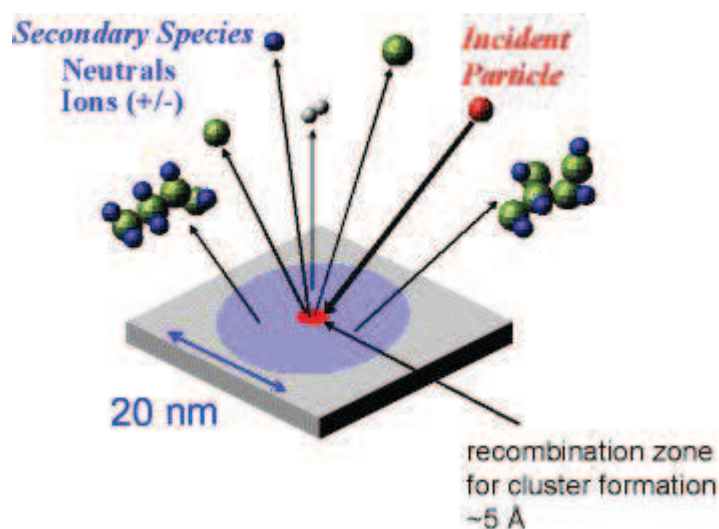


Figure 27 : Principe de la source SIMS

L'utilisation d'ions primaires polyatomiques tels que l'or, le bismuth ou le  $C60^+$  a également été une voie d'étude. Ces développements ont considérablement amélioré l'utilisation du SIMS pour la MSI en fournissant des ions primaires polyatomiques qui peuvent être focalisés de façon à réaliser des images à haute résolution avec une bonne sensibilité et qui présentent des ions intacts. Ceci offre un réel avantage pour la réalisation d'images à partir d'échantillons biologiques, comme le montre l'exemple développé dans la **figure 28**. Mais le plus grand avantage de ce type d'imagerie est la résolution spatiale que l'on peut obtenir comme en témoigne la **figure 29**, il est ainsi possible de descendre à l'échelle du nm et donc à l'échelle cellulaire au niveau d'une coupe histologique.

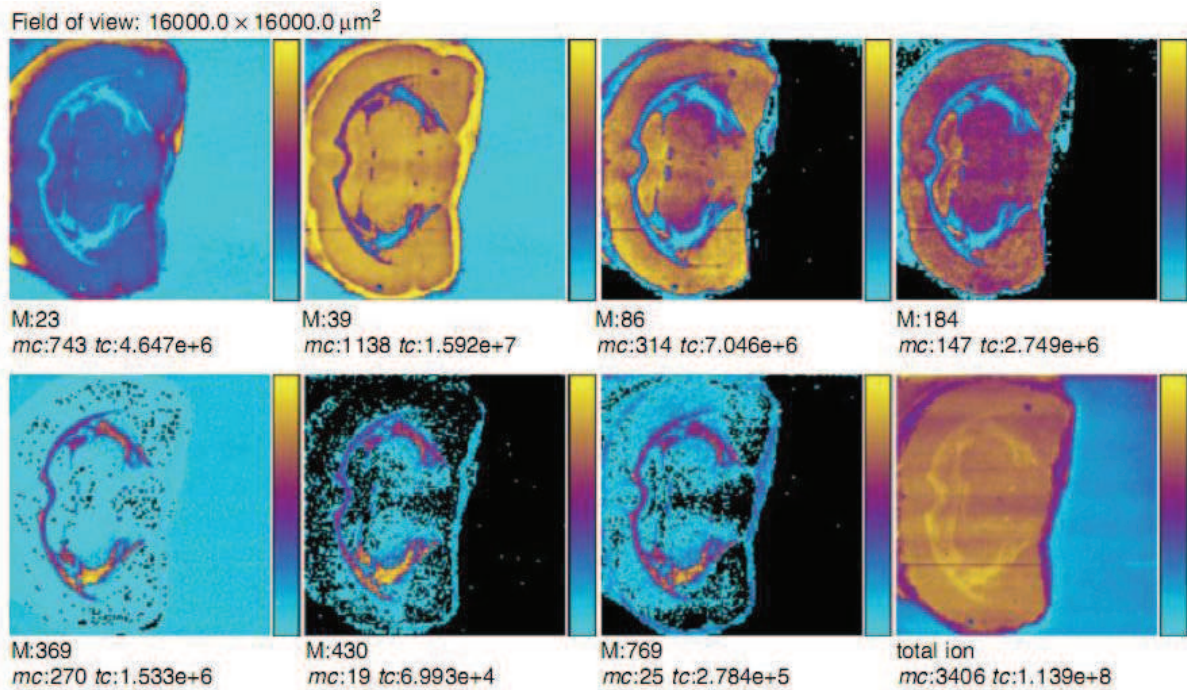


Figure 28 : Images obtenues à partir d'une coupe de cerveau de rat des ions Na<sup>+</sup> (m/z 23), K<sup>+</sup> (m/z 39), choline (m/z 86), phosphocholine (m/z 184), cholestérol [M+H<sub>2</sub>O]<sup>+</sup> (m/z 369), vitamine E (m/z 430), phospholipides (m/z 769) et du tic. Images réalisées à 62,5μm et avec une source d'ions primaires Au<sup>3+</sup>. [162]

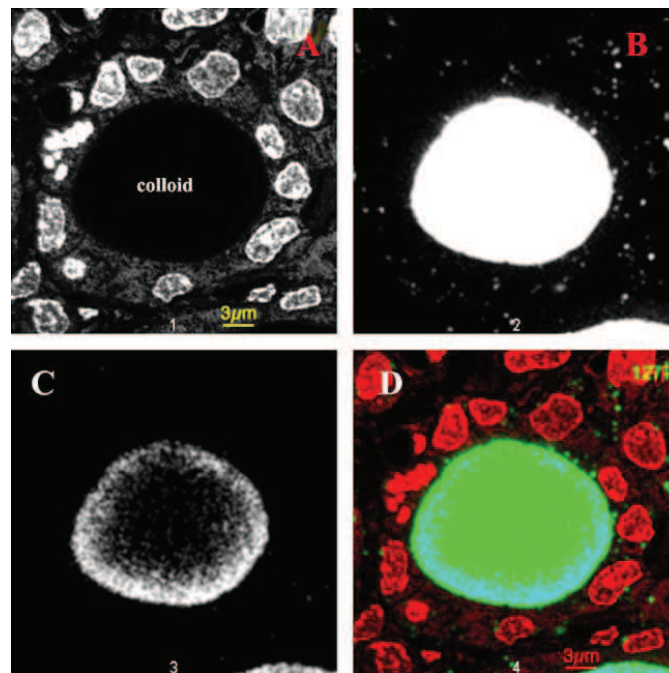


Figure 29 : Images SIMS d'ions secondaires négatifs d'une coupe de thyroïde de rat. (A) Image P<sup>-</sup> (m=31); (B) image <sup>127</sup>I (m =127); (C) image <sup>129</sup>I (m =129); (D) superposition (P<sup>-</sup>: rouge, <sup>127</sup>I: vert, <sup>129</sup>I: bleu). [163]

## L'imagerie DESI.

En DESI (desorption electrospray ionisation) [64], un spray fin de gouttelettes chargées vient frapper une région à analyser (tissus, peau, vêtements, drogues...) se trouvant à pression atmosphérique. Ces gouttelettes vont extraire de la surface de petits composés ou des biomolécules de haute masse, ces derniers vont être ionisés et emmenés jusqu'au spectromètre de masse où ils arriveront sous forme désolvatée (**Figure 30**). Tout comme l'électrospray, le DESI est une méthode douce, où les ions secondaires ont peu d'énergie interne, causant peu de fragmentations et permettant donc de détecter des molécules intactes. Grâce au mécanisme de désolvatation identique pour ces deux techniques, les spectres obtenus en ESI et DESI sont similaires. Pour les peptides, protéines ou tout autre composé de masse importante, les ions formés sont multichargés (protonés ou déprotonés) ou/et détectés sous forme d'adduits (Na<sup>+</sup>, K<sup>+</sup>...).

Ainsi, le DESI peut être considéré comme une version du SIMS à pression atmosphérique, où les clusters d'eau sont considérés comme des faisceaux d'ions primaires. Les avantages du DESI sont multiples : aucune préparation d'échantillon, grande sensibilité, technique applicable sur des produits complexes... Les ions détectés en DESI dépendent de plusieurs paramètres : la nature de l'échantillon, le solvant utilisé (polarité, pH, concentration et nature du sel), et le substrat ou le support sur lequel est déposé l'analyte. Celui-ci peut être très variable : plastique, métal, papier etc. En fonction de la substance à détecter, l'inclinaison du spray, la tension pour réaliser ce dernier ou la distance entre l'analyte et l'aiguille peuvent varier ; par exemple, une tension de 5kV, une distance de 5mm et un angle de 40° seront idéals pour la détection de petites molécules. L'angle sera plus important et la distance raccourcie pour la détection de protéines. Ces conditions sont directement en rapport avec le phénomène de d'ablation des composés de la surface solide par les gouttelettes chargées.

Comme en ESI avec un analyseur adéquat (triple quadripôle, trappe ionique, Q/TOF, FTMS...) tous ces composés peuvent être analysés en mode MS puis en mode MS/MS.



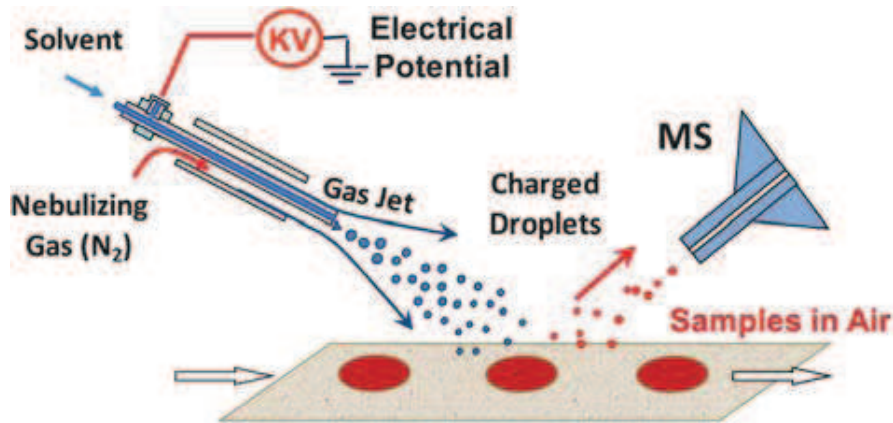


Figure 30 : Principe de la source DESI [164].

L'utilisation de la source DESI pour la MSI permet d'obtenir la distribution de molécules intactes au sein d'une surface telle qu'une coupe de tissus dans le cas d'analyse de biomolécules. Les derniers développements apportés aux sources DESI ont permis d'obtenir une résolution spatiale d'environ 40 $\mu$ m. Un exemple de l'utilisation de la source DESI en MSI est présenté **figure 31** avec l'étude de la distribution de lipides et d'acides gras au niveau de la moelle épinière de rat [165].

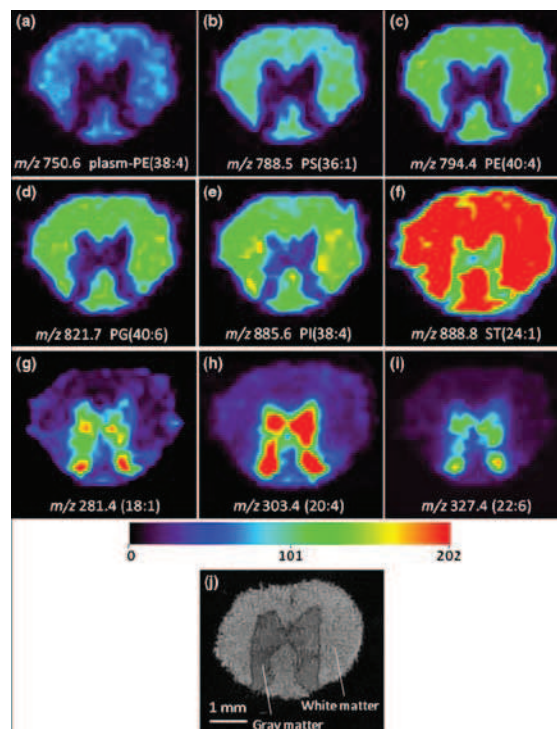


Figure 31 : Images DESI d'ions négatifs de lipides et d'acides gras au niveau de la moelle épinière de rat. Images des ions : (a) plasm-PE(38:4), (b) PS(36:1), (c) PE(40:4), (d) PG(40:6), (e) PI(38:4), (f) ST(24:1), (g) (18:1), (h) (20:4), and (i) (22:6). (j) Image optique de la coupe de moelle épinière. [165]

#### 2.1.3.4 Gestion et traitement des données d'imagerie par spectrométrie de masse

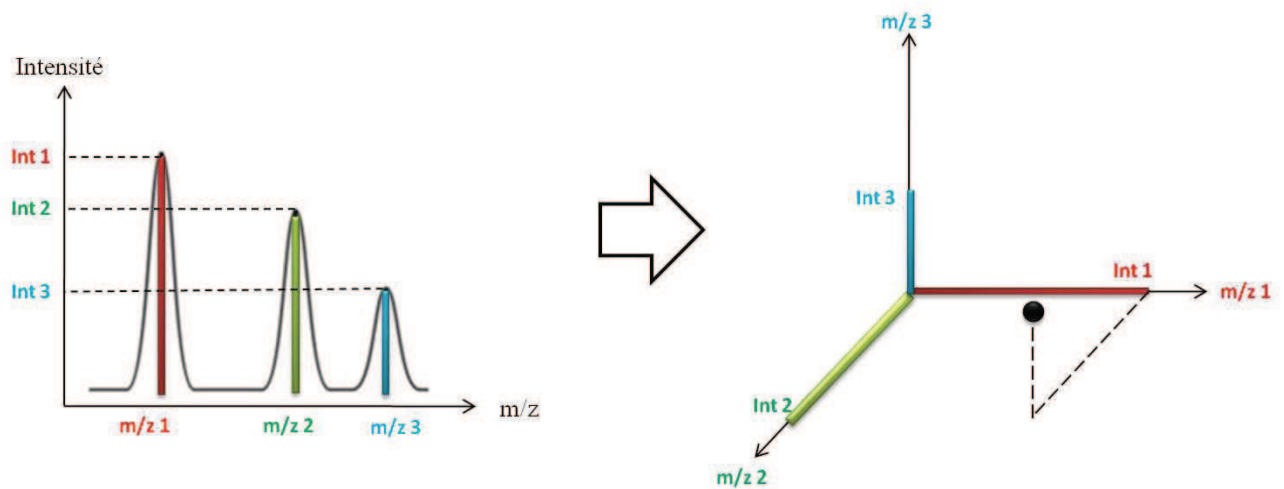
L'imagerie MALDI est capable de mettre en évidence des différences moléculaires entre plusieurs zones histologiques indétectables par des techniques histologiques classiques. Actuellement, pour comprendre des résultats d'imagerie MALDI, il est encore nécessaire d'étudier au préalable les zones histologiques du tissu à l'aide de coloration à un niveau microscopique. Le principal challenge de la MSI dans les années à venir reste donc de corréler les données histologiques et les données de MSI en une seule expérience.

Comme nous l'avons vu précédemment, la MSI va engendrer des jeux de données très importants par image et donc par tissu (10 000 spectres/image en moyenne). Chaque spectre contenu dans ces jeux de données possède l'information de la zone d'impact laser d'où il provient, et donc de l'ensemble des molécules ionisées dans cette zone. On imagine donc facilement la multitude d'informations qu'il nous faut traiter et surtout la difficulté à trouver l'information précise et de manière spécifique au milieu de ces données. Jusqu'à présent, il était très difficile de correctement analyser la totalité des données et de ne pas passer à côté d'informations qui pouvaient s'avérer très intéressantes. C'est pourquoi des méthodes statistiques, comme les analyses multivariées, permettant la simplification des données tout en gardant la totalité des informations ont présenté un réel intérêt en MSI. La méthode la plus adaptée à ce genre de problématique est l'analyse en composante principale (PCA). Cette approche a tout d'abord été utilisée en imagerie SIMS et connaît une popularité grandissante en imagerie MALDI ces dernières années. Il s'agit d'une méthode fiable qui peut être utilisée de manière supervisée ou non-supervisée. Pour permettre une meilleure analyse à l'échelle histologique, la classification hiérarchique (CH) permet la reconstruction d'images basée sur la similarité spectrale. Cette méthode qui permet la combinaison de l'analyse histologique et moléculaire d'un tissu est appelée histologie moléculaire.

Comme nous l'avons vu, les images MALDI ne peuvent réellement être interprétées en les associant avec des colorations histologiques. C'est pourquoi la possibilité d'effectuer la MSI et la coloration histologique sur la même coupe de tissu s'est avéré d'une grande aide pour l'analyse des résultats [9, 136, 166]. Pour ces raisons, le dépôt de la coupe de tissu d'intérêt sur une lame transparente et conductrice de type indium tin oxyde (ITO) permet à la fois l'acquisition en MSI via la conductivité électrique et l'analyse microscopique via la transparence optique.

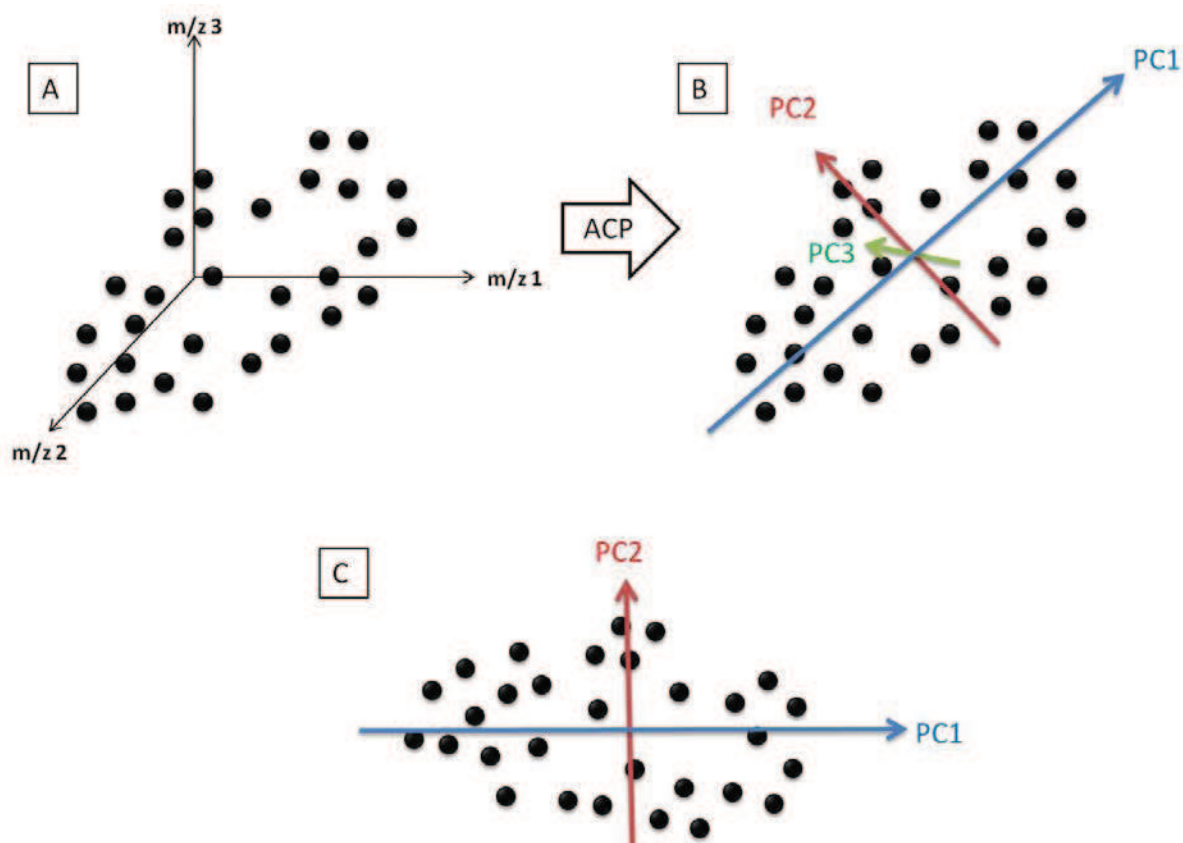
## L'Analyse en Composante Principale pour la MSI

Un spectre de masse est constitué d'un nombre donné de pics. Si ce spectre est constitué de  $n$  pics, alors un espace multidimensionnel à  $n$  dimensions sera créé où chaque axe représente un pic. La valeur de l'intensité de chaque pic au sein de chaque spectre sera reportée sur ces axes (**figure 32**). Cette étape permet la représentation d'un spectre dans un espace multidimensionnel, les données d'imagerie génèrent une grande quantité de spectres qui seront donc représentés par un nuage de points (**figure 33A**).



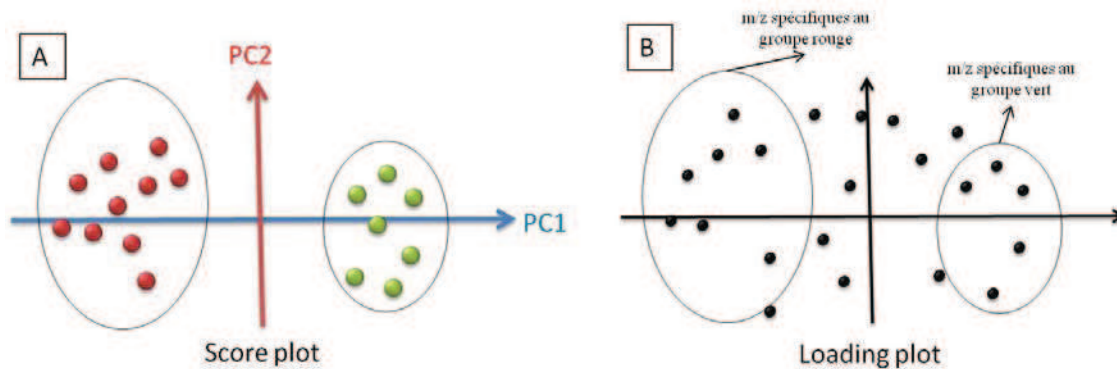
**Figure 32 : Représentation simplifiée d'un spectre de masse à l'aide d'un point dans un plan multidimensionnel.**

Chaque type histologique présent dans un tissu possède le même pattern de molécules, dès lors tous les spectres provenant du même type histologique seront pratiquement identiques au niveau de leur signal moléculaire et de leur intensité, et les points les représentant dans l'espace multidimensionnel seront donc proches car ayant les mêmes coordonnées. L'étape suivante consiste en la modification et en la simplification des axes de l'espace en appliquant de nouvelles coordonnées basées sur les composantes principales (**figure 33B**). La composante principale 1 (PC1) sera fixée comme étant l'axe couvrant la plus grande échelle de valeurs dans le nuage de point, puis se sera PC2, puis PC3, etc. Le plus souvent, cet espace avec ces nouvelles coordonnées sera juste représenté en fonction de PC1 et PC2 qui permettent la meilleure séparation et donc la meilleure comparaison des spectres les uns par rapport aux autres (**figure 33C**). Ce type de graphique comprenant les spectres est appelé « score plot ».



**Figure 33 : Simplification des données de MSI par PCA.**  
 (A) : schéma d'un jeu de données de MSI dans un plan multidimensionnel. (B) : Simplification de l'espace multidimensionnel à l'aide des composantes principales, exemple avec PC1, PC2 et PC3. (C) : représentation classique des résultats en fonction de PC1 et PC2.

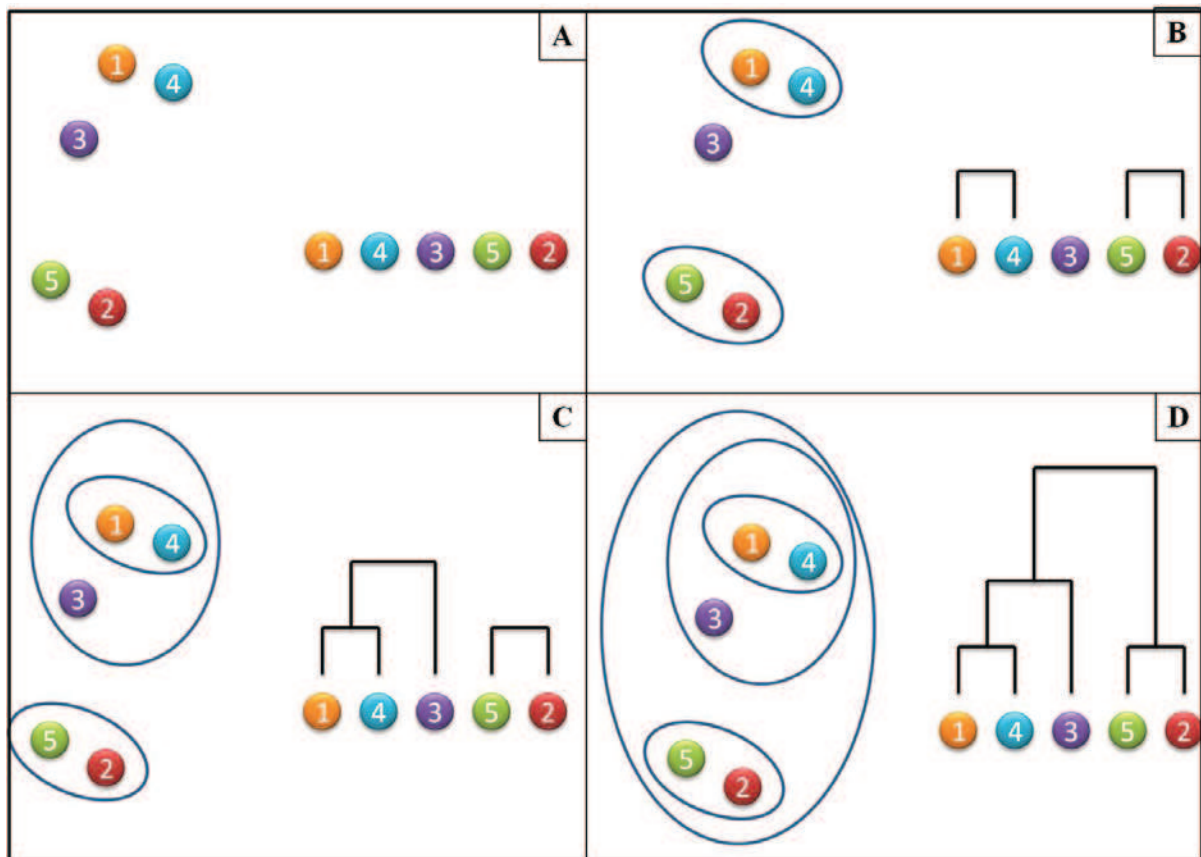
A partir du « loading plot », si des groupes de spectres se séparent suivant les composantes principales, c'est donc que les coordonnées dans l'espace sont différentes et par conséquent les  $m/z$  et leurs intensités respectives sont différents (**figure 34A**). Le loading plot est une représentation de la variance des  $m/z$ , il est alors possible de détecter et de lister les  $m/z$  spécifiques aux groupes de spectres d'intérêt (**figure 34B**).



**Figure 34 : Analyse des résultats PCA en mode supervisé, formation de deux groupes de spectres au niveau du loading plot (A) avec la possibilité de détecter les  $m/z$  spécifiques (représentés par des points) à ces groupes sur le loading plot (B).**

## La classification hiérarchique pour la MSI

La classification hiérarchique (CH) permet la création de clusters qui représentent des groupes de spectres similaires issus de données d'imagerie. La comparaison des spectres s'effectue à partir de l'espace à  $n$  dimensions discuté précédemment (**figure 33A**). Le résultat de la CH se présente sous la forme d'un dendrogramme qui se calcule de la manière suivante : les deux points les plus proches forment un premier cluster qui est alors considéré comme un nouveau point. Les deux points les plus proches suivants forment ensuite un autre cluster. Ce processus est répété jusqu'à ce que tous les points aient été regroupés et forment le dendrogramme (**figure 35**). La longueur des branches du dendrogramme représente la distance entre deux points (ou entre deux clusters).



**Figure 35 : Principe de la classification hiérarchique aboutissant à la formation d'un dendrogramme.**




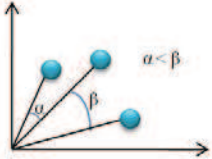
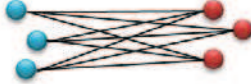
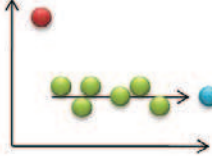
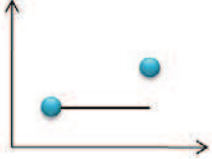

La création du dendrogramme dépend donc de deux valeurs : la distance entre les points et la liaison entre les clusters. Ces deux valeurs peuvent être calculées à partir de différentes méthodes qui sont schématisées dans le **tableau 8** et qui sont les suivantes :

- Euclidien : simple distance entre deux points
- Minkowski : la distance est définie par les valeurs de distance entre les points suivant les coordonnées.
- Cosine : la distance est définie par l'angle entre les vecteurs des points par rapport à l'origine du plan multidimensionnel. Le plus petit angle déterminera la plus petite distance.
- Corrélation : La distance est définie par la corrélation d'un point par rapport à un cluster.
- Chebychev : il s'agit de la distance maximum suivant une seule coordonnée.
- Spearman (non représenté dans le **tableau 8**) : au plus les patterns des spectres seront identiques, au plus ces derniers seront proches.

Les méthodes de calcul des liaisons entre les clusters sont (**tableau 8**):

- Simple : la liaison se fait entre les deux points les plus proches de deux clusters.
- Moyenne : la liaison s'effectue au niveau de la distance moyenne entre tous les points.
- Ward : une fonction erreur de Gauss est définie, elle détermine la distance moyenne entre tous les points des clusters par rapport au centre de gravité de chaque cluster. Ce qui revient à soustraire la fonction erreur de chaque cluster à la valeur de la fonction erreur de tous les points.

L'analyse en composante principale et la classification hiérarchique se basent donc toutes les deux sur des données représentées dans un espace multidimensionnel. Il est important de noter qu'il est possible d'utiliser la classification hiérarchique directement sur les données réduites de PCA, ce qui permet ainsi une simplification de la classification hiérarchique.

Distance	Liaison
Euclidien 	Simple 
Minkowski 	
Cosine 	Moyenne 
Corrélation 	
Chebychev 	Ward 

**Tableau 8: Schématisation des méthodes de calcul des distances entre les points et des liaisons entre les clusters pour la classification hiérarchique.**

## Les proprotéines convertases

Les différents outils utilisés en protéomique que nous venons de décrire ci-dessus connaissent un essor considérable car ils permettent des analyses poussées du protéome d'un individu à des temps donnés. Nous savons aujourd'hui que le nombre de peptides et de protéines qui résultent d'un génome donné dépend de multiples facteurs qui engendrent la diversité et la spécificité au sein du vivant. La plupart de ces molécules sont modifiées et maturées directement à partir du produit de traduction issu de l'ARNm, on parle alors de précurseurs des protéines. Les premiers événements gouvernant les modifications sur les chaînes d'acides aminés des protéines sécrétées sont la N-glycosilation et le clivage du peptide signal. Ces étapes sont suivies des glycosilations et des changements conformationnels permettant à la protéine d'aboutir à sa forme finale au sein du trans golgi network (TGN). Au cours des années 1960, la communauté scientifique s'est rendu compte que la plupart des protéines sécrétées subissaient au moins un clivage au cours de leur voie de sécrétion dans le réticulum endoplasmique (RE) ou le Golgi. Ce premier événement de clivage induit des changements conformationnels au niveau des protéines, leur conférant ainsi une activité biologique. On dénombre aujourd'hui de nombreuses classes de protéines subissant ce mécanisme de maturation. On retrouve des peptides endocriniens, des hormones, des facteurs de croissance et leurs récepteurs, des facteurs de transcription, des molécules d'adhésion, des protéines de la matrice extracellulaire, des protéases et certaines glycoprotéines.

A la suite du séquençage du génome humain, il a été déterminé que 566 protéases étaient produites, contre 644 chez la souris (**Figure 36**), dont seule une minorité était impliquée dans les phénomènes de maturation intracellulaire [167, 168]. La famille des proprotéines convertases de type subtilisine/kexine (PCs) compose la majorité de ces protéases.



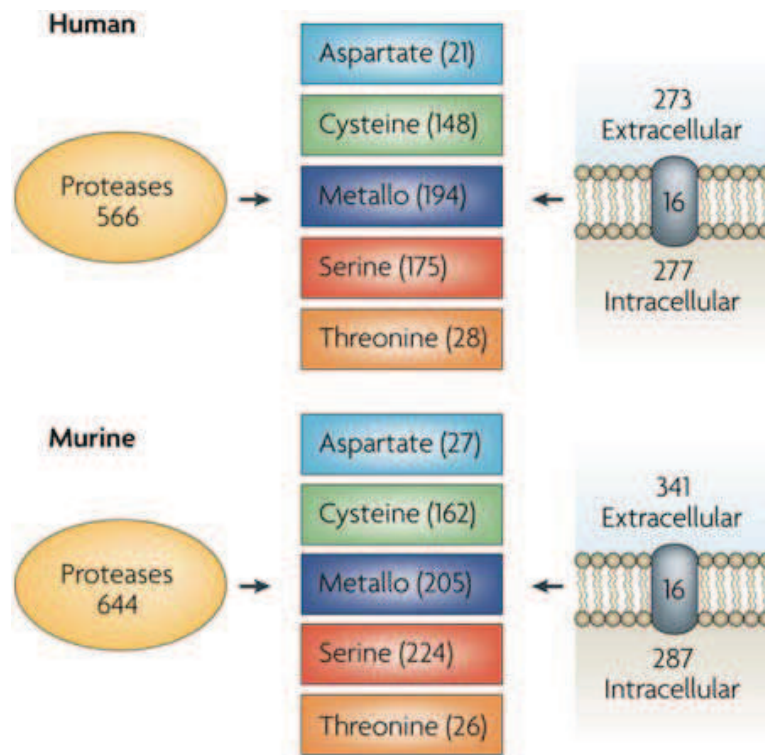


Figure 36 : le répertoire des protéases humaines et murines [168]

#### 2.1.4 Origine et nomenclature des PCs.

Tout démarre en 1967 par la mise en évidence par gel filtration de l'existence d'un précurseur de l'insuline par l'équipe de Steiner [169]. Ils suggèrent alors le nom de proinsuline pour cette protéine ayant une masse moléculaire plus élevée que l'insuline. Ainsi l'idée que les hormones et les protéines provenaient de molécules ayant une chaîne en acides aminés plus longue et impliquaient des étapes de protéolyses au niveau du réticulum endoplasmique ou des vésicules de sécrétion fit son chemin. La découverte des protéases impliquées dans ces mécanismes commença dans les années 1980, et c'est suite à la découverte d'une enzyme ayant une similarité de séquence avec la kexine, nommée furine, que les enzymes de maturation protéolytique chez les mammifères furent identifiées [170, 171]. En raison de la grande homologie de leur domaine catalytique avec celui de la subtilisine, une protéase à sérine bactérienne, ces enzymes furent nommées subtilisin-like proprotein convertases (PCs) [172]. Jusqu'à aujourd'hui, sept membres de la famille des PCs ont été identifiés chez les mammifères. Ces enzymes qui clivent leurs substrats au niveau des résidus basiques sont : Furine (PACE), PC1 (PC3), PC2, PC4, PACE4, PC5 (PC6, possédant deux isoformes : A et B) et PC7 (LPC, PC8, SPC7) (Tableau 9). Il existe deux autres PCs

distinctes de la catégorie précédente : SKI-1/SP1 (subtilisin/kexin-like isoenzyme-1 ou Site-1 protease) et NARC-1/PCSK9 (neuralapoptosis-regulated convertase-1 ou proprotein convertase subtilisin kexin 9). Elles furent récemment identifiées et sont apparentées respectivement à la pyrolysine et à la protéinase K [173]. Le nombre de noms associés aux PCs est lié aux découvertes souvent simultanées de ces enzymes par différents groupes de recherche. C'est pourquoi, lors du congrès *Gordon research conference : pro-protein processing, trafficking and secretion* de 2004, une nomenclature internationale fut choisie (**Tableau 9**).

Gènes (Homo sapiens)	Enzymes	Noms alternatifs
PCSK3	Furin	PACE, PC1, SPC1
PCSK2	PC2	SPC2
PCSK1	PC1/3	PC1, PC3, SPC3
PCSK6	PACE4	PACE4, SPC4
PCSK4	PC4	SPC5
PCSK5	PC5/6	PC5, PC6, SPC6
PCSK7	PC7	PC8, LPC, SPC7
<i>PCSK9</i>	<i>PCSK9</i>	NARC1

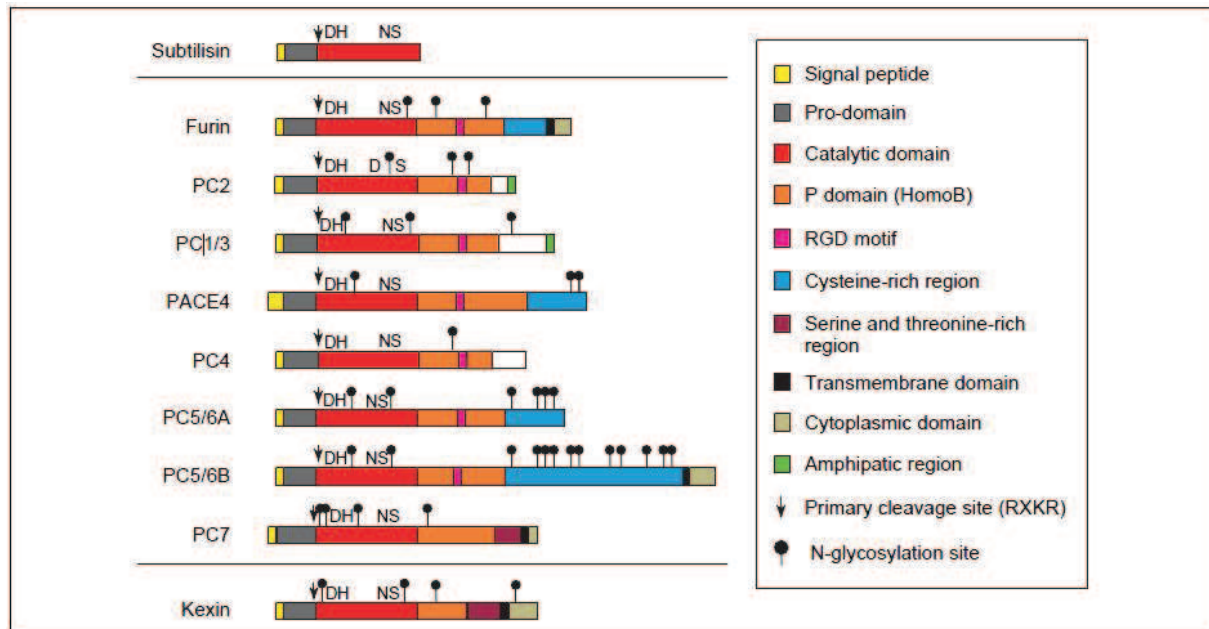
**Tableau 9 : Nomenclature des PCs chez l'Homme d'après Fugère et al [174].**

La base de données MEROPS (<http://merops.sanger.ac.uk>) utilise et propose une classification des protéases [175]. Chaque enzyme est associée à une famille sur la base de son homologie de séquence en acide aminé par rapport aux autres, et les familles sont regroupées en clans suivant leur homologie.

### 2.1.5 Structure et mode d'action des PCs.

Les PCs ont une structure homologue qui est composée d'un peptide signal en N-ter impliqué dans la voie de sécrétion ; d'un pro-domaine qui joue un rôle de chaperon intramoléculaire pour le transport ; un domaine catalytique qui contient un site actif responsable de l'interaction et du clivage des protéines cibles ; un domaine P qui est une unité structurale indépendante essentielle au maintien de la cohésion structurale de l'enzyme et à la reconnaissance du substrat en interagissant avec le domaine catalytique, mais

également impliqué dans la régulation de la stabilité, de la dépendance au calcium et au pH de l'enzyme ; et enfin un domaine C-ter propre à chaque enzyme et permettant, dans certains cas, une interaction avec la membrane cellulaire (**Figure 37**).

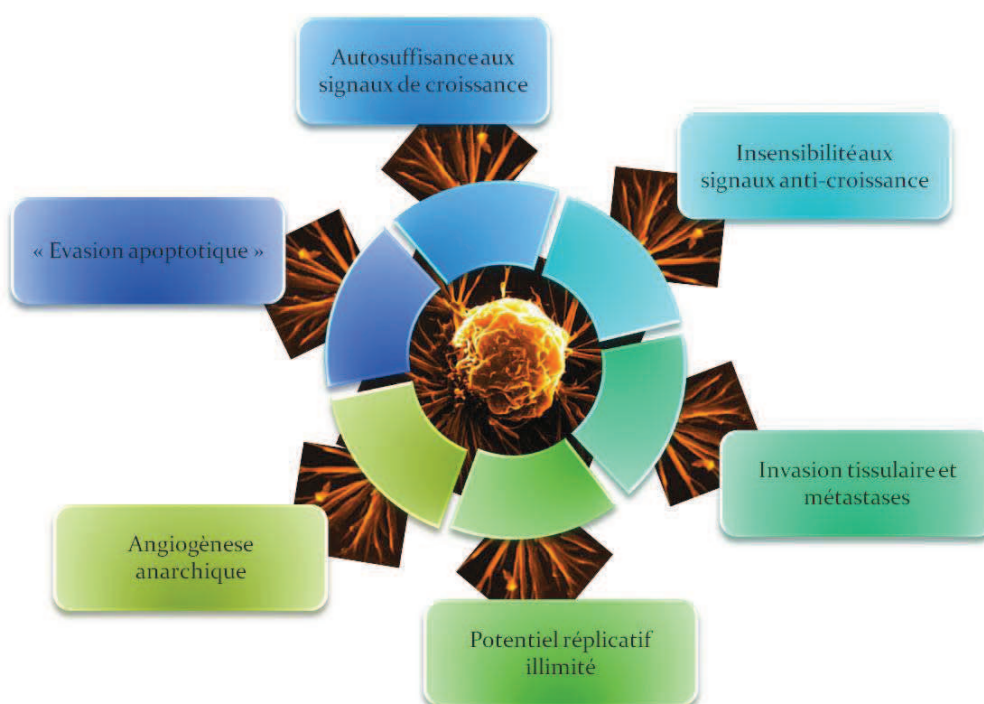


**Figure 37 : Structure des PCs. [174]**

Ces enzymes clivent les précurseurs de protéines au niveau des résidus basiques et reconnaissent le motif  $(K/R)-(X)_n-(K/R) \downarrow$  où  $n = 0, 2, 4$  ou  $6$  et où  $X$  peut représenter n'importe quel acide aminé à l'exception d'une cystéine [176, 177]. Elles possèdent ce qu'on appelle la « triade » catalytique propre aux protéases à sérine : l'aspartate, l'histidine et la sérine. Ces enzymes sont également dépendantes du calcium et du pH pour générer leur forme active obtenue à partir de la proenzyme. Certaines PCs, comme les PC1, PC2 et PC5A, sont activées dans la voie de sécrétion des cellules, contrairement aux protéines transmembranaires que sont la furine, PACE4, PC5B et PC7 qui semblent agir entre le TGN et la membrane des cellules. Tout comme leurs substrats, le pro-domaine des PCs est éliminé suite à un clivage au niveau d'un site en acide aminé basique, ce qui peut être considéré comme une auto-activation des PCs. Il est important de noter que PCSK9, qui n'est pas représentée dans la **figure 37**, n'a pas de substrat connu jusqu'à ce jour et semble pour l'instant impliquée dans l'interaction avec les récepteurs au LDL (low density lipoprotein) empêchant ainsi leurs « recyclage » et aboutissant à une augmentation significative des risques de maladies liées au cholestérol [178-180].

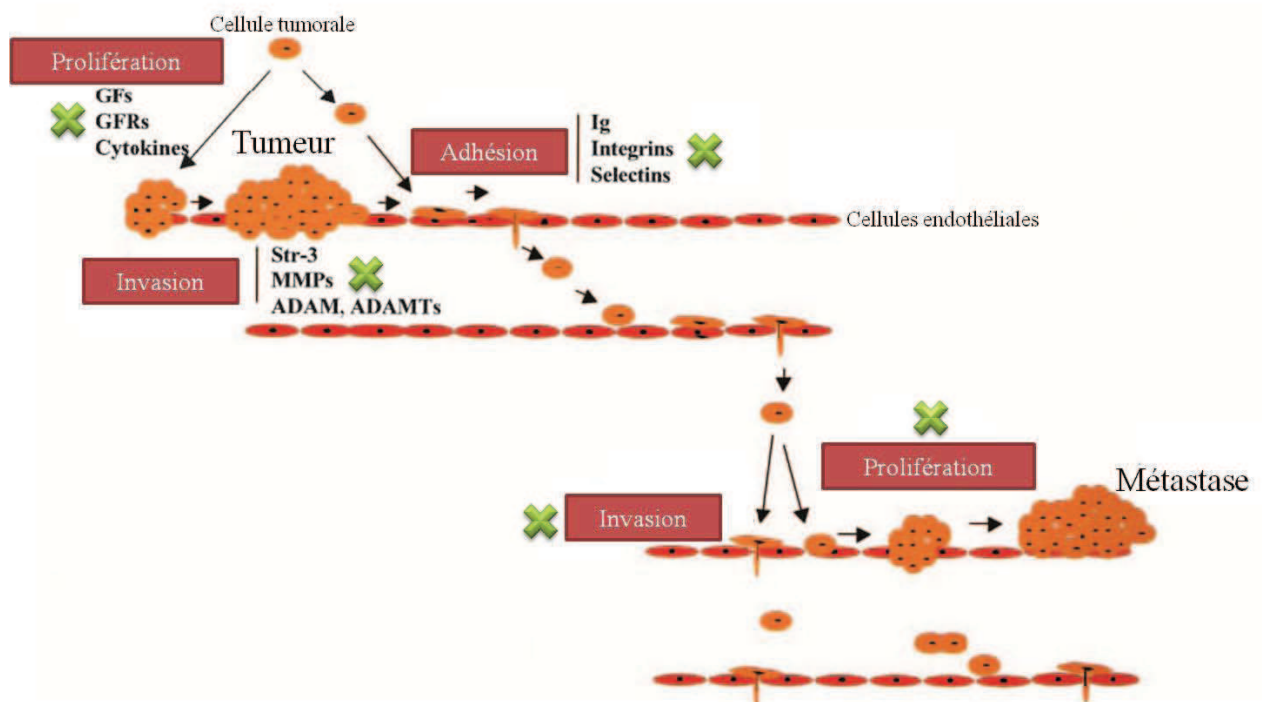
## 2.1.6 Proprotéines convertases et Cancers

La progression tumorale est caractérisée par l'acquisition d'aberrations phénotypiques des cellules au cours du temps. Parmi ces changements phénotypiques, on retrouve une autosuffisance aux signaux de croissance, une insensibilité aux signaux anti-croissance, un potentiel réplcatif illimité, une angiogénèse anarchique, une évation apoptotique et enfin un phénomène d'invasion tissulaire et de métastases (**Figure 38**).



**Figure 38 : Capacités acquises des cellules cancéreuses au cours du temps.**

La formation des métastases résulte en un détachement de cellules provenant d'une tumeur « primaire » envahissant ainsi les autres tissus de l'individu. Suite à l'implantation de ces cellules cancéreuses, une nouvelle zone se forme et peut également être une source de dissémination de cellules et engendrer des métastases secondaires puis tertiaires. Ces processus complexes sont aujourd'hui connus pour impliquer des facteurs de croissance et des protéines d'adhésion. Hors, ces molécules ont été décrites comme étant activées par des protéolyses impliquant les proprotéines convertases (**Figure 39**) [181].



**Figure 39 : Schéma des événements impliquant les PC dans la croissance tumorale et les métastases. Les croix vertes représentent les sites d'action potentiels des PC. (tiré de [181]).**

### 2.1.6.1 Rôle des PC dans la prolifération tumorale

La prolifération cellulaire est finement contrôlée entre l'activation et l'inhibition de la prolifération et ce, par l'action de facteurs croissance dans un tissu normal. La production et l'activation de ces facteurs sont contrôlées et régulées pour maintenir l'intégrité du tissu et la fonction de l'organe. A l'aide de plusieurs mécanismes, les cellules malignes échappent à ce contrôle conduisant ainsi à une croissance aberrante des cellules. Il est maintenant parfaitement établi que la croissance tumorale résulte en une aberration de nombreux facteurs de croissance et en une surexpression de ces molécules dans les tumeurs aboutissant malheureusement à des pronostics cliniques difficiles.

Les PDGFs, IGFs, PTH et les endothelines sont connus pour être maturés au niveau de leurs précurseurs par protéolyses impliquant les PCs.

Les PDGFs (platelet-derived growth factor) sont des facteurs de croissance qui stimulent plusieurs fonctions cellulaires comme la croissance, la prolifération et la différenciation. De nombreuses études ont rapporté l'expression de ces molécules dans différents types de tumeurs. Selon le type de tumeur, le rôle biologique des PDGFs peut aller

d'une stimulation autocrine à des interactions paracrines permettant à la fois la progression tumorale, l'angiogenèse et la création de métastases. Les PDGFs sont des dimères liés par des ponts disulfides composés de deux chaînes polypeptidiques, appelées A et B, et sont représentés par cinq membres : PDGF-AA, PDGF-AB, PDGF-BB, PDGF-C et PDGF-D dont le rôle est modulé par deux récepteurs de type tyrosine kinase : PDGFR- $\alpha$  et PDGFR- $\beta$ . Certains PDGFs, comme PDGF-A et PDGF-B ont été décrits comme étant activés par les PCs, PDGF-C et PDGF-D, découverts plus récemment. Ils présentent, eux, des sites potentiels de clivage par les PCs, mais qui restent encore à démontrer.

Habituellement, les cellules normales ont nécessairement besoin des facteurs de croissance pour leur prolifération. Généralement, plus d'un facteur de croissance est nécessaire. In vivo, le développement tumoral dépend grandement de l'activation des voies impliquant les facteurs de croissance [182]. Beaucoup de ces facteurs de croissance sont synthétisés sous forme de proprotéines et sont activés par les PCs (**Tableau 10**) [181].

Une augmentation de l'expression de facteurs de croissance et de PCs en même temps favorise la croissance tumorale. D'ailleurs, l'effet anti-prolifératif de  $\alpha 1$ -PDX, un inhibiteur des proprotéines convertases, entraîne une inhibition des fonctions et des processus impliquant des facteurs de croissance [183, 184]. Kayo et al. ont montré que du milieu conditionné provenant d'une lignée MIN6 surexprimant la furin stimulait la croissance de la lignée contrôle, alors que du milieu conditionné provenant de cellules exprimant  $\alpha 1$ -PDX la réduisait. Cette observation permet de conclure qu'une surexpression de la furin stimule la croissance via des voies autocrine ou paracrine [185]. Les implications de PDGF-A [186], de TGF- $\beta 1$  [187] et de VEGF-C [188] ont également été décrites dans la tumorigenèse.

Facteurs de croissance	P <sub>6</sub> P <sub>5</sub> P <sub>4</sub> P <sub>3</sub> P <sub>2</sub> P <sub>1</sub> ↓P <sub>1</sub> 'P <sub>2</sub> '	NCBI, Accession
Insulin site-1	TPKTRREA	XP028180
site-2	GSLQKRG	
IGF-1	PAKSARSV	P01343
IGF-2	PAKSERDV	XP028189
PDGF-A	PIRRKRSI	NP002598
PDGF-B	LARGRRSL	NP148937
PDGF-C	FGRKSRVV	NP057289
PDGF-D	HDRKSKVD	AAK56136
VEGF-C	HSIIRSL	P49767
VEGF-D	YSIIRSI	NP_004460
TGF- $\alpha$ 1	SSRHRRAL	XP_008912
Leftyprotein	RSRGKRFS	O00292
BMP-2 site-1	HVRISRSL	
site-2	HKREKRQA	NP_058874
BMP-4 site-1	HVRISRSL	
site-2	RRRAKRSP	NP_570912
BMP-7	HFRSIRST	NP_001710
Pancreatic polypeptide	PRYGKRHK	P01298
Gastrin	ASHHRRQ	P01350
GHRH site-1	PFRMQRHV	
site-2	RARLSRQE	NP_034415
FGF-23	PRRHTRSA	Q9GZV9
EGF site-1	HHYSVRNS	P01133
site-2	KWWELRHA	
Endothelin-1	LRRSKRCS	P05305
PTHrP	SRRLKRAV	P12272
Parathyroid hormone	KSVKKRSV	XP031173
Neurotrophin-3	TSRRKRYA	P20783
Neurotrophin-4	NRSRRGVS	A42687
NGF	THRSKRSS	XP002122
BDNF	SMRVRRHS	XP006027
APRIL	RSRKRAV	O75888
BAFF	NSRNKRAV	Q9Y275
HB-EGF	RDRKVRDL	Q99075
HGF	KTKQLRVV	XP052260
LEAP-2	LCKRRC	NP_694709
Virokinin site-1	LQRIARRP	
site-2	GLMGKRDA	NM_013998

Tableau 10 : Facteurs de croissance et sites de clivage pour les PCs. Tiré de [189]

Comme nous venons de le voir, une surexpression des PCs semble entraîner une augmentation de la prolifération tumorale dans de nombreux cancers via l'augmentation de la quantité de facteurs de croissance. Mais il est également important de préciser que beaucoup de récepteurs de ces facteurs de croissance possèdent des sites de clivage pour les PCs, et sont donc impliqués dans le cancer, comme par exemple HGF-R [190], le récepteur à l'insuline [191], le récepteur à IGF-1 [192] (**Tableau 11**).

Récepteur de facteur de croissance	P <sub>6</sub> P <sub>5</sub> P <sub>4</sub> P <sub>3</sub> P <sub>2</sub> P <sub>1</sub> ↓P <sub>1</sub> 'P <sub>2</sub> '	NCBI, Accession
Insulin receptor	PSRKRRSL	XP048347
IGF-1 receptor	PERKRRDV	IGHUR1
HGF receptor	EKRKKRST	P08581
DREG	KIKVKRSL	NP_001027567
sorCS1 Site-1	SGRRRRSG	
Site-2	ASRSPRGV	NP_443150
SorLA receptor	PLRRKRSA	U60975
Ldl-related protein	SNRHRRQI	Q07954
Leptin receptor	QVRGKRLD	P48357
Notch-1-receptor	SRKRRRQH	AAG33848
R-PTP-kappa	PHRTKREA	NP_002835

**Tableau 11 : Sites potentiels de clivage des récepteurs de facteur de croissance par les PCs. Tiré de [189]**

### 2.1.6.2 PCs et adhésion cellulaire

Les molécules d'adhésion cellulaire contrôlent le trafic cellulaire, les migrations au niveau de l'endothélium et l'inflammation, mais sont aussi impliquées dans la colonisation tumorale. Parmi les protéines impliquées, on retrouve les intégrines, les immunoglobulines, ou encore les cadhérines.

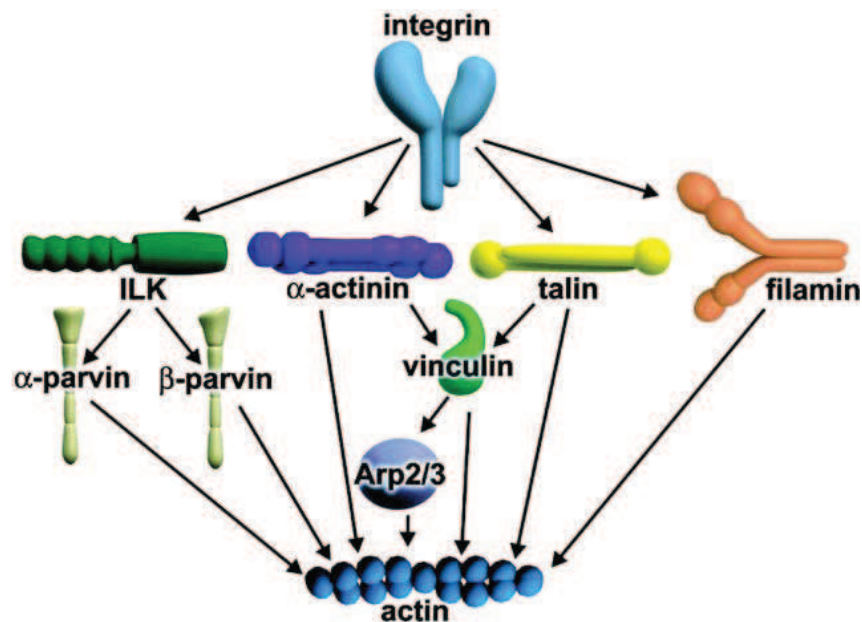
#### Les intégrines

Les intégrines sont constituées de 2 sous-unités : la chaîne  $\alpha$  qui fixe des cations divalents et la chaîne  $\beta$  qui possède un domaine riche en cystéine. On peut différencier 3 domaines : l'un extracellulaire (extrémité N-terminale), l'un transmembranaire et le dernier intracellulaire (en C-terminal). Il existe 18 sous-unités  $\alpha$  et 9 sous-unités  $\beta$  différentes, et leur



combinaison détermine la spécificité de l'intégrine au ligand. A ce jour, on dénombre 24 couples (dimères)  $\alpha\beta$  fonctionnels [193, 194].

La plupart d'entre elles se lient aux molécules, comme la fibronectine, vitronectine, le collagène, et la laminine de la matrice extracellulaire (par leur région extra-cellulaire) et aux microfilaments d'actine via un certain nombre de protéines de liaison, comme l'ILK, l' $\alpha$ -actine, la taline et les filamines, qui s'associent à leur région intracellulaire (**Figure 40**). Ces processus font de ces récepteurs des molécules jouant un rôle prépondérant dans l'adhésion cellulaire. Il existe des exceptions comme le dimère  $\alpha6\beta4$  qui est relié à des filaments intermédiaires et constitue les héli-desmosomes au sein des épithéliums. Certaines intégrines jouent aussi un rôle dans l'adhésion cellule-cellule (par exemple les intégrines médiant l'adhésion des cellules du système immunitaire aux cellules endothéliales avant leur sortie des vaisseaux sanguins sur le site d'une inflammation).



**Figure 40 :**Schéma représentant les voies par lesquelles les intégrines interagissent avec l'actine du cytosquelette [195].

Le domaine extracellulaire de chaque sous-unité est nécessaire pour la liaison au ligand et le domaine cytoplasmique interagit donc avec le cytosquelette entraînant des changements dans la structure cellulaire, la mobilité ou dans la transduction des signaux de croissance et/ou de survie [196, 197].

Parmi les dix-huit sous unités  $\alpha$  connues, neuf possèdent un site de clivage potentiel pour les PC5 (Tableau 12), et  $\alpha3$ ,  $\alpha4$ ,  $\alpha5$ ,  $\alpha6$  et  $\alpha v$  ont été décrites comme étant des substrats de PC5 et de furin [198, 199].

<b>Integrin</b>	<b>P<sub>6</sub>P<sub>5</sub>P<sub>4</sub>P<sub>3</sub>P<sub>2</sub>P<sub>1</sub>↓P<sub>1</sub>'P<sub>2</sub>'</b>	<b>NCBI, Accession</b>
Integrin α3	PQRRRRQL	XP008432
Integrin α4	HVISKRST	XP039011
Integrin α5	HHQQKREA	AAH08786
Integrin α6	NSRKKREI	NP000201
Integrin α7	RDRRRREL	Q13683
Integrin α8	HLVRKRDV	AAA93514
Integrin α <sup>E</sup>	TARQRRAL	XP008508
Integrin αIIb	HKRDRRQI	P08514
Integrin αv	HLITKRDL	XP002379

**Tableau 12 : Sites potentiels de clivage intégrines par les PCs. Tiré de [189]**

L'activation ou l'altération de l'expression des intégrines furent démontrées comme étant liées au cancer. En effet, les MMPs (Matrix metalloproteinases) et l'activité urokinase de nombreuses cellules cancéreuses sont liés à l'activation des intégrines comme pour le cancer du côlon ou dans certains mélanomes [200-202]. En effet, Une fois liées à leur ligand elles activent des voies de transduction intracellulaires faisant intervenir des protéines kinases. L'une des principales protéines activées est FAK (Focal Adhesion Kinase), une tyrosine kinase cytoplasmique [199]. Elle se lie aux intégrines et s'autophosphoryle lorsque les intégrines forment des regroupements appelés point d'adhésion focaux qui sont des points d'ancrage de la cellule à la matrice extracellulaire. FAK phosphoryle alors sur des tyrosines d'autres protéines-cibles tels l'oncogène Src ou paxilline.

### Les immunoglobulines.

La famille des immunoglobulines (Ig) comporte des molécules d'adhésion intercellulaires ICAM-1, ICAM-2 et ICAM-3, la molécule d'adhésion des cellules vasculaires (VCAM-1), et la Mucosal addressin cell adhesion molecule (MadCAM-1). Aucune de ces Ig n'est le substrat des proprotéines convertases, néanmoins, les PCs semblent être nécessaires à leur expression et à leur activation. ICAM-1 et VCAM-1 ont des expressions induites par des cytokines et des facteurs de croissance au niveau des cellules endothéliales. On retrouve l'interferon-γ (INF-γ), interleukin-1 (IL-1), tumor necrosis factor (TNF-α), insulin-like growth factor (IGF-1) et les endothelins [203-208]. Certaines de ces molécules sont directement maturées par les PCs comme pour l'IGF-1 et les endothelins. Les autres sont indirectement liées aux PCs, par exemple tumor necrosis factor-α converting enzyme (TACE) [209] et ADAM10 [210], qui interviennent dans le processus du proTNF-α, sont maturées par les PCs.

### Les cadhérines.

Plus de 80% des cancers sont d'origine épithéliale. Dans les tissus normaux, les cellules épithéliales sont interconnectées via différents types de jonctions qui sont associés avec l'actine et donc le cytosquelette [211, 212]. La cadhérine épithéliale (E-cadhérine) est la cadhérine majoritaire présente dans ces jonctions, elle est liée à l'actine à partir d'interactions impliquant les caténines. Au cours du développement des cancers épithéliaux, la fonction d'adhésion des E-cadhérines est souvent modifiée et corrélée avec la croissance tumorale et les métastases [213, 214]. D'autres cadhérines ont également été étudiées, par exemple la N-cadhérine favorise l'invasion cellulaire [215] et la cadhérine-6 favorise les métastases [216], contrairement à la H-cadhérine dont l'expression diminue dans le cancer du sein [217]. Toutes ces cadhérines sont synthétisées sous forme de précurseurs qui nécessitent un clivage protéolytique via les PCs pour leurs maturations.

### 2.1.6.3 Implications des PCs dans l'invasion tumorale

Trois étapes ont été identifiées dans le processus d'invasion cellulaire [218] :

- L'attachement à la membrane basale ou à la matrice stromale
- La protéolyse des composants du stroma comme le collagène ou/et les laminines
- L'augmentation de la mobilité vers les systèmes sanguin et lymphatique.

### Les métalloprotéinases (MMPs)

Les MMPs catalysent la dégradation du collagène, de la gélatine, des laminines, ayant pour conséquence une déstructuration de l'ECM (extra cellular matrix) et donc du tissu. Ces enzymes sont régulées sous des conditions physiologiques normales, et peuvent s'autoprotéolyser en cas d'activités anormales engendrant en même temps une restriction temporelle et spatiale de leurs expressions qui fait intervenir les tissus inhibiteurs of metalloproteinases (TIMPs) [219] et RECK [220].

Les MMPs sont synthétisées sous forme de zymogènes inactifs qui nécessitent une étape de protéolyse pour être activés. Plusieurs protéases sont responsables de cette maturation comme la plasmin ou les MMPs elles-mêmes. La stromelysin-3 et les membranes type metalloprotéinases (MT-MMPs) contiennent le motif de reconnaissance en acide aminé des PCs suggérant ainsi leurs implications dans le clivage et l'activation des MMPs.

L'invasion tumorale est possible grâce à une dégradation de l'ECM impliquant des protéases. C'est pourquoi l'implication des MMPs dans l'invasion tumorale et les métastases fut rapidement initiée [221]. Ainsi il a été observé qu'une surexpression des MMPs dans de nombreux types de cancers (sein, pancréas, colon, ovaire, peau) était souvent associée à un pronostic pessimiste [222, 223]. Dans certaines tumeurs, cette surexpression aboutissant à la forme active des enzymes indique que leur activité protéolytique est nécessaire à l'initiation du processus d'invasion [224].

La famille des MT-MMPs contient six membres possédant un domaine transmembranaire qui permet à leurs formes actives de se trouver au niveau de la membrane [225, 226]. Ces enzymes catalysent MMP-2, collagène I, lumican et decorine, les intégrines, la lamine  $\gamma 2$  et syndecan [227-231]. MT1-MMP est le membre de cette famille qui a été le plus étudié. Il est synthétisé sous la forme d'une proenzyme de 63 kDa qui est activée suite à l'intervention de la furine ou de PC5 sous sa forme active à 60kDa [232]. L'inhibition de la furine avec un inhibiteur synthétique (CMK) ou une approche antisens inhibe l'activation de cette enzyme et l'activité de MMP-2 dans les fibroblastes cardiaques et utérins [233] mais pas au niveau des fibroblastes du derme [234]. Ces résultats suggèrent que l'implication des PCs dans la maturation de MT1-MMP, et par conséquent au niveau de la régulation de la modulation collagénolytique, est espèce, tissu ou cellule dépendant.

La rupture de la barrière formée par l'ECM est nécessaire pour les processus de remodelage du tissu. Cette dégradation prend part dans les événements physiologiques mais elle représente également une étape cruciale dans l'invasion tumorale. Le rôle des PCs dans la dégradation de l'ECM est maintenant bien connu, en effet l'inhibition des PCs dans différentes tumeurs résulte d'une diminution du pouvoir invasif. Cette réduction est due à l'arrêt des processus des protéines impliquées dans l'invasion comme les MMPs ou les facteurs de croissance et/ou les intégrines qui induisent l'expression des protéines dégradant l'ECM comme l'urokinase ou les MMPS eux-même. Ces MMPs sont les MT-MMPs, adamalysin metalloproteinases (ADAMs), adamalysine metalloproteinases with thrombospondin motifs (ADAMTS) et stromelysin-3.

Impliqués dans les métastases, les MT-MMPs sont surexprimés dans de nombreux carcinomes comme le cancer du côlon [235]. MT1-MMP mature proMMP2 en MMP-2 qui est une enzyme qui dégrade le collagène IV, impliqué dans la structure des membranes basales. MT1-MMP possède deux motifs de reconnaissance des PCs, nommés ArgArgProArg<sup>92</sup> et ArgArgLysArg<sup>111</sup> qui sont clivés par des enzymes comme la furine sur

ces deux sites [236]. L'analyse des séquences MMPs révèle l'existence de nombreux sites de clivage potentiel pour les PCs (**Tableau 13**). Néanmoins l'implication des PCs n'a pas encore été démontrée dans les processus de maturation de ces enzymes.

MMPs	P <sub>6</sub> P <sub>5</sub> P <sub>4</sub> P <sub>3</sub> P <sub>2</sub> P <sub>1</sub> ↓P <sub>1</sub> 'P <sub>2</sub> '	NCBI, Accession
MMP-1	VMKQPRCG	P03956
MMP-2	TMRKPRCG	P08253
MMP-8	MMKKPRCG	XP006273
MMP-9	AMRTPRCG	XP029934
MMP-13	VMKKPRCG	XP040746
MMP-21	RARSRRSP	NP_671724
MMP-23	APRRRRYT	NP_004650
MMP-28	MRRKKRFA	NP_766385

**Tableau 13 : sites de clivage potentiels des MMPs par les PCs. Tiré de [189]**

Rozanov DV et al. ont montré que la forme active de MT1-MMP était capable d'induire sa propre protéolyse à partir du zymogène [237]. Ce processus autocatalytique génère six formes de MT1-MMP et le séquençage N-terminal de ces fragments révèle que MT1-MMP est capable de cliver son propre pro-domaine au niveau du site de clivage de la furine. Ce nouveau mécanisme peut constituer une voie alternative pour l'activation de MT1-MMP.

ADAMS sont des membres de la famille « membrane-associated multidomain zinc-dependent metalloproteinase »[238]. A ce jour, seize des trente protéines ADAMs sont des enzymes en référence à la séquence conservée liant le zinc (HEXXH) alors que les quatorze autres ne semblent pas être des protéases. Le prodomaine de plusieurs ADAM, comme ADAM12, est clivé par les PCs (**Tableau 14**) [239].

ADAMs	P <sub>6</sub> P <sub>5</sub> P <sub>4</sub> P <sub>3</sub> P <sub>2</sub> P <sub>1</sub> ↓P <sub>1</sub> 'P <sub>2</sub> '	NCBI, Accession
ADAM1	PPRSRKPD	AAA74920
ADAM8	PSRETRYV	XP005675
ADAM9	LLRRRAV	NP003807
ADAM10	LLRKKRTT	XP007741
ADAM12	ARRHKRET	XP005838
ADAM15	HIRRRRDV	Q13444
ADAM17	VHRVKRRA	P78536

**Tableau 14 : sites de clivage potentiels des ADAMs par les PCs. Tiré de [189]**

Cette famille joue un rôle dans plusieurs processus biologiques comme la fertilisation, la myogenèse, la neurogenèse et les protéolyses de la surface des cellules [209, 210, 240-245]. Les protéines cibles des ADAMS sont les cytokines, les facteurs de croissance comme transforming growth factor- $\alpha$  (TGF- $\alpha$ ), epidermal growth factor (EGF), heparin-binding epidermal growth factor (HBEGF), tumor necrosis factor- $\alpha$  (TNF- $\alpha$ ), c-Kit-ligand-1 (KL-1), colony-stimulating factor-1 (CSF-1), Fas-ligand (Fas-L), TNF receptor-I et II (TNFR1, TNFR2), p75 nerve growth factor receptor (p75NGFR), interleukin 6 receptor (IL6R), thyroid-stimulating hormone receptor (TSHR), L-selectin, tyrosine phosphatase (PTP), protein tyrosine phosphatase LAR (LAR), amyloid precursor protein (APP) et angiotensin-converting enzyme (ACE). Un précurseur peut être clivé par plusieurs enzymes, par exemple TNF- $\alpha$  peut être maturé par ADAM17 (TACE) et ADAM10. Aucune séquence consensus propre à ces protéases n'a encore été identifiée. En plus de leur fonction de protéase, certaines ADAMs, comme ADAM12 et ADAM2, peuvent se lier aux intégrines via leur domaine disintegrin [244, 246, 247].

On peut donc en conclure que les ADAMs représentent une nouvelle classe de protéases membranaires jouant un rôle important dans les interactions cellulaires, dans des conditions physiologiques et pathologiques comme le cancer. En effet, dans des cas de tumeurs du sein, de la prostate, des poumons et du colon, ces protéines étaient surexprimées.

ADAM-TSs sont sans doute également maturées par les PCs (**Tableau 15**) et sont des enzymes également impliquées dans de nombreux mécanismes biologiques comme le clivage des protéoglycanes, la dégradation de la matrice extracellulaire, l'inhibition de l'angiogenèse, le développement des gonades et l'organogenèse. ADAM-TS1 a été clonée à partir d'une lignée cellulaire d'adénocarcinome de colon [248]. ADAM-TS1 est capable de former un complexe covalent avec l' $\alpha$ 2-macroglobuline démontrant ainsi qu'elle est une protéase active [249].

ADAM-TS	$P_6P_5P_4P_3P_2P_1\downarrow P_1'P_2'$	NCBI,Accession
ADAM-TS1	SIRKKRFV	Q9UHI8
ADAM-TS2 site1	GVRTRRAA	P79331
site2	RRRMRRHA	
ADAM-TS3	TMRRRRHA	O15072
ADAM-TS4	PRRAKRFA	XP042446
ADAM-TS5/11	WRRRRRSI	Q9UNA0
ADAM-TS13	RQRQRRAA	CAC83682

**Tableau 15 : sites de clivage potentiels des ADAM-TSs par les PCs. Tiré de [189]**

## 2.1.7 Inhibiteurs des PCs : vers une nouvelle voie thérapeutique ?

Pour pallier à la croissance anarchique des tumeurs, de nombreuses stratégies ont été proposées pour inhiber la prolifération cellulaire, l'invasion tissulaire et l'adhésion cellulaire que nous avons vues précédemment. Les découvertes de l'implication des PCs dans ces processus liés aux cancers suggèrent que l'utilisation d'inhibiteurs des proprotéines convertases constituerait une excellente approche pharmaceutique. Les meilleurs inhibiteurs de PCs sont des peptides possédant le motif de reconnaissance des enzymes et donc rentrent en jeu comme un inhibiteur compétitif [250-252]. Néanmoins, la spécificité d'une PC n'est pas clairement établie en raison du motif commun de reconnaissance de ces enzymes. De plus, les PCs sont très souvent co-exprimées dans une cellule à temps donné et chaque PCs ne répondra pas de la même manière à l'inhibiteur, on peut donc imaginer qu'un cocktail d'inhibiteurs serait l'approche thérapeutique la plus adaptée. En effet, Fugère et al. ont montré l'efficacité et la spécificité de certains inhibiteurs sur les PCs (**Figure 41**), La puissance d'inhibition (échelle logarithmique) est représentée par la moyenne des  $K_i$  déterminée *in vitro*. Plus la valeur est faible, plus la puissance d'inhibition est élevée. Les diagonales indiquent que la puissance d'inhibition n'a pas été déterminée.

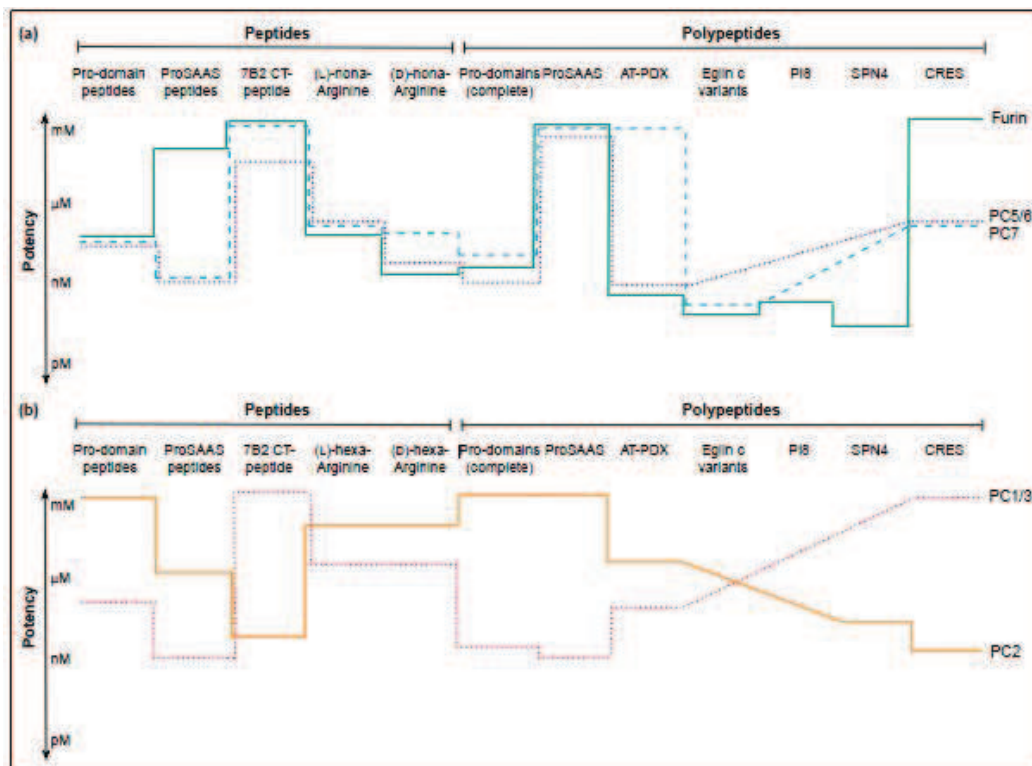


Figure 41 : Potentiel et spécificité des inhibiteurs de PCs [174].

Les inhibiteurs les plus intéressants sont ceux pouvant utiliser les voies de synthèses et les voies métaboliques. Par conséquent, plusieurs inhibiteurs des PCs ont été élaborés ces dernières années sous la forme de protéines, de peptides et de non-peptides [253]. Une liste de ces molécules est résumée dans le **tableau 16**. Des études se poursuivent sur l'impact et l'action de ces inhibiteurs sur des lignées cancéreuses [254].

<b>Protéines</b>	<b>Noms</b>
<b>Macromolécules</b> Endogènes	Pro7B2 ProSAAS $\alpha$ 1-PI Cystatin
Variants	$\alpha$ 1-ATPortland ( $\alpha$ 1-PDX) $\alpha$ 2-Mhexapeptide OMTKY3variant EglinCvariant
<b>Petites molécules</b> Peptides	Dérivés des pro-domaines de PCs Dérivé de Pro7B2 Dérivé de ProSAAS Polyarginine (6–9 résidus) Peptides transportables dans la cellule
Analogues de peptides	Acide aminé synthétique substitué en P1' C-ter modifié avec une fonction nucléophile
Non-peptides	SEA Quinoline bis ou tirs, ou pryridine

**Tableau 16 : Liste des inhibiteurs des PCs [253].**



## **Chapitre 3: Applications des analyses multivariées et de la classification aux données d'imagerie MALDI et application au cancer de la prostate et de l'ovaire.**

---

### **Analyses Multivariées et classification des données**

Les données d'imagerie MALDI peuvent atteindre plusieurs milliers de spectres contenant plusieurs centaines de pics sur une seule coupe de tissu. Il est donc difficile de pouvoir comparer sur une même coupe de tissu deux à trois populations cellulaires différentes dont l'analyse MALDI génèrent des populations d'ion différents tout comme la comparaison de plusieurs coupes issues d'un même bloc de tissu ou de tissu issu de patients différents par exemple. Il faut nécessairement passer par l'analyse multivariée. Celles-ci permettent essentiellement de simplifier des jeux de données tout en gardant la totalité des informations. Il est alors possible de cibler les variables les plus intéressantes et de hiérarchiser les données les unes par rapport aux autres. Les premières études en imagerie faisant appel à l'analyse multivariée ont été par McCombie et al. (2005) [255] et Maddalo et al. (2005) [256] puis de Van de Plas et al (2007) [257] et ensuite Trim et al (2008) [258] et McDonnell et al. (2008) [259]. Cette analyse multivariée a ensuite été intégrée directement dans les logiciels d'imagerie à partir de 2008 [260, 261].

Notre démarche quant à elle a été de valider cette approche multivariée sur des tissus ayant subis des micro-digestions enzymatiques à la trypsine afin de valider l'approche « bottom-up » d'identification des marqueurs mis en évidence par PCA (principal component analysis) mais aussi dans le cadre de l'utilisation de tissu FFPE. Nous avons voulu apprécier le taux de délocalisation lié à la digestion mais aussi par des traitements de type dérivation N-terminale pour l'identification ou pour des approches de type In Source Decay (ISD) lors de l'application de matrice spécifique. Nous proposons également une démarche qui combine à la fois l'interprétation des résultats de hiérarchisation et de PCA pour la recherche fine de marqueurs dans des échantillons présentant une grande complexité histologique : la PCA-SDA (symbolic discriminant analysis).

## Résultats : Publication 1.

### MULTIVARIATE ANALYSES FOR BIOMARKERS HUNTING AND VALIDATION THROUGH ON TISSUE BOTTOM-UP OR IN SOURCE DECAY IN MALDI-MSI

David BONNEL<sup>1,2</sup>, Rémi LONGUESPEE<sup>1,2</sup>, Julien FRANCK<sup>1</sup>, Morad  
ROUDBARAKI<sup>3</sup>, Pierre GOSSET<sup>4</sup>, Robert DAY<sup>2</sup>, Michel SALZET<sup>1\*</sup> and Isabelle  
FOURNIER<sup>1\*</sup>

<sup>1</sup> Université Nord de France, CNRS, MALDI Imaging team, Laboratoire de  
Neuroimmunologie et Neurochimie Evolutives, FRE 3249, Université Lille 1, 59650  
Villeneuve d'Ascq, France

<sup>2</sup> Institut de pharmacologie de Sherbrooke, Faculté de médecine et des Sciences de la  
santé, Université de Sherbrooke, Sherbrooke, (Québec), Canada, J1H 5N4

<sup>3</sup> Université Nord de France, INSERM, U-800, Equipe labellisée par la Ligue  
Nationale contre le cancer; Université Lille 1, 59650 Villeneuve d'Ascq, France

<sup>4</sup> Faculté Libre de Médecine, Laboratoire d'Anatomie et de Cytologie Pathologique du  
groupement hospitalier de l'Institut Catholique de Lille, Lille, France

Soumise dans Molecular & Cellular Proteomics.

**Abstract**

The large amount of data generated using matrix-assisted laser desorption/ionization mass spectrometric imaging (MALDI-MSI) poses a challenge for data analysis. Thus, the tremendous amount of information requires processing techniques using dedicated informatics tools. This work proposes and summarizes an approach that utilizes a multivariable analysis of MSI data. The multivariate analysis, such as principal component analysis-symbolic discriminant analysis (PCA-SDA), can remove and highlight specific  $m/z$  from the spectra in a specific region of interest (ROI). This approach facilitates data processing and provides better reproducibility, and thus, broadband acquisition for MALDI-MSI should be considered an effective tool to highlight biomarkers of interest. Additionally, we demonstrate the importance of the hierarchical classification of biomarkers by analyzing studies of clusters obtained either from digested or undigested tissues and using bottom-up and in-source decay strategies for in-tissue protein identification. This provides the possibility for the rapid identification of specific markers from different histological samples and their direct localization in tissues. We present an example from a prostate cancer study using formalin-fixed paraffin-embedded tissue.

**Key words :** MALDI Mass Spectrometry Imaging, Principal Component Analysis, Symbolic Discriminant analysis, Hierarchical Clustering, Bottom-Up, In Source Decay, biomarkers, prostate cancer

**Running Title:** PCA-SDA & MSI

\*: co-corresponding author

## INTRODUCTION

Thirteen years after its original publication (1), matrix-assisted laser desorption/ionization mass spectrometric imaging (MALDI-MSI) has become a commonly utilized tool for imaging endogenous biomolecules (2)(3-5) and exogenous compounds (6-9) from a diverse array of tissue types. Since then, many improvements in MALDI MSI performance have been achieved, including improvements related to tissue preparation (10-12), data processing and imaging software (13, 14) and spatial resolution (15, 16). Biomarker discovery and identification has been a field of intense research over the past ten years, and MALDI MSI has been proven to provide an interesting and promising strategy for such investigations (17-22). However, tremendous amounts of data are generated from imaging, and the manual determination of fine variations in marker regulation has become an impossible task with respect to the need to generate a number of spectra for a single image and the number of images required to obtain a statistical overview that accounts for the sample variability between different experiments for a specific tissue and the intrinsic variability between tissues from different organisms. Therefore, an understanding of the fine regulation of biomolecules to highlight markers of a specific physiological stage or cell phenotype requires the use of bioinformatics tools for fast and reliable analyses.

Multivariate statistics is a form of statistics that simultaneously encompasses the observation and analysis of more than one statistical variable. Methods of bivariate statistics, for example, ANOVA and t-tests, are special cases of multivariate statistics that involve two variables. There are several other types of multivariate analyses that are divided into two groups: descriptive methods and explorative methods. Descriptive methods include principal component analysis (PCA), factor analysis of correspondence (FAC), multiple correspondence analysis (MCA) and

classification methods such as hierarchical clustering. Explorative methods include multiple regression, discriminatory analysis and segmentation.

Considering the large amounts of data that are analyzed using the MALDI MSI approach, data reduction and multivariable analysis has become a necessity in facilitating extraction of the information of interest. Among the different multivariable analytical tools that can be used for this purpose, principal component analysis (PCA) has been proven to provide an interesting method for automatic feature extraction (23, 24). Several studies have posited arguments that multivariable analysis though PCA is a good approach for this purpose and can be used with MALDI MSI for pathological disease applications (18, 19, 25-27). However, some limitations are associated with the use of PCA. These limitations encompass different factors, such as data analysis complications. In fact, many spectra can be similar for one principal component, whereas they can differ for another principal component, and thus, it is difficult to define the best data representation. This problem can be resolved by linking hierarchical clustering directly to PCA (18, 19, 25, 26).

In this manuscript, we developed a strategy, designated principal component analysis-symbolic discriminant analysis (PCA-SDA), to facilitate the identification of biomarkers. This concept of data mining was also applied to data sets generated for identification. In fact, protein identification using direct imaging strategies has been shown to be possible either by enzymatic digestion using tissue bottom-up approaches (10, 28, 29) or by in-source decay (ISD) with top-down approaches (30). Such strategies also provide tremendous amounts of data, and markers of interest can be masked within the forest of non-relevant data. Here, we explored the use of PCA-SDA analysis with data generated using both bottom-up and ISD. Finally, the strategy was applied to a biologically relevant case, namely prostate cancer, using samples from hospital libraries, i.e., FFPE tissues.

## EXPERIMENTAL PROCEDURE

### Materials

Alpha-cyano-4-hydroxycinnamic acid (CHCA), sinapinic acid (SA), Trifluoroacetic acid (TFA), ethanol (EtOH), acetonitrile (ACN), chloroform, aniline (ANI), N-succinimidylloxycarbonylmethyltris(2,4,6-trimethoxyphenyl)phosphonium bromide (TMPP), water CHROMASOLV PLUS for HPLC (H<sub>2</sub>O) were purchased from Sigma-Aldrich (Saint Quentin Fallavier, France). Sequencing grade modified Trypsin, porcine enzyme was from Promega (Charbonnières-les-Bains, France)

### Tissue preparation & treatment

#### *Frozen tissues*

Tissue sections (10 µm) were obtained from frozen rat brains using a Leica CM1510S cryostat (Leica Microsystems, Nanterre, France) and applied to ITO-coated conductive glass slides (Bruker Daltonics, Bremen, Germany). It has been demonstrated that the quality of the MS spectra depends generally on the chemical washing step. Lemaire and co-workers introduced the use of organic solvents such as chloroform or acetone (11) to remove the majority of phospholipids, which can interfere with the spectra and result in poor crystallization for the analysis of peptides or proteins. Caprioli et al. described the use of alcohol to fix the tissue and to remove salts and lipids (31). Thus, for the peptide/protein analysis, the tissue sections were submitted to a washing step with cold 95% EtOH for 15 seconds, followed by a washing step with chloroform for 1 minute to remove abundant lipids.

#### *FFPE tissues*

**Prostate Tissue Specimens.** Prostate tissue specimens were obtained from patients undergoing radical prostatectomy for clinically localized prostate carcinoma. All patients gave informed consent according to a protocol

approved by the Internal Review Board for research involving human subjects at Saint Philibert Hospital. None of the patients had any treatment for prostate cancer before surgery. Prostate tissues were fixed by placing the samples for 24h in neutral 3,7% (w/v) formalin solution, rinsed in PBS buffer and then were paraffin-embedded which allows tissues to be cut into thin (i.e. a few microns) sections. For MALDI imaging experiments, 10µm-thick tissue sections were cut from the the formalin-fixed paraffin embedded (FFPE) whole-mount prostate tissue block, placed on the ITO-coated slides, and heated for 60 min at 58 °C. Tissue was also counterstained with Mayer's hematoxylin, dehydrated through graded ethanol solutions, and air-dried for the histological examinations by the anatomopathologists. According to the anatomopathology examinations, the tissues appeared heterogenous and contained cancer, hyperplasiae and normal regions in addition to stromal tissues in each region.

**Tissue dewaxing.** Formalin-fixed embedded prostate carcinoma tissues (FFPE) that had been stored for less than one year were used in this study. Tissue sections (10 µm) were generated using a microtome and applied to conductive glass slides that were coated with ITO (indium teen oxide) on one side. The paraffin was removed by submerging in toluene two times for 5 min, followed by a light rehydration in ethanol (100°C, 96°C; 70°C and 30°C) baths before drying in a dessicator at room temperature.

**Antigen retrieval.** FFPE tissues have a well-conserved tissue morphology and the capacity to prevent the degradation of molecules. However, due to the occurrence of methylene bridge formation during the course of conservation, trypsin digestion (32) is necessary to retrieve molecular information regarding peptides and proteins. Because cross-linking is still performed in the FFPE tissue block, protein accessibility becomes difficult over time, and enzyme cleavage can be affected and can lead to poor enzymatic yields.

Therefore, it is necessary to establish processes that facilitate access to proteins. Similar challenges are encountered for HIC of FFPE samples. Over time, many different protocols have been developed to unmask antigen epitopes (33-43). These so-called antigen retrieval (AR) processes are still not well understood in terms of the associated molecular changes. The most reasonable hypothesis is that AR allows changes in the conformation of cross-linked proteins and thus facilitates access to sites within the proteins. To date, many different AR protocols have been published. However, according to the nature of the tissue used and the protein (antigen) under consideration, some protocols work better than others and there is no real understanding. Consequently, systematic tests of different AR protocols are required to define the best protocol. Here, several protocols were tested. Citric acid antigen retrieval (CAAR) was found to provide good results for prostate tissue sections. The unmasking of proteins by CAAR was performed by immersing the slides in 10 mM of citric acid for 20 min at 90°C and then drying them in a dessicator for 10 min. Prior to the enzymatic digestion, the slides were incubated in 10 mM NH<sub>4</sub>HCO<sub>3</sub> two times to remove the remaining AR solution and to condition the tissue for good enzyme activity.

### ***Enzymatic digestion***

Automatic trypsin digestions were performed using a high position accuracy chemical inkjet printer (CHIP-1000, Shimadzu Biotech, Kyoto, Japan) (10). The printed array of the CHIP-1000 on the tissue section comprised spots spaced 250  $\mu$ m center-to-center. A total of 20 nL of a solution containing 25  $\mu$ g/mL of trypsin in pure H<sub>2</sub>O was applied to each spot. Five droplets of approximately 100 pL were deposited on each spot per cycle. Thus, 40 iterations were necessary to obtain the final volume. The tissue sections were then incubated at 37°C for 2 h in a box with an atmosphere of saturated MeOH/H<sub>2</sub>O (1:1, v/v) to avoid drying the enzyme droplets.

### ***N-terminal derivatization***

Automatic N-succinimidylloxycarbonylmethyl) tris (2,4,6-trimethoxyphenyl) phosphonium bromide (TMPP) derivatization was performed using the automatic microspotting device (10). A CHIP-1000 was used to deposit TMPP on the tryptic spots according to the procedure that has been previously described for trypsin digestion. A total of 10 nL of solution containing 1 mg/mL of TMPP in ACN/H<sub>2</sub>O (3:7, v/v) was deposited on each spot. Five droplets of approximately 100 pL were deposited on each spot per cycle. Thus, twenty iterations were necessary to obtain the final volume for each spot. A solution of ACN/H<sub>2</sub>O containing 1% TEA was then spotted, thus initiating the derivatization reaction. The automated microspotting was performed either according to a global square area of points or by dividing this area into smaller squares, with each including 10-by-10 spots to increase the reaction yield (10).

### **MALDI-MSI**

Depending on whether the bottom-up or the ISD approaches were used, solutions specific for the matrices were deposited using the CHIP-1000 printer (10) or the ImagePrep sprayer.

#### ***Bottom-up approach***

A solid ionic matrix CHCA/aniline was used as the matrix to image peptides and was prepared according to a previously established procedure (32).

***Matrix microspotting.*** Forty nanoliters of ionic matrix solution containing 10 mg/mL of CHCA, and 7.2  $\mu$ L of aniline in 0.1% ACN/TFA (6:4, v/v) was spotted on each spot. Five droplets of approximately 100 pL were deposited on each spot per cycle. Thus, 20 iterations were necessary to reach the total final volume.

***Matrix microsyringing.*** Five milliliters of a solution containing 10 mg/mL of the matrix CHCA/aniline in 0.1% ACN/TFA (3:2, v/v)

was prepared. The matrix was deposited on the tissue section with the ImagePrep automatic sprayer (Bruker Daltonics, Bremen, Germany), which creates an aerosol via vibrational vaporization and allows one to monitor the different stages of matrix deposition, i.e., extraction, drying and matrix layer thickness. Indeed, the method used to spray CHCA/aniline included an initialization phase to deposit the first layer of the matrix on the slide. This process was followed by phases that permitted partial moisturization over several cycles of the tissue section before it had dried completely.

#### ***MS/MS identification of digested tissues.***

MS/MS experiments on digested tissue were performed using the UltraFlex II MALDI-TOF/TOF instrument equipped with LIFT III cell using FlexAnalysis 3.0. For MS/MS experiments, parameters were set as follow: laser repetition rate was 100 Hz, ion source voltages were respectively 8 kV and 7.3 kV on MALDI sample plate and 1<sup>st</sup> electrode; LIFT cell was pulse from ground for electrode 1 and 2 to 19 kV and in the last step, electrode 3 was decrease to 3.2 kV; reflector end voltage was set to 29.5 kV and mid-grid to 13.85 kV. For each MS/MS spectrum 10000 total shots were averaged including 1000 for parent ions and 9000 for fragments. Laser fluence was increase over the experiments for generating lower  $m/z$  fragments.

***Data Analysis.*** Protein identification in databanks were performed using Biotoools 3.0 (Build 1.88) software (Bruker Daltonics, Bremen, Germany) connected to Mascot 2.2.0.3 search engine and interrogating the Swiss-Prot Protein Knowledgebase Release 56.1 of 02-Sep-2008 (397539 sequence entries, comprising 143289088 amino acids abstracted from 172934 references) with Oxidation as variable modification and no fixed modification. Taxonomy was specified to be *human*. Trypsin was selected as enzyme and one missed cleavages were selected. The mass tolerance was set at 0.4 and 0.4 Da respectively for the MS and MS/MS.

#### ***In-source decay approach***

The 1.5-diaminonaphtalen (1.5-DAN)/camphor was used as the ionic matrix to image ISD fragments of proteins. It was prepared by adding 1 eqM of cyclic terpenoid camphor to 5 mg/mL 1,5 DAN in 2% ACN/TFA0 (1:1, v/v).

***Matrix microspotting.*** Twenty nanoliters of a solution containing 5 mg/mL of the matrix 1.5-DAN/camphor in 0.2% ACN/TFA(1:1, v:v) was deposited on an array of spots that were 100  $\mu\text{m}$  in size and spaced 200  $\mu\text{m}$  in a hemi rat brain section, using a chemical inkjet printer CHIP 1000 (Shimadzu, Kyoto, Japan).

***Matrix microsyringing.*** Five milliliters of a solution containing 5 mg/mL of the solid ionic matrix 1.5-DAN/camphor in ACN/TFA 0.2% (1:1 v/v) was prepared. The matrix was deposited on a rat brain section using the ImagePrep automatic sprayer (Bruker Daltonics, Bremen, Germany), which creates an aerosol by vibrational vaporization and allows monitoring of the different stages of matrix deposition, i.e., extraction, drying and matrix layer thickness. The technique employed to spray the 1.5-DAN/camphor consisted of an initialization phase to deposit the first layer of the matrix on the slide. This procedure was followed by phases that allowed partial moisturization of the tissue section before it had dried completely. Each step comprised several cycles.

***Image acquisition.*** Images were acquired using the UltraFlex II MALDI-TOF/TOF instrument. Imaging was performed in the positive reflectron mode, and MALDI-MS spectra were acquired in the  $m/z$  range from 550 to 5000. A total of 300 spectra were acquired for each spot at a laser frequency of 100 Hz. The images were recorded and reconstructed using Fleximaging 2.1 software.

***Statistical analysis.*** The data set obtained with Fleximaging v2.1 software (Bruker Daltonics, Bremen) was loaded in the soft ClinProTools v2.2 (Bruker Daltonics, Bremen) to conduct a principal component analysis and hierarchical clustering of the data. The unsupervised method was selected for the PCA after standardization of the data. Thus, PC1 and PC2 were found to have the largest variance.

After the hierarchical clustering, the same data were exported to the image in Fleximaging 2.1 to view areas with different profiles.

## RESULTS AND DISCUSSION

Determining specific markers of a physiological stage or cell phenotype using MALDI MSI requires the use of data processing and data analysis using multivariable analytical bioinformatics tools. The most widely used multivariable analysis is principal component analysis (PCA). PCA provides access to  $m/z$ , which demonstrates differences between or within samples. PCA is often combined with hierarchical clustering (HC), because HC classifies the mass spectra according to similarities between their profiles and thus provides the ability to highlight regions containing differences in molecular content. Figure 1 describes the general application of the analytical workflow to MALDI-MSI data sets.

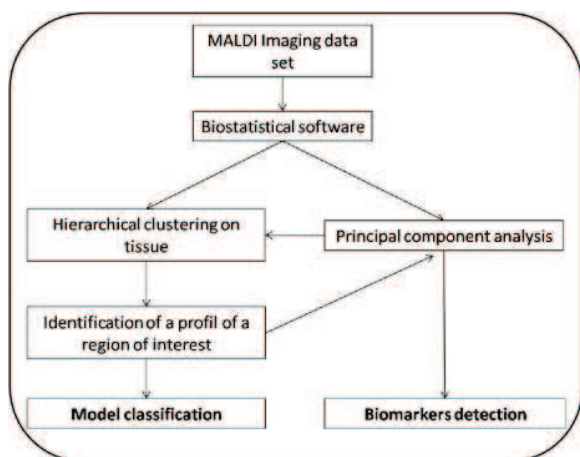


Figure 1 : General workflow to search for biomarkers and model classification by MALDI-MSI.

The first approach of MALDI MSI data sets is based on the supervised or non-supervised PCA mode, which allows the detection of  $m/z$  that demonstrate important variations within the tissue. PCA is conducted using Pareto scaling, which uses the square root of the standard deviation as a scaling factor to

reduce the dominance of large-scale intensity changes in the matrix and other high-abundance ions, which may mask the variation in lower abundance ions during PCA.

The overall outcome of the PCA is greatly affected by the ability of ions related to the matrix coating and other endogenous molecules to mask the underlying, relevant information. For the unsupervised PCA, each spectrum is classified individually. Thus, the selected PCs account for the greatest separation of each individual spectrum. At the top of each, the scores and loading plots are shown. In the second approach, we propose the use of hierarchical clustering (Figure 1) as an initial step to identify major spectral difference between each histological type. This method determines the discriminatory signals and builds functions using these signals that distinguish sample populations based on their classification. Peak lists from MALDI spectra, which were obtained using data processing software (Flex Analysis v3.0 and Flex Control v2.5), were used for the statistical analyses and were clustered by similarity discrimination based on presence vs. absence as criteria. Statistical analyses were conducted using PCA with ClinProTools v2.2 software. In this case, the mass spectra were internally recalibrated on common peaks (also known as a spectral alignment) and normalized to the total ion count. An average spectrum created from all of the individual spectra was used to perform peak selection and to define the integration ranges. The integration ranges were used to determine intensities or areas of single spectra. The signal intensities were used for all calculations. For PCA and HC, the individual peak intensities were standardized across the database.

### PCA-SDA based on hierarchical clustering for biomarker identification.

PCA-SDA based on HC was first tested using the data sets from an image generated from a rat brain tissue section under analytical



conditions to detect peptides and small proteins. Following MALDI imaging acquisition, the data were loaded in the analytical software for HC (Figure 2). HC was prepared using a Euclidean distance method and a ward linkage method based on the PCA results previously obtained. The dimensions were reduced to 85% of the explained variance. A dendrogram was then obtained (Figure 2A), which could be translated into a molecular image by allocating a specific color plot to each of the different clusters of the HC dendrogram, as presented in Figure 2B. The molecular image generated from the HC results could be superimposed on the optical image of the tissue section (Figure 2C) and showed a clear correlation of the different clusters to the histological tissue types. Each group of spectra delimitating a cluster differed from other groups based on spectral profiles and thus demonstrated a unique  $m/z$  (Figure 2B).

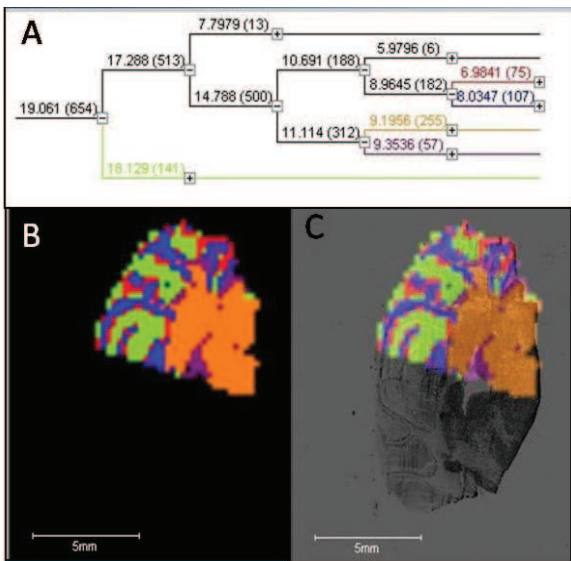


Figure 2 : Hierarchical clustering applied to a frozen rat brain MALDI image. Colors correspond to the different clusters found in the dendrogram. (A) Dendrogram obtained from the image data set. (B) Image representation of clusters of the spectra generated by hierarchical clustering based on the PCA results. (C) Superposition of the clustering on the rat brain tissue section optical image.

This finding suggested the presence of potentially relevant markers for this area of interest (cluster of interest). Based on this observation, a region of interest (ROI) could be created (in the rat brain section for the 4th ventricle in this case) by determining a slightly larger area to obtain spectra with different profiles that served as negative controls (Figure 3A). Spectra for this region were then selectively loaded for PCA analysis (Figure 3B). After PCA analysis, the PC1 versus PC2 loading was performed to compare  $m/z$  that were differentially detected in the selected ROI, because PC1/PC2 provided the best discrimination. Several different ions presented a high level of discrimination along the PC1 axis. To perform a study case, two ions with a high level of discrimination and that presented a very different distribution within the tissue sections were selected: ions at  $m/z$  2030 and  $m/z$  2793. 7

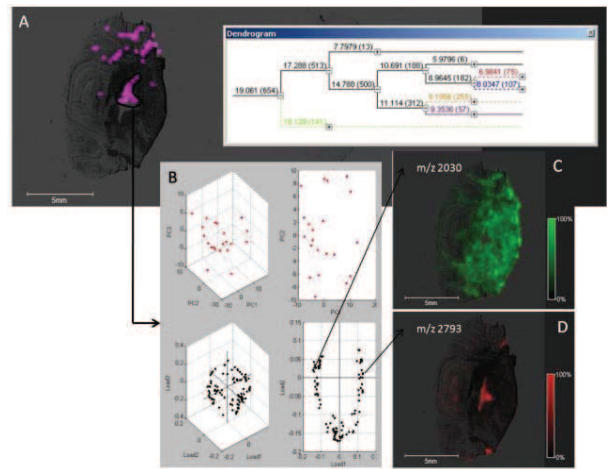
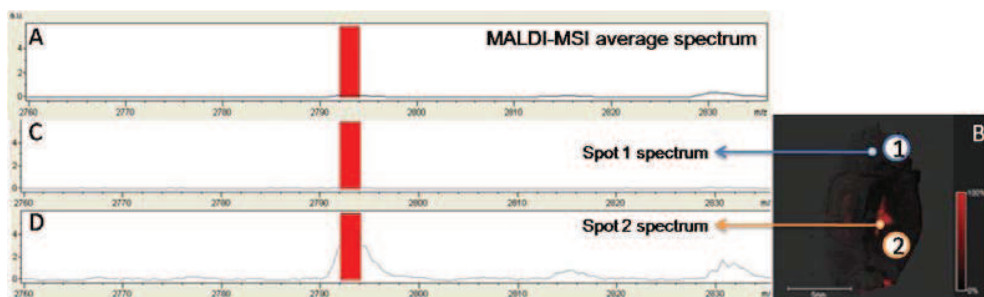


Figure 3 : Selection of a cluster representing the ROI. (A) The spectra for this region were loaded for a new PCA analysis to provide the different  $m/z$  variations. (B) Examples of images at  $m/z$  with different distributions, as demonstrated by the PCA analysis. (C-D), Images of the distributions at  $m/z$  2030 and  $m/z$  2793, respectively.

As demonstrated by the superimposition of the MALDI images of these ions onto the optical image of the tissue section (Figures 3C & D), these ions had a very different distribution. In fact,  $m/z$  2030 was specific to the 4th ventricle (Figure 3C), whereas  $m/z$

2793 (**Figure 3D**) demonstrated a relatively large distribution within the tissue section that corresponded to other morphological structures in the tissue.  $m/z$  2793 was absent in the 4th ventricle but present in the area obtained from the ROI around the structure, which served as a negative control. The MALDI-MSI results are usually obtained from the average mass spectrum, which is reconstructed from all of the spectra recorded during the image acquisition using the software. Therefore, by using such a representation of the data, the  $m/z$  present in a small region (i.e., in a few spectra) are not properly represented in the average mass spectrum, even if they are very intense. This lack of information is exemplified by the signal at  $m/z$  2793, is not observed in the average mass spectrum (**Figure 4A**) and precludes the reconstruction of an image. However, as shown in **Figure 4B**, it was possible to detect this ion specifically in the 4th ventricle using the PCA-SDA analytical approach described herein (**Figure 4C**), which is in contrast to other regions (**Figure 4D**).



**Figure 4 :** (A) MALDI images of the average spectrum with a zoom in the  $m/z$  range of the ions at  $m/z$  2793, which are non-detectable. (B) Image of the  $m/z$  2793 ion distribution. (C-D) Single MALDI spectra recorded in the region where  $m/z$  2793 is found (C) and not found (D).

Thus, without the PCA-SDA approach, it would not have been possible to detect the presence of this marker, which is highly specific to the 4th ventricle. This shows that PCA-SDA facilitates the direct tracking of specific markers in an ROI defined by the HC analysis. As compared to other proteomic

techniques, this will greatly affect the MSI. In fact, the great potential of MSI coupled with PCA-SDA is reflected by the ability to quickly detect specific markers that are present in a pathological region or in regions containing cells with diverse phenotypes while conserving their spatio-temporal localization in the tissue. These signals would be diluted using classical proteomic strategies in which whole organs are used for extraction. The next step is the identification of the detected markers in tissues. Therefore, we were interested in testing the PCA-SDA strategy using data generated from identification strategies.

### Bottom-up approach.

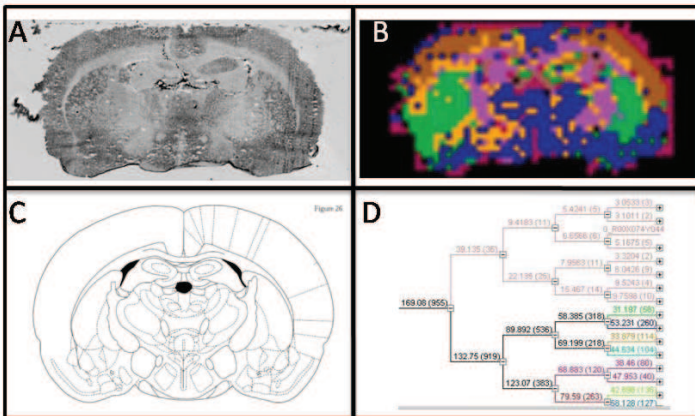
The bottom-up strategy was the first technique explored for on-tissue protein identification (10, 28, 32, 44). This is explained principally based on two factors: it is difficult to use a proper top-down analysis with ion selection on MALDI instruments, and archived clinical samples (formalin-fixed paraffin-embedded FFPE) cannot be used for MSI without a digestion step (32, 45, 46). Therefore,

marker screening from a classical MSI sequence is performed. However, bottom-up and classical approaches have been shown to lead to complementary data with proteins specifically observed in one strategy. Moreover, marker screening for archived FFPE samples can only be performed using digestion data. Therefore, the validity of statistical analyses from such data, which could be hampered by a slight delocalization of markers due to the type of tissue preparation, must be investigated.

### *PCA & HC from enzymatic digestion data.*

To evaluate the digestion data, a rat brain tissue section (interaural, 7.20 mm, bregma, 1.80 mm) was subjected to trypsin digestion prior to matrix deposition (**Figure 5A and 5C**).

After PCA analysis, hierarchical clustering was applied to the data set. The results of the clustering are presented in **Figure 5B**, and the dendrogram is shown in **Figure 5D**. The image reconstructed from the clustering demonstrated a good correlation with the different brain areas determined in the brain section (**Figure 5A, C**). The presented images had a low resolution due to one of the drawbacks of the software used (the number of data points accepted by the software). Therefore, HC imaging analysis data sets with a poor resolution (962 points at 250  $\mu\text{m}$ ) were recorded. This confirmed that a bottom-up strategy could be used for protein identification in tissues and was statistically relevant.



**Figure 5 :** (A) Optical image and (C) image from the brain atlas of the rat brain tissue section used in the bottom-up experiment with on-tissue trypsin digestion. (D) HC dendrogram showing the differentiation of the different histological images and (B) HC image reconstructed from HC dendrogram data.

#### *HC and N-terminal derivatization.*

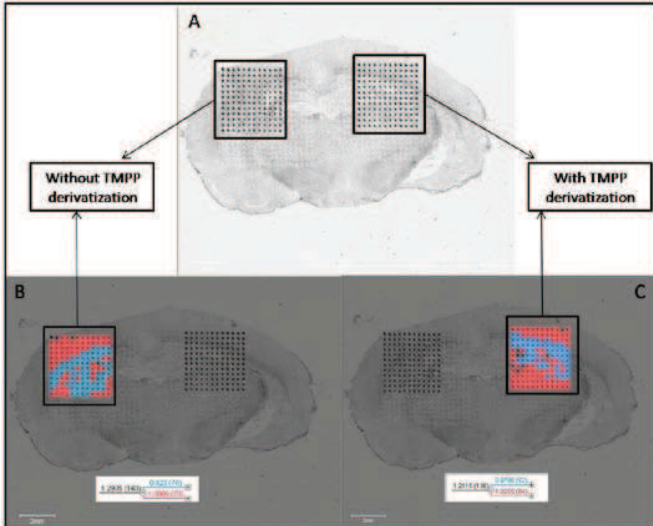
To simplify the protein identification procedure, derivatization strategies were developed in the group to be performed on digestion peptides (29). In fact, N-terminal chemical derivatization of the peptides using specific agents, such as the addition of a negative charge, can lead to a simplification of the MS/MS spectra (for trypsin digestion peptides) by providing a fragmentation orientation toward the  $y_{i+}$  fragment ion series.

Several derivatizations were studied, and TMPP N-terminal derivatization of tryptic peptides was found to be the easiest for use in a micro-spotting manner to allow the generation of MALDI images even after the derivatization step. In the present study, we tested PCA and HC on MALDI imaging data recorded from a rat brain tissue section (interaural, 4.84 mm, bregma, 4.16 mm) following micro-spotting trypsin digestion and then micro-spotting TMPP derivatization, as compared to micro-spotting of trypsin alone (**Figure 6**). In this experiment, two distinct areas that were symmetrical according to the symmetry axis of the rat brain were digested (**Figure 6B**) and then derivatized (**Figure 6C**). As demonstrated, the classification resulting from HC discriminated two regions, each of which was represented by a branch of the dendrogram. The first branch (red) corresponds to the cortex, whereas the second branch (blue) corresponds to the spectra depicting the corpus callosum. From the two dendrograms, it is clear that enzymatic digestion alone and enzymatic digestion followed by N-terminal chemical derivatization provide the same data classification. This result demonstrates that all types of data can be used for statistical analysis if the data is of sufficient quality, i.e., if the particular treatment used does not induce delocalization of the molecules.

#### **In-source decay approach**

The bottom-up approach is the most advanced and most widely used technique for protein identification and characterization because most instruments permit fragmentation of the amino acid sequence in the mass range of 500-3500. Nevertheless, there are several fundamental and practical limitations to the bottom-up strategy. Most importantly, only a fraction of the total peptide population for a given protein is identified, and therefore, only a portion of the information for the protein sequence is provided. It is clear from genomic studies that each open-reading frame can give rise to many protein isoforms that originate from alternatively spliced products, and each

isoform can present different types and locations of posttranslational modifications (PTMs).



**Figure 6 :** (A) Optical image of a rat brain tissue section presenting two patches that are symmetrically distributed and that correspond to a specific tissue preparation: enzymatic digestion (left) and enzymatic digestion followed by TMPP derivatization (right). HC was applied to the MALDI-MSI data for the two conditions, and the corresponding dendrogram and reconstructed images are presented (B) for trypsin digestion and (D) for trypsin digestion followed by TMPP derivatization.

PTMs such as phosphorylation and glycosylation are known to be important in the regulation of protein function and cell metabolism. A consequence of the limited sequence coverage in bottom-up proteomics is the global loss of information, including information concerning PTMs. In contrast, the top-down strategy provides potential access to a large part, and sometimes all, of the protein sequence, including its PTMs (47-53). In addition, the time-consuming protein digestion step required for the bottom-up method is eliminated.

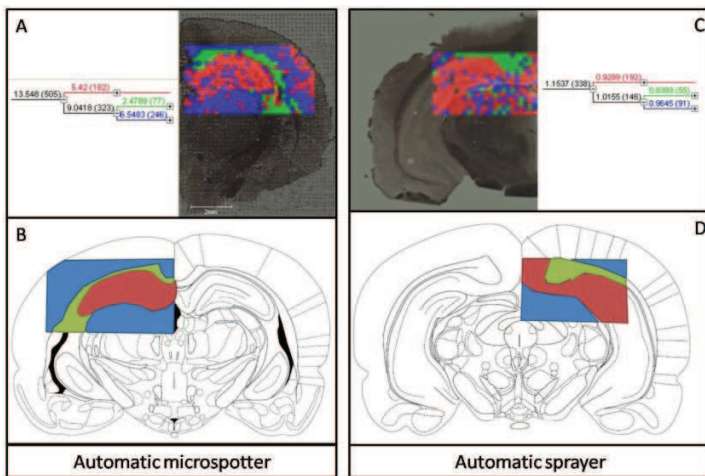
Classical top-down strategies are not possible using MALDI instruments because the instrumentation used for such purposes is not compatible with the very low charge states observed in MALDI. However, the top-down method without ion selection can be used with MALDI to rapidly generate fragmentation

using in-source decay (ISD) (47-54). ISD has been demonstrated previously (55) for tissue sections. However, because ISD is not used with ion selection, the resulting mass spectra are extremely difficult to interpret. Here, we tested the ability of HC and PCA to facilitate interpretation of the ISD data. A previous publication has demonstrated the advantage of using a 1,5 diaminonaphthalene (1.5 DAN) matrix for protein sequence and glycosylation characterization. The 1,5 diaminonaphthalene (1.5 DAN) matrix is a good proton donor and tends to fragment proteins from their C-terminal region to produce  $ci^+$  and  $zi^+$  fragment ions (56). In the present study, we developed a specific solid ionic matrix (1,5-DAN/camphor) to achieve better reproducibility, sensitivity and deposition facility by micro-spraying or micro-spotting of the matrix for ISD from tissues. We found that adding the camphor cyclic terpenoid to the classical 1,5-DAN matrix resulted in average spectra with more intense peaks, which were found to be ISD ions. This process facilitated the detection of ions that corresponded to the tags used for protein identification in the images (**Supplementary data 1**). This ionic matrix was also used for instrumentation preservation, and it demonstrated greater solubility in solvents that are classically used with automated micro-spraying and micro-spotting devices; 1,5 DAN alone was shown to be a matrix with a high level of deterioration for the deposition instruments.

#### *PCH & HC for ISD-MALDI-MSI.*

Clustering images (**Figure 7**) were generated from the obtained MALDI images using 1.5-DAN/camphor, and subsequently, prompt fragments were obtained from whole proteins in the section at each position. Two different deposition methods for the matrix, the micro-spotter and the micro-sprayer, were used. The hierarchical clustering results are shown in **Figures 7A** and **7C**. The histological regions corresponding to the clusters are demonstrated symmetrically in **Figures 7B** and **7D**. We note that the classification of the spectra correlated

with the histological areas of the tissue section. These clusters would not be detected if there were no differences between the spectra determined for each area. Therefore, it is possible to classify the specific ions by PCA and to reassemble them with greater ease to identify proteins of interest while suppressing common ions of the different class (common proteins). Unfortunately, we identified only myelin in the *cingulum* region of the rat brain section (supplementary data 1) using this technology.



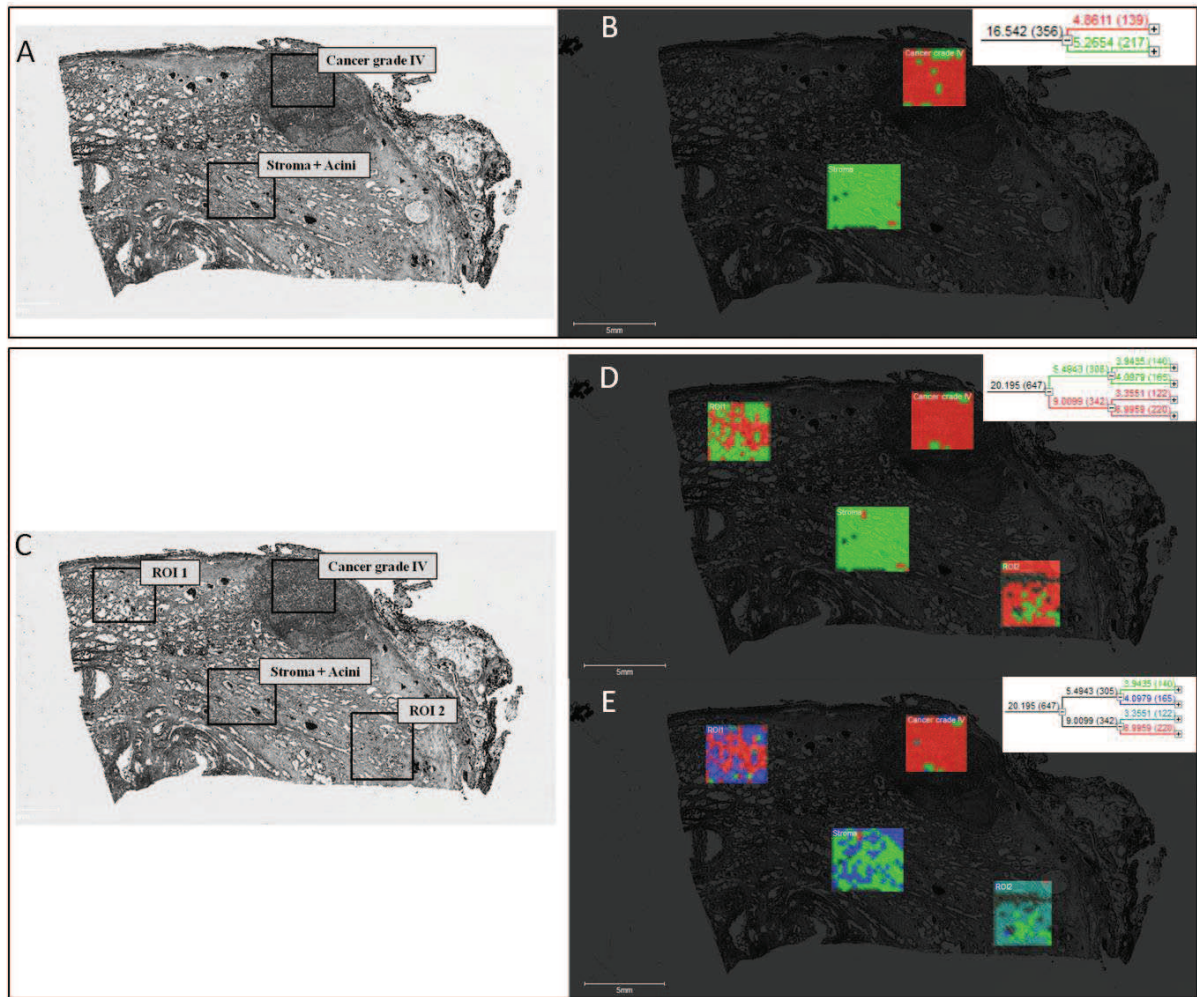
**Figure 7 :** HC analysis performed for MALDI-MSI data recorded in ISD from a rat brain tissue section using a 1.5 DAN/camphor ionic matrix deposited with an automatic micro-spotter (A) and with an automatic micro-sprayer (C). (B & D) Schematic representation of the spectral classification for each cluster. Green: cingulum, red: dentate gyrus, blue: combination of cerebral cortex and hippocampus.

Moreover, it was often difficult to analyze properly one of the amino acid sequences within a spectrum ISD due to the large number of peaks that were detected for the different proteins. Thus, using HC, we were able to identify differences between histological regions and PCA, which permitted the efficient detection of differences in the complexity of the ISD spectra. Importantly, ISD can be only performed using frozen tissue (never formalin-fixed paraffin-embedded tissues), and only major proteins with a high abundance can be detected. The developments of on-tissue depletion methods are clearly necessary to further the uses of such techniques.

### PCA & HC in the bottom-up approach: application to FFPE prostate cancer tissues

The opportunity to investigate archived FFPE (formalin-fixed paraffin-embedded) tissues will pave the way toward a gold mine of tissue samples that are stored in pathological institutes. Consequently, many efforts by others and by our group (32, 45) have been implemented to find a strategy to analyze tissues using MALDI-MSI (27, 45, 57, 58). The major difficulty associated with such tissue is the occurrence of protein cross-linking due to tissue fixation. Indeed, formalin creates methylene bonds between proteins and prevents direct tissue analysis by MALDI-MSI. To circumvent this problem, enzymatic digestion using, for example, trypsin is used to obtain and to identify protein fragments, and therefore, a bottom-up approach is used.

In this context, we performed a bottom-up analysis of FFPE tissue from grade IV prostate cancer samples using trypsin as the enzyme and HCCA/ANI as the matrix, with a micro-spraying device. In a grade IV carcinoma section, several histological cell types were observed that resembled tumor or stroma with normal acini (close to normal) regions, as classified by the pathologist. Thus, the spectra of these two specific regions were collected before they were exported for the HC and PCA analyses (**Figure 8A**). The dendrogram from the hierarchical clustering divides the spectra into two branches. The locations of these branches are shown in **Figure 8B**. Classification of the spectra correlated with their histological origin (cancer grade IV in red vs. stroma with acini in green) (**Figure 8B**). This ranking suggests that it is possible to use the MS profiles (in this context, peptides and proteins) as a "signature" for cell phenotype classification or MSI histology, and it may allow the development of a new diagnostic approach based on the molecular signature and not on a specific biomarker alone. To illustrate this idea, two ROIs (ROI1 and ROI2) that were determined in different regions of the tissue were added to the analysis (**Figure 8C**) together with the previous regions. HC analysis



**Figure 8 :** HC analysis of FFPE prostate cancer tissue after on-tissue trypsin digestion. (A) Delimitation of stroma with acini and cancer grade IV regions on the optical image of the tissue. (B) HC results (dendrogram and image) for the cancer and stroma with acini spectra. (C) Delimitation of the new ROIs on the optical image of the tissue. (D-E) HC results (dendrogram and image) for four regions (stoma + acini, cancer, ROI1 and ROI2).

was performed for these 4 regions. The resulting HC dendrogram revealed an initial separation into two branches and then a sub-division of each of these branches into two others. The first two branches (first division) were represented as an image on the tissue (**Figure 8D**). In **Figure 8D**, grade IV tumor and stroma with acini clusters can still be observed as distinct areas, with each area corresponding to a branch. However, it should be noted that the ROI1 spectra appear close to the stroma and tumor profiles, whereas ROI2 has a profile that is closer to the tumor alone. However, based on the sub-branch dendrogram analysis (**Figure 8E**), we note that the turquoise blue cluster of ROI2 differs from the red cluster of the tumor

and correspond to a sub-division. In contrast, ROI1 consistently shows a signature that is closer to the tumor and the stroma with acini. Therefore, it is possible to use the same approach for different biopsies and to correlate the signatures of the clusters obtained for the examined histological areas using a bank of peptide/protein signatures generated from tissues for which the pathologic stages are known.

The spectral classification for the grade IV tumor and stroma with acini area also supports the conclusion that specific m/z are related to the histologic classification. The PCA-SDA strategy can be performed to track markers for each cell phenotype (**Figure 9A**).

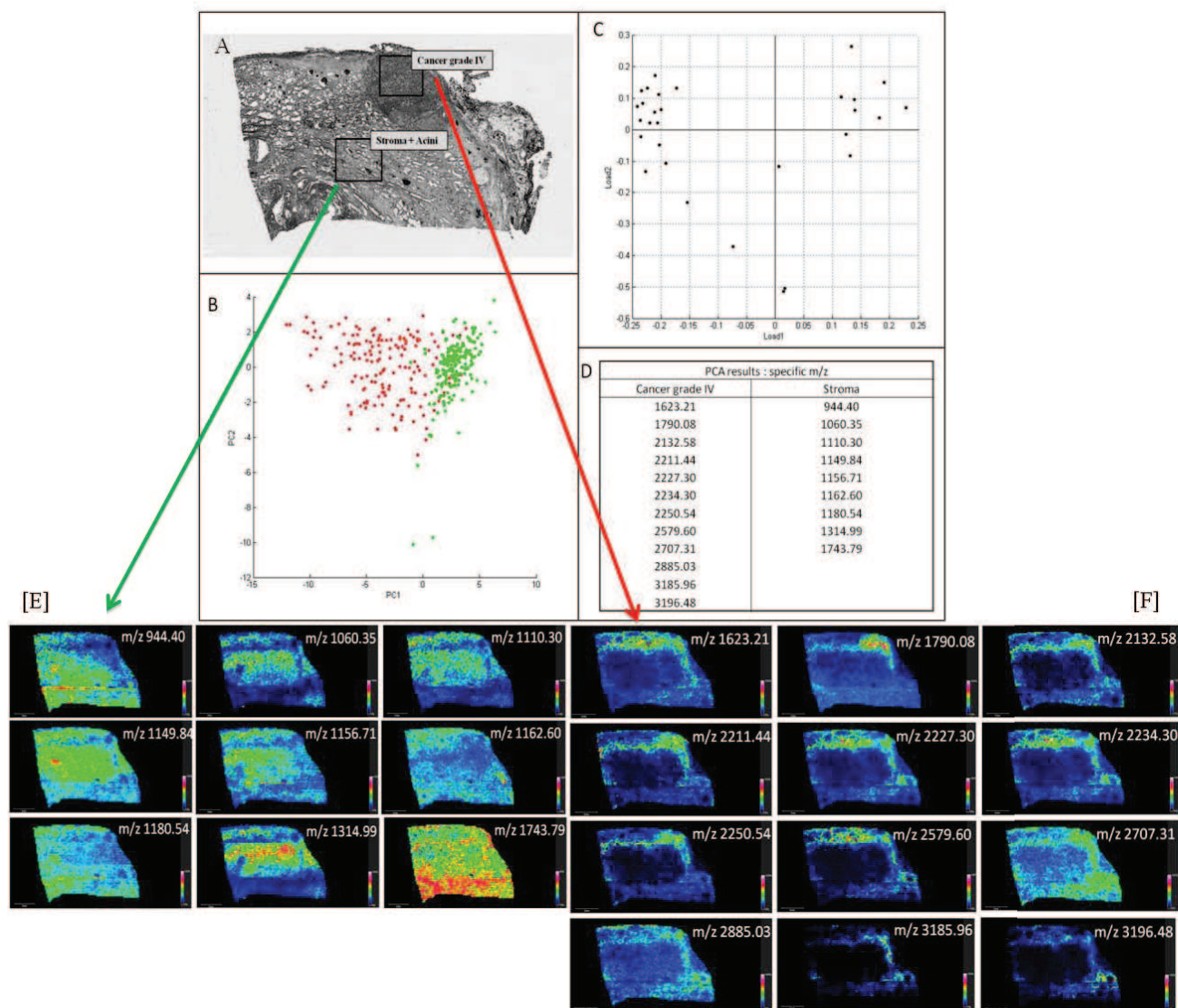


Figure 9 : Supervised PCA analysis of prostate cancer grade IV FFPE tissue after on-tissue trypsin digestion. (A) Optical images indicating the stroma with acini and cancer grade IV regions. (B) Score plot (green: stroma with acini spectra, red: cancer grade IV spectra). (C) Loading plot. (D) List of m/z specific for each region. (E-F) MALDI-MSI images of the ions determined by PCA in the stroma with acini (E) and in the tumor region (F).

The score plot (**Figure 9B**) and loading plot (**Figure 9C**) obtained after supervised PCA analysis provides a list of  $m/z$  that are specific to each area (**Figure 9D**), determining clusters. Images of the ions detected in the stroma with acini are presented in **Figure 9E**, and those detected in the tumor region are shown in **Figure 9F**. To further understand the progression of cancer, it is of primary importance to identify the ions detected for each cell phenotype, i.e., tumor *vs.* stroma + acini. As described previously, enzymatic digestion does not lead to delocalization and allows a hierarchical classification. The PCA analysis related to the tissue type to be performed to highlight the presence of specific markers in the different cell types. Therefore, the bottom-up approach could be useful for analyses of FFPE tissues from pathologist libraries. Some ions obtained using PCA-SDA were directly identified in the tissue section after the CAAR antigen retrieval procedure by MS<sup>2</sup> using the MALDI TOF/TOF instrument. In this context, MS<sup>2</sup> was performed at  $m/z$  944.4, 1180.54, 1623.21, 1743.79, 1790.08, 2227.30 and 3196.48 (**Figures 10 & 11**). These ions correspond to actin fragments (**Figure 10B**) or to histone H4, histone 2A type 1-A, and histone 2B type 1-B (**Figure 11**). Interestingly, all of the actin fragments were specifically localized in the tumor region and not in the stroma with normal acini, which is consistent with the PCA-SDA results (**Figure 10A**), whereas the histone fragments were localized in the stroma with acini (**Figure 8E**). The ions related to actin present in the tumor region presented a sequence coverage of 21.5% and were localized to the C-terminal portion of the protein, which suggested that the N-terminal portion of the protein was less accessible due to the presence of methylene bonds or to protein truncation (**Figure 10C**). All the identified proteins, related tryptic peptide sequences, identification scores and Mw are summarized **Table 1**. Ions attribution for the different MS/MS spectra corresponding to the different tryptic fragments studied are summarized in **Supplementary data 2**. The

presence of actin in the tumor region is consistent with the cytoskeletal remodeling that occurs during prostate cancer development together with muscle contraction. The detection of histone types in the stroma with acini seems to indicate mitotic and apoptotic indices for these cells, which reflects the possibility that their phenotype is more similar to benign than to healthy cells. These data are consistent with the results obtained for other cancers, such as breast (27) or ovarian (18) cancers. Interestingly, we showed that based on the 150  $m/z$  detection using PCA-SDA in the tumor region, only 4 were related to actin. This result supports the finding that not only major proteins can be detected using the bottom-up strategy. These findings highlight the necessity to develop new approaches to gain access to the minor  $m/z$  detected using PCA-SDA. In fact, AR followed by enzymatic digestion is not sufficient for the identification of these markers or for other minor ones. Nevertheless, several MS/MS studies have been performed, and identification problems have been encountered due to the methylene bridges that occur between proteins. In fact, in some cases, MS/MS spectra were uninterpretable and did not provide any identification results. This phenomenon was presumably attributed to the presence of cross-linked peptides that were linked by a methylene bridge. Thus, fundamental investigations are necessary to fully understand the chemistry of fixation and the formation of methylene bridges to develop the appropriate software for MS/MS spectral analysis from FFPE tissues.



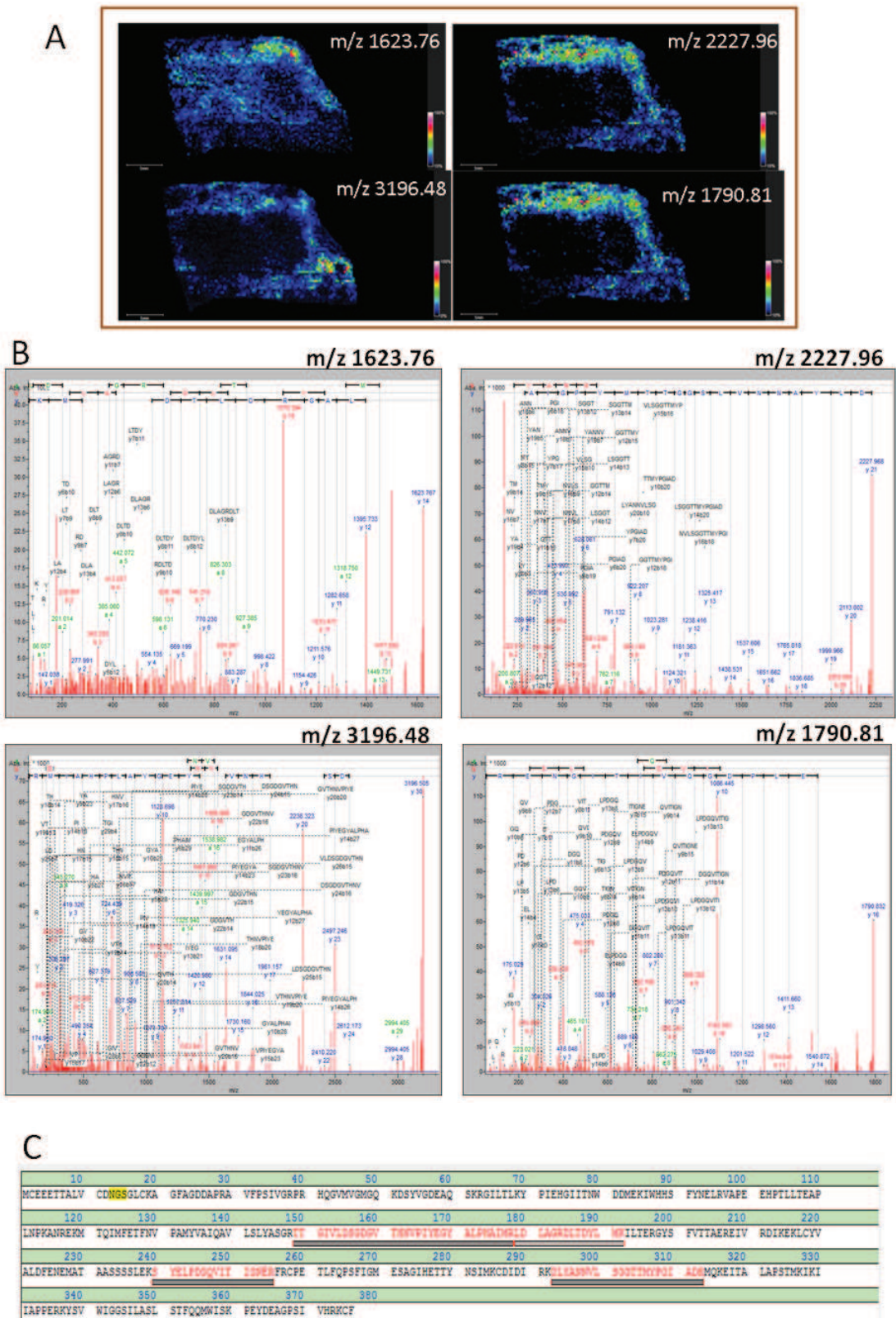
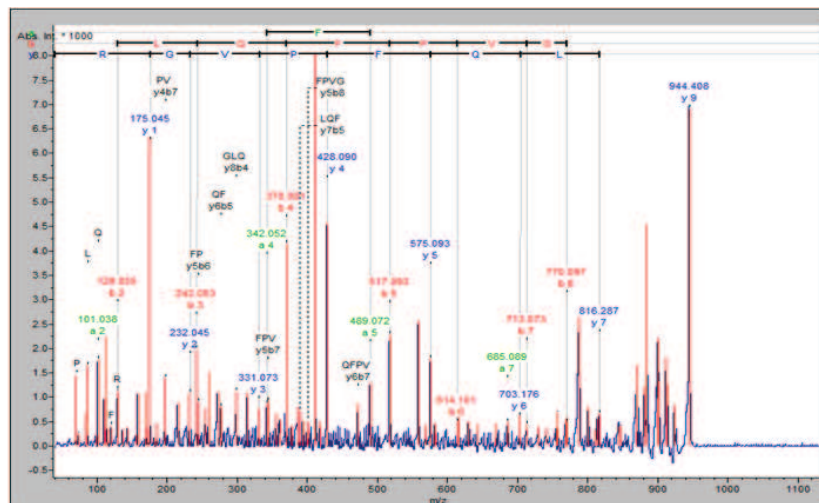
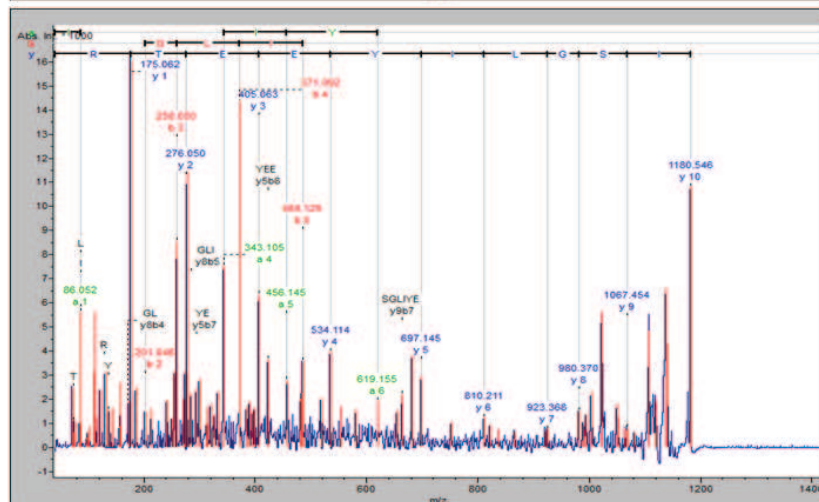


Figure 10 : (A) MALDI-MSI molecular images of several fragments of actin tryptic fragments detected by PCA-SDA from the tissue section after AR and trypsin digestion, (B) MS/MS spectra of these actin tryptic fragments and (C) sequence coverage if these four fragments on the actin protein sequence.

**[A]**  
**Histone H2A type 1-A**  
 m/z 944.40  
 AGLQFPVGR



**[B]**  
**Histone H4**  
 m/z 1180.54  
 ISGLIYEETR



**[C]**  
**Histone H2B type 1-B**  
 m/z 1743.79  
 AMGIMNSFVNDIFER

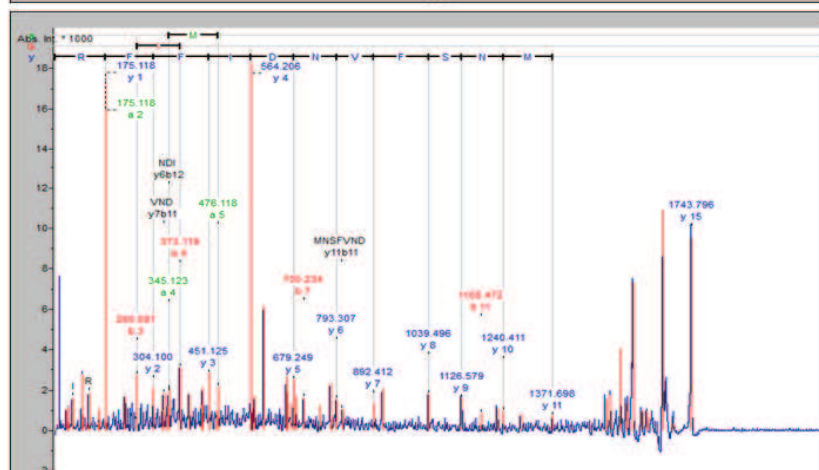


Figure 11: MS/MS spectra recorded from prostate cancer FFPE tissue sections after AR and trypsin digestion of three different ions (three histones) found by PCA-SDA to be specific of the stroma with acini region.

Measured m/z	Mw (Experimental)	Mw (calculated)	Delta	z	Miss cleavage	Mascot score	Expect	Peptide sequence	Protein Name
<b>944.40</b>	943.39	943.52	-0.13	+1	0	49	0.00075	AGLQFPVGR	H2A1A_HUMAN
<b>1180.54</b>	1179.53	1179.61	-0.08	+1	0	54	0.0002	ISGLIYEETR	H4_HUMAN
<b>1623.76</b>	1622.76	1622.83	-0.08	+1	1	43	0.0019	LDLAGRDLTDYLMK	ACTA_HUMAN
<b>1743.79</b>	1742.78	1742.81	-0.03	+1	0	88	5.8e-08	AMGIMNSFVNDIFER	H2B1B_HUMAN
<b>1790.81</b>	1789.80	1789.88	-0.08	+1	0	91	2.5e-08	SYELPDGQVITIGNER	ACTA_HUMAN
<b>2227.97</b>	2226.96	2227.06	-0.10	+1	0	127	5.8e-12	DLYANNVLSGGTTMY PGIADR	ACTA_HUMAN
<b>3196.48</b>	3195.48	3195.60	-0.12	+1	0	132	1.1e-12	TTGIVLDSGDGVTHN VPIYEGYALPHAIMR	ACTA_HUMAN

Mw(exp): experimental Mw

Mw(cal): calculated Mw

**Table 1** : Identified proteins by on-tissue MS/MS after enzymatic digestion of a FFPE section of prostate cancer.

## CONCLUSION

The PCA-SDA based on hierarchical clustering represents an excellent, fast and reproducible technique for the analysis of MSI data to search for markers. We showed that this approach is useful after different tissue treatments, such as enzyme digestion and N-terminal derivatization, which are used in bottom-up strategies for the direct identification of ions of interest in the tissue. Thus, it is possible to combine the MALDI-MSI and bottom-up approach for protein identification. This can be conducted via enzymatic digestion, derivatization and PCA-SDA based on hierarchical clustering for protein fragment detection and identification. All of these steps can be performed for the same tissue and in only one experiment. Data obtained using the top-down method with in-source decay (ISD) can also be used for PCA-SDA following on-tissue depletion to highlight proteins other than the major ones. In the near future, the most promising development will be the emergence of databases that are based on approaches using these statistical models. The creation of mathematical models that represent histological groups could enable a significant advance for molecularly based diagnostic and prognostic studies.

## ACKNOWLEDGMENTS

This research was supported by grants from Centre National de la Recherche Scientifique (CNRS), Ministère de L'Éducation Nationale, de L'Enseignement Supérieur et de la Recherche, Agence Nationale de la Recherche (ANR PCV to IF), Institut du Cancer (INCA to IF), Région Nord-Pas de Calais (PhD financing to DB & RL), the Canadian Institutes of Health Research (CIHR to RD) and the Ministère du Développement Économique, de l'Innovation et de l'Exportation (MDEIE to RD) du Québec.

## REFERENCES

1. Caprioli, R. M., Farmer, T. B., and Gile, J. (1997) Molecular imaging of biological samples: localization of peptides and proteins using MALDI-TOF MS. *Anal Chem* 69, 4751-4760.
2. McLean, J. A., Ridenour, W. B., and Caprioli, R. M. (2007) Profiling and imaging of tissues by imaging ion mobility-mass spectrometry. *J Mass Spectrom* 42, 1099-1105.
3. Sugiura, Y., Shimma, S., Konishi, Y., Yamada, M. K., and Setou, M. (2008) Imaging mass spectrometry technology and application on ganglioside study; visualization of age-dependent accumulation of C20-ganglioside molecular species in the mouse hippocampus. *PLoS One* 3, e3232.
4. Landgraf, R. R., Prieto Conaway, M. C., Garrett, T. J., Stacpoole, P. W., and Yost, R. A. (2009) Imaging of lipids in spinal cord using intermediate pressure matrix-assisted laser desorption-linear ion trap/Orbitrap MS. *Anal Chem* 81, 8488-8495.
5. Taban, I. M., Altelaar, A. F., van der Burgt, Y. E., McDonnell, L. A., Heeren, R. M., Fuchser, J., and Baykut, G. (2007) Imaging of peptides in the rat brain using MALDI-FTICR mass spectrometry. *J Am Soc Mass Spectrom* 18, 145-151.
6. Sugiura, Y., and Setou, M. Imaging mass spectrometry for visualization of drug and endogenous metabolite distribution: toward in situ pharmacometabolomes. *J Neuroimmune Pharmacol* 5, 31-43.
7. Khatib-Shahidi, S., Andersson, M., Herman, J. L., Gillespie, T. A., and Caprioli, R. M. (2006) Direct molecular analysis of whole-body animal tissue sections by imaging MALDI mass spectrometry. *Anal Chem* 78, 6448-6456.

8. Prideaux, B., Staab, D., and Stoeckli, M. Applications of MALDI-MSI to pharmaceutical research. *Methods Mol Biol* 656, 405-413.
9. Stoeckli, M. S., D.; Schweitzer, A. (2007) Compound and metabolite distribution measured by MALDI mass spectrometric imaging in whole-body tissue sections. *Int J Mass Spectrom* 260, 195-202.
10. Franck, J., Arafah, K., Barnes, A., Wisztorski, M., Salzet, M., and Fournier, I. (2009) Improving tissue preparation for matrix-assisted laser desorption ionization mass spectrometry imaging. Part 1: using microspotting. *Anal Chem* 81, 8193-8202.
11. Lemaire, R., Wisztorski, M., Desmons, A., Tabet, J. C., Day, R., Salzet, M., and Fournier, I. (2006) MALDI-MS direct tissue analysis of proteins: Improving signal sensitivity using organic treatments. *Anal Chem* 78, 7145-7153.
12. Seeley, E. H., and Caprioli, R. M. (2008) Molecular imaging of proteins in tissues by mass spectrometry. *Proc Natl Acad Sci U S A* 105, 18126-18131.
13. Jardin-Mathe, O., Bonnel, D., Franck, J., Wisztorski, M., Macagno, E., Fournier, I., and Salzet, M. (2008) MITICS (MALDI Imaging Team Imaging Computing System): A new open source mass spectrometry imaging software. *J Proteomics*.
14. Stoeckli, M., Farmer, T. B., and Caprioli, R. M. (1999) Automated mass spectrometry imaging with a matrix-assisted laser desorption ionization time-of-flight instrument. *J Am Soc Mass Spectrom* 10, 67-71.
15. Chaurand, P., Schriver, K. E., and Caprioli, R. M. (2007) Instrument design and characterization for high resolution MALDI-MS imaging of tissue sections. *J Mass Spectrom* 42, 476-489.
16. Koestler, M., Kirsch, D., Hester, A., Leisner, A., Guenther, S., and Spengler, B. (2008) A high-resolution scanning microprobe matrix-assisted laser desorption/ionization ion source for imaging analysis on an ion trap/Fourier transform ion cyclotron resonance mass spectrometer. *Rapid Commun Mass Spectrom* 22, 3275-3285.
17. Caprioli, R. M. (2008) Perspectives on imaging mass spectrometry in biology and medicine. *Proteomics* 8, 3679-3680.
18. El Ayed, M., Bonnel, D., Longuespee, R., Castelier, C., Franck, J., Vergara, D., Desmons, A., Tasiemski, A., Kenani, A., Vinatier, D., Day, R., Fournier, I., and Salzet, M. (2010) MALDI imaging mass spectrometry in ovarian cancer for tracking, identifying, and validating biomarkers. *Med Sci Monit* 16, BR233-245.
19. Franck, J., Arafah, K., Elayed, M., Bonnel, D., Vergara, D., Jacquet, A., Vinatier, D., Wisztorski, M., Day, R., Fournier, I., and Salzet, M. (2009) MALDI imaging mass spectrometry: state of the art technology in clinical proteomics. *Mol Cell Proteomics* 8, 2023-2033.
20. McDonnell, L. A., Corthals, G. L., Willems, S. M., van Remoortere, A., van Zeijl, R. J., and Deelder, A. M. Peptide and protein imaging mass spectrometry in cancer research. *J Proteomics*.
21. Wisztorski, M., Croix, D., Macagno, E., Fournier, I., and Salzet, M. (2008) Molecular MALDI imaging: An emerging technology for neuroscience studies. *Dev Neurobiol* 68, 845-858.
22. Wisztorski, M., Lemaire, R., Stauber, J., Menguelet, S. A., Croix, D., Mathe, O. J., Day, R., Salzet, M., and Fournier, I. (2007) New developments in MALDI imaging for pathology proteomic studies. *Curr Pharm Des* 13, 3317-3324.
23. Norris, J. L., Cornett, D. S., Mobley, J. A., Andersson, M., Seeley, E. H., Chaurand, P., and Caprioli, R. M.

- (2007) Processing MALDI Mass Spectra to Improve Mass Spectral Direct Tissue Analysis. *Int J Mass Spectrom* 260, 212-221.
24. Van de Plas, R., Ojeda, F., Dewil, M., Van Den Bosch, L., De Moor, B., and Waelkens, E. (2007) Prospective exploration of biochemical tissue composition via imaging mass spectrometry guided by principal component analysis. *Pac Symp Biocomput*, 458-469.
  25. Deininger, S. O., Ebert, M. P., Futterer, A., Gerhard, M., and Rocken, C. (2008) MALDI Imaging Combined with Hierarchical Clustering as a New Tool for the Interpretation of Complex Human Cancers. *J Proteome Res*.
  26. Walch, A., Rauser, S., Deininger, S. O., and Hofler, H. (2008) MALDI imaging mass spectrometry for direct tissue analysis: a new frontier for molecular histology. *Histochem Cell Biol* 130, 421-434.
  27. Djidja, M. C., Francese, S., Loadman, P. M., Sutton, C. W., Scriven, P., Claude, E., Snel, M. F., Franck, J., Salzet, M., and Clench, M. R. (2009) Detergent addition to tryptic digests and ion mobility separation prior to MS/MS improves peptide yield and protein identification for in situ proteomic investigation of frozen and formalin-fixed paraffin-embedded adenocarcinoma tissue sections. *Proteomics* 9, 2750-2763.
  28. Franck, J., Ayed, M. E., Wisztorski, M., Salzet, M., and Fournier, I. On tissue protein identification improvement by N-terminal Peptide derivatization. *Methods Mol Biol* 656, 323-338.
  29. Franck, J., El Ayed, M., Wisztorski, M., Salzet, M., and Fournier, I. (2009) On-tissue N-terminal peptide derivatizations for enhancing protein identification in MALDI mass spectrometric imaging strategies. *Anal Chem* 81, 8305-8317.
  30. Debois, D., Bralet, M. P., Le Naour, F., Brunelle, A., and Laprevote, O. (2009) In situ lipidomic analysis of nonalcoholic fatty liver by cluster TOF-SIMS imaging. *Anal Chem* 81, 2823-2831.
  31. Seeley, E. H., Oppenheimer, S. R., Mi, D., Chaurand, P., and Caprioli, R. M. (2008) Enhancement of protein sensitivity for MALDI imaging mass spectrometry after chemical treatment of tissue sections. *J Am Soc Mass Spectrom* 19, 1069-1077.
  32. Lemaire, R., Desmons, A., Tabet, J. C., Day, R., Salzet, M., and Fournier, I. (2007) Direct analysis and MALDI imaging of formalin-fixed, paraffin-embedded tissue sections. *J Proteome Res* 6, 1295-1305.
  33. Gustafsson, J. O., Oehler, M. K., McColl, S. R., and Hoffmann, P. Citric Acid Antigen Retrieval (CAAR) for tryptic peptide imaging directly on archived formalin-fixed paraffin-embedded tissue. *J Proteome Res*.
  34. Tian, Y., and Zhang, H. Isolation of proteins by heat-induced extraction from formalin-fixed, paraffin-embedded tissue and preparation of tryptic peptides for mass spectrometric analysis. *Curr Protoc Mol Biol* Chapter 10, Unit 10 26 11-17.
  35. Xu, H., Yang, L., Wang, W., Shi, S. R., Liu, C., Liu, Y., Fang, X., Taylor, C. R., Lee, C. S., and Balgley, B. M. (2008) Antigen retrieval for proteomic characterization of formalin-fixed and paraffin-embedded tissues. *J Proteome Res* 7, 1098-1108.
  36. Shi, S. R., Liu, C., Perez, J., and Taylor, C. R. (2005) Protein-embedding technique: a potential approach to standardization of immunohistochemistry for formalin-fixed, paraffin-embedded tissue sections. *J Histochem Cytochem* 53, 1167-1170.

37. D'Amico, F., Skarmoutsou, E., and Stivala, F. (2009) State of the art in antigen retrieval for immunohistochemistry. *J Immunol Methods* 341, 1-18.
38. Yamashita, S. (2007) Heat-induced antigen retrieval: mechanisms and application to histochemistry. *Prog Histochem Cytochem* 41, 141-200.
39. Shi, S. R., Cote, R. J., and Taylor, C. R. (2001) Antigen retrieval techniques: current perspectives. *J Histochem Cytochem* 49, 931-937.
40. Shi, S. R., Cote, R. J., and Taylor, C. R. (2001) Antigen retrieval immunohistochemistry and molecular morphology in the year 2001. *Appl Immunohistochem Mol Morphol* 9, 107-116.
41. Shi, S. R., Datar, R., Liu, C., Wu, L., Zhang, Z., Cote, R. J., and Taylor, C. R. (2004) DNA extraction from archival formalin-fixed, paraffin-embedded tissues: heat-induced retrieval in alkaline solution. *Histochem Cell Biol* 122, 211-218.
42. Shi, S. R., Liu, C., and Taylor, C. R. (2007) Standardization of immunohistochemistry for formalin-fixed, paraffin-embedded tissue sections based on the antigen-retrieval technique: from experiments to hypothesis. *J Histochem Cytochem* 55, 105-109.
43. Shi, S. R., Liu, C., Young, L., and Taylor, C. (2007) Development of an optimal antigen retrieval protocol for immunohistochemistry of retinoblastoma protein (pRB) in formalin fixed, paraffin sections based on comparison of different methods. *Biotech Histochem* 82, 301-309.
44. Groseclose, M. R., Andersson, M., Hardesty, W. M., and Caprioli, R. M. (2007) Identification of proteins directly from tissue: in situ tryptic digestions coupled with imaging mass spectrometry. *J Mass Spectrom* 42, 254-262.
45. Stauber, J., Lemaire, R., Franck, J., Bonnel, D., Croix, D., Day, R., Wisztorski, M., Fournier, I., and Salzet, M. (2008) MALDI imaging of formalin-fixed paraffin-embedded tissues: application to model animals of Parkinson disease for biomarker hunting. *J Proteome Res* 7, 969-978.
46. Wisztorski, M., Franck, J., Salzet, M., and Fournier, I. MALDI Direct Analysis and Imaging of Frozen Versus FFPE Tissues: What Strategy for Which Sample? *Methods Mol Biol* 656, 303-322.
47. Reiber, D. C., Brown, R. S., Weinberger, S., Kenny, J., and Bailey, J. (1998) Unknown peptide sequencing using matrix-assisted laser desorption/ionization and in-source decay. *Anal Chem* 70, 1214-1222.
48. Reiber, D. C., Grover, T. A., and Brown, R. S. (1998) Identifying proteins using matrix-assisted laser desorption/ionization in-source fragmentation data combined with database searching. *Anal Chem* 70, 673-683.
49. Brown, R. S., and Lennon, J. J. (1995) Sequence-specific fragmentation of matrix-assisted laser-desorbed protein/peptide ions. *Anal Chem* 67, 3990-3999.
50. Hardouin, J. (2007) Protein sequence information by matrix-assisted laser desorption/ionization in-source decay mass spectrometry. *Mass Spectrom Rev* 26, 672-682.
51. Takayama, M., and Tsugita, A. (2000) Sequence information of peptides and proteins with in-source decay in matrix assisted laser desorption/ionization-time of flight-mass spectrometry. *Electrophoresis* 21, 1670-1677.
52. Takayama, M. (2001) In-source decay characteristics of peptides in matrix-assisted laser desorption/ionization time-of-flight mass spectrometry. *J Am Soc Mass Spectrom* 12, 420-427.

53. Suckau, D., and Resemann, A. (2003) T3-sequencing: targeted characterization of the N- and C-termini of undigested proteins by mass spectrometry. *Anal Chem* 75, 5817-5824.
54. Resemann, A., Wunderlich, D., Rothbauer, U., Warscheid, B., Leonhardt, H., Fuchser, J., Kuhlmann, K., and Suckau, D. Top-down de Novo protein sequencing of a 13.6 kDa camelid single heavy chain antibody by matrix-assisted laser desorption ionization-time-of-flight/time-of-flight mass spectrometry. *Anal Chem* 82, 3283-3292.
55. Debois, D., Bertrand, V., Quinton, L., De Pauw-Gillet, M. C., and De Pauw, E. MALDI-in source decay applied to mass spectrometry imaging: a new tool for protein identification. *Anal Chem* 82, 4036-4045.
56. Demeure, K., Quinton, L., Gabelica, V., and De Pauw, E. (2007) Rational selection of the optimum MALDI matrix for top-down proteomics by in-source decay. *Anal Chem* 79, 8678-8685.
57. Groseclose, M. R., Massion, P. P., Chaurand, P., and Caprioli, R. M. (2008) High-throughput proteomic analysis of formalin-fixed paraffin-embedded tissue microarrays using MALDI imaging mass spectrometry. *Proteomics* 8, 3715-3724.
58. Ronci, M., Bonanno, E., Colantoni, A., Pieroni, L., Di Ilio, C., Spagnoli, L. G., Federici, G., and Urbani, A. (2008) Protein unlocking procedures of formalin-fixed paraffin-embedded tissues: application to MALDI-TOF imaging MS investigations. *Proteomics* 8, 3702-3714.



## SUPPLEMENTARY DATA

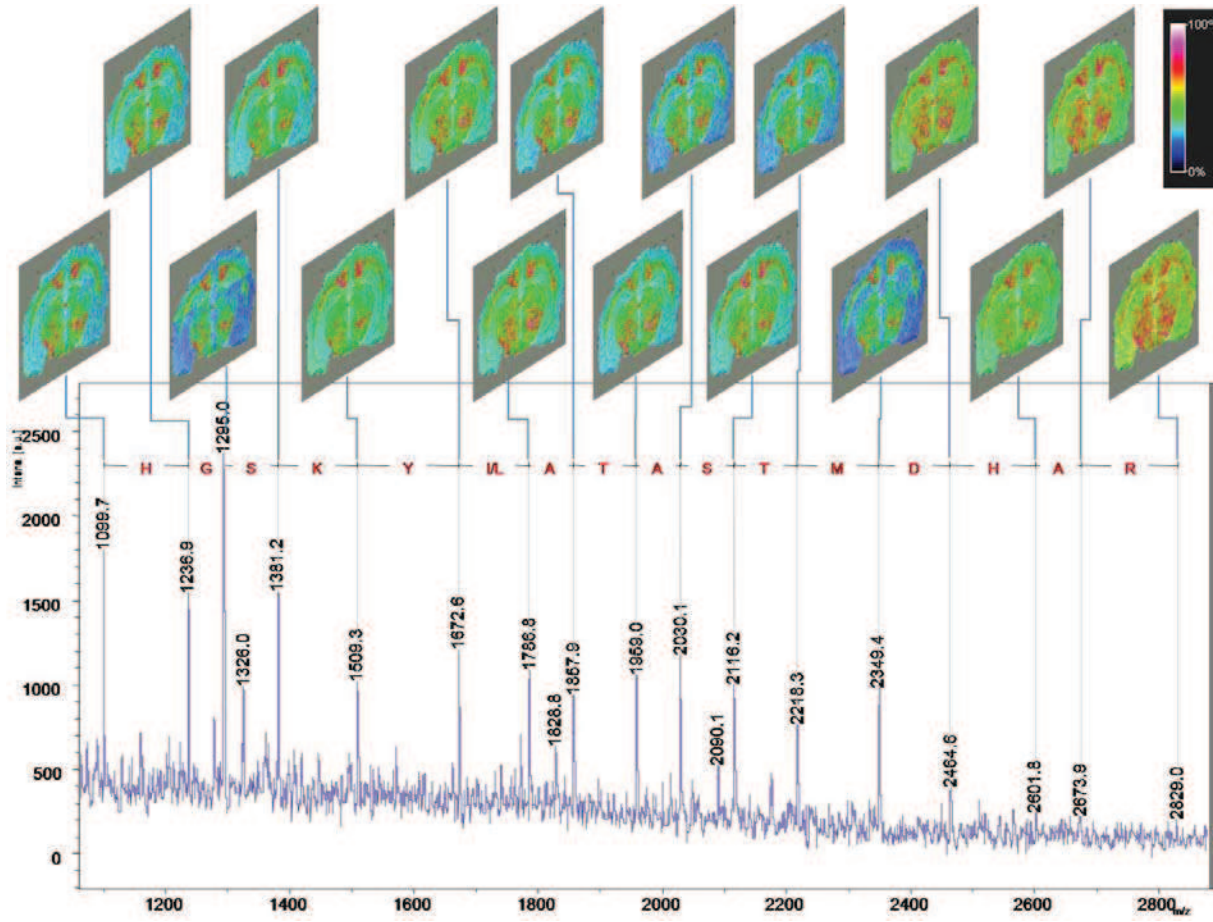


Figure S1: Myelin identification by ISD-MALDI-MSI in cingulum region of rat brain section with MALDI-images of each ISD-fragments.



## Discussion et perspectives

L'utilisation de la classification hiérarchique sur des données d'imagerie MALDI à partir de tissu frais ou FFPE nous montre que les étapes de préparation du tissu, influent certes sur le type de molécules détectées, mais n'entraînent pas de délocalisation de l'information. En effet, les images des clusters obtenus par Hiérarchisation (CH) coïncident avec les zones histologiques du tissu après chaque type de traitement, démontrant qu'elles ont conservé leur profil spectral. Cette information est primordiale, puisqu'habituellement le seul moyen de vérifier s'il n'y avait pas de délocalisation consistait en la réalisation d'images de  $m/z$  choisis arbitrairement et ayant une localisation spécifique ainsi qu'une validation par immunohistochimie.

Comme nous venons de le décrire dans la publication, l'apport de la PCA-SDA permet d'éviter toutes pertes d'informations. Contrairement à l'exemple sur le cancer de la prostate ci-dessus, certains tissus peuvent présenter une grande complexité histologique, comme le cerveau par exemple, et les informations d'une zone histologique très petite vont très vite se « noyer » dans l'ensemble des données. Cette perte d'information diminue grandement l'efficacité de l'imagerie par spectrométrie de masse puisque tout l'intérêt de cette technique est de pouvoir détecter un très grand nombre de molécules et d'aboutir à leurs localisations. C'est pourquoi, suite à une classification hiérarchique permettant de distinguer les plus fines zones histologiques, il est possible de sélectionner et d'exporter les spectres de cette zone (en prenant d'autres spectres comme contrôle) et d'appliquer une nouvelle analyse PCA sur ces données pour trouver précisément les  $m/z$  d'intérêt et ne perdre aucune information. De la même manière, il est possible de sélectionner plusieurs zones d'intérêt détectées suite à la CH et de comparer ces zones les unes par rapport aux autres.

Cette approche peut s'avérer très utile dans le cadre d'étude sur des cancers à un stade précoce où le doute persiste sur une région du tissu. En effet, suite à la mise en évidence d'une zone potentiellement cancéreuse par classification hiérarchique, il sera possible de récupérer les spectres de cette zone qui ne pourrait contenir que quelques dizaines de cellule et de mettre en évidence des marqueurs spécifiques préalablement identifiés.

Cette stratégie de PCA-SDA s'inscrit donc dans une problématique de protéomique clinique de recherche de biomarqueurs et de leurs validations par des approches sur tissus de type bottom-up (**Figure 42**).

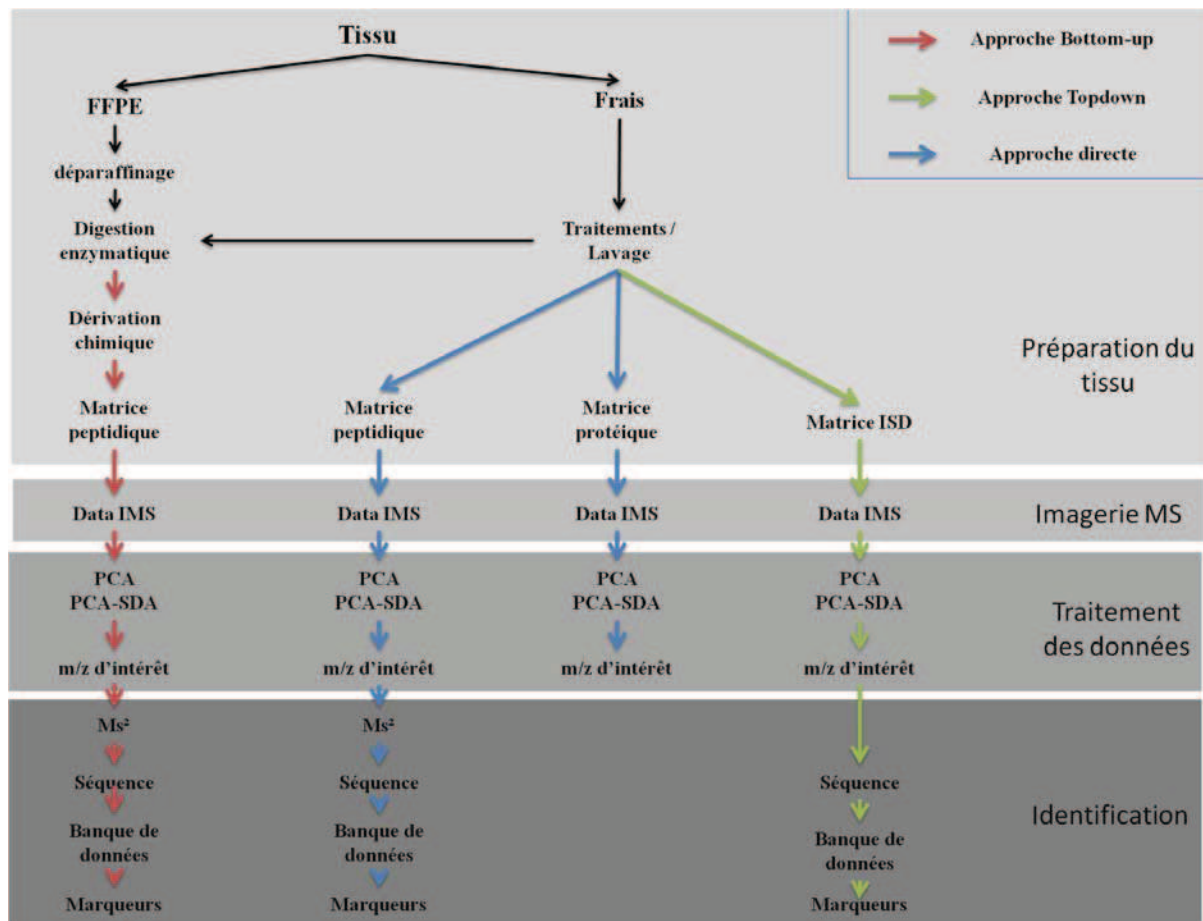


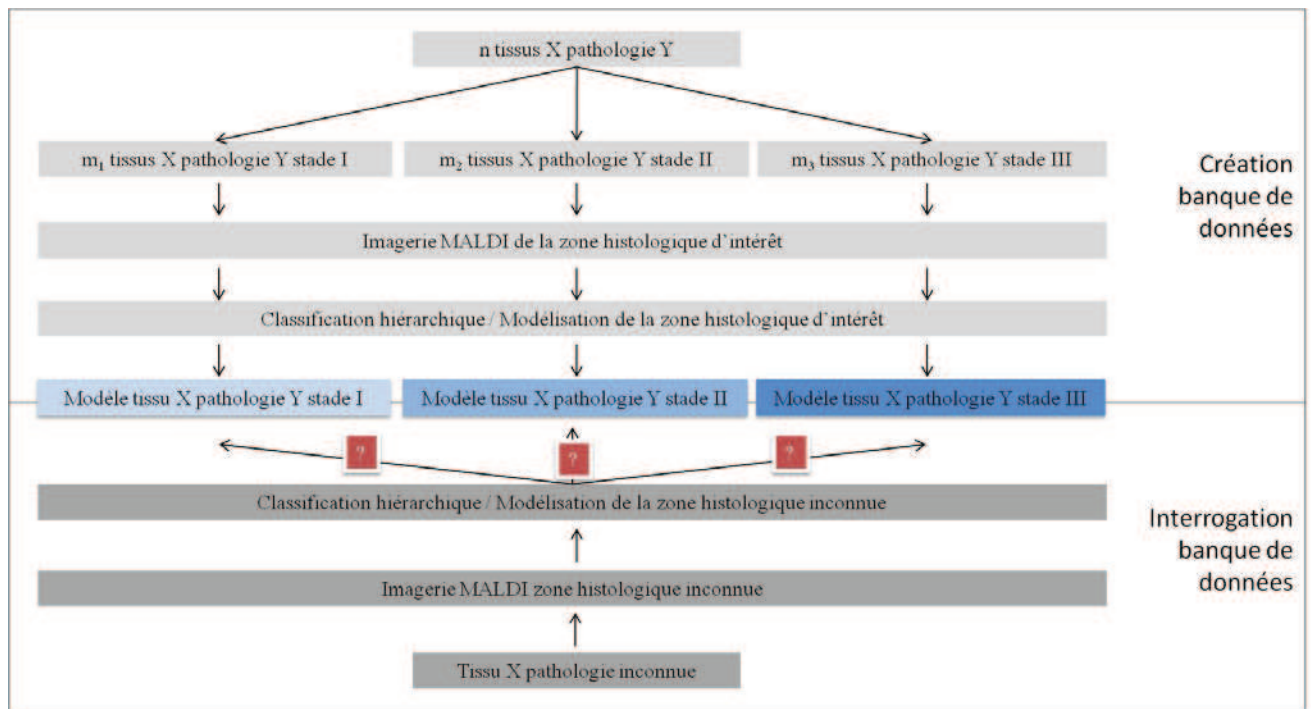
Figure 42 : Recherche et identification de biomarqueurs par imagerie MALDI en protéomique clinique.

Cette stratégie d'analyse multivariées et de modélisation appliquées aux données d'imagerie MALDI pourrait servir également comme nouvelle méthode diagnostic. La **figure 43** illustre l'idée générale de cette approche et de sa mise au point éventuelle. Comme nous l'avons vu, la classification hiérarchique permet de générer des groupes de spectre ayant le même profil général peptidique et/ou protéique en accord avec la zone histologique et physiologique d'un tissu à un stade donné de son développement. Ce travail nécessite une profonde collaboration entre les chercheurs spécialisés en MSI, les anatomo-pathologistes et les chirurgiens pour créer une banque de donnée qui reposerait sur l'utilisation de modèle, de signature peptidique et/ou protéique pour chaque type de tissu pathologique. Une fois la banque de données générée, les anatomo-pathologistes pourraient s'appuyer sur cette technologie en vue de réaliser un diagnostic moléculaire précis basé sur ces développements. Ainsi, en accord avec les anatomo-pathologistes, il serait possible de corréliser certains groupes (clusters) de spectres avec une zone histologique précise du tissu (comme le stade pour le cancer par exemple) par rapport aux résultats obtenus en immuno-histochimie comme le fait déjà l'équipe du Professeur A. Walch à l'hôpital de Munich [9, 261-263].

Ces spectres d'intérêt seraient ensuite exportés et permettraient de générer un modèle mathématique (que l'on peut assimiler à une signature) du profil général peptidique et/ou protéique. Plusieurs algorithmes de classification pourraient être utilisés comme le support vector machine (SVM), le quickclassifier (QC), le genetic algorithm (GA) ou encore le supervised neural network (SNN).

- Le GA mime l'évolution du groupe de spectres et permet la sélection de combinaisons de pics qui sont les plus influents sur la séparation.
- SVM détermine des plans de séparation entre les différentes classes et à partir de ces plans, un classement et une modélisation des groupes de spectres obtenus sont effectués.
- SNN est un algorithme de classification basé sur les prototypes. Le SNN identifie des spectres caractéristiques de chaque classe, appelés prototypes, qui peuvent être considérés comme des représentants de chaque groupe.
- Le QC trie les classes des pics dans un modèle qui reprend également des données statistiques comme les p-values ou les positions de certains pics dans l'espace multi-dimensionnel. Pour la classification, le rapport aire sur intensité de ces pics est trié et une valeur moyenne pour tous les pics est calculée.

Chaque modèle d'une pathologie donnée à un stade donné pourrait facilement être référencé dans une banque de données. Notons que pour aboutir à un modèle valide, un nombre important de tissus provenant d'une pathologie précise serait nécessaire issus de centres différents. De même, d'importantes étapes de validation et de validation croisée devront être envisagées pour confirmer définitivement le modèle et juger de sa pertinence. Dès lors, suite à la biopsie d'un tissu suspect aux yeux du médecin, il sera possible d'effectuer une acquisition en MSI suivi d'une classification hiérarchique des données puis d'exporter les spectres de la zone d'intérêt (idéalement ceux-ci formeraient un cluster), modéliser ces données et interroger la banque de données en comparant la signature de la zone suspecte avec les signatures enregistrées pour aboutir à un diagnostic. Les résultats obtenus se présenteraient sans doute sous la forme d'un pourcentage de similitude associé à des p-values et des T-Test.



**Figure 43 : Utilisation de la modélisation pour les données d'MSI : hypothèse de son application pour le diagnostic clinique.**

L'atout majeur de cette technique serait la méthode de diagnostic basée sur une signature moléculaire et non sur un marqueur seul. C'est pourquoi, il est fondamental d'identifier les marqueurs liés au stage, au type d'une pathologie et de comparer ces marqueurs à ceux obtenus par d'autres équipes afin de juger la robustesse de ceux-ci.

La signature moléculaire comporterait à la fois l'information générale de la composition en peptides et/ou protéines de la zone histologique et leurs abondances relatives basée sur des bouquets de biomarqueurs préalablement caractérisés.

## Conclusion

Comme nous l'avons vu l'apport des analyses multivariées et de la classification hiérarchique a modifié notre manière d'analyser les données d'imagerie MALDI. Cette approche robuste, puissante et très rapide nous permet aujourd'hui d'exploiter finement et pleinement la totalité de nos données. De la même manière, la CH nous a permis de valider les différents protocoles et approches, à la fois sur tissu frais et tissu FFPE, utilisés en MSI.

La PCA-SDA combine à la fois l'interprétation des résultats de CH et de PCA pour la recherche fine de marqueurs dans des échantillons présentant une grande complexité histologique. Enfin, nous ouvrons l'hypothèse d'une nouvelle voie de diagnostic qui demanderait sans doute plusieurs années de mise en place et d'élaboration.

## **Application des analyses multivariées au cancer de l'ovaire**

Les cancers ovariens épithéliaux comprennent cinq sous-types histologiques : séreux, à cellules claires, endométrieux, mucineux et transitoire avec une prévalence plus élevée pour les adénocarcinomes séreux. 90% des tumeurs ovariennes proviennent de l'épithélium de surface ovarien (OSE). Actuellement, l'origine précise des EOC (epithelium ovarian cancer) n'est pas très bien définie en raison du manque de compréhension de l'étiologie du cancer de l'ovaire. A l'inverse de plusieurs autres cancers humains, les modèles murins des EOC humains n'ont été développés que depuis peu. Un modèle d'adénocarcinome ovarien a été décrit chez le poulet (*Gallus domesticus*) et des cas de tumeurs sporadiques ont été décrits chez certaines souches de souris et de rats. Néanmoins, ces modèles présentent des incidences tumorales faibles et des périodes de latence longue, rendant les modèles peu attractifs pour la recherche et la compréhension des EOC. Les connaissances actuelles ne parviennent pas à déterminer les modèles de croissance histologique des différents sous-types tumoraux. Des études épidémiologiques tendent à démontrer qu'une grossesse, la lactation et la prise orale d'un contraceptif durant plus de 5 ans seraient liées à une diminution des risques d'apparition du cancer de l'ovaire. Après l'ovulation, les OSE se divisent et migrent pour réparer le tissu suite à la sortie du follicule. Les OSE peuvent rester encapsulés dans le stroma et développer des kystes d'inclusion. La grossesse, la lactation et la prise de contraceptif, empêchant et ralentissant l'ovulation chez la femme, conduisent à une réduction de l'ovulation et donc une moins forte migration des OSE, réduisant ainsi les risques de formation des kystes. Les facteurs impliqués dans cette formation de kystes ne sont pas encore très bien identifiés mais une hypothèse serait que ces cellules, une fois présentes dans le stroma, sont soumises à de fortes concentrations en molécules extracellulaires, augmentant ainsi le risque de transformation en néoplasie. D'autres hypothèses pourraient également impliquer les cellules provenant des trompes de Fallope et de l'épithélium endométrial.

On utilise la classification de la Fédération Internationale de Gynécologie Obstétrique (FIGO), qui se rapproche de celle du TNM (détaillée dans le chapitre 5), pour classer les différents stades physiologiques du cancer de l'ovaire :

- STADE I : tumeur limitée aux ovaires
  - Ia : un seul ovaire, capsule intacte
  - Ib : 2 ovaires, capsule intacte
  - Ic : rupture capsulaire, tumeur à la surface de l'ovaire ou dans l'ascite
- STADE II : tumeur ovarienne étendue au pelvis
  - IIa : extension à l'utérus et / ou aux trompes
  - IIb : extension organes pelviens
  - IIc : extension avec ascite
- STADE III : métastases péritonéales extra pelviennes
  - IIIa : métastases péritonéales microscopiques
  - IIIb : métastases macro < 2 cm
  - IIIc : métastases macro > 2 cm
- STADE IV : métastases à distance

Le marqueur sérique CA 125 (Cancer Antigen 125) est le plus utilisé de nos jours pour le diagnostic des EOC. Il s'agit d'un marqueur de haut poids moléculaire aussi appelé MUC16 (mucinous glycoprotein) qui fut décrit pour la première fois en 1981. CA 125 a rapidement été utilisé pour le screening de cette pathologie [264]. Au fil des années, d'autres marqueurs potentiels ont été découverts seuls ou en association avec CA 125 pour leur rôle potentiel dans l'EOC [265], comme CA 19.9 et LPA (acide lysophosphatidique lipidique). Des algorithmes complexes associés à la compréhension de l'expression des marqueurs ont augmenté l'efficacité des tests, permettant le pronostic ou le diagnostic du cancer de l'ovaire [266]. Le dépistage pourrait apparaître comme une solution pour améliorer la précocité du diagnostic d'autant plus qu'avec l'échographie on dispose d'un moyen très sensible pour détecter des anomalies ovariennes. Malheureusement, la spécificité de ce test même couplé au dosage du marqueur biologique CA 125 est très insuffisante au regard de la faible prévalence du cancer de l'ovaire par rapport aux lésions bénignes. La généralisation des échographies conduirait donc à un beaucoup trop grand nombre d'explorations chirurgicales (cœlioscopie) pour parvenir à un diagnostic histologique de certitude. Le dépistage de masse n'est donc pas



préconisé actuellement et en dehors des formes familiales où il existe une prédisposition génétique (5 à 10 % des cas) aucun groupe à risque ne peut être véritablement individualisé.

Notons que 75% des cancers de l'ovaire sont diagnostiqués tardivement, *i.e.* aux stades III ou IV. A ce stade le temps de survie des patientes est alors estimé à cinq ans et le taux de survie reste très faible.

Dans ce contexte, la recherche de nouveaux marqueurs et la compréhension des mécanismes impliqués dans le développement du cancer de l'ovaire sont donc des priorités auxquelles nous avons essayé de répondre à l'aide d'une approche protéomique et plus particulièrement en utilisant la technique d'imagerie MALDI.

En effet, L'imagerie MALDI développée depuis maintenant plus de dix ans montre un intérêt croissant pour la « chasse » aux biomarqueurs dans le cadre de la recherche clinique que ce soit pour l'étude de cancer du sein [261], de la prostate [136, 267], du colon [140], de lymphomes [268] ou du pancréas [269]. A partir du moment où il est possible pour l'expérimentateur de discriminer les zones tissulaires saines des zones tissulaires cancéreuses sur une coupe de tissu ou suite à l'analyse histologique d'un anatomo-pathologiste, alors la MSI devient un outil extrêmement puissant pour mettre en évidence des peptides et/ou protéines d'intérêt. Il est possible de détecter les molécules présentes sur le tissu, d'obtenir leur localisation et d'aboutir à leur identification dans le même appareil ou en utilisant des spectromètres de masse complémentaires. Le type d'étude basée sur cette méthodologie est rapide en raison du temps de préparation des échantillons et d'acquisition. Ce gain de temps permet d'introduire un nombre important d'échantillons dans l'analyse pour valider la ou les molécules identifiées. Il reste également possible de valider un biomarqueur potentiel à l'aide des techniques classiques comme le western blot.

Cette étude, initiée par Rémi Lemaire durant sa thèse, a été possible grâce à un partenariat avec une équipe de chirurgiens en gynécologie à l'hôpital Jeanne de Flandre de Lille (Pr D. Vinatier, Dr J. Lucot et Dr P. Collinet) via l'analyse de profils de biopsies ovariennes cancéreuses humaines que nous avons ensuite comparées à des profils obtenus sur des kystes bénins ou des ovaires sains (descente d'organes). Cette collaboration a permis la mise en place d'une procédure en relation avec les blocs opératoires : le tissu et le sang ont pu être prélevés, étiquetés et échantillonnés pour être conservés dans de la glace sur les lieux de l'opération dans des conditions reproductibles, avant d'être stockés au laboratoire à -80°C. Le prélèvement comprend le tissu ovarien bénin ou malin.

Cette étape de conditionnement du tissu est réellement primordiale. En effet, chaque prélèvement doit suivre un protocole standardisé pour permettre une étude reproductible et la plus représentative possible. Le temps nécessaire pour réaliser les analyses doit être équivalent pour chaque prélèvement afin de réduire les variabilités inter-échantillons dues à des dégradations des protéines par les protéases. De la même manière, le stockage des échantillons doit être identique et doit permettre de les maintenir dans des conditions les plus proches du profil moléculaire physiologique.

Ensuite, un soin tout particulier est apporté à la classification des tissus en accord avec les chirurgiens et les anatomo-pathologistes. Nous référençons les biopsies suivant leur état physiologique (sain, borderline et carcinome) et chaque tumeur est classée en fonction de son stade en utilisant la classification FIGO.

Enfin, il est important de noter l'existence d'une variabilité inter-patiente. Cette variabilité est la plus difficile à palier. Elle peut être liée à la variabilité naturelle existant entre chaque individu ou encore à des traitements préalables suite à la détection d'une masse sur leur ovaire. Ces traitements ont pu modifier les patterns moléculaires des ovaires et peuvent parfois expliquer certaines différences. L'utilisation d'un nombre conséquent de biopsies et la validation croisée des résultats permettent de supprimer certains faux positifs. De plus, nous n'avons considéré que les marqueurs absents/présents entre les biopsies cancéreuses et les tumeurs/kystes bénins ou ovaires sains (prolapsus) ; les pics présentant simplement une modulation d'intensité n'ont pas été considérés.

Le but de ce travail était donc de valider les résultats obtenus à l'aide de profiling MALDI sur des coupes de biopsies ovariennes il y a trois ans dans le laboratoire [138]. Mais cette fois, nous avons utilisé l'imagerie MALDI et la totalité des données issues de cette technique pour la recherche de biomarqueurs potentiels (et non plus comme un outil de validation de la localisation). Ainsi nous souhaitons mettre en évidence de nouveaux marqueurs potentiels à l'aide de données obtenues sur des coupes de biopsies de cancer de l'ovaire en utilisant de nouvelles applications des méthodes statistiques pour ainsi mieux comprendre les phénomènes biologiques liés au processus de développement du cancer.

Received: 2010.03.24  
 Accepted: 2010.04.12  
 Published: 2010.08.01

## MALDI imaging mass spectrometry in ovarian cancer for tracking, identifying, and validating biomarkers

BR

**Authors' Contribution:**

- A** Study Design
- B** Data Collection
- C** Statistical Analysis
- D** Data Interpretation
- E** Manuscript Preparation
- F** Literature Search
- G** Funds Collection

**Mohamed El Ayed**<sup>1,2,3,4</sup>, **David Bonnel**<sup>\*1,3,4</sup>, **Remi Longuespée**<sup>1,3,4</sup>,  
**Céline Castellier**<sup>1,4</sup>, **Julien Franck**<sup>1,4</sup>, **Daniele Vergara**<sup>1</sup>, **Annie Desmons**<sup>1</sup>,  
**Aurélien Tasiemski**<sup>1</sup>, **Abderraouf Kenani**<sup>2</sup>, **Denis Vinatier**<sup>1</sup>, **Robert Day**<sup>3</sup>,  
**Isabelle Fournier**<sup>1,4</sup>, **Michel Salzet**<sup>1,4</sup>

\* join first authors

<sup>1</sup> Université Nord de France, CNRS, MALDI Imaging Team, Laboratoire de Neuroimmunologie et Neurochimie Evolutives, Université Lille 1, Lille, France

<sup>2</sup> Laboratoire de Biochimie – Unité de Recherches “Mécanismes Moléculaires et Pathologies”, Faculté de Médecine de Monastir, Monastir, Tunisie

<sup>3</sup> Institut de Pharmacologie de Sherbrooke, Faculté de Médecine et des Sciences de la Santé, Université de Sherbrooke, Sherbrooke, Québec, Canada

<sup>4</sup> Clinique de Gynécologie, Hôpital Jeanne de Flandre, CHRU Lille, Lille, France

**Source of support:** Supported by grants from Centre National de la Recherche Scientifique (CNRS), Ministère de l'Éducation Nationale, de l'Enseignement Supérieur et de la Recherche, Agence Nationale de la Recherche (ANR PCV to IF), Institut du Cancer (INCA to IF), Institut de Recherche en Santé du Canada (ISRC to MS & RD), Région Nord-Pas de Calais (to DB & RL), Fonds de la Recherche en Santé du Québec (FRSQ to RD)

**Background:**

Among biomarkers, cancer-antigen 125 (CA-125) is the most studied. We propose an analytical tool to track ovarian carcinoma biomarkers, that is, the MALDI mass spectrometry imaging.

**Material/ Methods:**

Ovarian carcinomas and benign ovaries were directly analyzed by MALDI-TOF-MS. After automatic profiling and mass spectrometry imaging analyses, hierarchical clustering based on principal component analysis in nonsupervised mode was carried out. On the same samples, preparations were performed to investigate peptides, then proteins, followed by high mass proteins, in an automatic profiling to specific signatures for diagnosis. Using tissue bottom-up strategy on tissue digestion, and mass spectrometry imaging after by shotgun sequencing by nano-LC-IT-MS in MS/MS mode from washing samples from on tissue digested peptides, several biomarkers were found.

**Results:**

A list of specific biomarkers from the ovarian carcinoma regions was obtained and classified as proteins associated with cell proliferation, involved in immune response modulation, signaling to the cytoskeleton, and tumor progression. These specific biomarkers were then validated by immunocytochemistry using Tag-mass technology, cell biology, Western blot, and by PCR (using SKOV-3 ovarian epithelial cancer cells). A link between the immune regulation (innate immunity, tolerance) and virus cause is also discussed.

**Conclusions:**

From the biomarkers identified, proteins involved in immune response modulation and cell proliferation have been pointed out in this study. Two new markers have been identified using such a strategy, that is, fragment C-terminal of the PSME1 (Reg-Alpha) and mucin-9.

**key words:**

**MALDI imaging • pathologies • innate immunity • proteomics • PCA • hierarchical clustering**

**Full-text PDF:**

<http://www.medscimonit.com/fulltxt.php?ICID=881095>

**Word count:**

5539

**Tables:**

4

**Figures:**

7

**References:**

76

**Author's address:**

Michel Salzet, Laboratoire de Neuroimmunologie et Neurochimie Evolutives, FRE CNRS 3249, 59650 Villeneuve d'Ascq, France, e-mail: michel.salzet@univ-lille1.fr

## BACKGROUND

Ovarian cancer is the fourth leading cause of cancer death among women in Europe and the United States. Among biomarkers, cancer-antigen 125 (CA-125) is the most studied. CA-125 has a sensitivity of 80% and a specificity of 97% in epithelial cancer (stage III or IV). However, its sensitivity is around 30% in stage I cancer, its increase is linked to several physiological phenomena, and it is also detected in benign situations [1]. CA-125 is particularly useful for at-risk population diagnosis and for following illness evolution during therapeutic treatment. In this context, CA-125 is insufficient as a single biomarker for ovarian cancer diagnosis. Thus, other biomarkers are required to use a proteomic strategy for diagnosis [2–10].

At this time, 2 strategies have been undertaken by researchers. Several, several groups have tried to identify ovarian cancer markers in plasma or serum using SELDI-TOF profiling or chromatography coupled to mass spectrometry [3,11–16]. Other groups have developed a classic proteomic strategy using comparative 2D-gels and mass spectrometry [10,17–19].

Here, we propose another strategy based on direct tissue analysis and peptide profiling followed by MALDI profiling and imaging. Ovarian carcinomas (stages III and IV) and benign ovaries were directly analyzed by MALDI-TOF-MS. Hierarchical clustering based on principal component analysis (PCA) was carried out using ClinProTools software to classify tissues. Principal component analysis was used in the unsupervised mode to differentiate tumorous and healthy spectra based on their proteomic composition as determined by MALDI-MSI. Two stage IV tumor regions as well as 2 healthy regions were identified. To validate the procedure, 2 biomarkers identified in the 2 carcinoma regions were characterized using tissue MS-MS and nanolc-IT-MS. Validation was performed by specific MALDI imaging using the Tag-mass concept, and PCR was performed on mRNA extracted from patients or from an epithelial cancer cell line (SKOV3). The first biomarker was identified as a fragment of the immunoproteasome Reg-Alpha [20], and the second was identified as orosomucoid.

Taken together with recent results from gastric cancer diagnosis [21,22], our results suggest that MALDI-MSI is a promising technology for tumor classification and for identifying specific biomarkers in various types of tumors.

## MATERIAL AND METHODS

### Materials

$\alpha$ -cyano-4-hydroxycinnamic acid (HCCA), sinapinic acid (SA), 3-acetylpyridine (3AP), 1,1,1,3,3,3-Hexafluoro-2-propanol (HFIP) ammonium bicarbonate ( $\text{NH}_4\text{CO}_3$ ), trisma base, xylene, methanol (MeOH), ethanol (EtOH), acetonitrile (AcN), angiotensin II, Des-Arg-bradykinin, substance P, ACTH 18-39, ACTH 7-38, and bovine insulin were obtained from Sigma-Aldrich and used without any further purification. Trypsin was from Promega. AspN, LysC enzyme was from Roche. Trifluoroacetic acid (TFA) was purchased from Applied Biosystems. Acetonitrile p.a. and methanol p.a. were from J.T. Baker.

### Samples

Tissues, ascites, and cyst fluids were obtained with informed consent and institutional review board approval (CCPPRBM Lille: CP 05/83) from patients undergoing any ovarian tumor resection at Hospital Jeanne de Flandre. A total of 48 tissue samples from 25 patients with grade III and IV ovarian cancer, and 23 benign tumors, were analyzed. Patient information was collected, including sex, age, treatment received before and after surgery, extent of surgery, current status (alive, alive with progressive disease, deceased, and cause of death), and survival from the time of original pathologic diagnosis. Samples were collected at the time of surgery, immediately frozen, and stored at  $-80^\circ\text{C}$  until analysis. Typically, 10- to 12- $\mu\text{m}$  thick sections were cut using a cryostat and thaw-mounted on flat, electrically conductive sample slices. Histopathologic diagnoses were made by an anatomopathology blinded to the original clinical diagnosis from subsequent H&E-stained sections.

### Tissue preparation

Frozen ovary sections were immediately transferred onto a conductive Indium-Tin Oxide (ITO) glass (Bruker Daltonics, Wissenbourg, France). After drying the sections for 5 minutes at room temperature, tissues were heated to  $37^\circ\text{C}$  for 20 seconds to adhere to slides and placed under a vacuum for 10 minutes. They were then rinsed in chloroform and analyzed in MALDI ms for MALDI imaging analyses [23].

### MALDI Imaging and Specific MALDI Imaging

#### Automated tissue profiling

For automated profiling assays, markers of 3 mass ranges were screened for each section. After the tissue treatment, 10  $\mu\text{L}$  of HCCA 10 mg/mL in ACN/TFA 0.1% 7:3 were dropped to perform the analysis of peptides. The laser settings for these analyses were power 40%, offset 65%, range 20%, small focus, and 150 ns of pulsed ion extraction. A methanol 100% wash was then used to remove the matrix, and 10  $\mu\text{L}$  per section of SA 20 mg/mL in AcN/TFA 0.1% were dropped to perform the analysis on the proteins mass range. The laser settings were power 45%, offset 65%, range 20% small focus, and 150 ns of pulsed ion extraction. Ten  $\mu\text{L}$  of SA 10 mg/mL in HFIP 100%, then 10  $\mu\text{L}$  of SA 20 mg/mL in ACN/TFA 0.1% 7:3 were then dropped to perform the analysis on high-mass proteins after their extraction from the tissue section [24]. For each analysis, 1000 shots were accumulated for 50 spectra per tissue section.

#### MALDI Imaging

A solid ionic matrix (sinapinic acid SA/3AP) was prepared just before use and was deposited using a Chemical Inkjet Printing CHIP-1000 device (Shimadzu, Kyoto, Japan) after a raster of spots spaced by 250  $\mu\text{m}$  center to center. Briefly, 4.8  $\mu\text{L}$  (1 equivalent) of 3AP was added to a solution containing 10 mg/mL of SA in ACN/aqueous TFA 0.1% (6/4, v: v). The solution was agitated for several minutes before use. Five droplets of approximately 100 pL were deposited at each spot per cycle. Twenty iterations were necessary to reach the total final volume of 10 nL. Images were acquired using an UltraFlex II MALDI-TOF/TOF instrument

(Bruker Daltonics, Bremen, Germany) equipped with a smart beam laser with a repetition rate of 200 Hz [25] and controlled by FlexControl 2.5 software (Bruker Daltonics). Images were performed in positive reflectron mode and MALDI-MS spectra were acquired in the mass range from 550 to 5000 Da. Six-hundred spectra were acquired at each spot using a laser frequency of 100 Hz. The images were recorded and reconstructed using Fleximaging II 2.5 (Bruker Daltonics) software.

### Specific MALDI Imaging

For specific MALDI imaging studies, ovary sections were incubated at room temperature with 500  $\mu$ L buffer (0.1 M PBS/1% BSA / 1% normal goat serum /0.05% triton  $\times$ 100) for 30 minutes [26,27]. The same buffer was used to dilute the Reg alpha antibody (1/100) (Zymed Laboratories, Invitrogen, ref. 38–2400), and incubation was performed overnight at 4°C. After 3 washes in PBS, sections were incubated overnight with anti-Human IgG photocleavable tagged rabbit antibody (1/100) (Eurogentec) at 4°C [28,29]. The tissues were then rinsed 3 times for 5 minutes with ultrapure water to remove salts, and sections were dried at room temperature before matrix application. In the case of the orosomucoid (alpha 1 glycoprotein) biomarker, a polyclonal antibody (MCA3312Z) was purchased from AbD Serotec and used at a dilution of 1/50. The secondary antibody is an anti-human IgG photocleavable tagged mouse antibody (1/100) (Eurogentec).

### Automatic trypsin digestions

The printed array of the CHIP-1000 on the tissue section was composed of spots spaced by 250  $\mu$ m center-to-center. A total of 40 nL of solution containing 25  $\mu$ g/mL of trypsin in water was applied to each spot. Five droplets of approximately 100 pL were deposited at each spot per cycle. Forty iterations were necessary to obtain the final volume.

### MS/MS of digested and derived tissues

MALDI-TOF MS/MS experiments on digested and derived tissue sections were performed using an Ultraflex II TOF-TOF instrument equipped with a LIFT III cell. For MS/MS experiments, the following parameters were set: laser repetition rate of 100 Hz with 33% attenuation, ion source voltages of 8 kV and 7.3 kV on the MALDI sample plate and first electrode; LIFT cell pulse from ground for electrode 1 and 2 to 19 kV; and in the last step, electrode 3 was decreased to 3.2 kV. Reflector end voltage was set to 29.5 kV and mid-grid to 13.85 kV. For each MS/MS spectrum, 1000 total shots were averaged, including 200 for parent ions and 800 for fragments. Laser fluence was constant over the experiments.

Protein identification in databanks was performed using Biotoools 3.0 software (Bruker Daltonics) connected to the Phenyx search engine interrogating the NCBI, Swissprot, EST, or MSDB databases. Methionine oxidation was set as the variable modification, and no fixed modification was used. Taxonomy was specified to be human. Trypsin was selected as the enzyme, and 2 missed cleavages were settled. The mass tolerance was set at 1 and 0.5 Da for the MS and MS/MS.

### Bottom-up analyses

#### Peptide tissue extractions

Slices were washed with acetonitrile (60%) acidified with HCl 1N (w/5v). The collected supernatants were incubated overnight at 4°C with gentle rocking. After centrifugation at 12 000 rpm for 30 minutes at 4°C, the supernatants were loaded on Sep-Pak C18 cartridges (500  $\mu$ L extract/cartridge; Waters). After washing with 5 mL acidified water (0.05% TFA, Pierce), samples were eluted with 5 mL 60% AcN in acidified water (0.05% TFA). Sixty percent of the eluted fraction was reduced in a vacuum centrifuge (Savant). This fraction was resuspended in 100  $\mu$ L acidified water (0.05% TFA) and fractionated on a C18 reversed-phase HPLC column (4.6 mm $\times$ 25 cm, Interchim) equilibrated with acidified water (0.05% TFA). Elution was performed with a linear gradient of acetonitrile in acidified water (0.05%) from 0% to 70% at a flow rate of 500  $\mu$ L/min. Each fraction was collected manually before being evaporated in a SpeedVac vacuum and resuspended with 50  $\mu$ L of HPLC grade water. Each fraction was analyzed by MALDI-TOF-MS before trypsin digestion.

#### Trypsin digestion

After drying, samples (extracted peptides) were placed on ice for 30 minutes in 50  $\mu$ L of protease solution (sequence grade-modified trypsin, Promega, at 0.02 mg/mL in 25 mM (NH<sub>4</sub>HCO<sub>3</sub>). Digestion was performed overnight at 37°C. Peptide extraction was performed twice for 15 minutes with 50% acetonitrile, 1% TFA for further MALDI-MS analysis, or with 50% acetonitrile, 1% formic acid for further ESI-MS/MS analysis. Trypsin digests were then lyophilized in a SpeedVac concentrator and resuspended in 5  $\mu$ L of 0.1% formic acid.

#### For nanoLC-ESI MS samples

Ten patients with grade III and IV ovarian cancer and 10 benign tumors were analyzed.

The cancer tissue sample contains serous cystadenoma, mucinous cystadenoma, and borderlines tissue.

**For nanoLC-ESI MS analysis**, on a section of 2 cm<sup>2</sup>, *in situ* enzymatic digestion is performed by adding 15  $\mu$ L of trypsin enzyme (0.033  $\mu$ g/ $\mu$ L in 25 mM Tris buffer pH 7.4) for 1 hour at room temperature. After enzymatic digestion, purification of resulting digestion peptides was achieved by using reverse phase C<sub>8</sub> coated silica magnetic beads (ClinProts, Bruker Daltonics) according to the protocol of the manufacturer modified for tissues. Then, this, 15  $\mu$ L of binding solution was directly applied onto the tissue during 1 minute, and then 15  $\mu$ L of magnetic bead was added on the section. Extraction occurred for 10 minutes. During this step, beads and digested products were mixed 3 times using a micropipette directly onto the tissue. Digestion solution and beads were then deposited in a polypropylene tube and washed 3 times using 500  $\mu$ L of H<sub>2</sub>O/TFA 0.1%. Peptides were eluted from the beads with 30  $\mu$ L of ACN/H<sub>2</sub>O (1:1, v/v) and solution was dried by vacuum centrifugation. For the nano LC-MS/MS identification, peptides were redissolved in H<sub>2</sub>O/MEOH 0.1% formic acid (9:1 v/v) after elution and evaporation.

**Table 1.** Forward and reverse primer sequences used in RT-PCR are shown 5' to 3'.

Primer designation	Forward primer	Reverse primer
Orosomucoid	CTGGGAGAGTTCTACGAAGC	CCTCTGTTCCTCTCCTT
Actin	AGCGCAAGTACTCCGTGTG	GACTGGGCCATTCTCCTAG
GAPDH	ACCACAGTCCATGCCATCAC	TCCACCACCTGTGTGCTGTA

### Nano LC-nanoESI-IT MS and MS/MS

Analyses were performed on an ion trap mass spectrometer (Esquire 3000 plus, Bucker Daltonics) equipped with a nano ESI ion source and on-line coupled to a nano HPLC system. An injection of 0.5  $\mu$ L of digest was made with a Switchos Autosampler (Dionex corporation) and separation was performed on a C18 silica bonded stationary phase (75  $\mu$ m id, 150 mm long, 3  $\mu$ m 100 Å pore size, Dionex). Samples were washed for 2 minutes at 10  $\mu$ L/min with 100% mobile phase A (95% H<sub>2</sub>O, 5% ACN 0.1% formic acid). Peptides were then eluted using a linear gradient of 1%/minute mobile phase B (ACN 80%, H<sub>2</sub>O 20%, formic acid 0.08%) for 70 minutes at a flow rate of 0.2  $\mu$ L/min. The Esquire was operated in a data-dependent MS/MS mode in which 1 MS full scan was followed by 1 MS/MS scan on the most-abundant peptide ion. Collision energy was set to 35%. The heated capillary temperature and electrospray voltage were 160°C and 1.5kV respectively.

Protein identification was performed under MASCOT sequence query search program using SwissProt database filtered for the taxonomy "human." A tolerance of 1 Da for peptide and 0.5 Da for MS/MS was set. Only protein sequences with MOWSE score higher than 20 (indicating significant homology or identity) and identified in several samples representing at least 4 significant MS/MS were considered. Methionine oxidation and acetylation of N-terminal were defined as variable modification.

### Statistical data analysis

For statistical analyses, mass spectra were internally recalibrated on common peaks (also known as *spectral alignment*) and normalized to the total ion count. An average spectrum created from all single spectra was used for peak selection and to define integration ranges. These integration ranges were used to obtain the intensities or areas of single spectra. Signal intensities were used for all calculations. For the principal component analysis and hierarchical clustering, the individual peak intensities were standardized across the data set. The Principal Component Analysis (PCA) was carried out using Pareto scaling, which uses the square root of the standard deviation as a scaling factor to reduce the dominance of large-scale intensity changes in the matrix and other high-abundance ions, as these may mask variations in lower abundance ions during PCA. The overall outcome of PCA is greatly affected by the masking of the underlying relevant information by ions related to matrix coating and other endogenous molecules. Under unsupervised PCA, each spectrum is classed as an individual so the principal components are selected and account for the greatest separation of each of the individual spectra.

### RNA isolation and PCR analysis

mRNA from biopsies or the SKOV-3 cell line was extracted in 2 mL tubes prefilled with Qiazol reagent (Qiagen, France) and 1.4 mm ceramic beads. The mixture was shaken twice for 45 seconds at 6500 rpm in a Precellys 24 homogenizer (Bertin distributed by Ozyme, France). RNA extraction was performed according to the manufacturer's instructions, and the total extracted RNA was treated with RQ1DnaseI (Promega, France) to prevent genomic DNA contamination. First strand cDNA was generated from 2  $\mu$ g of total RNA using random primers (Promega, France) and the Superscript III reverse transcriptase (RT) kit (Invitrogen, France) in a final volume of 60  $\mu$ L.

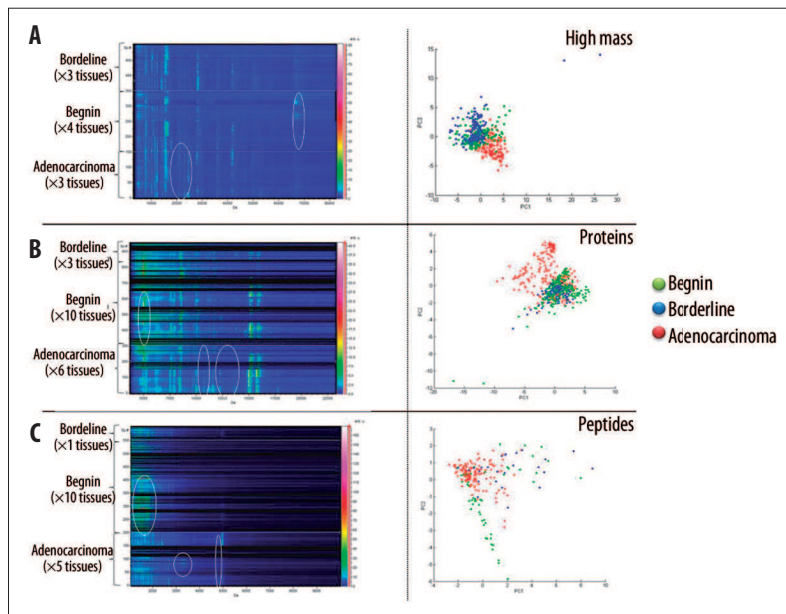
Omitting RT or RNA from the reaction mixture resulted in nonamplification and no-template controls. cDNA was treated with RNaseH (Promega, France) to optimize the amplification. For the TLR, Reg-alpha fragment and orosomucoid genes, forward and reverse primers (Table 1) were designed with the Primer3 Input software ([http://frodo.wi.mit.edu/cgi-bin/primer3/primer3\\_www.cgi](http://frodo.wi.mit.edu/cgi-bin/primer3/primer3_www.cgi)). Toll Like Receptor (TLR) primer pairs were obtained from Krug et al. [29] Actin, GAPDH and 18S were used as reference genes. In each case, PCRs were performed for 30 cycles using Advantage 2 polymerase (Clontech) with an elongation time of 2 minutes. All PCR products were subcloned into the pGEM-T easy vector (Promega) and cDNA clones were sequenced with an ABI Prism 310 genetic analyzer (Applied Biosystems).

### qPCR analyses for virus detection genes

Epstein-Barr Virus R-gene quantification kit and CMV HHV6,7,8 R-Gene diagnostic kit (CE-IVD labeled are purchased at Argene and performed on Applied Biosystems 7500 as recommended by the manufacturer.

### SKOV-3 cells stimulation and immunohistochemistry

The human epithelial ovarian cancer cell line SKOV-3 was grown in RPMI-1640 medium with glutamine, supplemented with 10% FBS, 100 U/mL penicillin, and 100  $\mu$ g/mL streptomycin. Cells were treated with TGF- $\beta$  (10 ng mL<sup>-1</sup>) for 24 hours in the absence of serum (B). Control cells were left untreated (A). For confocal microscopy analysis, cells were fixed for 5 minutes with 3.7% formaldehyde in phosphate-buffered saline (PBS) solution, permeabilized with a 0.1% solution of Triton X-100 in PBS, followed by 30 minutes incubation at room temperature with phalloidin-TRITC (Sigma). Slices were then mounted in glycerol and examined using a confocal microscope (Zeiss LSM 510). Control cells have a typical epithelial-like morphology in culture flask/Petri with the tendency to form dense colonies.



**Figure 1.** Automatized MALDI profiling on three ovarian tissues classes: adenocarcinoma, borderline and benign. (A) Pseudogel and supervised PCA loading plot obtained for a high mass procedure. (B) Protein procedure. (C) Peptide procedure.

## Western blotting

Whole cell lysates (in lysis buffer as described above) containing 50 µg of total proteins (COF1, PROF1) (Sigma) were diluted in cold RIPA buffer, resolved on 12% sodium dodecyl sulphate-polyacrylamide gels (mini protean II electrophoresis system, BioRad) and blotted onto nitrocellulose membranes. The membranes were blocked with 5% w/v nonfat dry milk in TBS containing 0.1% Tween-20 (Sigma) and probed with selected primary antibodies. After incubation with secondary antibodies, immunoblots were visualized with the ECL detection kit (Amersham Biosciences).

## RESULTS

### Automatic profiling and PCA for biomarkers tracking

Automatic profiling was performed on 20 samples (10 benign, 6 carcinoma, 3 borderline) laid on the same slides and submitted to peptides, then small proteins, and finally high mass proteins tracking using 3 in-house methods, 1 after 1 based on classical matrix deposition and solvent treatments [23,24,30–32]. The analyses of a MALDI imaging data set by molecular mass images can be illustrated as a pseudogel (Figure 1). In peptides conditions (with HCCA as a matrix), carcinoma specific signatures can be found for  $m/z$  around 3300 and 4800, whereas in a benign larger zones for  $m/z$  ranged from 1200 to 2000 can be found. As can be seen, the borderline profile resembles the benign one. The principal component analysis (PCA) analyses confirm the observation (Figure 1). For small proteins (with SA as matrix), 2 signatures can be seen for  $m/z$  comprised between 9500 and 14 000 and 1 around 17 500 in carcinoma samples.

It is difficult to discriminate between borderline and benign profiles except for a specific signature at  $m/z$ , ranging between 4900 and 5200 which is, specifically detected in benign samples, and is absent in both carcinoma and borderline profiles (Figure 1). In high-mass proteins, a specific zone comprised between  $m/z$  23 000 and 25 000 in carcinoma whereas in benign and borderline profiles a zone between

$m/z$  65 000 and 68 000 is detected (Figure 1). Principal component analyses (PCA) confirmed the eyes detected zones but 2 others are detected through statistics in carcinoma; that is, 1 around  $m/z$  9000 and 12 000; 1 between  $m/z$  23 000 and 24 000, and the last 1 around  $m/z$  35 000–37 000 (Figure 1). In a borderline profile PCA analyses detected a zone between  $m/z$  5300 and 5000 and in begin a cross-zone with the malignant 1 at  $m/z$  10 000 and 11 000 (Figure 1).

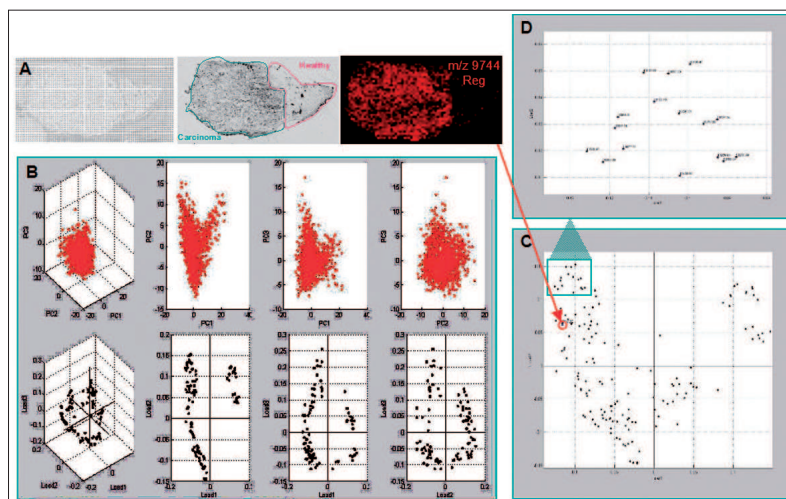
All together, the data shows that it is necessary to perform on same samples the 3 procedures, helped with PCA analyses to have a clear overview of different classes of biomolecules present in the samples and to drive the biomarker hunting. In this context, as it can be seen in Figure 1B, a classic procedure using SA as matrix for peptides and small proteins did not allow one to detect high-mass proteins, only polar ones can be obtained. Thus, the HFIP procedure [24] was applied, and the number of peaks detected in the  $m/z$  20 000 to 50 000 range was clearly increased (Figure 1A). Interestingly, several of these peaks are consistent with previously identified biomarkers using classic 2D proteomic procedures (Table 2), suggesting that the chemical treatments could enable known biomarkers to be recorded directly from tumor regions in tissue biopsies. These show that MSI can be more complementary to the classic proteomic for biomarker tracking.

### MALDI Imaging Mass Spectrometry and Principal Component Analyses for biomarkers tracking

MALDI imaging mass spectrometry analyses were performed on carcinomas tissues (stage III or IV) followed by PCA analyses using SA/3AP as matrix (Figure 2A). Unsupervised PCA was used to attempt to identify hidden variables between spectra taken from various regions of ovarian biopsy (Figure 2B). Figure 2B shows the score plots for unsupervised PCA. Plot PC1/PC2 shows the best separation, and 2 groups can be separated; that is, tumor versus healthy regions, which is in line with the histologic data (Figure 2C). A zoom of the most-separated ions from the PC1/PC2 plot is presented in Figure 2D, and a group of

**Table 2.** Molecular masses of known ovarian cancer protein biomarkers and masses found by MALDI MSI after using the HFIP high-mass protein preparation procedure compared to molecular masses of previously identified ovarian cancer markers.

Protein Name	Mw (Da) MALDI MSI with high mass procedure	Mw (Da) Protein previously identified in OVC	References
Tetranectin (CAA45860)	17 775	17 776	[67–70]
Neutropil Gelatinase-associated lipocalin precursor	22 576	22 571	[71]
Plasma retinol-binding protein precursor	22 986	22 990	[72]
Metalloproteinase inhibitor 1 precursor	23 152	23 153	[73]
Kallikrein 5 Precursor	26 842	26 838	[74]
Isoform 1 of Urokinase plasminogen activator surface receptor precursor	36 940	36 949	[75,76]

**Figure 2.** MALDI MSI analyses of a mucinous adenocarcinoma stage IV. (A) mucinous ovarian carcinoma section containing tumor and healthy parts submitted to automatic matrix deposition using a micro-spotter (CHIP-1000, Shimadzu). Location of the C-terminal part of Reg-alpha (m/z 9744) biomarker confirming the location of the tumor portion of the biopsy slice. (B) PCA analyses through PC1, PC2 and PC3 of the mucinous ovarian carcinoma. (C) Statistical PCA analyses with PC1 as tumor and PC2 as healthy tissue. Location of tumor biomarker vs protein present in the healthy part. (D) Zoomed view of the most-differentiated m/z corresponding to part of the specific tumor biomarkers.

specific biomarkers was detected. Their locations in the tumor region are presented in Frank et al. (2009) [33]. From the PCA results, we detected a fragment of the (PSME 1: Proteasome activator complex subunit 1, named the Reg alpha fragment), previously identified [20], is present in the cancer group with other biomarkers and not present in healthy tissue.

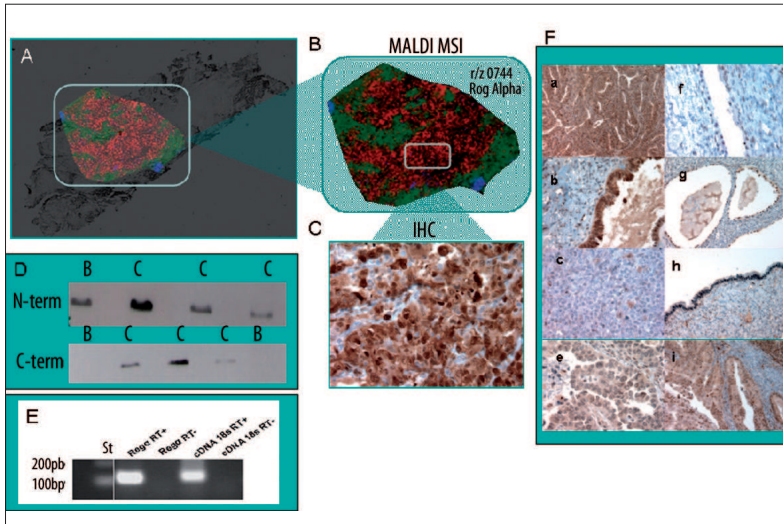
### Biomarker identification

From the list of biomarkers detected by PCA analyses, we detect the ion at m/z of 9744, which corresponds to the fragment C-terminal of Reg alpha [20]. We validate it here through MSI (Figures 2A, 3A,B), immunocytochemistry (Figure 3C,F), Western blot (Figure 3D), and molecular biology (Figure 3E) confirmed its specificity to ovarian carcinoma. Ion at m/z of 9744 corresponding to the C-terminal part of PSME1 is detected in the carcinoma area by MSI (Figure 3A,B), which is confirmed by immunocytochemistry, with an antibody raised against the C-terminal part of the protein (Figure 3C). Western blot analyses of benign and carcinoma samples were performed either with an antibody raised against the N-terminal of PSME1 or the 1 directed against the C-terminal part (Figure 3D).

The data reflect the antibody raised against the N-terminal part recognizes the protein in both carcinoma and benign samples, whereas the antibody raised against the C-terminal recognizes its epitope only in carcinoma samples, which is in line with previous data [20] and the immunocytochemical results (Figure 3F). Figure 3F shows that this C-terminal fragment is over-present in epithelial cells of carcinoma (Figures 3F a, b, e, i). It can be detected in endometrioid nondifferentiated (Figure 3F c) or differentiated (Figure 3F a) carcinoma, sero-mucinous adenocarcinoma (Figure 3F e), and clear-cell adenocarcinoma (mesonephroma) (Figure 3F i). In benign tissues, the immunolabeling is also at the level of the epithelial cells, but more nucleus than in carcinoma can be observed in (Figures 3F f, h) the adenofibromatous tumor (Figure 3F g). All together, the data is in line with the transcriptomic results performed on the SKOV-3 epithelial carcinoma cells (Figure 3E), confirming overexpression of PSME1 gene in this carcinoma cell line. These data confirm that Reg-alpha is a specific ovarian biomarker. We recently discovered it in ascites liquid (data not shown).

For the other biomarkers detected by PCA analyses, we could detect that some have already been found in genomic

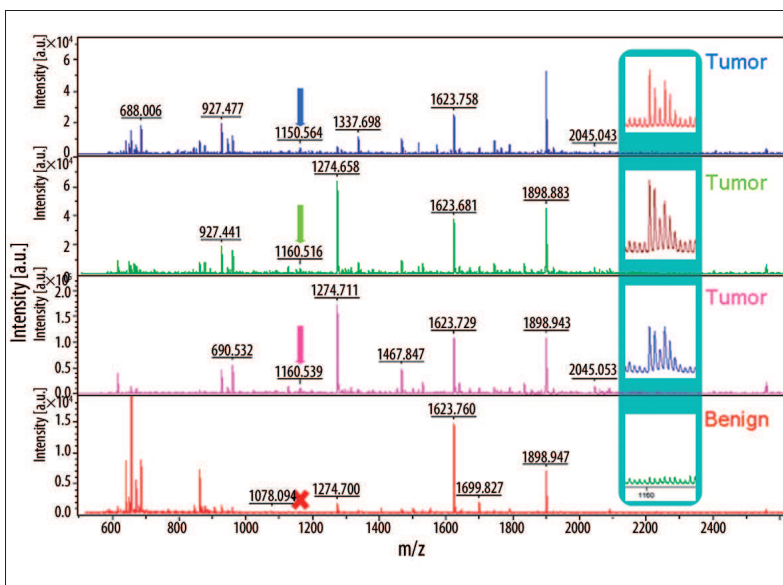




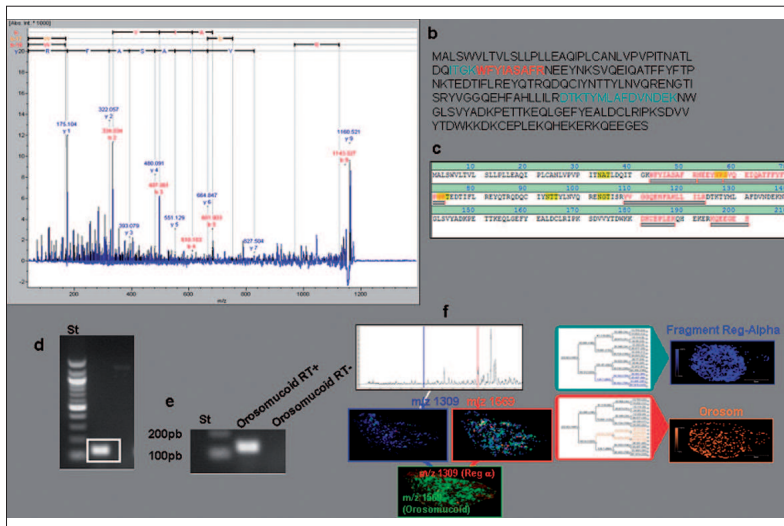
**Figure 3.** (A) Ovarian carcinoma section stained with Hematoxylin, Eosin, Safran (H&S) colorants. (B) MSI image of the ion at  $m/z$ : 9744, reg alpha fragment. (C) Immunocytochemical analyses with antibody raised against the C-terminal fragment on stage IV carcinoma sections. Most of the epithelial cells are labeled. (D) Western blot analyses with benign (B) or carcinoma from different patients (C) with antibodies directed against either the N-terminal or C-terminal parts of PSME1. (E) RT-PCR analyses of PSME1 of mRNA extracted from SKV03. (F) Immunocytochemical studies with polyclonal antibody raised against the c-terminal part of Reg alpha. (a) Epithelial cells of immunolabeled differentiated endometrioid carcinoma, (b) Epithelial cells of immunolabeled in carcinoma region. (c) Cytoplasmic epithelial cells immunolabeling of nondifferentiated endometrioid carcinoma. (d) Epithelial cells of immunolabeled in clear cells adenocarcinoma (mesonephroma). (e) Nuclear epithelial cells immunolabeling of benign tumor. (f) Nuclear epithelial cells immunolabeling of adenofibromatous tumor. (g) Nuclear epithelial cells immunolabeling of seromucous area in benign tumor. (h) Cytoplasmic epithelial cells immunolabeling of sero-mucinous adenocarcinoma.

studies, and share the same mass, for example, protein S100-A12 ( $m/z$  10568) and apolipoprotein A1 ( $m/z$  10155) [34]. We decided to remove the MALDI matrix from the slides through a washing procedure and analyze samples by MALDI before and after trypsin digestion [32] (Figure 4). The differentiated ions then were subjected to MS/MS analysis by MALDI before analysis by nanoLC-IT/MS in MS/MS mode. For example, the ion at  $m/z$  1160.5 detected by MALDI in a patient sample after trypsin digestion (Figure 4) was subjected to MS/MS analysis (Figure 5A). The obtained sequence of WFYIASAFR with a score of 42 (Figure 5B) was confirmed by the IT-MS/MS data, and it belongs to orosomuroid 1. Five others fragments were characterized and gave 30.3% sequence coverage (Figure 5B). The presence in the biopsy of the mRNA coding for this protein was confirmed by RT-PCR from patient with ovarian carcinoma (Figure 5C). A 150-bp fragment coding for orosomuroid was amplified (Figure 5D). Similarly, RT-PCR performed on RNA extracted from SKOV-3 epithelial carcinoma cells confirmed the expression of the orosomuroid (Figure 5E), further supporting the biochemical data. Finally, antibodies raised against the C-terminus of Reg alpha and a

monoclonal against orosomuroid were used for immunocytochemistry using the specific MALDI imaging procedure [28] (Figure 5F). Both biomarkers were detected using the



**Figure 4.** MALDI-MS profiles of 3 ovarian carcinomas after tissue trypsin digestion using an automatic micro-spotter (CHIP-1000, Shimadzu) versus benign tumor; a zoom of the peptide 1160.5 is performed. This peptide is present in 3 ovarian carcinomas and absent in the benign tumor. (Inset corresponds to the zoom of the 1160.5 region.)



**Figure 5.** (A) MALDI MS/MS sequencing of 1160.5 using MALDI TOF-TOF. (B) Sequence (WFYIASAFR) in red corresponding to the fragment identified by MS-MS and localized in the orosomuroid-1 protein. (C) Sequence of the orosomuroid obtained by nano-LC-ITMS after trypsin digestion and by tissue MALDI TOF/TOF analyses. (D) RT-PCR amplification of orosomuroid from mRNA extracted biopsy (5 carcinoma versus 5 benign). (E) RT-PCR amplification of orosomuroid from SKOV-3 cell lines. (F) (a)MALDI specific imaging using the TAG-MASS concept with a polyclonal anti-Ct Reg alpha antibody tagged with a reporter (with a mass of 1309) and a monoclonal anti-orosomuroid tagged with another reporter (with a mass of 1569). (b) Merge of specific MALDI imaging molecular images of the C-terminal fragment of Reg alpha and orosomuroid. (c) MALDI imaging molecular images of the C-terminal fragment of Reg alpha and orosomuroid with the correspondence of the 2 subclass's location in the tumor part; each biomarker is characteristic of 1 region.

tag-mass procedure in the carcinoma region with specific localization (Figure 5F b), which was consistent with the clustering distribution (Figure 5F c).

A combination of MALDI TOF-TOF and nanoLC-IT MS/MS analyses (Table 3) allowed us to characterize 15 biomarkers (Table 3) from 20 samples (10 patients with ovarian cancer and absent in 10 benign tumors). Some of these have been previously reported as secreted proteins in large-scale analyses of human plasma from ovarian cancer patients [35], human ovarian cancer tumors [36] or ovarian ascites analyses [37]. Moreover, their location was determined by MSI using a bottom-up strategy. After tissue trypsin digestion, proteins were reconstructed based on their

*in silico* trypsin digested peptides detected on the digested tissue (Table 4, Figure 6). MALDI imaging mass spectrometry images from digested fragments for the same protein were cumulative and gave the molecular image of its

**Table 3.** Biomarkers identified by nanoLC-IT-MS in MS/MS mode from 10 patients with ovarian cancer absent in 10 benign tumors.

Access number	Protein name	Molecular mass	Mascot score (average)	Sequence coverage (average)
P08670_CHAIN_0	Vimentin	53651	89.6±22.3	27±10%
P51884_CHAIN_0	Lumican Precursor	38428	63.6±5.7	19.3±4%
P02647_CHAIN_1	Apolipoprotein 1	11183	49.9±4	21.6±4.9%
P51888_CHAIN_0	Prolargin	43809	52.5±0.7	17±0.8%
Q12889_CHAIN_0	Oviductin (mucin-9)	75372	47.8±1.3	13.3±1.2%
Q01995_CHAIN_0	Transgelin	22610	49.5±2.48	26±2.44%
P02765	Orosomuroid	23511	45±2.44	20.8±1.8%
P02787_CHAIN_0	Siderophilin	77049	38.1±2.2	12±4.8%
P01009_ISOFORM_2	Alpha 1 antiprotease	39066	29.8±1.6	21±3.2%
P08727	Cytokeratin 19	44635	29.3±1.75	10±1.6%
P30086_CHAIN_0	Phosphatidyl Ethanolamine Binding Protein	21056	28.1±0.9	23±2%
P31949	Protein S100 A11	11740	22.6±0.25	24.5±0.5%
P02790_CHAIN_0	Hemopexin	51676	21.0±0.4	6.5±1.2%
P07737_CHAIN_0	Profilin -1	15054	18.3±0.9	15.4±4.7%

**Table 4.** Trypsin digested fragment of proteins identified by NanoLC-IT-MS in MS/MS and found on tissue after bottom-up strategy and characterized by MS/MS.

Orosomucoide		Oviductal glycoprotein 1		Alpha anti-protease 1	
Mass	Fragment	Mass	Fragment	Mass	Fragment
1160.59	WFYIASAFR	1118.68	FIASVISLLR	1015.6	SVLGQLGITK
1559.83	ITGKWFYIASAFR	1120.6	LIMGIPTYGR	1110.59	LSITGTYDLK
1752.83	YVGGQEHAHLLILR	1288.62	FTTMLSTFANR	1568.78	ECCGGVFTDPKGIFK
		1248.54	SSAYAMNYWR		
		1508.77	TLSIGGWNFGTSR		
Siderophilin		Reg alpha C-terminal		Profilin-1	
Mass	Fragment	Mass	Fragment	Mass	Fragment
1577.8	TAGWNIPMGLLYNK	1500.98	QLVHELDEAEYR	1643.9	TFVNITPAEVGVLV GK
		1518.2	IEDGNNGVAV QEK	1470.7	SSFYVNGLTGGQK
Prolargin		Transgelin		Hemopexin	
Mass	Fragment	Mass	Fragment	Mass	Fragment
1064.59	SFPNLAFIR	1204.7	TLMALGSLAVTK	1141.58	GGYTLVSGYPK
1309.7	LPGLVFLYMEK	1530.7	TDMFQVDLFEKG	1120.599	NFPPSPVDAEFR
1352.71	NQLEEVPALPR	1221.62	QMEQVAQFLK		
1549.856	NLMQLNLAHNLR				
Vimentin		S100		Lumican	
Mass	Fragment	Mass	Fragment	Mass	Fragment
1254.5	LGDLYEEEMR	1060.48	DGNYNTLSK	1178.63	LKEDAVSAAFK
1169.70	ILLAELEQLK	1849.89	TEFLSFMNTELAFTK	1024.55	FNALQYLR
1428.7	SLYASSPGGVYATR			1297.66	SLEDLQLTHNK
1533.84	KVESLQEEIAFLK			1180.65	RFNALQYLR
1570.88	ISLPLPNFSSLNLR				
Cytokeratin 19		Phosphatidylethanolamine-binding		Apolipoprotein 1	
Mass	Fragment	Mass	Fragment	Mass	Fragment
1064.09	SFPNLAFIR	1560.8	LYTLVLTPDPAPSR	1031.51	LSPLGEEMR
1365.7	SRLEQEIATYR	1632.79	NRPTSISWDGLDSGK	1301.64	THLAPYSDEL R
1389.67	AALEDTLAETEAR	1949.93	GNDISSGTVLSDYVSGPPK	1386.7	VSFLSALEEYTK
1586.83	LEHLYLNNSIEK			1230.7	QGLLPVLESFK

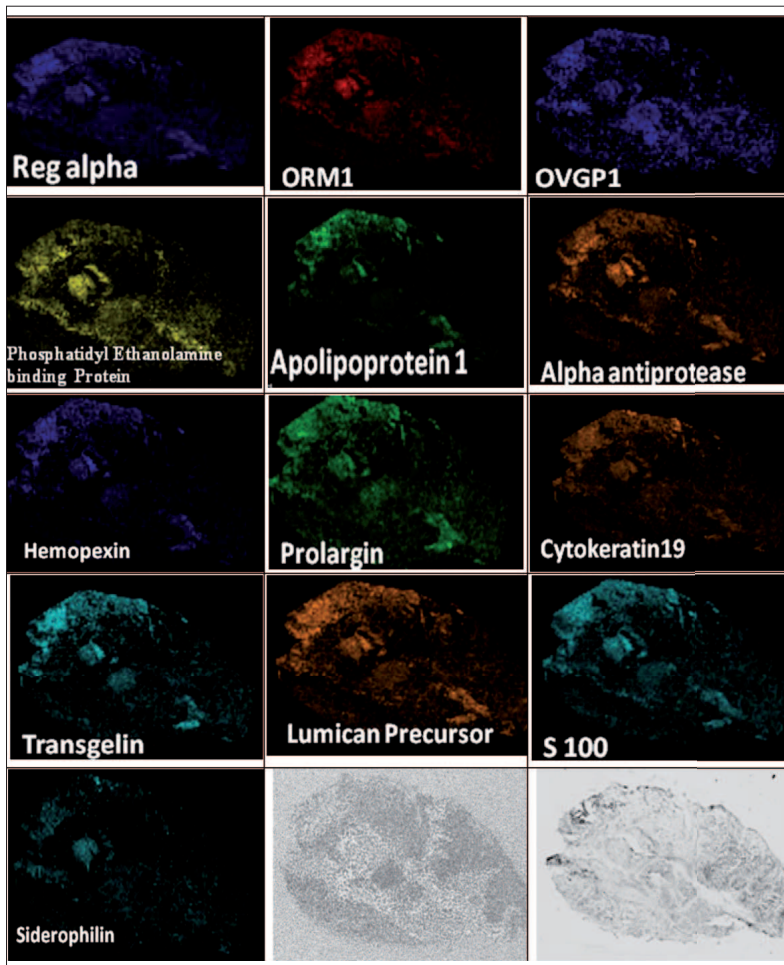
tissue repartition. Interestingly, all the detected proteins have the same location.

## DISCUSSION

These characterized proteins can be grouped into functional categories such as cell proliferation, immune response modulation, signaling to the cytoskeleton, and tumor progression.

## Proteins associated with cell proliferation

The S100 protein family has been previously detected in aggressive ovarian tumors [16]. In our study, we detected S100 A11 (Table 3) and S100 A12 (PCA analyses) proteins in tumors. S100 A11 has already been detected in ovarian ascites [37]. S100 A11 (or calgizzarin) is known to regulate cell growth by inhibiting DNA synthesis [38,39]. S100 A12 is known to contribute to leukocyte migration in chronic inflammatory responses [40].



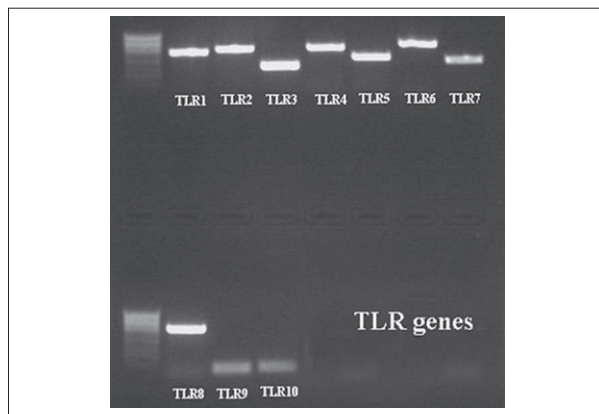
**Figure 6.** MALDI molecular images of the identified proteins by nanoLC-ITMS reconstructed on their trypsin digested fragment. Directly detected on tissue.

#### Proteins involved in immune response modulation

Recent studies have shown that ovarian cancer-associated ascites may provide an immunosuppressive environment [41]. In fact, a high CD4/CD8 ratio, which may indicate the presence of regulatory T-cells, is associated with poor outcomes. Reduced IL-2 and elevated IL-6 and IL-10 levels favor a Th2 inhibitory immune response. This immunosuppressive climate may explain the observation of non-responsiveness in ascites-derived T cells [41]. Considering these data, we tried to identify biomarkers from our pool that could be involved in such immunosuppression. In fact, we detected 5 factors involved in immune response modulation in our mucinous tumors: a C-terminal fragment of the 11S immunoproteasome (Reg-alpha), orosomucoid, apolipoprotein A1, hemopexin, and lumican.

PSME1 (proteasome activator complex subunit 1, 11S regulator complex [syn.: PA28 alpha]) cleaved into the Reg-alpha fragment could lead to default self-antigen presentation [20]. In fact, PA28 is a regulatory complex associated with 20S proteasome that consists of 3 subunits: alpha, beta, and gamma [42]. Binding of the 11S regulator complex to the 20S proteasome does not depend on ATP hydrolysis, and unlike the 19S regulatory subunit, the 11S regulator complex does not catalyze degradation of large proteins. Rather, it is responsible for MHC-class I antigen processing [43–45], which is greatly improved by interferon gamma-induced expression of the alpha and beta subunits [46].

Several viral proteins that interact with these proteasome subunits have been reported, and they may interfere with host anti-viral defenses, thereby contributing to cell transformation [47]. The manner in which they bind to the core particle via its subunits' C-terminal tails, and induce an  $\alpha$ -ring conformational change to open the 20S gate, suggests a mechanism similar to that of the 19S particle [42]. No role in ovarian cancer has been demonstrated for the 11S regulator complexes. Our data demonstrate a high level of expression of PA28 in carcinomas, especially in epithelial cells. The PA28 activator belongs to the antigen processing machinery (APM). Its alteration by cleavage in ovarian carcinomas may be a mechanism to evade immune recognition. Such a hypothesis has already been proposed for the case of APM chaperones such as TAP, LMP2, LMP10, and tapasin in colon carcinoma, small cell lung carcinoma, and pancreatic carcinoma cell lines. In fact, IFN- $\gamma$  treatment of these carcinoma cell lines corrects the TAP, LMP, and tapasin deficiencies and enhances PA28  $\alpha$ , LMP7, calnexin and calreticulin expression, which is accompanied by increased levels of MHC class I antigens [48]. Recently, PSEM2 (proteasome activator complex subunit 2, PA28 Beta) has also been detected in ascites fluid, which implicates it in immune cell tolerance toward carcinoma cells and confirms the dysregulation of self-antigen processing in ovarian tumors [37]. Moreover, PA28 alpha seems to be a target for Epstein-Barr virus (EBV) and herpes virus (HV), as we preliminary detected by qPCR (data not shown). In fact, Pudney and colleagues [49] have shown that, as EBV-infected cells move



**Figure 7.** RT-PCR amplification of Toll-like receptors from SKOV-3 cell lines.

through the lytic cycle, their susceptibility to EBV-specific CD8<sup>+</sup> T-cell recognition falls dramatically, concomitant with a reduction in transporter associated with antigen processing (TAP) function and surface human histocompatibility leukocyte antigen (HLA) class I expression. Implication of virus in the cause of ovarian cancer is also sustained by the overexpression of furin enzyme (data not shown), which is known to be implicated in glycoprotein B cleavage through a motif R-X-K/R-R in both EBV and HV [50,51].

Among the 4 other factors that might participate in the tolerance phenomenon by inhibiting immune activation, the acute phase protein, orosomucoid (ORM, also known as alpha1-acid glycoprotein or AGP), is normally increased in infection, inflammation, and cancer, and it seems to have immunosuppressive properties in ovarian carcinoma ascites through inhibition of IL-2 secretion by lymphocytes [52]. Similarly, apolipoprotein A1 has been detected in conjunction with transthyretin and transferrin in early-stage mucinous tumors [53]. ApoA-I is known to decrease expression of surface molecules such as CD1a, CD80, CD86, and HLA-DR in dendritic cells, and it stimulates the production of IL-10 [54].

Interestingly, hemopexin has recently been demonstrated to reduce TNF and IL-6 from macrophages during inflammation, and it limits TLR4 and TLR2 agonist-induced macrophage cytokine production [55]. We demonstrate that in SKOV-3 epithelial ovarian carcinoma cells, all TLRs are overexpressed with the exception of TLR9 and TLR10 (Figure 7). This point is also in line with the overexpression of lumican, which is a small LRR proteoglycan in the extracellular matrix. Along with other proteoglycans, such as decorin, biglycan, and prolargin, lumican is known to be overexpressed in breast cancer and to play a role in tumor progression [56,57]. However, as demonstrated for biglycan, which interacts with TLR2/4 on macrophages [58,59], we speculate that lumican is also involved in the activation of the inflammasome through TLR2/4 interaction. The activation of all danger-sensing receptors in carcinoma cells can be explained by regulation of inflammation by carcinoma cells to facilitate tumor progression. It seems that ovarian cancer cells act as parasites and use molecular mimicry [60] to escape the immune response, and they produce immunosuppressors to achieve tolerance.

### Signaling to the cytoskeleton

Several candidate proteins, including profilin-1, cofilin-1, vimentin, and cytokeratin 19 are involved in intracellular signaling to the cytoskeleton. Changes in cell phenotype, such as the conversion of epithelial cells to mesenchymal cells, are integral not only to embryonic development but also to cancer invasion and metastasis. Cells undergoing the epithelial-mesenchymal transition (EMT) lose their epithelial morphology, reorganize their cytoskeleton, and acquire a motile phenotype through the up- and down-regulation of several molecules, including tight and adherent junction proteins and mesenchymal markers. TGF- $\beta$  has been described to induce EMT in ovarian adenocarcinoma cells [61].

In the human lung adenocarcinoma cell line A549, this differentiation is accompanied by modification in the expression of several cytoskeleton proteins including  $\beta$ -actin, cofilin 1, moesin, filamin A and B, heat-shock protein beta-1, transgelin-2, S100 A11, and calpactin. These changes presumably increase migratory and invasive abilities [62]. We recently demonstrated that treatment of the ovarian cancer cell line SKOV-3 with TGF- $\beta$  (10 ng/mL, 24 h) increases the expression of cofilin and profilin-1 at mRNA and protein level, and modifies its cytoskeletal organization as assessed by confocal microscopy analysis [63]. After binding to its receptor, TGF- $\beta$  stimulates the reorganization of the actin cytoskeleton and triggers the formation of stress fibers and cellular protrusions [63].

### Tumor progression

In conjunction with S100 proteins and cytoskeleton modifying proteins, we also detected expression of oviduct-specific glycoprotein (OGP, Mucin-9), a marker of normal oviductal epithelium. Our data are in line with that of Woo and associates, who found that OGP is a tubal differentiation marker and may indicate early events in ovarian carcinogenesis [64,65].

### CONCLUSIONS

In summary, PCA and clustering, performed after MALDI imaging, analysis allow molecular tissue classification and will be of great help for pathological diagnoses. We confirmed that biomarkers detected in our tissue samples such as orosomucoid and lumican are highly glycosylated, which is in line with the mucinous phenotype of ovarian cancers. Thus, clustering analysis in conjunction with specific biomarker detection can enhance tumor tissue classification and subclassification and lead to better diagnoses. Moreover, based on the biomarkers identified in this study, we propose that ovarian cancer cells act to suppress immune responses. These findings also suggest that ovarian cancer can have a viral cause. Viruses are known to orient the immune response to immunosuppression [66]. Further studies are now in progress to explore this possibility to consider as potential biomarkers for ovarian cancer diagnosis Reg alpha fragment, mucin-9 in conjunction with oncoviral proteins specific from EBV and HHV6 viruses

### Competing interests

The authors declare that they have no competing interests.

## Authors' contributions

AM performed NanoLC-ITMS and MALDI imaging analyses; DB: performed PCA analyses RL and CC: performed automatic profiling studies; JF design the new ionic matrices and on tissue bottom-up strategy for biomarker identification; DV performed the EMT studies; AT carried out the molecular studies; AD performed immunocytochemistry and Western blotting, DV realized the surgeries, and the follow-up of the study; AK and RD follow-up the study though the codirection of AM and DB theses; and IF and MS conceived of the study and participated in its design and coordination. All authors read and approved the final manuscript.

## REFERENCES:

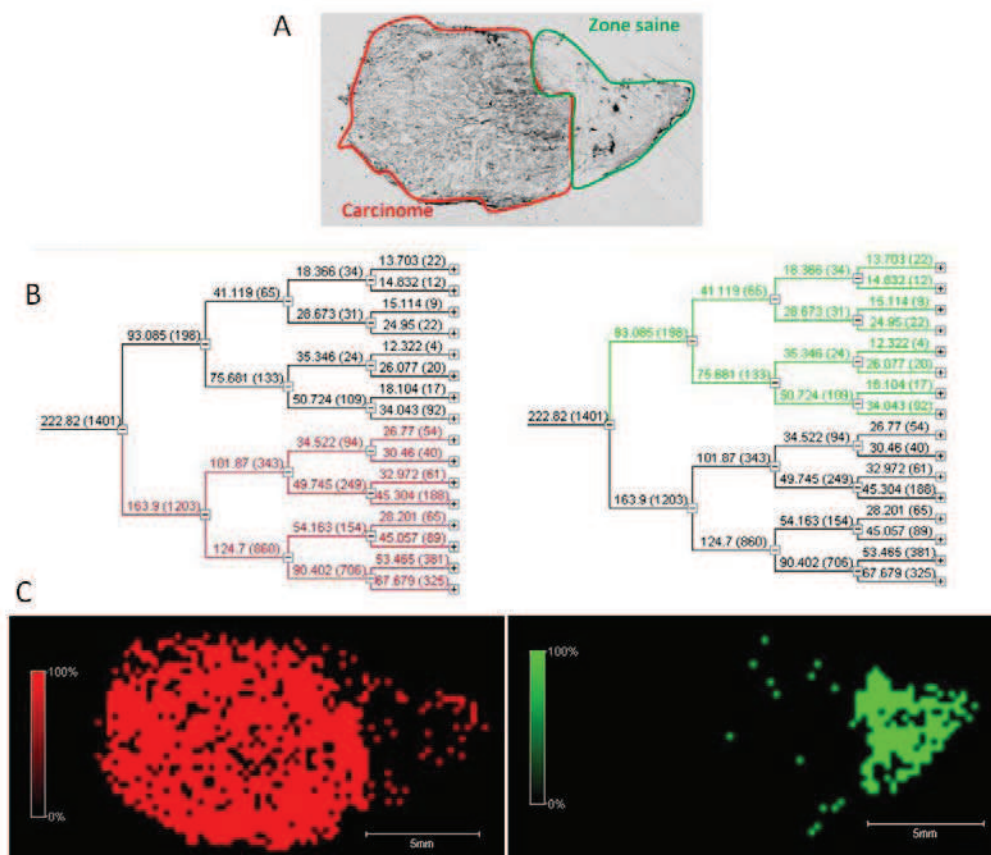
- Lambaudie E, Collinet P, Vinatier D: [Ovarian cancers and CA 125 in 2006]. *Gynecol Obstet Fertil*, 2006; 34: 254–57
- Edwards BK, Brown ML, Wingo PA et al: Annual report to the nation on the status of cancer, 1975-2002, featuring population-based trends in cancer treatment. *J Natl Cancer Inst*, 2005; 97: 1407–27
- Ardekani AM, Liotta LA, Petricoin EF III: Clinical potential of proteomics in the diagnosis of ovarian cancer. *Expert Rev Mol Diagn*, 2002; 2: 312–20
- Bandera CA, Tsui HW, Mok SC, Tsui FW: Expression of cytokines and receptors in normal, immortalized, and malignant ovarian epithelial cell lines. *Anticancer Res*, 2003; 23: 3151–57
- Conrads TP, Fusaro VA, Ross S et al: High-resolution serum proteomic features for ovarian cancer detection. *Endocr Relat Cancer*, 2004; 11: 163–78
- Conrads TP, Zhou M, Petricoin EF III et al: Cancer diagnosis using proteomic patterns. *Expert Rev Mol Diagn*, 2003; 3: 411–20
- Fields MM, Chevlen E: Ovarian cancer screening: a look at the evidence. *Clin J Oncol Nurs*, 2006; 10: 77–81
- Johann DJ Jr, McGuigan MD, Patel AR et al: Clinical proteomics and biomarker discovery. *Ann NY Acad Sci*, 2004; 1022: 295–305
- Kohn EC, Mills GB, Liotta L: Promising directions for the diagnosis and management of gynecological cancers. *Int J Gynaecol Obstet*, 2003; 83(Suppl.1): 203–9
- Rapkiewicz AV, Espina V, Petricoin EF III, Liotta LA: Biomarkers of ovarian tumours. *Eur J Cancer*, 2004; 40: 2604–12
- Petricoin EF, Ardekani AM, Hitt BA et al: Use of proteomic patterns in serum to identify ovarian cancer. *Lancet*, 2002; 359: 572–77
- Bergen HR III, Vasmataz G, Cliby WA et al: Discovery of ovarian cancer biomarkers in serum using NanoLC electrospray ionization TOF and FT-ICR mass spectrometry. *Dis Markers*, 2003; 19: 239–49
- Diamandis EP: Proteomic patterns in serum and identification of ovarian cancer. *Lancet*, 2002; 360: 170; author reply 171
- Engwegen JY, Gast MC, Schellens JH, Beijnen JH: Clinical proteomics: searching for better tumour markers with SELDI-TOF mass spectrometry. *Trends Pharmacol Sci*, 2006; 27: 251–59
- Fung ET, Yip TT, Lomas L et al: Classification of cancer types by measuring variants of host response proteins using SELDI serum assays. *Int J Cancer*, 2005; 115: 783–89
- Kikuchi N, Horiuchi A, Osada R et al: Nuclear expression of S100A4 is associated with aggressive behavior of epithelial ovarian carcinoma: an important autocrine/paracrine factor in tumor progression. *Cancer Sci*, 2006; 97: 1061–69
- Rai AJ, Zhang Z, Rosenzweig J et al: Proteomic approaches to tumor marker discovery. *Arch Pathol Lab Med*, 2002; 126: 1518–26
- Xiao Z, Prieto D, Conrads TP et al: Proteomic patterns: their potential for disease diagnosis. *Mol Cell Endocrinol*, 2005; 230: 95–106
- Zhu Y, Wu R, Sangha N et al: Classifications of ovarian cancer tissues by proteomic patterns. *Proteomics*, 2006; 6: 5846–56
- Lemaire R, Menguellet SA, Stauber J et al: Specific MALDI imaging and profiling for biomarker hunting and validation: fragment of the 11S proteasome activator complex, Reg alpha fragment, is a new potential ovary cancer biomarker. *J Proteome Res*, 2007; 6: 4127–34
- Deininger SO, Ebert MP, Futterer A et al: MALDI Imaging Combined with Hierarchical Clustering as a New Tool for the Interpretation of Complex Human Cancers. *J Proteome Res*, 2008; 7(12): 5230–36
- Walch A, Rausser S, Deininger SO, Hofler H: MALDI imaging mass spectrometry for direct tissue analysis: a new frontier for molecular histology. *Histochem Cell Biol*, 2008; 130: 421–34
- Lemaire R, Wisztorski M, Desmons A et al: MALDI-MS direct tissue analysis of proteins: Improving signal sensitivity using organic treatments. *Anal Chem*, 2006; 78: 7145–53
- Franck J LR, Wisztorski M, Van Remoortere A et al: MALDI mass spectrometry imaging of proteins exceeding 30000 Da. *Med Sci Monit*, 2010; In press
- Holle A, Haase A, Kayser M, Hohendorf J: Optimizing UV laser focus profiles for improved MALDI performance. *J Mass Spectrom*, 2006; 41: 705–16
- Lemaire R, Lucot JP, Collinet P et al: New developments in direct analyses by MALDI mass spectrometry for study ovarian cancer. *Mol Cell Proteomics*, 2005; 4: S305–8
- Stauber J, Lemaire R, Wisztorski M et al: New developments in MALDI imaging mass spectrometry for pathological proteomic studies; Introduction to a novel concept, the specific MALDI imaging. *Mol Cell Proteomics*, 2006; 5: S247–S49
- Lemaire R, Stauber J, Wisztorski M et al: Tag-mass: specific molecular imaging of transcriptome and proteome by mass spectrometry based on photocleavable tag. *J Proteome Res*, 2007; 6: 2057–67
- Krug A, Towarowski A, Britsch S et al: Toll-like receptor expression reveals CpG DNA as a unique microbial stimulus for plasmacytoid dendritic cells which synergizes with CD40 ligand to induce high amounts of IL-12. *Eur J Immunol*, 2001; 31: 3026–37
- Franck J, Arafah K, Barnes A et al: Improving tissue preparation for matrix-assisted laser desorption ionization mass spectrometry imaging. Part I: using microspotting. *Anal Chem*, 2009; 81: 8193–202
- Lemaire R, Tabet JC, Ducoroy P et al: Solid ionic matrices for direct tissue analysis and MALDI imaging. *Anal Chem*, 2006; 78: 809–19
- Fournier I, Wisztorski M, Salzet M: Tissue Imaging Using MALDI Mass Spectrometry: The New Frontier of Histopathology Proteomics Expert Review Proteomics, 2008; 5(3): 413–24
- Franck J, Arafah K, Elayed M et al: MALDI imaging mass spectrometry: state of the art technology in clinical proteomics. *Mol Cell Proteomics*, 2009; 8: 2023–33
- Hibbs K, Skubitz KM, Pambuccian SE et al: Differential gene expression in ovarian carcinoma: identification of potential biomarkers. *Am J Pathol*, 2004; 165: 397–414
- Schweigert FJ: Characterisation of protein microheterogeneity and protein complexes using on-chip immunoaffinity purification-mass spectrometry. *Brief Funct Genomic Proteomic*, 2005; 4: 7–15
- Bengtsson S, Krogh M, Szgyarto CA et al: Large-scale proteomics analysis of human ovarian cancer for biomarkers. *J Proteome Res*, 2007; 6: 1440–50
- Gortzak-Uzan L, Ignatchenko A, Evangelou AI et al: A proteome resource of ovarian cancer ascites: integrated proteomic and bioinformatic analyses to identify putative biomarkers. *J Proteome Res*, 2008; 7: 339–51
- Makino E, Sakaguchi M, Iwatsuki K, Huh NH: Introduction of an N-terminal peptide of S100C/A11 into human cells induces apoptotic cell death. *J Mol Med*, 2004; 82: 612–20
- Sakaguchi M, Miyazaki M, Sonogawa H et al: PKCalpha mediates TGFbeta-induced growth inhibition of human keratinocytes via phosphorylation of S100C/A11. *J Cell Biol*, 2004; 164: 979–84
- Yang Z, Tao T, Raftery MJ et al: Proinflammatory properties of the human S100 protein S100A12. *J Leukoc Biol*, 2001; 69: 986–94
- Giuntoli RL II, Webb TJ, Zoso A et al: Ovarian cancer-associated ascites demonstrates altered immune environment: implications for anti-tumor immunity. *Anticancer Res*, 2009; 29: 2875–84
- Yang Y, Fruh K, Ahn K, Peterson PA: *In vivo* assembly of the proteasomal complexes, implications for antigen processing. *J Biol Chem*, 1995; 270: 27687–94
- Kloetzel PM: The proteasome system: a neglected tool for improvement of novel therapeutic strategies? *Gene Ther*, 1998; 5: 1297–98
- Rivett AJ, Gardner RC: Proteasome inhibitors: from *in vitro* uses to clinical trials. *J Pept Sci*, 2000; 6: 478–88
- Rotem-Yehudar R, Groettrup M, Soza A et al: LMP-associated proteolytic activities and TAP-dependent peptide transport for class I MHC molecules are suppressed in cell lines transformed by the highly oncogenic adenovirus 12. *J Exp Med*, 1996; 183: 499–514

46. Kuckelkorn U, Ruppert T, Strehl B et al: Link between organ-specific antigen processing by 20S proteasomes and CD8(+) T cell-mediated autoimmunity. *J Exp Med*, 2002; 195: 983–90
47. Regad T, Saib A, Lallemand-Breitenbach V et al: PML mediates the interferon-induced antiviral state against a complex retrovirus via its association with the viral transactivator. *EMBO J*, 2001; 20: 3495–505
48. Delp K, Momburg F, Hilmes C et al: Functional deficiencies of components of the MHC class I antigen pathway in human tumors of epithelial origin. *Bone Marrow Transplant*, 2000; 25(Suppl.2): S88–95
49. Pudney VA, Leese AM, Rickinson AB, Hislop AD: CD8+ immunodominance among Epstein-Barr virus lytic cycle antigens directly reflects the efficiency of antigen presentation in lytically infected cells. *J Exp Med*, 2005; 201: 349–60
50. Sorem J, Jardtzyk TS, Longnecker R: Cleavage and secretion of Epstein-Barr virus glycoprotein 42 promote membrane fusion with B lymphocytes. *J Virol*, 2009; 83: 6664–72
51. Sorem J, Longnecker R: Cleavage of Epstein-Barr virus glycoprotein B is required for full function in cell-cell fusion with both epithelial and B cells. *J Gen Virol*, 2009; 90: 591–95
52. Elg SA, Mayer AR, Carson LF et al: Alpha-1 acid glycoprotein is an immunosuppressive factor found in ascites from ovaria carcinoma. *Cancer*, 1997; 80: 1448–56
53. Nosov V, Su F, Amneus M et al: Validation of serum biomarkers for detection of early-stage ovarian cancer. *Am J Obstet Gynecol*, 2009; 200: 639 e1–5
54. Kim KD, Lim HY, Lee HG et al: Apolipoprotein A-I induces IL-10 and PGE2 production in human monocytes and inhibits dendritic cell differentiation and maturation. *Biochem Biophys Res Commun*, 2005; 338: 1126–36
55. Liang X, Lin T, Sun G et al: Hemopexin down-regulates LPS-induced proinflammatory cytokines from macrophages. *J Leukoc Biol*, 2009; 86: 229–35
56. Leygue E, Snell L, Dotzlaw H et al: Expression of lumican in human breast carcinoma. *Cancer Res*, 1998; 58: 1348–52
57. Leygue E, Snell L, Dotzlaw H et al: Lumican and decorin are differentially expressed in human breast carcinoma. *J Pathol*, 2000; 192: 313–20
58. Babelova A, Moreth K, Tsalstra-Greul W et al: Biglycan: A danger signal that activates the NLRP3 inflammasome via toll-like and P2X receptors. *J Biol Chem*, 2009; 284(36): 24035–48
59. Schaefer L, Babelova A, Kiss E et al: The matrix component biglycan is proinflammatory and signals through Toll-like receptors 4 and 2 in macrophages. *J Clin Invest*, 2005; 115: 2223–33
60. Salzet M, Capron A, Stefano GB: Molecular crosstalk in host-parasite relationships: schistosome- and leech-host interactions. *Parasitol Today*, 2000; 16: 536–40
61. Kitagawa K, Murata A, Matsuura N et al: Epithelial-mesenchymal transformation of a newly established cell line from ovarian adenocarcinoma by transforming growth factor-beta1. *Int J Cancer*, 1996; 66: 91–97
62. Keshamouni VG, Michailidis G, Grasso CS et al: Differential protein expression profiling by iTRAQ-2DLC-MS/MS of lung cancer cells undergoing epithelial-mesenchymal transition reveals a migratory/invasive phenotype. *J Proteome Res*, 2006; 5: 1143–54
63. Vergara D, Merlot B, Lucot JP et al: Epithelial-mesenchymal transition in ovarian cancer. *Cancer Lett*, 2010; 291: 59–66
64. Woo MM, Alkushi A, Verhage HG et al: Gain of OGP, an estrogen-regulated oviduct-specific glycoprotein, is associated with the development of endometrial hyperplasia and endometrial cancer. *Clin Cancer Res*, 2004; 10: 7958–64
65. Woo MM, Gilks CB, Verhage HG et al: Oviductal glycoprotein, a new differentiation-based indicator present in early ovarian epithelial neoplasia and cortical inclusion cysts. *Gynecol Oncol*, 2004; 93: 315–19
66. Wahl A, McCoy W, Schafer F et al: T Cell Tolerance for Variability in a Class I HLA Presented Influenza A Virus Epitope. *J Virol*, 2009; 83(18): 9206–14
67. Begum FD, Hogdall CK, Kjaer SK et al: The prognostic value of plasma soluble urokinase plasminogen activator receptor (suPAR) levels in stage III ovarian cancer patients. *Anticancer Res*, 2004; 24: 1981–85
68. Deng X, Hogdall EV, Hogdall CK et al: The prognostic value of pretherapeutic tetranectin and CA-125 in patients with relapse of ovarian cancer. *Gynecol Oncol*, 2000; 79: 416–19
69. Hogdall EV, Hogdall CK, Tingulstad S et al: Predictive values of serum tumour markers tetranectin, OVX1, CASA and CA125 in patients with a pelvic mass. *Int J Cancer*, 2000; 89: 519–23
70. Lundstrom MS, Hogdall CK, Nielsen AL, Nyholm HC: Serum tetranectin and CA125 in endometrial adenocarcinoma. *Anticancer Res*, 2000; 20: 3903–6
71. Lim R, Ahmed N, Borregaard N et al: Neutrophil gelatinase-associated lipocalin (NGAL) an early-screening biomarker for ovarian cancer: NGAL is associated with epidermal growth factor-induced epithelial-mesenchymal transition. *Int J Cancer*, 2007; 120: 2426–34
72. Gericke B, Raila J, Sehoul J et al: Microheterogeneity of transthyretin in serum and ascitic fluid of ovarian cancer patients. *BMC Cancer*, 2005; 5: 133
73. Rauvala M, Puustola U, Turpeenniemi-Hujanen T: Gelatinases and their tissue inhibitors in ovarian tumors; TIMP-1 is a predictive as well as a prognostic factor. *Gynecol Oncol*, 2005; 99: 656–63
74. Diamandis EP, Borgono CA, Scorilas A et al: Immunofluorometric quantification of human kallikrein 5 expression in ovarian cancer cytosols and its association with unfavorable patient prognosis. *Tumour Biol*, 2003; 24: 299–309
75. Pedersen N, Schmitt M, Ronne E et al: A ligand-free, soluble urokinase receptor is present in the ascitic fluid from patients with ovarian cancer. *J Clin Invest*, 1993; 92: 2160–67
76. Sier CF, Stephens R, Bizik J et al: The level of urokinase-type plasminogen activator receptor is increased in serum of ovarian cancer patients. *Cancer Res*, 1998; 58: 1843–49

## Discussion et perspectives.

Dans la partie précédente, nous avons vu l'apport de la classification hiérarchique permettant d'aboutir à l'observation de groupes de spectres ayant un profil commun et spécifique à une zone histologique. L'utilisation de cette méthode juste après l'acquisition des données, permet de conclure ou non en la présence de m/z spécifiques à une zone d'intérêt et de choisir s'il est nécessaire de poursuivre l'analyse multivariée de type PCA pour mettre en évidence les ions qui peuvent représenter des biomarqueurs potentiels. Une fois l'analyse PCA effectuée, les loading-plots permettent de pointer ces ions d'intérêt. Dans ce chapitre, nous nous axerons sur le cancer de l'ovaire et utiliserons l'analyse multivariée sur des biopsies, comme celle présentée en exemple dans les **figures 44A et 45A**, présentant une zone dite saine et un carcinome.

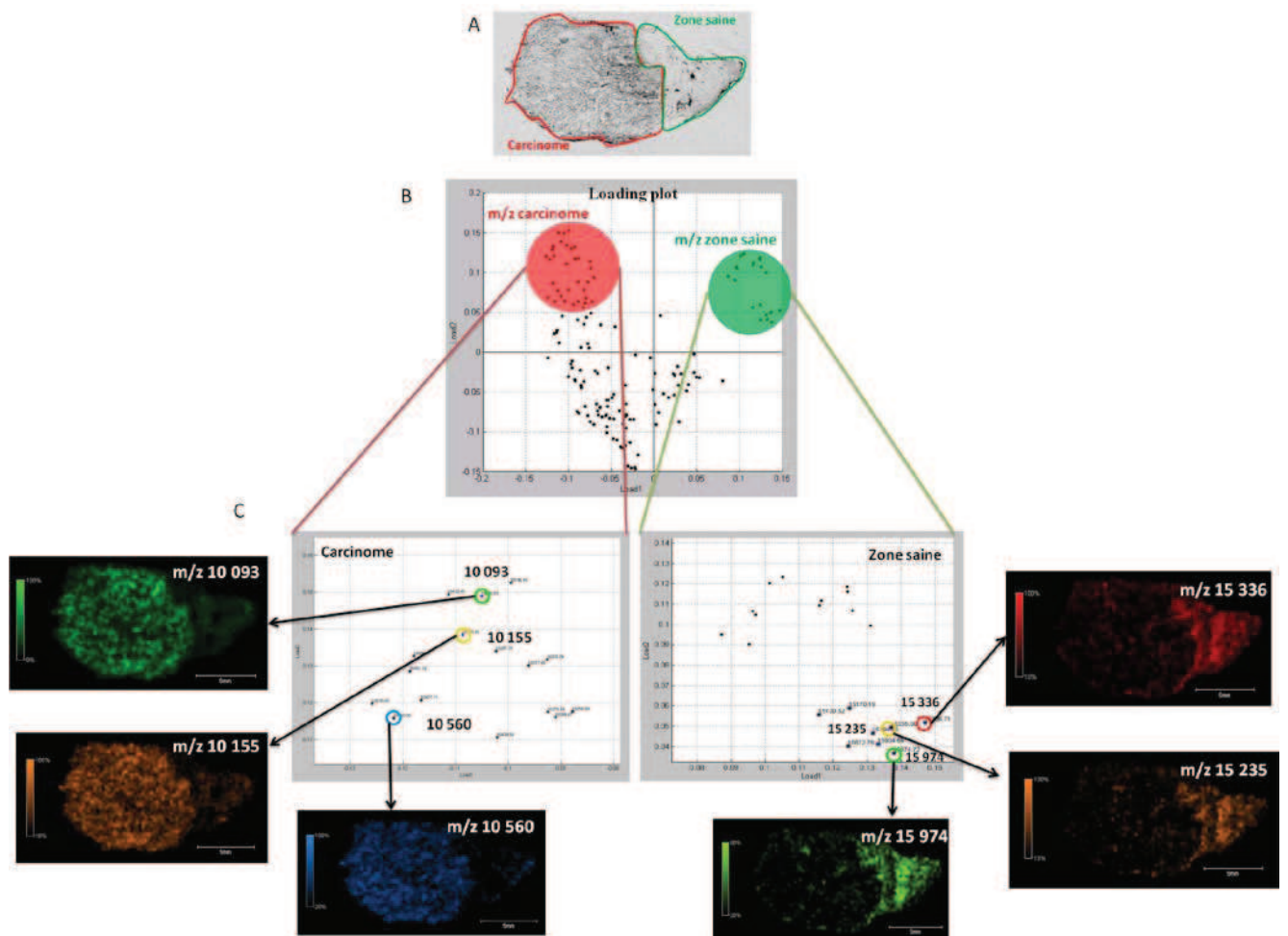
La classification des spectres par une méthode de calcul des distances de type euclidien et des liaisons de type ward a permis d'obtenir un dendrogramme (**Figure 44B**) dont les images des groupes de spectres des deux premiers embranchements coïncident parfaitement avec le carcinome (rouge) et la zone dite saine (vert) (**Figure 44C**).



**Figure 44 : Analyse par classification hiérarchique d'une coupe de cancer d'ovaire (A). (B) dendrogramme obtenu après classification des spectres. Les deux premiers embranchements sont indiqués en rouge et en vert. (C) images des localisations des deux groupes de spectres obtenus.**



Ainsi chaque région est classifiée et séparée en raison des  $m/z$  et de leur abondance relative spécifique. L'intérêt est donc de pouvoir détecter et lister ces ions. C'est pourquoi la totalité des spectres obtenus a été analysée par PCA en mode non-supervisé (**figure 45B**). L'analyse du loading plot révèle les  $m/z$  spécifiques à chaque zone histologique dont certains sont donnés en exemple (**figure 45C**).



**Figure 45 : Utilisation de la PCA pour la mise en évidence de  $m/z$  d'intérêt. (A) coupe histologique de cancer d'ovaire présentant un carcinome et une zone saine analysée en imagerie MALDI. (B) loading plot obtenu par PCA à partir des données d'imagerie MALDI. Les  $m/z$  se séparent en deux groupes distincts. (C) ces deux groupes distincts correspondent aux  $m/z$  spécifiques aux deux régions du tissu comme l'illustrent les images d'ions données en exemple.**

Les ions spécifiques de la région cancéreuse détectée par PCA ont ensuite été caractérisés par plusieurs stratégies décrites dans l'article [95]. Ce qui est notoire dans cette étude est la caractérisation de plusieurs populations de marqueurs protéiques liés à la régulation de la réponse immunitaire, à la dynamique et au remodelage du cytosquelette ainsi qu'à la progression tumorale. Parmi ces marqueurs, 2 sont plus remarquables i.e. le fragment reg alpha et la mucine 9 qui n'avaient pas été découverts auparavant.

En effet, parmi les ions propres à la zone cancéreuse retrouvés par PCA, le m/z 9744 correspond au fragment C-terminal de l'activateur de l'immunoprotéasome 11S, PA28 REG- $\alpha$  (PSME1 : proteasome activator complex subunit 1) qui a préalablement été identifié par Lemaire et al. [138]. Cette protéine joue le rôle d'activateur du complexe protéasome.

Le protéasome est un complexe de 700 kDa responsable de la dégradation naturelle des protéines intracellulaires chez les eucaryotes, régulant ainsi la présence des protéines synthétisées à partir du génome. Comme détaillé dans la **figure 46**, deux voies de dégradation sont envisageables : dans le premier système, ATP dépendant, les protéines à dégrader sont marquées par de l'ubiquitine et régulées par une protéine appelée PA700 (ou 19S). Dans le deuxième cas, les protéines ne sont pas ubiquitinées, et les complexes activateurs sont les protéines REG (PA28 ou 11S) composées de différentes sous-unités (REG- $\alpha$ , REG- $\beta$ , REG- $\gamma$ ). La dégradation protéasomale est un élément essentiel de nombreux processus cellulaires, notamment le cycle cellulaire, l'expression génétique, la réponse au stress oxydatif et l'apoptose.

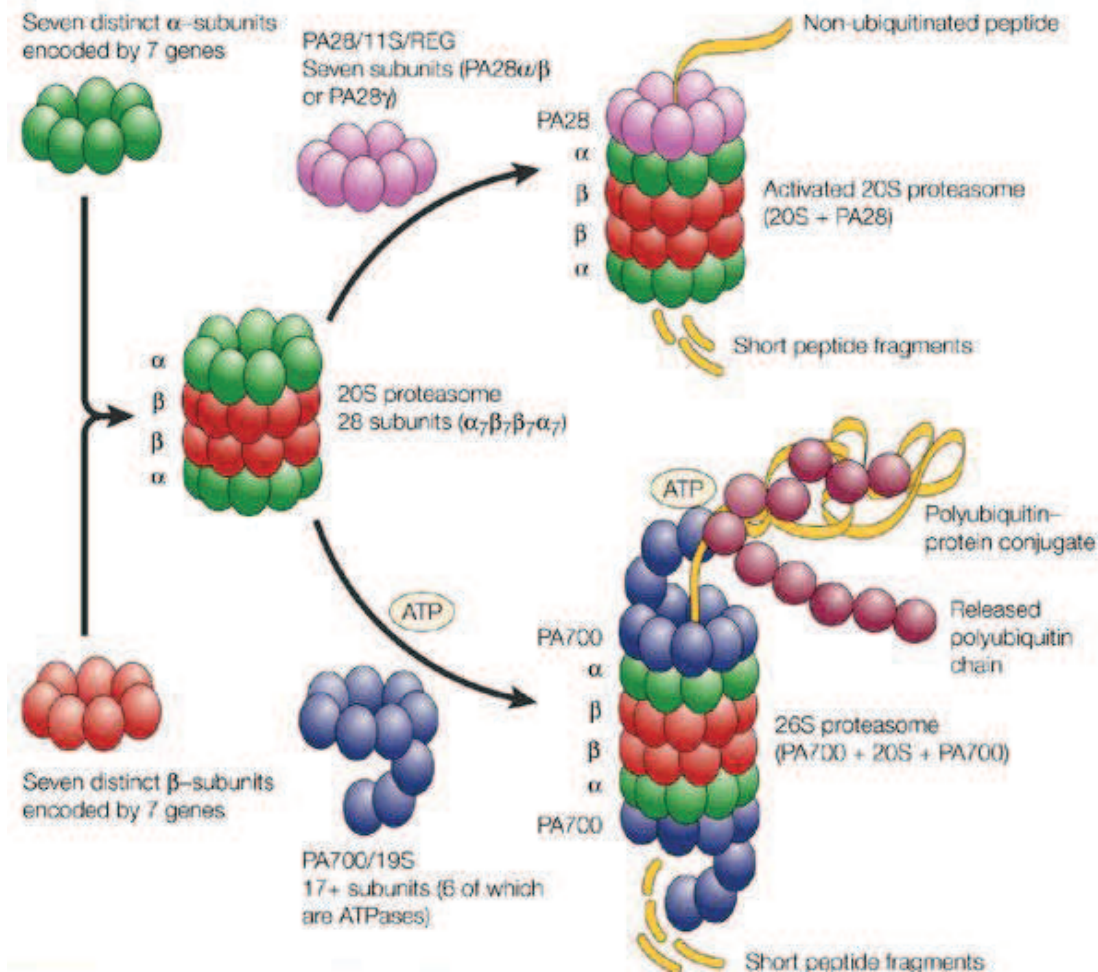


Figure 46 : Composition et organisation du protéasome avec ses différents activateurs intracellulaires.

[270]

Le complexe 11S (PA28 ou REG) heptamérique dit « activateur » peut s'associer au 20S. Il stimule l'activité peptidase du protéasome mais est incapable de reconnaître l'ubiquitine et ne possède pas d'activité ATPase. Ce complexe se lie au 20S à partir de son C-ter, ouvrant ainsi l'accès au cœur du protéasome et favorisant une modification de la conformation des anneaux  $\alpha$ . L'expression du 11S est induite par l'interf  $\gamma$  et, en association avec les sous-unités  $\beta$  de l'immunoprotéasome, provoque la génération de peptides qui iront se lier au complexe majeur d'histocompatibilité (CMH) de classe I.

Hors, nous avons démontré que REG- $\alpha$  était dégradé dans le cadre du cancer de l'ovaire [138]. Cette dégradation, dont les effecteurs ne sont pas encore définis, pourrait être une capacité acquise pour échapper au système immunitaire en entraînant une dérégulation du processus de présentation des antigènes. Les derniers développements sur ce fragment réalisés au laboratoire ont porté sur la recherche de celui-ci dans les fluides biologiques par le Dr. C. Castellier (Interne Médecine) et Mr. M. El Ayed (doctorant). Leurs résultats de co-immunoprécipitation avec les anticorps dirigés contre les parties N et C terminales montrent l'absence de la protéine entière dans le liquide d'ascite chez les patients cancéreux *versus* sa présence dans les cas bénins. En ce qui concerne le fragment avec un m/z de 9744, celui-ci a disparu chez les patients cancéreux, certainement suite à son clivage en un fragment plus petit. Par contre, un fragment plus grand de 14 kDa a été détecté chez les patients positifs lié au clivage de la protéine entière PA28. La poursuite de ce travail sera la recherche de ce fragment dans le sérum en lien avec la recherche de marqueurs pathogéniques de type protéines oncovirales qui ont également été détectées lors de nos études sur le tissu. Nos premières investigations révèlent une possible étiologie virale de type EBV et HHV6. Il nous faut donc associer clivage de l'immunoprotéasome et cause virale de la pathologie. C'est dans cette optique qu'il faudra donc rechercher dans les fluides ces marqueurs liés aux virus et à l'inhibition de la réponse immunitaire. De récentes études ont confirmé l'association de EBV et des virus de types Herpes dans des pathologies de type cancer et l'inhibition de la réponse immunitaire par attaque au niveau du complexe ubiquitine et protéasome [271].

Le second marqueur spécifique identifié est la mucine 9. Les mucines peuvent être des candidats intéressants pour la détection de cancer de l'ovaire stade précoce [272, 273]. Les mucines sont de grandes protéines extracellulaires fortement glycosylées avec des oligosaccharides et sont généralement connus pour fournir une protection aux tissus épithéliaux dans les conditions physiologiques normales [274]. Les mucines sont habituellement sécrétées par les tissus épithéliaux des voies aériennes, de l'estomac, de l'intestin et du carcinome canalaire du foie, du pancréas, vésicule biliaire, les glandes

salivaires, la glande lacrymale, etc. Dans ces tissus, les cellules épithéliales sont exposées à une variété de micro-organismes, toxines, protéases, lipases, glycosidases et des fluctuations micro-environnementales diversifiées tel que le pH, la force ionique, l'oxygénation [274]. Toutes les mucines partagent des caractéristiques générales. Par exemple, elles ont des domaines peptidiques répétitifs riches en sérine, en thréonine et en proline. Les sérines et les thréonines sont des sites pour O-et N-glycosylation. La présence de la répétition en tandem des domaines varie en nombre, longueur et en O-glycosylation, ce qui est la caractéristique commune des structures de toutes les mucines [275]. Leur structure générale et la composition biochimique fournit une protection pour la surface des cellules. De plus, les mucines communiquent les informations de l'environnement externe aux cellules épithéliales via la signalisation cellulaire [275]. Il semble donc que celles-ci aient la capacité de servir de récepteurs de surface cellulaire et de capteurs de signaux en réponse à des stimuli externes pour une variété de réponses cellulaires : prolifération et croissance cellulaires, différenciation et apoptose. Ces données suggèrent donc que l'expression aberrante des mucines pourrait être impliquée dans le développement et la progression du cancer de l'ovaire.

Le **tableau 17** décrit les différentes études sur les mucines réalisées au niveau de l'ovaire et montre clairement l'absence de la détection de la mucine 9 par les techniques classiques alors que par MSI couplée à la MS, nous avons pu détecter cette mucine 9. Or des études récentes par immuno-histochimie et ELISA ont montré que MUC9 est détectée au niveau sérique [276]. Sur 135 sérums normaux, 21 de tumeurs bénignes, 12 de lésions frontières, et 87 cancers de l'ovaire (18, degré 1-2 séreuse, 44, de grade 3 séreuse; 10, mucineux; 10, à cellules claires, et 5, endométrioïde), les résultats ont montré des concentrations élevées de MU9 dans le sérum de 75% des femmes avec des tumeurs borderline et 80% des femmes avec mucineux, 60% des cellules claires, 59 % avec grade 1 et 2 séreuse, 22% de grade 3 séreuse, et 0% des carcinomes endométrioïde [276].

Gene	Normal Ovary	Borderline (Mucinous)	Low Stage (Stage 1-2)	High Stage (Stage 3-4)	Detection method
MUC1	+/-	++	+ to +++ (in all histological types i.e. C, M, E, S)	+ to +++ (in all histological types i.e. C, M, E, S)	ISH, NB, IHC
MUC2	ND	+++	+++ (all histological types, primarily in mucinous type)	+ to ++	ISH, NB, IHC
MUC3	ND	+++ (primarily in intestinal phenotype)	+++ (E, M)	- to +	ISH, NB
MUC4	-	+++ (primarily in endocervical phenotype)	+++ (all types i.e. C, M, E, S)	- to ++	ISH, NB, IHC
MUC5AC	ND	++ (primarily in gastric surface cell or mucinous type)	++ (E, M, S)	++	ISH, NB,
MUC5B	ND	++ (Express primarily in endocervical phenotype)	++ (C, S)	- to +	ISH, NB
MUC13	ND	+	+++ (S, M)	++ (S, M)	OMA, TMA, IHC
CA125/MUC16	-	- (express in non-mucinous borderline tumors)	- to +++ (rarely express in mucinous tumors)	+ to +++ (rarely express in mucinous tumors)	IHC
MUC17	-	+	-	-	

**Tableau 17 : comparaison des profils d'expression des mucines dans différents stades et types histologiques du cancer de l'ovaire. C, cellules claires ; M, mucineux ; E, endometroïde ; S, séreux ; ISH, hybridation in-situ ; NB, northern blot ; IHC, immunohistochimie ; TMA, tissue microarray ; OMA, oligonucleotide microarray. [277]**

## Conclusion

Comme nous venons de le voir, suite à l'utilisation de l'imagerie MALDI couplée à l'analyse multivariée, il nous a été possible de détecter des molécules d'intérêt et d'aboutir à leur identification. Il est donc possible de mettre en évidence des marqueurs spécifiques du stade, du type de cancer afin de réaliser un diagnostic précis. Parmi l'ensemble des protéines identifiées, nous nous sommes plus focalisés sur le fragment de l'activateur de l'immunoprotéasome et la mucine 9 car ceux-ci semblent spécifiques de la pathologie de l'ovaire.

Dans cette étude, il nous a été possible de mettre en avant certaines protéines marqueurs nous permettant de mieux comprendre les mécanismes biologiques altérés qui s'opèrent lors du processus de cancérisation dans le cadre du cancer de l'ovaire, comme l'échappement à la réponse immunitaire de l'individu. Dans ce contexte, nous avons également mis en évidence, lié à cette modulation de la réponse immunitaire, la surexpression des membres de la famille des proprotéines convertases au sein des cellules cancéreuses (voir **chapitre 4**). Des travaux au sein de l'unité semble montrer l'importance des PCs dans la régulation de la réponse inflammatoire [278] mais aussi dans la virulence et les différents mécanismes infectieux utilisés par certains virus comme EBV et HV [279]. Il serait donc intéressant d'évaluer l'impact de knock-down ou knock-out sur celles-ci dans des lignées de cancer d'ovaire (SKOV-3 par exemple). En effet, il est aujourd'hui établi que les PCs jouent un rôle important dans la virulence et les différents mécanismes infectieux utilisés par les virus [280-283]. On peut donc se demander si elles ne sont pas également impliquées dans ces mécanismes immunosuppresseurs. Par la même occasion et suite aux observations faites sur les protéines de l'ECM dans le cancer de l'ovaire, ces lignées knock-down pourraient servir pour des études d'adhésion et d'invasion pour juger si ces dérèglements ne seraient pas également dus à l'implication des PCs.

Enfin, une étude plus approfondie des différents complexes du protéasome dans ce type de cancer permettrait de vérifier si d'autres sous-unités ne sont pas également modifiées. L'action d'inhibiteurs ou/et d'activateur du protéasome permettrait également de comparer le phénotype cancéreux de lignées, de xénogreffes ou de tissus animaux cancéreux.

Ces résultats pointent donc l'importance des proprotéines convertases dans les processus de cancérisation, ce que nous avons pu montrer dans le **chapitre 4** au niveau du cancer de la prostate.

## Chapitre 4 : Implications de la proprotéine convertase PACE4 dans le cancer de la prostate.

---

### Introduction

Le cancer de la prostate est actuellement un problème majeur de santé publique dans les pays développés. Il s'agit du premier cancer de l'homme et du deuxième cancer responsable de décès chez l'homme en France. Aux États-Unis, avec 198 000 nouveaux cas par an (soit 31 % des cancers de l'homme), il est le premier cancer en incidence (hommes et femmes confondus) et le deuxième cancer responsable de décès chez l'homme [284]. En 2000, son incidence était de 75 pour 100 000 en France, avec, chaque année, 40 000 nouveaux cas environ et 10 000 décès. Son incidence continue d'augmenter en France et cela est en partie dû au vieillissement de la population et à la pratique du dépistage par le *prostate specific antigen* (PSA) sérique. Après une longue période durant laquelle le traitement du cancer de la prostate était mal codifié, les dernières années ont apporté des résultats d'études comparatives permettant de mieux rationaliser la prise en charge des patients. C'est en particulier le cas avec la démonstration d'un bénéfice de survie obtenu par les associations hormonothérapie-radiothérapie dans les formes localisées à risque [285, 286], par l'hormonothérapie immédiate [287], pour les formes métastatiques [288], par la chimiothérapie pour les formes hormono-réfractaires [289]. À côté des thérapeutiques classiques, des efforts de recherche intensifs portent sur la compréhension des mécanismes biologiques du cancer de la prostate, avec l'espoir de mettre au point de nouvelles thérapeutiques.

Comme tous les cancers, il est possible de classer les tumeurs suivant leurs stades histologiques. La classification TNM (Tumors, Nodes, Metastasis) des cancers définie par l'OMS s'applique également au cancer de la prostate :

- T1 et T2 : cancers localisés à la prostate et ne s'étendant pas au-delà de la capsule (absence d'atteinte des ganglions, absence de métastase, c'est-à-dire N0M0).
- T1 est un stade dont la découverte est uniquement histologique (non visible à l'imagerie et non palpable).
  - T1a : moins de 5% de cellules cancéreuses sur les prélèvements

- T1b : plus de 5 %
- T2 : cancers palpables au toucher rectal.
  - T2a : cancers occupant moins de 50 % d'un des deux lobes prostatiques
  - T2b : plus de 50 %
  - T2c : atteinte des deux lobes.
- T3 : extension du cancer aux tissus périphériques (franchissement de la capsule prostatique).
- T4 : extension aux organes adjacents : vessie, rectum, paroi pelvienne).
- N0 ou N1 : absence ou présence d'atteinte(s) ganglionnaire(s).
- M0 ou M1 : absence ou présence de métastase(s)

Les trois lettres TNM symbolisent la propagation du cancer sur le site de la tumeur primitive (T), dans les ganglions lymphatiques voisins (N) et à distance pour d'éventuelles métastases (M).

La mise en évidence par Wang et al., en 1979 [290], d'une enzyme prostatique exclusivement produite par la glande, l'antigène prostatique spécifique (PSA), a entraîné une nouvelle approche diagnostique et thérapeutique du cancer de la prostate. Le PSA est une glycoprotéine de 34 KDa, composée de 240 acides aminés. Le gène codant la molécule a été localisé sur le chromosome 19 et séquencé complètement [291].

Le PSA présente des homologies avec les protéases de la famille des kallikréines et jouerait un rôle dans l'hydrolyse du sperme et la fertilité masculine [292]. La molécule est produite par les cellules épithéliales de la prostate et est considérée comme spécifique de cet organe. Cependant, la mise en évidence de PSA au niveau des urines de patients ayant subi une ablation complète de la glande prostatique laisse suspecter une production par les glandes périurétrales [293] rendant ce marqueur assez controversé. De plus, certains facteurs peuvent influencer le taux sérique du PSA chez individu comme les manipulations endorectales, certains traitements médicaux ou encore dans certains cas d'infection ou de rétention urinaire.

Néanmoins aujourd'hui la majorité des méthodes de détection du cancer de la prostate repose sur le dosage du PSA dans le sang, les techniques d'imagerie médicale ou le touché rectal. Plusieurs études ont montré qu'un nombre important d'hommes



touchés par le cancer de la prostate avaient un taux de PSA inférieur au seuil considéré comme anormal de 4.0ng/ml dans le sang [294, 295]. Suite à ces découvertes, l'idée d'abaisser le seuil critique de PSA à 2.5ng/ml pour le diagnostic du cancer de la prostate a entraîné un important débat dans le monde médical. Néanmoins, l'abaissement de ce seuil imposerait à de nombreux hommes une biopsie de la prostate qui n'est pas sans risque. La fragilité de ce marqueur explique pourquoi de nombreux groupes de recherche continuent à étudier cette pathologie dont l'incidence ne cesse d'augmenter en raison du vieillissement de la population mais dont la thérapie est de plus en plus efficace. Ainsi il reste très important de poursuivre les recherches de biomarqueurs potentiels de cette pathologie et de compréhension des mécanismes biologiques impliqués.

Jusqu'à ce jour, une seule étude a été menée sur l'implication de l'activité des proprotéines convertases dans des lignées de cellules de cancer de la prostate [296]. Cette équipe a déterminé que les PCs, comme la furin, étaient capable de moduler la maturation du prostate derived factor (PDF) dans les cellules de cancer de prostate humain. Le PDF est un facteur de croissance (membre de la superfamille des TGF- $\beta$ ) fortement exprimé dans la prostate, où il joue des rôles importants dans la différenciation prostatique et la régulation de la croissance des cellules cancéreuses. Suite au traitement de lignées modèles du cancer de la prostate (PC3, DU145 et LNCaP) par un inhibiteur des PCs appelé dec-RVKR-cmk, les auteurs ont remarqué que les LNCaP n'arrivent plus à produire le PDF mature et retrouvent un phénotype similaire aux cellules épithéliales basales de la prostate. Néanmoins, les lignées PC3 et DU145 n'expriment pas de PDF et la seule différence par rapport aux LNCaP est leur phénotype androgène-indépendant. C'est pourquoi, le Dr François D'Anjou a initié et géré le projet sur l'étude de l'impact des proprotéines convertases sur des lignées androgène-indépendantes dans le cadre du cancer de la prostate.

## Résultats : Publication 3.

### **ROLE OF PACE4 IN PROSTATE CANCER: A CELLULAR STUDY USING HUMAN DU145 CELL LINE**

François D'Anjou<sup>1</sup>, David Bonnel<sup>1,2</sup>, Sophie Routhier<sup>1</sup>, Isabelle Brochu<sup>1</sup>, Alain Latil<sup>3,4</sup>, Isabelle Fournier<sup>2</sup>, Michel Salzet<sup>2</sup>, Jean-Pierre Perreault<sup>5</sup> and Robert Day<sup>1,6</sup>

Départements de <sup>1</sup>Pharmacologie et de <sup>5</sup>Biochimie, Université de Sherbrooke, Sherbrooke, Québec, J1H 5N4, Canada. <sup>2</sup>Université Nord de France, CNRS, Laboratoire de Neuroimmunologie et Neurochimie Evolutives, FRE 3249, Université Lille 1, 59650 Villeneuve d'Ascq, France. <sup>3</sup>Urogene, Génopole, Evry Cedex, France. <sup>4</sup>*Present adress*: Laboratoire de microbiologie, Institut de Recherche Pierre Fabre, Toulouse, France.

Running title: PACE4 in prostate cancer

Key words: proprotein convertases, prostate cancer, proliferation, ribozyme, microarray, proteomic, apoptosis

## **Abstract**

The involvement of proprotein convertases in tumor cell proliferation and metastasis has been suggested for various cancers. In this study, We demonstrate the specific contribution of PACE4 for the proliferation of the human prostate cancer cell line DU145. Based on quantitative PCR analysis in human prostate tumor tissues from different clinical stages, we decided to target the highly expressed PACE4 for subsequent cellular studies. The functional activity was reduced by establishing a stable transfected cell line expressing a specifically designed SOFA-HDV ribozyme targeting PACE4 mRNA. Northern blots confirmed the specific reduction of PACE4 mRNA levels into selected clonal cell lines. The slower cell proliferation was assessed by cell counts and colony formation assays. Subcutaneous inoculation of DU145 cells into nude mice confirmed the major contribution of PACE4 for tumoral growth. Gene expression profiling confirmed the importance of PACE4 activity to support the high tumoral aggressiveness of DU145 cells. Thus, these results emphasize the specific contribution of PACE4 for the proliferation of prostate cancer cells DU145, confirming that major proteolytic events carried out by this proprotein convertase are essential to sustain this pathological state.

## INTRODUCTION

Cancer cells are characterized by multiple genetic alterations that confer physiological changes, leading to uncontrolled division and ability to invade other tissues [1]. These acquired capabilities, namely cell proliferation, tissue invasion and adhesion are essential for malignant growth and progression. Numerous studies have associated the family of enzymes known as the proprotein convertases (PCs) to cancer [2-8]. PCs are serine proteases that cleave substrates at R-X-K/R-R↓ motif [9, 10]. These processing events, resulting in the activation of protein precursors, occur at multiple levels of cell secretory pathways, and even at the cell surface. In mammalian cells, nine members of this family have been identified: furin, PACE4, PC1/3, PC2, PC4, PC5/6, PC7, PCSK9 and subtilisin kexin isoenzyme 1. Those genes have shown differential expression levels within tissues, ranging from ubiquitous (e.g. furin) to an endocrine restricted expression (e.g. PC1/3 and PC2).

The first association of PCs with cancer was done by comparative studies of normal and cancerous cells showing higher expression of PCs in small cell lung cancer [11], non-small cell lung carcinoma [12], breast [13], colon [14], and head and neck [15] tumor cells, or lower expression in ovarian cancer [16]. A correlation between expression of some PCs, namely furin and PACE4, and tumor cell aggressiveness has been established for different cell types [8, 15, 17-20]. Thus, it is hypothesized that aberrant processing events provide cancer cells a higher capacity: (i) to remodel the extracellular matrix (e.g. MT1-MMP and MT2-MMP; [17, 21, 22]); (ii) to interact with their host micro-environment to favor tumor cell adhesion (e.g. IL-1 $\alpha$ , TNF- $\alpha$  and E-selectin [7]); and (iii) to modulate their proliferation and differentiation (e.g.

IGF-1 receptor [7], VEGF-C,-D [23, 24], PDGF-A, -B [24, 25], TGF- $\beta$  [26]). One possible mechanism underlying these observations could be on the basis of the precursors activation by overexpressed PCs. Alternatively, PCs overexpression is required to sustain these pathophysiological functions to maintain cancer cells immortality. Nevertheless, these studies suggest that PCs processing events are directly linked to tumor development by allowing cancer cells to maintain some of their acquired capabilities.

The situation becomes more complex as different PCs can be used in different cancer cells. If we are to understand the specific contribution of each PC in tumorigenesis, the necessity for potent, specific and cell effective inhibitors, either pharmacologic or molecular, for each member of this enzyme family is crucial. Until now, these pharmacological tools are limited and may lack specificity for single PCs [9, 27-30]. However, molecular approaches targeting specific PC mRNAs are closer to achieve this goal, and would be relevant methods for the evaluation of the potential implication of each enzyme in tumoral growth. There are many sequence-specific approaches taking advantage of either endogenous ribonucleases (e.g. small interfering RNA (siRNA) and phosphorothioate antisense oligonucleotide) or their intrinsic catalytic activity (e.g. deoxyribozyme and ribozyme) to drive the degradation of the targeted RNA [31].

In this study, we firstly evaluated the implication of each PC on different types of prostate cancer biopsies. We detected the unique PACE4 overexpression in human androgen-independent prostate cancer tissues. Secondly, we created a DU145 cell line model, using the SOFA-HDVRz as a RNA silencing tool, to

investigate PACE4 contribution in androgen-independent prostate cancer cells. By combining, gene expression profiling using DNA microarrays and *in vivo* experiments, our results demonstrate the important role of PACE4 for tumoral progression in prostate cancer. Moreover, a proteomic approach led us to the identification of a proapoptotic factor, namely TRPS1, whose detection is linked to lower PACE4 levels.

## EXPERIMENTAL PROCEDURES

**Patients and samples.** Prostate tumor samples were obtained from patients either St Louis and Bichat Hospital (Paris, France), or Tournan's clinic (Tournan en Brie, France) as previously described [32]. Each patient included has signed the Sample Informed Consent. Samples tissues from the 34 patients with clinically localized prostate tumors were obtained by removing clinically localized tumors by radical prostatectomy. The surgical specimens were first sliced thickly, and samples were then cut from suspect areas. Part of the selected tissue was immediately placed in liquid nitrogen for RNA extraction, while adjacent sections were stained with H/E (hematoxylin and eosin) for histopathological examination. The sample tissues from hormone-refractory recurrent prostate carcinoma were obtained from patients with metastatic disease at diagnosis. Since these patients were not amenable to radical surgery, they received endocrine therapy, either by classical androgen deprivation (orchidectomy or luteinizing-hormone-releasing hormone (LHRH) agonist administration); or, by maximal androgen blockade (castration combined with antiandrogen therapy). These patients relapsed, and their tumors became clinically androgen-independent. Thirteen samples of these

hormone-refractory recurrent prostate carcinomas were obtained during transurethral resection.

**Selected tissues.** Suspect areas were examined histopathologically in the surgery suite, and a thick shave section was taken for research purposes. This pre-selected tumor specimen section was then sliced on each side and again subjected to pathological examination. Samples were considered suitable for molecular studies when all epithelial cells were neoplastic. Confirmed malignant areas were carefully dissected using a scalpel. This process yields a homogeneous cell population and thereby avoids dilution of tumor-specific genetic changes by nucleic acids from normal and reactive cells present in the same specimen. The histological diagnosis and clinical stage based on the TNM system were determined during a routine work-up after surgery [33, 34]. Thirteen tumors were hormone-refractory, 18 of the clinically localized tumors were at stage pT2, and 16 at stage pT3. Nine well-characterized matched normal prostate specimens from the thirty-four patients with clinically localized prostate who underwent radical prostatectomy were used to assess basal target-gene mRNA expression. Normal-looking areas of each surgical specimen were examined histologically for the absence of cancer cells and selected upon its microscopic pathological criteria to avoid including areas with benign hyperplasia.

### Real-time PCR on tissues

(i) *RNA extraction and cDNA synthesis.* Total RNA was extracted from tissue specimens by using the acid-phenol guanidium method, as described earlier [32]. The quality of RNA samples was determined by electrophoresis through agarose gels, staining with ethidium bromide, and visualization of the 18S and 28S RNA

bands under ultraviolet light. RNA was reverse-transcribed as previously described [32].

(ii) *Real time PCR amplification.* All PCR reactions were performed using an ABI Prism 7900HT Sequence Detection System (Applied Biosystems) and the SYBR<sup>®</sup> Green PCR Core Reagents kit (Perkin-Elmer Applied Biosystems), as previously described [32]. Briefly, the thermal cycling conditions comprised an initial denaturation step at 95°C for 10 min and 45 cycles at 95°C for 15 s and 65°C for 1 min. A genomic DNA and non-template control was included in each experiment. Samples and controls were tested in duplicate. Primers were chosen with the assistance of the computer programs Oligo 4.0 (National Biosciences, Plymouth, MN) and Primer Express (Perkin-Elmer Applied Biosystems), as previously described [32]. Primer sequences for endogenous control genes PPIA (the peptidyl prolyl isomerase A gene encoding cyclophilin A) were described earlier [32]. The PACE4 primer sequences are: sense, 5'-CAAGAGACCCAGGAGCATCCC-3' and, antisense, 5'-ACCCGCTGGTCCGAGTGCT-3'. The threshold cycle (Ct) numbers obtained from PCR amplifications were expressed as fold change in target gene expression relative to PPIA expression, as described earlier [32].

***In situ hybridization.*** The *in situ* hybridization protocol using digoxigenin-labeled cRNA probes have been described in details previously [35].

***SOFA-HDVRz design, cloning and in vitro cleavage assays.***

(i) *Bioinformatic analysis.* Human PACE4 cDNA sequence (M80482) was used for the determination of the optimal SOFA-HDVRz targeting site. The selection

criteria were: (i) the first nucleotide (nt) should be a G, in order to form the required wobble bp within the P1 stem (**figure 2A**);

(ii) after the 6 subsequent nt of the P1 stem, a 5 nt spacer was added; and (iii) a 12 nt biosensor stem was used to complete the design of our SOFA-HDVRz. In summary, the targeted site on cDNA was: 5'(GN<sub>6</sub>)<sub>P1</sub>-(N<sub>5</sub>)<sub>Spacer</sub>-(N<sub>12</sub>)<sub>Biosensor</sub> -3' (**figure 2A**). We focused our search for optimal cleavage site within the first 1200 nt of the PACE4 cDNA and tested several potential targeting sites with the Ribosubstrates online software (<http://www.riboclub.org/ribosubstrates>; [36]). This integrated software searches in selected cDNA data bases all potential substrates for a given SOFA-HDVRz. These potential substrates include mRNAs with perfect matches with the catalytic RNA tested, but also the wobble bp and mismatches. Only SOFA-HDVRz with a PACE4-specific biosensor and harbouring a minimum of wobble bp were considered.

(ii) *Plasmid DNA constructs.* We used the previously described expression vector, ptRNA<sup>Val</sup>/hygromycin, containing the RNA polymerase III promoter tRNA<sup>Val</sup> promoter for cellular applications [37]. A PCR strategy was used to create a DNA template containing a 5'-*KpnI* restriction site and a 3'-blunt end. The sequences of the two complementary and overlapping DNA oligodeoxynucleotides (ODNs) used were: sense, 5'-ATCCATCGGGTACCGGGCCAGCTAGTTT(GGCCTCTGCTAC)<sub>BS</sub>(CAAC)<sub>BL</sub>CAGGGTCCACC-3' and, antisense, 5'-CCAGCTAGAAAGGGTCCCTT-AGCCATCCGCGAACGGATGCCCA(ATCAAC)<sub>P1</sub>ACCGCGAGGAGGTGGACCCTG(GTTG)<sub>BL</sub>-3'. The underlined nt correspond to the *KpnI* restriction site, and those in parenthesis to the

PACE4 specific biosensor (BS), blocker (BL) and P1 stem (P1) of the PACE4-SOFA-HDVRz. The purified and *KpnI*-digested PCR product was cloned in the expression vector previously digested with *KpnI* and *EcoRV* restriction enzymes. The vector used to restore PACE4 mRNA levels contained the full length PACE4 cDNA and a neomycin resistance gene was kindly provided by Dr. J.W. Creemers [38].

(iii) *RNA synthesis.* As previously described, radiolabeled PACE4 RNA was obtained from transcription of a *XhoI*-digested pcDNA3 vector containing a chimeric cDNA composed of the PC5/6A signal peptide linked to proPACE4 coding sequence [38] using T7 RNA polymerase with 50  $\mu$ Ci of [ $\alpha$ - $^{32}$ P]GTP [37]. The catalytic RNAs were synthesized using a PCR-based strategy with the expression vectors to generate DNA templates containing a 5'-T7 RNA polymerase promoter. We used either the following sense primer 5'-TTAATACGACTCACTATAGGGACCGTTGGTTTCCGTAG-3' or 5'-TTAATACGACTCACTATAGGGCCAGCTAGTTT-3', complementary to either the tRNA<sup>Val</sup> promoter or the PACE4-SOFA-HDVRz, respectively (the underlined nucleotides correspond to the T7 RNA polymerase promoter sequence). The antisense ODN sequence used for both PCR was 5'-CCAGCTAGAAAGGGTCCCTTA-3'. After PCR, the purified products were used as templates for T7 RNA polymerase transcription of tRNA<sup>Val</sup>-PACE4-SOFA-HDVRz or PACE4-SOFA-HDVRz, as described previously [37]. All products were purified on denaturing 5% (PACE4 transcript) or 7.5% (for PACE4-SOFA-HDVRz transcripts) PAGE.

(iv) *Cleavage assay.* The SOFA-HDVRz cleavage assays under single turnover conditions ([SOFA-HDVRz] > [PACE4 RNA]) were done at 37°C for

3 h in a 10  $\mu$ L reaction containing trace amount of radiolabeled PACE4 RNA and 1  $\mu$ M of SOFA-HDVRz in reaction buffer containing 50 mM Tris-HCl, pH 7.5, and 10 mM MgCl<sub>2</sub>. The reactions were stopped by the addition of loading buffer (97% formamide, 1 mM EDTA (pH 8.0), 0.025% xylene cyanol and 0.025% bromophenol blue), electrophoresed on denaturing 5% PAGE gel, and analysed with a PhosphorImager (Amersham Biosciences).

**Cell culture and DNA transfection.** Human cancer prostate cell lines DU145 were obtained from ATCC. Cells were maintained in Roswell Park Memorial Institute medium (RPMI 1640) supplemented with 5% fetal bovine serum (Wisent Bioproducts). Cells were grown at 37°C in a water-saturated atmosphere in air/CO<sub>2</sub> (5%). Cells were transfected using lipofectamine2000 (Invitrogen) as per manufacturer's instruction, and were selected for resistance to hygromycin B (Invitrogen) at 125  $\mu$ g/mL, with 200  $\mu$ g/mL of neomycin for double-transfected cells. The stable cell line transfected with the SOFA-HDVRz expression vector was named 4-2, while the 4-2 cell line stably transfected with the PACE4 expression vector was named 4-2+PACE4.

**Northern blot hybridizations.** Total RNA was isolated from DU145 cells using guanidinium isothiocyanate followed by lithium chloride precipitation, as described previously [37]. RNA migration (5  $\mu$ g) on denaturing agarose gel, membrane transfer and  $^{32}$ P-labeled RNA probe transcriptions were performed as previously described [37]. Linearized vectors were used as DNA template for complementary RNA probe transcription using either T7 or SP6 RNA polymerase. The 1066-bp cDNA

for human furin probe was obtained by digestion of the full-length clone with *XhoI* enzyme [39]. A 456-bp cDNA fragment of PACE4 was cloned in pGEM-T easy vector system (Promega) by RT-PCR reaction on DU145 total RNA with specific primers [13]. This vector was subsequently used for probe transcription. For PC7 probe, we used a 285-bp rat cDNA [40], and for bovine 18S ribosomal RNA probe, a 600-bp cDNA [41] was used. The ImageJ software 1.37v (<http://rsb.info.nih.gov/ij/>) was used for all densitometric analyses.

### Microarray analysis.

(i) *RNA extraction.* RNA from DU145 or 4-2 cell pellets (in duplicate) was extracted using TRIzol<sup>®</sup> reagent (Invitrogen) according to the manufacturer's instructions. RNA extractions were purified using RNeasy mini kit<sup>®</sup> (Qiagen). The quality of the RNA was assessed using a 2100 Bioanalyzer with the RNA 6000 Nano LabChip kit (Agilent Technologies, Mississauga, ON, Canada) according to the manufacturer's protocol.

(ii) *Microarray hybridization.* The hybridization assays and data collection of expression values were performed on the Functional Genomics Platform of the McGill University and Genome Quebec Innovation Centre (Montreal, Canada) using Illumina whole genome HumanWG-6 v3.0 beadchip<sup>®</sup> (38,275 genes).

(iii) *Processing of data.* All analyses were performed with the software FlexArray v1.2 (<http://genomequebec.mcgill.ca/FlexArray/>). The raw data were pre-processed and normalized for each array using the 'lumi' algorithm. Then, the mean fold changes of normalized data were calculated for DU145 vs 4-2. Statistical analysis of microarray data was performed using the Significance Analysis of Microarrays (SAM)

method, which utilizes a modified *t* test statistic and sample-label permutations to evaluate statistical significance [42].

(iv) *Real-time PCR.* The real-time PCR assays were carried out on the Mx3000P quantitative PCR system using the Brilliant II SYBR Green kits (Stratagene). Briefly, the RNA was DNase I treated (Invitrogen), reverse-transcribed using Superscript II reverse transcriptase (Invitrogen) and RNaseH treated (Ambion) before quantitative PCR. The reaction mixture combined 1 µg of cDNA with 12.5 µL of 2X master-mix containing primers at a 300 nM final concentration. The PCR profile was: 1 cycle for 10 min at 95°C, 40 cycles of 30 s at 95°C, 1 min at 60°C and 30 s at 72°C. Primers were designed using Beacon Designer<sup>™</sup> software ([www.premierbiosoft.com](http://www.premierbiosoft.com)). The sequences of primer used were:

SERPINE2 (NM\_006216): sense 5'-CCCACTTCAATCCTCTGTCTCTC-3', antisense 5'-CCCAGGACCGACGCAATC-3';  
 CDK6 (NM\_001259): sense 5'-GCCGCCAGCCAGAACAC-3', antisense 5'-CTCCAGATAGCAATCCTCCACAGC-3';  
 IGFBP6 (NM\_002178): sense 5'-GCGTGGAGGAGGAGGATGG-3', antisense 5'-CTTGGGCGGATGGCACTG-3';  
 TGFBI (NM\_000358): sense 5'-GTGCGGCTGCTGGCTCTC-3', antisense 5'-GCTGGTAGGGCGACTTGGC-3';  
 CCNA1 (NM\_003914): sense 5'-YGAAGAAGCAGCCAGACATC-3', antisense 5'-AAGTTGACAGCCAGATACAGG-3';  
 IGFBP3 (NM\_001013398): sense 5'-TTCTGCTGGTGTGTGGATAAG-3', antisense 5'-GGCGTCTACTTGTCTGCTGC-3';  
 THBS1 (NM\_003246): sense 5'-GGACTCGCTGTAGGTTATGATG-3', antisense 5'-CGGCTGCTGGACTGGTAG-3';



MUC1 (NM\_002456): sense 5'-CTGGTGCTGGTCTGTGTTTC-3', antisense 5'-GCTGCCCGTAGTTCTTTTCG-3'; ACTB (NM\_007393): sense 5'-GGGAAATCGTGCGTGACATCAAA G-3', antisense 5'-CATACCCAAGAAGGAAGGCTGGA A-3'. For all PCR reactions, standard curves, dissociation curves and migration of PCR products on acrylamide gels were done to confirm the specificity of the products. Moreover, two PCR reactions were performed as negative controls: one without cDNA template, and the other using reverse transcription reaction that never received Superscript II reverse transcriptase enzyme. Results were estimated as Ct values and fold change of the expression of target gene in DU145 compared to 4-2 cells was expressed as  $2^{\Delta\Delta Ct}$ . The  $\Delta Ct$  was obtained from the subtraction of the mean Ct for the target with the mean Ct for the internal control gene (beta-actine). The  $\Delta\Delta Ct$  is the mean difference between the  $\Delta Ct$  obtained from DU145 RNA extracts minus  $\Delta Ct$  obtained from 4-2 RNA extracts. All PCR reactions were repeated twice for each set of primers.

**Conditioned growth medium preparation.** Four 10-cm plates were seeded with  $1.0 \times 10^6$  cells (DU145 or 4-2 cells) in complete growth media. The next day, the cells were washed twice with PBS and the growth media was replaced with 5 mL serum-free RPMI for 1 h. The conditioned media were collected, filtered through 0.45  $\mu\text{M}$  Acrodisc<sup>®</sup> syringe filter units (Pall Corporation, USA) and kept on ice. The filtrates were concentrated to a final volume of 350  $\mu\text{L}$  with Amicon<sup>®</sup> Ultra centrifugal filter devices with a 3 kDa molecular weight cut-off. Fresh RPMI was treated the same way as a control for cell experiments.

**MTT proliferation assay.** The cell proliferation was measured by the colorimetric MTT assay (thiazolyl blue tetrazolium bromide; Sigma-Aldrich). Briefly, cells were seeded in 96-well plates (BD Biosciences) in triplicate with 100  $\mu\text{L}$  of a  $5.0 \times 10^4$  cells/mL cell suspension in complete growth medium. The following day, cells were carefully washed twice with PBS twice and were incubated with 50  $\mu\text{L}$  of either fresh RPMI, concentrated RPMI (RPMI conc), DU145 or 4-2 concentrated conditioned growth media (DU145 and 4-2 CM conc, respectively). After 48 h, 12,5  $\mu\text{L}$  of a MTT solution (5 mg/mL in PBS) was added to each well for 4.5 h at 37°C/5% CO<sub>2</sub>. The media was then carefully discarded and the cells were solubilized with 100  $\mu\text{L}$  isopropanol/0.04 N HCl solution. The absorbance was measured at a wavelength of 550 nm with a reference at 650 nm in microplate reader (SpectraMax190; Molecular Devices).

**Cell count assay.** Cells were plated at a density of  $5.0 \times 10^4$ /well in 6-well plates (BD Biosciences) in duplicates. Complete growth medium was changed after 48 h. After incubation, cells were washed in PBS, trypsinized and counted in after staining in 0.4% (w/v) trypan blue solution (Sigma). Only viable cells were counted in duplicate.

**Colony formation assay.** Cell lines were seeded in 6-well plates (BD Biosciences) at a density of 300 cells/well in triplicate. Cells were maintained in complete growth medium for 10 days. After colony formation, media was discarded and cells were washed once with PBS. Colonies were fixed and stained in 5 mg/mL methylene blue/50% methanol solution for 10 min. Excess of staining solution was removed carefully with distilled water

and the plates were dried overnight before scanning with Li-Cor Odyssey Infrared Imaging System (Li-Cor Biosciences). Scanned images were analyzed with ImageJ software 1.37v (<http://rsb.info.nih.gov/ij/>) to measure the total particulate area.

#### **Flow cytometric analysis of apoptosis without cell fixation.**

(i) *Propidium iodide (PI) DNA staining.* Cell lines were seeded in 10-cm plates ( $5.0 \times 10^5$  cells) in complete growth media and grown overnight. Afterwards, cells were washed three times with PBS, incubated in fresh complete growth for 12h and harvested with trypsin (including the supernatant). The cell pellets were washed once with PBS, centrifuged and resuspended in 50  $\mu$ L of PBS. A chromatin decondensation treatment was performed using 450  $\mu$ L of H-buffer/Triton-X 100 (20 mM HEPES, pH 7.2, 0.16 M NaCl, 1 mM EGTA and 0.05% Triton-X 100) warmed at 37°C for 1 min on ice. After this incubation, cells were gently centrifuged for 5 min and 450  $\mu$ L of the supernatant was discarded. Cell pellets were then gently resuspended with the remaining supernatant and the DNA staining was performed using 450  $\mu$ L of H-buffer containing 10  $\mu$ g/mL of RNaseA (Sigma) and 10  $\mu$ g/mL of PI (Sigma) for 30 min at room temperature in the dark. After this incubation, cells were analyzed on fluorescence-activated cell sorter (FACScan cytometer; Becton Dickinson).

(ii) *Fluorescence-activated cell sorter (FACS) analysis.* After the PI staining, cells were analyzed on a FACScan cytometer (Becton Dickinson). Forward and side scatter signals were used to establish the live gate to exclude debris and cell clumps. A minimum of 10,000 gated events by sample were acquired. The fluorescence of PI was collected in red channel on a

linear scale (FL3). A second live gate was set using the FL3-A and FL3-W parameters of the doublet discrimination module, allowing single cell measurements. Fluorescence intensity distribution was analyzed with the CellQuest software while the distribution of non apoptotic cells into the phases of the cell cycle was determined using the ModFit software.

**Human tumor xenograft models.** Four-week-old female athymic nude mice (NU/NU; Charles River Laboratories) were inoculated subcutaneously at the opposite sides of the flank with  $3.0 \times 10^6$  cells per inoculum (five mice/group). Cells were grown in complete media and harvested at their exponential growing state. Mice were housed under pathogen free conditions and the implantations were done under anesthesia conditions in laminar flow hood. Xenografts were measured three times per week and volume (V) was determined by this equation :  $V = (L \times W^2) \times 0.5$ , where L is the length and W is the width of a xenograft.

**Histology.** Thirty-one days old subcutaneous DU145 and 4-2 tumors were eradicated, formalin-fixed and paraffin-embedded. Hematoxylin and eosin staining procedures were performed according s standard protocol on 3- $\mu$ m sections.

#### **Mass spectrometry analysis of conditioned media.**

(i) *Sample preparation.* The DU145 and 4-2 conditioned growth media were prepared as described earlier in this section, except that the concentrated media were washed in centrifugal filter device with 12 mL of water before being collected and lyophilised. The samples were then solubilised with 100 $\mu$ l acidified water - 0.1% trifluoro acetic acid (TFA).

Acetone precipitation was performed with a ratio 4/1 (v/v). After a centrifugation at 13 000 rpm at 4°C for 30 min, pellets were solubilized in 50mM pH 8.0 ammonium bicarbonate (NH<sub>4</sub>HCO<sub>3</sub>), a reduction of cysteines residues with a 45 mM dithiothreitol (DTT) solution was performed for 15 min at 50°C and a subsequent alkylation was performed with a 100 mM iodoacetamide (IAA) solution for 15 min at room temperature in the dark.

(ii) *Trypsin Digestion.* The samples were placed on ice for 30 min in 50 µL of proteases solution (sequence grade-modified trypsin, Promega, at 0.02 mg/mL in 25 mM NH<sub>4</sub>HCO<sub>3</sub>). Digestion were performed overnight at 37°C. The samples were then desalted and concentrated using C18 ZipTips (Millipore) before mass spectrometry analyses.

(iii) *MALDI-TOF-TOF analyses.* The MALDI TOF/TOF mass spectrometer (Ultraflex II, Bruker Daltonik, Bremen, Germany) was used in this study, equipped with an Nd:YAG laser (355 nm) with a repetition rate of 200 Hz. Mass spectra were the sums of 8000 shots acquired in positive ion reflectron TOF mode. The acceleration potential was 25 kV. For MS<sup>2</sup> analysis, 2000 shots were acquired for parent ion selection and 10,000 shots for fragments selection. For MALDI analysis, the matrix used was alpha-Cyano-4-hydroxycinnamic acid (HCCA) ACN:H<sub>2</sub>O/TFA 0.1% (7:3). Calibration was performed with peptides calibration standard (Bruker Daltonik, Bremen, Germany). Samples underwent software processing with FlexAnalysis 2.4 and Biotools 2.2 (Bruker Daltonik, Bremen, Germany) for baseline subtraction and peaks detection.

(iv) *Protein identification.* Protein identification was performed under the MASCOT sequence query search program using the SwissProt data base filtered for the taxonomy

“*Homo sapiens*”. A tolerance of 1 Da for peptide and 0.6 Da for MS/MS was set. Only protein sequences with a MOWSE score higher than 36 (indicating significant homology or identity) and identified in several samples representing at least two significant MS/MS were considered. Methionine oxidation and acetylation of N-terminals were defined as variable modification.

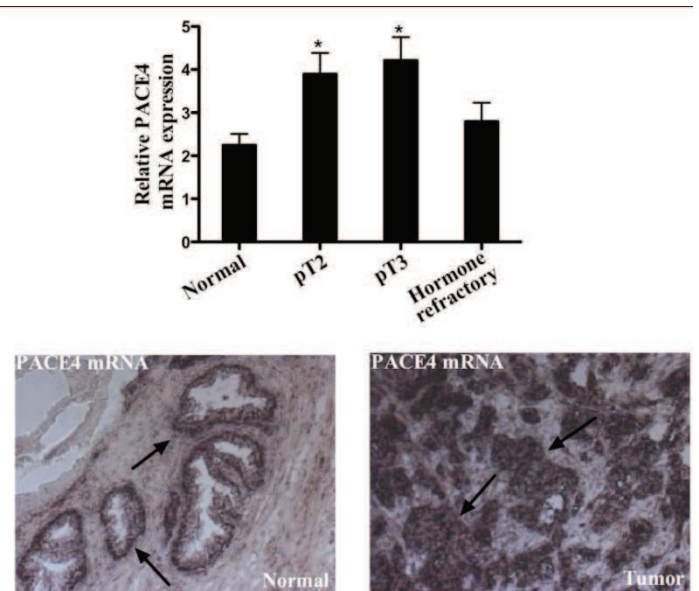
**Statistical analyses.** All experiments were repeated at least three times, except for human tumor xenografts experiment and microarray analysis. Comparisons between the test group and controls of this study were done using Student’s *t* test.

## RESULTS

### **PACE4 expression in clinically localized prostate tumors.**

We tested 47 primary prostate tumor samples obtained from patients undergoing surgery for PACE4 expression. The majority of the patients (i.e. 34) had clinically localized tumors and 13 had hormone-refractory recurrent prostate carcinoma. All collected samples were subjected to histopathological examination. Only tissues where all epithelial cells were neoplastic were dissected and used for further studies. A real-time PCR strategy was used to evaluate PACE4 mRNA expression levels in prostate tumor tissues using nine matched normal prostate tissues as a reference (**figure 1A**). The tissues were grouped into similar clinical stages based on TNM system as: 18 pT2 samples (tumors strictly confined to the organ), 16 pT3 samples (tumors with extracapsular extension), and 13 hormone-refractory samples (tumors no longer responsive to endocrine therapy). The mean relative PACE4 mRNA

expression levels (**figure 1A**) were significantly higher in both pT2 and pT3 groups ( $3.9 \pm 0.5$  and  $4.2 \pm 0.5$ , respectively), when compared to the mean level found in normal prostate tissues ( $2.2 \pm 0.3$ ). However, the mean PACE4 expression level measured in hormone refractory tissues ( $2.8 \pm 0.4$ ) was not significantly higher than the one measured in controls. Real-time PCR for the other PCs showed that furin, PACE4 and PC7 were the most expressed PCs in normal prostate tissues. However, only PACE4 mRNA levels increased in tumor tissues, while the others remained unchanged (unpublished data). This higher PACE4 expression, particularly in epithelial cells, was directly assessed by an *in situ* hybridization. Normal prostate tissues showed the expected epithelial cell distribution of PACE4 mRNA. However, tumors showed disorganization of the tissue structure, with a higher PACE4 expression and even cells invading the stroma (**figure 1B**). Because PACE4 was the only member of the PC family to be overexpressed in prostate tissues, we investigated its expression levels in the Oncomine data base ([www.oncomine.org](http://www.oncomine.org)). As summarized in the **Table I**, six independent studies reported a significant overexpression of PACE4 ( $p < 0.001$ ) in prostate cancer tissues compared with corresponding normal tissues. Moreover, in five of the six studies, the PACE4 overexpression rank was within the top 2%, while the sixth was in the top 7%. It is noteworthy to report that furin, PC5/6 and PC7 did not show an upregulation of their expression within the thirteen repertoried studies. Thus, these data combined to our real-time PCR results strongly suggest that PACE4 is the only member of the PC family associated to prostate cancer.



**Figure 1 : Relative PACE4 mRNA expression in normal and tumoral prostate tissues.** A. Real-time PCR from reverse-transcribed RNA from tissue specimens. PCR with PACE4 and PPIA specific primers were used to quantify the relative expression levels of PACE4 in each sample. Nine normal prostate tissues and 47 prostate tumors samples were used. Clinical stage classification, based on TNM system, was: pT2, tumor strictly confined to the organ (18 samples) and, pT3 (16 samples), tumor with extracapsular extensions. Hormone refractory prostate tumors tissues (13 samples) were issued from patients who relapsed after receiving endocrine therapy. Values are mean  $\pm$  SEM. \* $P < 0.05$ . B. PACE4 in situ hybridization in epithelial cells from normal and tumoral prostate tissues (arrows indicate purple staining corresponding to PACE4 mRNA).

Table I. PACE4 is up-regulated in prostate cancer according to Oncomine data base\*

Study	Description	Fold change	P-value	Number of measured genes	Overexpression gene rank	Reference
1	Prostate Carcinoma (8) vs. Prostate Gland (4)	1.951	1.58E-4	4,564	7 (in top 1%)	Magee JA <i>et al.</i> , Cancer Res. 2001 Aug 1;61(15):5692-6
2	Prostate Carcinoma (61) vs. Prostate Gland (39)	2.190	9.12E-15	10,021	35 (in top 1%)	Lapointe J. <i>et al.</i> , Proc Natl Acad Sci U S A. 2004 Jan 20;101(3):811-6
3	Prostate Carcinoma (65) vs. Prostate Gland (23)	1.246	4.65E-7	7,820	89 (in top 2%)	Yu TP <i>et al.</i> , J Clin Oncol. 2004 Jul 15;22(14):2790-9
4	Prostate Carcinoma (44) vs. Prostate Gland (13)	1.271	7.56E-5	12,427	113 (in top 1%)	Liu P <i>et al.</i> , Cancer Res. 2006 Apr 15;66(8):4011-9
5	Prostate Carcinoma (7) vs. Prostate Gland (6)	2.627	5.93E-4	19,079	250 (in top 2%)	Varambally S <i>et al.</i> , Cancer Cell. 2005 Nov;8(5):393-406
6	Prostate Adenocarcinoma (27) vs Prostate Gland (8)	1.349	5.30E-4	17,358	1202 (in top 7%)	Vanaja DK <i>et al.</i> , Cancer Res. 2003 Jul 15;63(14):3877-82

\*Data from www.oncomine.org

### Rational design of SOFA-HDVRz targeting PACE4 mRNA.

(i) *SOFA-HDVRz as a gene silencing tool.* Considering the high expression level of PACE4 mRNA in tumoral prostate tissues, we decided to evaluate the contribution of this enzyme to the maintenance of the tumorigenic state. While PACE4-specific and potent inhibitors are still under development [9, 43], we opted for a molecular approach targeting PACE4 mRNA. We previously showed that HDVRz could be a powerful and specific silencing tool in cultured cells [37]. Here, we tested a new generation of HDVRz harboring a Switch On/Off Adapter (SOFA), resulting in an enhanced *in vitro* specificity and potency [44, 45]. As shown in **figure 2A**, the SOFA-HDVRz has three new features over the classical HDVRz. First, a blocker stem of 4 nt prevents the formation of the P1 stem, essential for the cleavage event. A biosensor of 12 nt at the 5' end of the blocker stem compels the SOFA-HDVRz to maintain an inactive state ('OFF' conformation) until the right complementary RNA hybridizes. This sequence specific hybridization releases the P1 stem from

the blocker, allowing the formation of the P1 stem.

(ii) *SOFA-HDVRz design.* We used the complete sequence of human PACE4 cDNA (GenBank data base accession number M80482) to design a selective SOFA-HDVRz against PACE4 mRNA.

The **figure 2B** shows the *in vitro* cleavage of PACE RNA with a selected PACE4-SOFA-HDVRz cleaving into the prodomain coding sequence of the transcript at position 519. As shown on the autoradiogram of the PAGE gel, the reaction generated only one cleavage product that could be seen after the addition of the PACE4-SOFA-HDVRz. In order to confirm the specificity of this molecule, we firstly tested the same ribozyme against a radiolabeled human furin transcript and, secondly, we used a catalytically inactive mutant (C76A) of the SOFA-PACE4-HDVRz against the PACE4 transcript [44]. In both cases, no cleavage product could be observed (unpublished data), confirming the capacity of the PACE4-SOFA-HDVRz to selectively cleave its target *in vitro*.

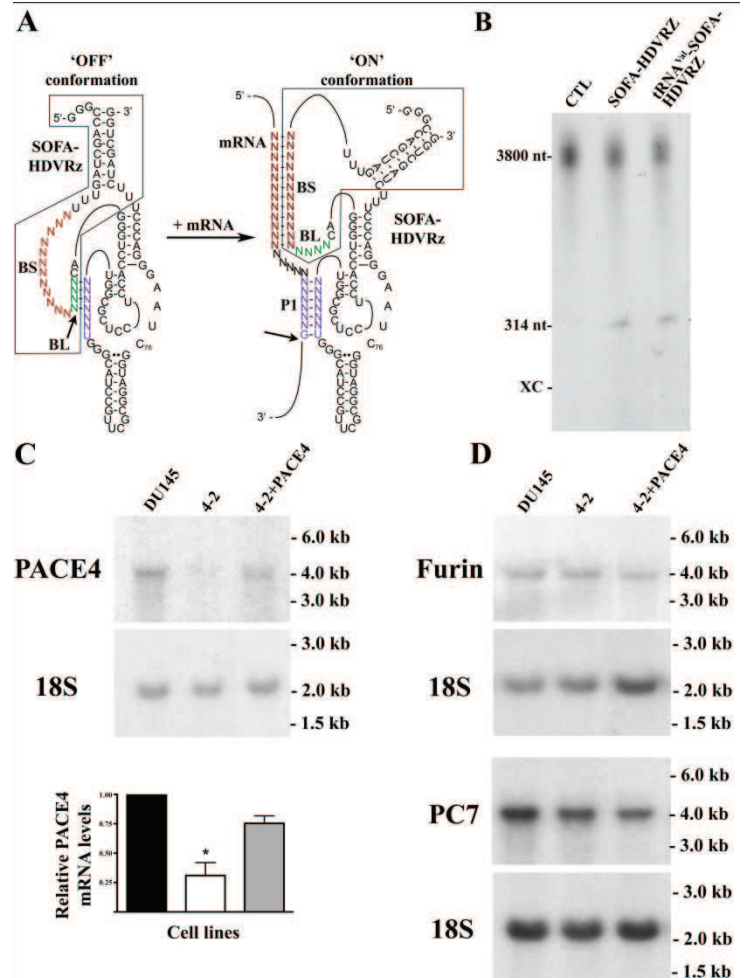
### Downregulation of PACE4 mRNA in DU145 cells.

(i) *Expression vector.* We used an expression vector containing the tRNA<sup>Val</sup> promoter to produce the PACE4-SOFA-HDVRz into transfected cells [37]. To insure that the tRNA<sup>Val</sup> part did not interfere with the catalytic activity of the associated SOFA-HDVRz, we performed an *in vitro* cleavage assay by incubating the chimeric transcript

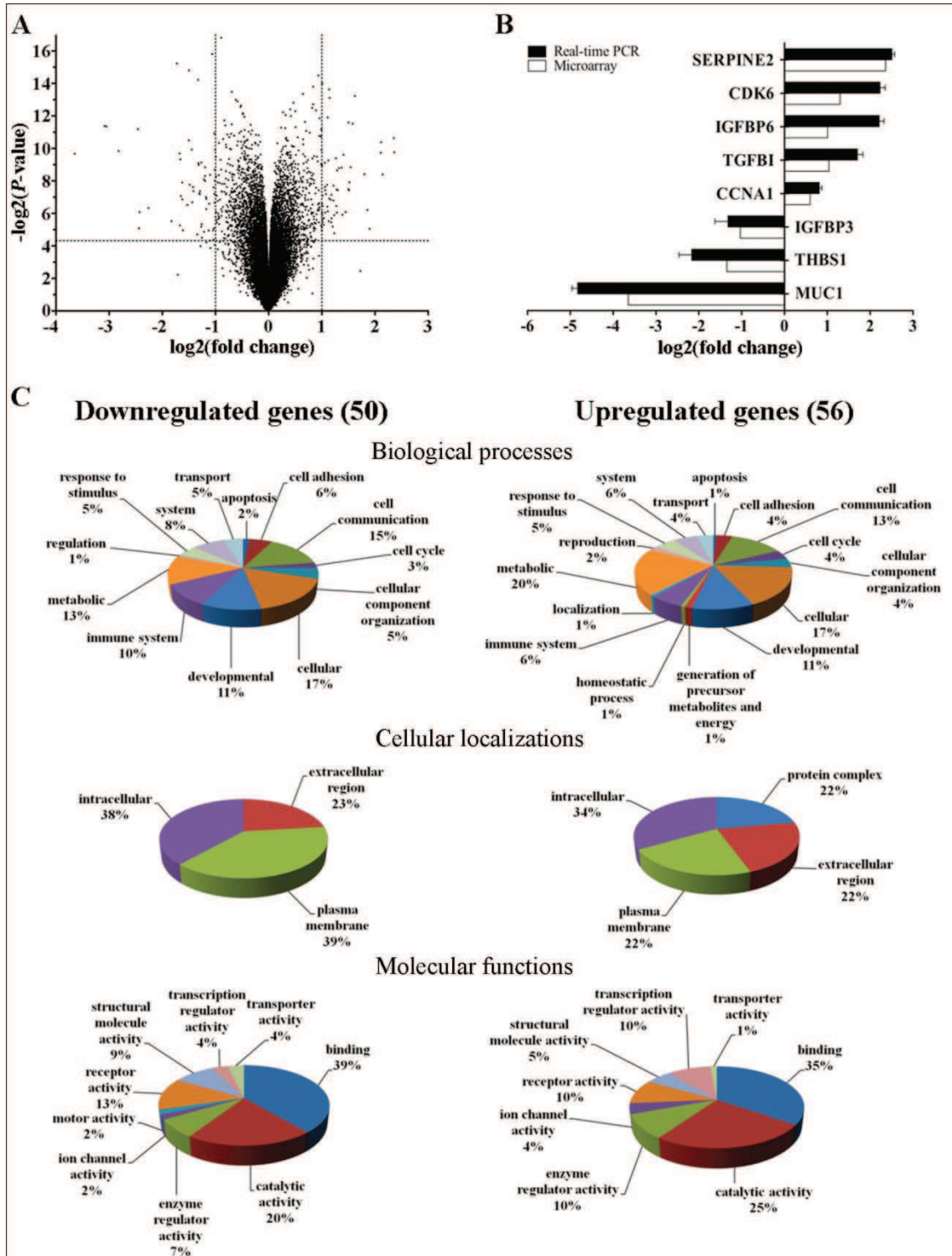
(tRNA<sup>Val</sup>-SOFA-HDVRz) with an internally radiolabelled PACE4 RNA (figure 2B). As seen on the autoradiogram of the PAGE gel, the addition of the tRNA<sup>Val</sup> motif did not modify neither the specificity nor the level of the cleavage activity of the PACE4-SOFA-HDVRz.

ii) Northern blots. We transfected the PACE4-SOFA-HDVRz expression vector (named ptRNA<sup>Val</sup>-PACE4-SOFA-HDVRz) into DU145, a highly invasive, androgen-independent prostate epithelial tumor cell line. Stable cell lines transfected were established by the selection of clones resistant to hygromycin B. Northern blot analyses on total RNA extracts were performed for wild-type DU145 (DU145), DU145 transfected with ptRNA<sup>Val</sup>-PACE4-SOFA-HDVRz (4-2) and on 4-2+PACE4 cells, which are 4-2 cells co-transfected with PACE4 cDNA expression vector. As seen in figure 2C, the PACE4 mRNA level in the SOFA-HDVRz transfected cell line is significantly reduced when compared to the untransfected cells. These levels were partially re-established by the overexpression of PACE4 cDNA. The graph shows a densitometric analysis using 18S ribosomal RNA as loading control performed to quantify the mRNA levels in these clonal cell lines using wild type DU145 cells as reference ( $0.3 \pm 0.1$  and  $0.75 \pm 0.06$  for 4-2 and 4-2+PACE4, respectively). The mRNA levels of two other endogenous expressed PCs confirmed the specificity of the PACE4-SOFA-HDVRz cleavage. Levels of furin and PC7 mRNAs (figure 2D) remained mostly unchanged in the 4-2 cells, confirming the reduction of PACE4 expression without significantly affecting the expression of other endogenous PCs. However, mRNA levels of these two PCs seemed reduced in the 4-2+PACE4 cells,

particularly for furin, suggesting a cellular adaptation following PACE4 re-introduction.



**Figure 2 : Downregulation of PACE4 expression in DU145 cells.** A. The secondary structure of both the 'OFF' and 'ON' conformation of the SOFA module (box) attached to the classical HDVRz. The biosensor (in red, BS) and the blocker (in green, BL) can be modified to hybridize on a specific mRNA target (N indicates A, C, G or U). After the recognition of the mRNA, the P1 stem (in purple, P1) issued from the sequence-specific hybridization of the HDVRz with the targeted RNA will allow the catalytic cleavage of the latter, as indicated by the arrow. B. Autoradiogram of *in vitro* cleavage assays performed with either the SOFA-HDVRz or the chimeric RNA transcript tRNA<sup>Val</sup>-SOFA-HDVRz both specific for PACE4 RNA. The control lane is the migration of the internally <sup>32</sup>P radiolabeled 3800 nt PACE4 RNA alone. The BS, BL and P1 nucleotide sequences of the PACE4-SOFA-HDVRz used are indicated in the *experimental procedures* section. XC indicate the position of xylene cyanol after migration on the denaturing 5% PAGE gel. C-D. Autoradiograms of Northern blot hybridizations showing levels of PACE4 (C), furin (D) and PC7 (D) mRNAs in untransfected DU145 cells (control cell line) or in cell lines either transfected with the expression vector coding for the tRNA<sup>Val</sup>-PACE4-SOFA-HDVRz alone (4-2) or co-transfected with a PACE4 cDNA expression vector (4-2+PACE4) are shown. The bands corresponding to the 18S ribosomal RNA are shown for all blots. The size markers are shown on the right. The expression levels of PACE4 mRNA in transfected cell lines relative to the untransfected DU145 cells were obtained by densitometric analyses using 18S ribosomal RNA as loading control (C). Values are mean ± SEM (n=3). \*P < 0.05.



**Figure 3 : Microarray analysis and real-time PCR.** *A.* Volcano plot of differentially expressed genes in DU145 vs 4-2 cells measured with Illumina human expression Beadchips™. Vertical dashed lines show threshold values for the fold change, while the horizontal dashed line indicates the threshold value for the *P*-value (0.05). Only genes with *P*-values higher than the threshold with a  $\log_2(\text{fold change})$  lower than -1 or higher than 1 were include in supplemental tables I and II, respectively. *B.* Confirmation of values obtained with microarrays by real-time PCR. Amplification and detection of cDNA were performed using SYBR Green I dye. The expression levels, measured by real-time PCR (black bars), were normalized using beta-actin as internal control. The white bars indicate values obtained from microarray analyses. *C.* Classification of differentially expressed genes in 4-2 cells compared to DU145 according to their biological processes, cellular localizations and molecular functions. The PANTHER classification system ([www.pantherdb.org](http://www.pantherdb.org)) had no attributed gene ontology grouping for nine downregulated genes and six upregulated genes.

### Gene expression profiling in DU145 cells.

With the aim of better understanding the general contribution of PACE4 on prostate cancer growth, we performed a comparative gene expression analysis by DNA microarray screening with DU145 and 4-2 cells (**figure 3A**). The gene expression levels measured by microarrays was validated by real-time PCR for eight genes covering a broad range of fold changes (**figure 3B**). For all the gene tested, the quantitative PCR results (black bars) were either similar or higher than microarray data, confirming that microarray results (white bars) were reliable or even underestimated, at least for those tested. Of the 38,275 genes tested, 59 genes (0.15%) showed a 2-fold decrease while 62 genes (0.16%) had a 2-fold increase in 4-2 cells ( $P$ -value  $< 0.05$ ). These genes correspond to dots found in the higher left and right quadrants of the volcano plot (**figure 3A**).

The PANTHER (Protein Analysis THrough Evolutionary Relationships; <http://www.pantherdb.org>) system was used to group the gene listed in supplemental **tables I and II** according to their biological processes, cellular localizations and molecular functions using the gene ontology classification (**figure 3C**). Among the 59 downregulated and 62 upregulated genes, 9 and 6, respectively, could not be associated with any group into the database. for the others, the classification reported three major biological processes associated to the downregulated genes: cell communication (15%), developmental (11%) and metabolic (13%) processes, while four majors were linked to the upregulated genes set: cell communication (13%), cellular (17%), developmental (11%) and metabolic (20%) processes. The analysis of the

cellular localization suggests a central role of PACE4 into the processing of various intracellular, membrane bound, protein complexes and extracellular proteins as it is also shown by the major biological functions affected by its downregulation in 4-2 cells, including binding and various catalytic (enzymatic) activity. Although their known implications in various biological functions, the majority of the proteins encoded by these genes were not potential PACE4 substrates. However, the analysis of CDH1 amino acid sequence revealed that it is a potential PACE4 substrate, since it necessitates a proteolytic processing event at the  $_{149}\text{LRRQKR}\downarrow\text{DW}_{156}$  site of its propeptide domain to be biologically active [46]. Moreover, PTHLH also carries a potential processing site at the  $_{31}\text{SRRLKR}\downarrow\text{AV}_{38}$  of its propeptide region [47].

### Comparison of gene expression profile in 4-2 cells with microarray-based study in prostate cancer.

To validate the gene expression profile obtained in the present study, we sought to compare our results to those reported into a previously published study exploring gene expression variations among 62 primary prostate tumors and 41 matched normal prostate tissues by using cDNA microarrays [48]. Based on an unsupervised hierarchical clustering, this study identified three distinct subclasses of prostate tumors based on distinct patterns of gene expression. Interestingly, high grade and advance stage tumors were more highly represented among tumor subgroups II and III, while the subtype I tumors were associated with least aggressive tumors.

Among the  $\approx 26,000$  genes measured, Lapointe *et al.* highlighted 5,153 cDNAs whose expression varied most across samples



Table II. Comparison of downregulated and upregulated genes in 4-2 cells with variations observed in different subclasses of prostate cancer tissues\*

Accession number	Gene symbol	Gene description	Tumor subclass I	Tumor subclass II	Tumor subclass III
			vs Normal	vs Normal	vs Normal
<b>Downregulated genes in 4-2 cells</b>					
NM_002456	MUC1	mucin 1	**↑	-	↓
NM_004360	CDH1	cadherin 1	-	↓	↑
NM_002354	TACSTD1	tumor-associated calcium signal transducer 1	↓	↓	↓
NM_020770	CGN	cingulin	↓	↓	-
NM_014220	TM4SF1	transmembrane 4 L six family member 1	-	↓	↑
NM_001993	F3	coagulation factor III	↑	↓	-
NM_001823	CKB	creatine kinase, brain	-	↑	↓
NM_001003794	MGLL	monoglyceride lipase	-	-	↓
NM_052886	MAL2	mal, T-cell differentiation protein 2	↓	↓	↓
<b>Upregulated genes in 4-2 cells</b>					
NM_001001390	CD44	CD44 molecule (Indian blood group)	-	-	↑
NM_006216	SERPINE2	serpin peptidase inhibitor, clade E, member 2	↑	↑	-
NM_002639	SERPINB5	serpin peptidase inhibitor, clade B, member 5	↑	↑	↑
NM_001387	DPYSL3	dihydropyrimidinase-like 3	↑	↑	↑
NM_002737	PRKCA	protein kinase C, alpha	↑	↑	-
NM_001964	EGR1	early growth response 1	-	↑	↑
NM_004163	RAB27B	clone 25194 mRNA sequence	↓	↓	↑
NM_000165	GJA1	gap junction protein, alpha 1	↑	↑	↑
NM_001889	CRYZ	crystallin, zeta (quinone reductase)	↑	↑	↑
NM_000104	CYP1B1	cytochrome P450, family 1, subfamily B, polypeptide 1	↑	-	↑
NM_006307	SRPX	sushi-repeat-containing protein, X-linked	↑	↑	↑

\*Data from Lapointe J. et al., Proc Natl Acad Sci U S A. 2004 Jan 20;101(3):811-6

\*\* ↓ and ↑ significate a fold change lower than 0.25 or higher than 1.25 in tumor tissues, respectively

[48]. When we compared this list to the microarray results issued from 4-2 cells (supplemental **tables I and II**), we retrieved 19 of the 59 repressed genes and 16 of the 62 overexpressed genes from our data. Moreover, 9 of these 19 downregulated genes and 11 of these 16 upregulated genes in 4-2 cells had a similar expression variation in high grade/advance stage tumors (either in tumor subclasses II, III or both of them) compared to matched normal tissues (**table II**). It should be noted that we used a less stringent fold change criteria than in our microarray study to attribute the genes into either the downregulated (0.75 or less instead of 0.5) or upregulated (1.25 or higher instead of 2.0) group. This decision was justified by the fact that we used RNA preparations issued from a culture cell lines, while the Lapointe *et al.* study was undertaken on RNA extracts from tissue specimens susceptible to contain more than one cell type, which might decrease the signal specific to tumors.

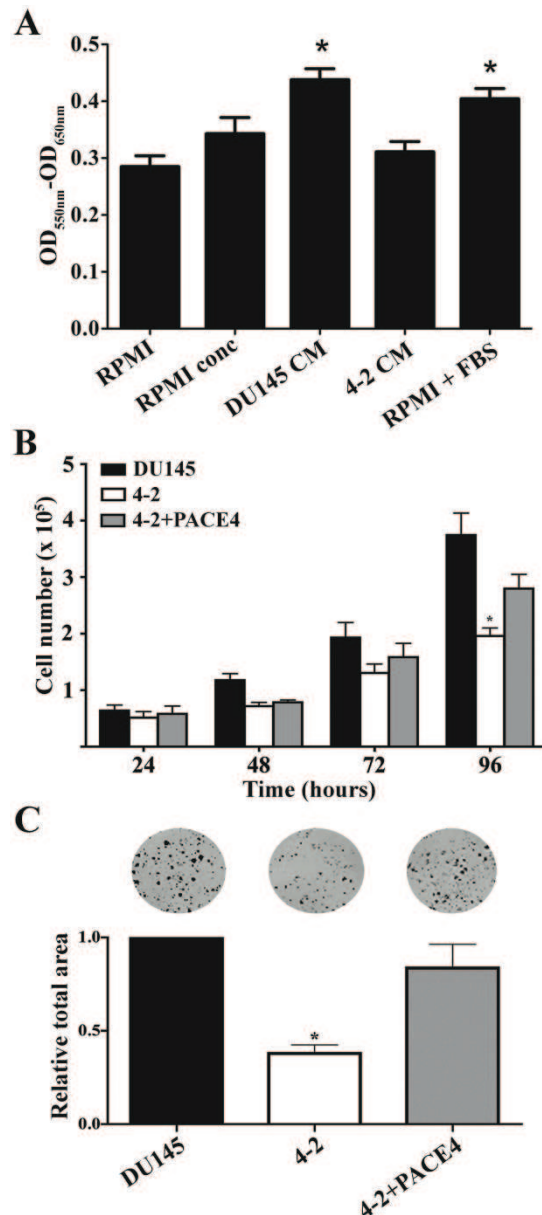
It is also noteworthy that some genes varied in the same way in tumors of subclass I than in tumors of subclasses II and/or III (e.g. CGN, MAL2, SERPINE2, PRKCA); while others varied in an opposite way in least aggressive vs aggressives tumors (e.g. MUC1, F3, RAB27B). Meaning that PACE4 might specifically contributes to the acquisition of a more aggressive phenotype in prostate cancers, as previously suggested in mouse keratinocytes [49]. Although these variations of genetic expression still remain to be better understood, the genes listed in **table II** clearly confirmed the importance of PACE4 on the expression of genes normally associated with a more aggressive prostate tumors.

**PACE4 is important for the generation of proliferation factors.**

Considering the previous links established between PCs expression and tumor progression, we measured cell proliferation by generating concentrated conditioned media issued from either untransfected DU145 or 4-2 cells. These growth media were incubated with DU145 cells for 48 h and proliferation state was measured by adding the MTT compound. As shown in **figure 4A**, concentrated conditioned growth medium issued from DU145 cells (DU145 CM) stimulated the proliferation of DU145 cells in a similar range than the complete growth medium RPMI+FBS did ( $OD_{550nm}-OD_{650nm} = 0.44 \pm 0.02$  and  $0.40 \pm 0.02$ , respectively). This higher proliferation was not measured in cells incubated with conditioned growth medium issued from 4-2 cells (4-2 CM), where the  $OD_{550}-OD_{650}$  value ( $0.31 \pm 0.02$ ) was similar to those obtained with cells incubated either with RPMI or concentrated RPMI (RPMI conc) treated in the same way than the concentrated conditioned media, i.e.  $0.28 \pm 0.02$  and  $0.34 \pm 0.03$ , respectively. Thus, these results suggest that PACE4 is important for the secretion of factors used by DU145 cells to maintain their high proliferative status.

**PACE4 downregulation slows DU145 proliferation *in vitro*.**

To further characterize the cell proliferation reduction observed with conditioned media experiments, we counted the total cell number of our stable cell lines at different times. As seen in **figure 4B**, the results showed a significant reduction of proliferation for the 4-2 cells ( $\approx 200\,000 \pm 10\,000$  cells) when compared to untransfected DU145 ( $\approx 380\,000 \pm 40\,000$  cells) 96 h after the initial plating.



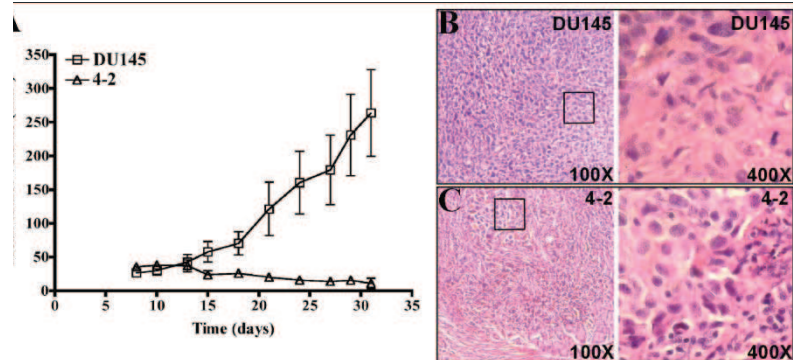
**Figure 4 : Proliferation of DU145 cells with reduced PACE4 expression.** A. MTT proliferation assay with conditioned media. DU145 cells ( $5.0 \times 10^3$ /well in triplicate) were incubated either with RPMI, concentrated RPMI (RPMI conc), DU145 conditioned medium (DU145 CM) or 4-2 CM for 48 h. After treatment, MTT solution was added for 4.5 h and  $OD_{550nm}-OD_{650nm}$  of solubilized cells were measured and compared using untreated cells (RPMI) as reference. Values are mean  $\pm$  SEM ( $n=7$ ). B. Time-dependant proliferation of untransfected DU145 cells (DU145) or expressing tRNA<sup>Val</sup>-PACE4-SOFA-HDVRz with or without partially re-established PACE4 expression (4-2 and 4-2+PACE4, respectively) was assessed by counting total cell number after 24, 48, 72 and 96 h incubation time. Values are mean  $\pm$  SEM ( $n=3$ ). C. Colony formation after 10 days incubation of DU145, 4-2 and 4-2+PACE4 cell lines. Culture plates with fixed and stained cells were scanned and the total particle area was measured. Typical result is illustrated for each cell line above in the upper part of the panel. The relative total area of the transfected cell lines compared to the control cell line is shown. Values are mean  $\pm$  SEM ( $n=9$  for DU145 and 4-2+PACE4 and  $n=7$  for 4-2). \* $P < 0.05$ .

This reduction was partially reversed in the cell line 4-2+PACE4 ( $\approx 280\,000 \pm 30\,000$  cells). We also performed an *in vitro* clonogenic assay on the same cell lines to detect the proportion of cells that retained the capacity to form colonies (**figure 4C**). The results of this assay correlates the proliferation of DU145 with PACE4 expression levels, as we observed a 68% reduction of the colony growth in 4-2 cells when compared to wild-type DU145 while this colony formation capacity was partially restored in 4-2+PACE4 cells (16% reduction when compared to the same cells).

#### PACE4 inhibition prevents tumor growth in xenograft tumor model.

We tested the ability of our experimental cell lines to grow as tumors in mouse model. We injected  $3 \times 10^6$  cells subcutaneously in nude mice flanks and measured tumor volumes one week post-injection. As shown in **figure 5A**, the reduction of PACE4 mRNA levels reduced dramatically the ability of 4-2 cells to induce tumor growth, while untransfected DU145 cells were able to develop into well-defined tumor masses.

After 31 days, subcutaneous tumors were eradicated from nude mice and stained with hematoxylin and eosin (**figure 5B and 5C**). At a lower magnification (100X), viable cells predominate in DU145 tumors, while 4-2 tumors showed only a small proportion of viable cells. The high proportion of necrotic cells in 4-2 tumors is shown at higher magnification (**figure 5C**, 400X), while this was not observed in DU145 tumors (**figure 5B**, 400X).



**Figure 5 : *In vivo* tumorigenicity assay.** A.  $3.0 \times 10^6$  DU145 and 4-2 cells were subcutaneously injected in four-week-old nu/nu female mice (2 injections/mice, 5 mice/group). The length and the width of the tumors were measured three times per week for 31 days. Values shown are mean  $\pm$  SEM of tumors volumes for all tumors measured per mice group. Morphologic patterns of DU145 (B) and 4-2 (C) tumors stained with hematoxylin and eosin at 100x and 400x magnification (boxes indicate the magnified regions).

#### Lower PACE4 expression induces the secretion of the proapoptotic factor TRPS1.

As the results observed earlier correlated PACE4 expression with the proliferation of DU145 cells, we decided to analyze the cell cycle of DU145, 4-2 and 4-2+PACE4 cells by flow cytometric analysis using the propidium iodide (PI) staining. As shown in **figure 6A**, the proportion of hypodiploid cells (characteristic of apoptotic cells) is higher in 4-2 cells (12.3%) than those found either in DU145 or 4-2+PACE4 cells (1.74% and 3.30%, respectively). Moreover, 78.8% of the non-apoptotic 4-2 cells were blocked into the G0/G1 phase of the cell cycle, and only 3.84% were into the G2-M phase. This is in contrast with the DU145 cells, where only 67.3% of non-apoptotic cells were in the G0-G1 phase, and 14.5% entered into the G2-M phase. However, 14.5% of the latest cells are in the G2-M phase, while only 3.84% of 4-2 cells are distributed in this phase. Thus, the lower cell proliferation observed in 4-2 cells can be attributed to a higher apoptosis rate and a blockade into the G0-G1 phase of the cell cycle. Interestingly, the study of the partial rescue cell line 4-2+PACE4 showed that a lower population of these cells were in apoptosis when compared to the

4-2 cells (3.30% and 12.3%, respectively), confirming an important role of PACE4 into this phenomenon. However, these cells with a partially restored PACE4 expression still seemed to be blocked into the G0/G1 phase of cell cycle as for the 4-2 cells (80.6% and 78.8% respectively) and like 4-2 cells, the 4-2+PACE4 cells had a low percentage of cells into the G2-M phase (3.61%). Thus, the partial re-establishment of PACE4 expression levels led to a lower apoptotic percentage of DU145 cells, but still results in a partial blockade of the cell cycle into the G0/G1 phase.

Since important differences in the cell population of DU145 and 4-2 cells were observed in regards to apoptosis and since PACE4 acts into the secretion pathways and into the extracellular environment of cells [10, 17], we sought for proapoptotic factors secreted by our cell lines by analyzing their secretome (**figure 6B**). We used triplicates of DU145 and 4-2 cells CM digested with trypsin to perform a MALDI-TOF-TOF analysis. We repeatedly found a biomarker of 1227.8 Da only into the 4-2 CM. The protein identification was then performed under the MASCOT sequence query search program (**figure 6B**). This biomarker corresponded to a fragment of the TRPS1 transcription factor, a protein issued from the apoptosis associated gene GC79/TRPS1 that encodes an atypical multitype zinc finger GATA-type transcription factor associated to prostate cancer that potently and specifically represses transcriptional activation mediated by other GATA factors [50, 51].

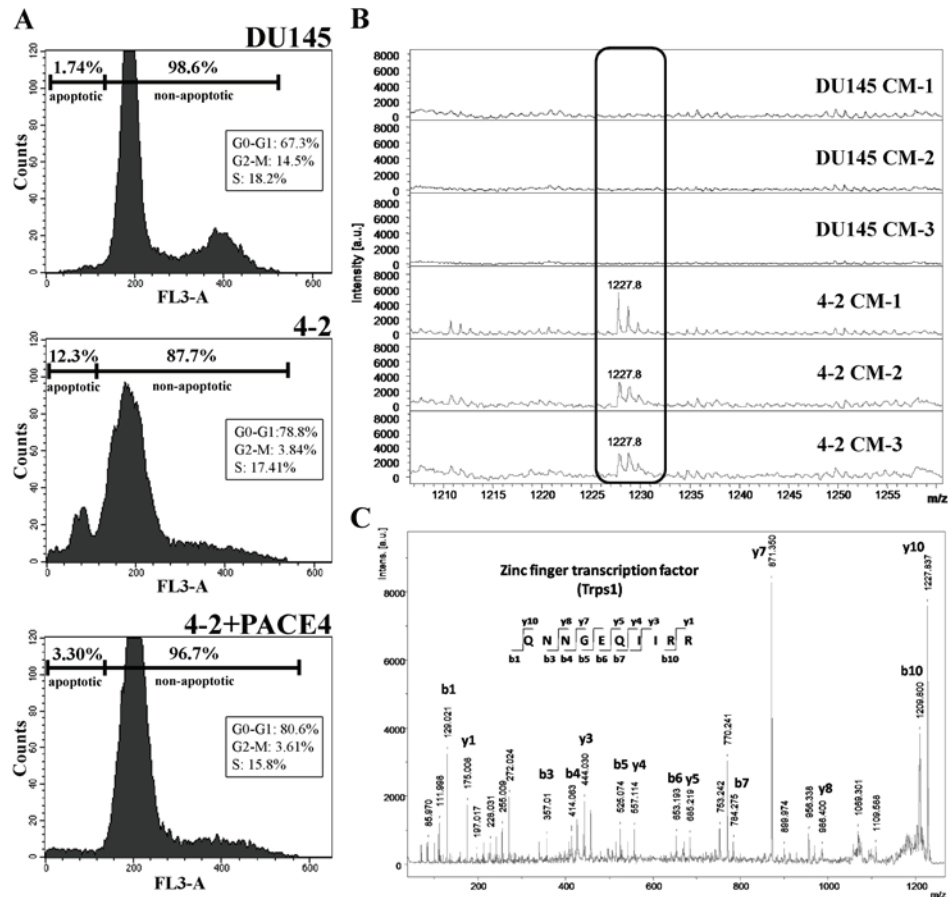
## DISCUSSION

The relationship between PCs and cancer has become stronger within the last few years. Since cancer

cell lines generally co-express multiple PCs, it still remains unclear whether one PC is more important or whether the cell simply establish multiple PC overexpression to assure redundancy. In this study, we report the overexpression of PACE4 in different clinical stages prostate cancer tissues, arguing for a specific contribution of this PC to the establishment or the maintenance of prostate cancer (**figure 1**). Moreover, six independent studies reported the selective PACE4 overexpression in prostate cancer (**table I**), arguing for the importance of this enzyme to sustain this disease. This finding now defines PACE4 as a potential therapeutic target for prostate cancer, since the use of a general PC inhibitor previously underlined the implication of PCs for the regulation of prostate epithelial cell differentiation [52].

Despite the fact that PC activity is essential to activate cancer-associated protein, it is impossible to predict which member of the PC family will cleave a given substrate in a cellular context. Since *in vitro* cleavage assays of a PC for a given substrate do not reproduce the cellular environment (including subcellular localization and the relative

amounts of each molecule), the results obtained may not reflect the actual cellular processing events. Therefore, targeted inhibition studies in tumoral cell lines with endogenously high expression levels of PCs are useful to understand the specific contribution of these enzymes into the generation of cancer related proteins, although functional redundancy might be observed for some substrates. For this study, we chose the androgen-independent prostate cancer cell line DU145. These cells were chosen for



**Figure 6 : Cell cycle analysis and identification of an apoptotic biomarker by MALDI-TOF-TOF.** *A.* DNA content of DU145, 4-2 and 4-2+PACE4 cells was determined by flow cytometric analysis using PI staining without cell fixation. The 4-2 cells showed a higher level of fragmented DNA than DU145 and 4-2+PACE4 cells (12.30%, 1.74% and 3.30%, respectively), indicating a higher population of hypodiploid cells for the 4-2 cell line, a characteristic of apoptotic cells. Insets show the distribution of non apoptotic cells into the phases of cell cycle. *B-C.* The detection and identification of the apoptotic biomarker TRPS1 was done by MALDI-TOF-TOF in positive reflectron mode from different secretomes coming from triplicates of DU145 and 4-2 conditioned culture media after trypsin digestions. The panel *B* represents the mass spectra comparison between conditioned media from both cell lines and the box shows a biomarker with a m/z of 1228 in 4-2 CM. The panel *C* shows the identification by fragmentation of this peak as the Zinc finger transcription factor TRPS1 by MALDI-TOF-TOF.

their high level of PACE4 expression, aggressiveness and facility to be transfected when compared to the androgen-dependant LNCaP or the androgen-independant PC-3 cells (unpublished data). After PACE4-SOFA-HDVRz transfection, a very low number of stable cells were analyzed, since transfected DU145 cells grew very slowly. This may be the consequence of lowered PACE4 level, thus arguing for the important role of this PC for DU145 cells proliferation. This link between PACE4 and cell proliferation could explain why we did not get clones with a lower expression levels, as previously observed in a

similar study using HDVRz to inhibit PC2 expression [37].

Considering the high selectivity potential of PACE4-SOFA-HDRz, we chose the cell line with the lowest levels of PACE4 mRNA levels for further studies (**figure 2C**). Northern blots performed for two other endogenous expressed PCs showed that this effect is specific to PACE4, although we cannot confirmed that the ribozyme activity is exclusive to the target RNA into this cell line (**figure 2D**). As a control, we transfected the line with lower PACE4 levels with an expression vector containing the PACE4 cDNA. We were surprised to see a re-established PACE4

RNA level, since this cell line still expressed the tRNA<sup>Val</sup>-PACE4-SOFA-HDRz (**figure 2C**). This suggests that within the cellular context, the catalytic activity of the ribozyme is not as high as required to affect large pools of mRNA issued from expression vectors. Nonetheless, this partially re-established cell line served our purpose as a control. We believed that this cell line was the best control to use since it is directly derived from cells used for all experiments, eliminating all possible cell line-to-cell line variations. However, further improvements of the SOFA-HDVRz are in development to reach the highest turnover without significantly affecting its specificity, thereby assuring an even better molecular inhibitor for further studies.

To confirm that the established cell line with lower PACE4 expression level was a suitable studying model, we undertook a comparative gene expression profiling study (**figure 3**). Among the downregulated genes, MUC1 is the one whose expression was the most affected by the PACE4 knockdown (supplemental **table I**). The mucin 1 is an oncoprotein expressed at high levels in diverse human carcinoma that is implicated in cellular transformation. Indeed, by its interactions with diverse effectors including EGFR,  $\beta$ -catenin, p53, I $\kappa$ N kinase and NF $\kappa$ B p65, MUC1 blocks the induction of apoptosis and necrosis in human carcinoma [53, 54]. A tissue microarray study reported the overexpression of MUC1 in 58% of human primary prostate cancer tissues [55], while a correlation of its expression level in human prostate carcinoma tissues with the most advanced stages of the disease was also reported [56]. The androgen-dependant prostate cancer cell line LNCaP do not express MUC1, while the androgen-independent PC-3 cells express low constitutive levels, while DU145 cells

highly and constitutively express MUC1 [53, 57]. Our study now strongly correlates MUC1 expression with PACE4 expression level. Interestingly, the inhibition of MUC1 in DU145 cells with a peptide inhibitor interfering with its oligomerization led to similar effects than those reported in this study for the 4-2 cell line with low PACE4 expression levels: inhibition of cell proliferation, induction of cell death pathways and tumor regression of DU145 xenografts growing in nude mice [53]. Thus, some of the effects reported in this study could be directly linked to the reduced MUC1 expression caused by PACE4 inhibition, arguing for a central role of this overexpressed PCs in prostate carcinoma. Although this hypothesis needs further investigations, the study of MUC1 transcription regulators could lead to a better understanding of PACE4 contribution to its expression control since the known transcriptional regulators of MUC1 include TNF- $\alpha$ , IL-6, IFN- $\gamma$ , TGF- $\beta$ 1 and HIF1- $\alpha$  [54]. In addition to TGF- $\beta$ 1, a known PC substrate, PACE4 could also regulate MUC1 expression indirectly by activating metalloproteinases such as ADAM17/TACE, which is directly involved into the regulation of TNF- $\alpha$  and IL-6 receptor [3].

The different subclasses of prostate tumor tissues found by gene expression profiling of prostate cancer tissues [48] allowed us to associate PACE4 expression to the highly aggressive tumor phenotype (**table II**). For example, some genes whose expression levels were either down- or upregulated solely in tumor subclasses II or III (e.g. MUC1, F3, RAB27B), did not show the same expression variation in the least aggressive subclass I tumors, thus correlating PACE4 expression to the higher aggressive tumors. On the other hand, genes whose expression in tumor subclass I varied in

the same way as in subclasses II and/or III might could be associated with a more general cellular defect common to all prostate cancer cells in which PACE4 could also be implicated (e.g. TACSTD1, CGN, MAL2, SERPINE2, SERPINB5, DPYSL3, PRKCA, GJA1, CRYZ, CYP1B1, SRPX). Thus, these microarray results confirmed that the cell line 4-2 is a suitable studying model for prostate cancer. However, considering the broad spectrum of potential endogenous substrates of PACE4, further investigations of its direct or indirect actions on the transcriptional control of these gene will be needed to better understand the gene expression profile of 4-2 cells.

The consequences of lowered levels of PACE4 are well illustrated by the reduced cell proliferation rate and the incapacity of these cells to form subcutaneous tumors in nude mice (**figures 4 and 5**, respectively). The partial restoration of PACE4 expression levels in this cell line allowed a partial recovery of the *in vitro* proliferation rate, suggesting that PACE4 is a key player for tumoral growth and its levels have to be high to achieve this function (**figure 4**). Moreover, the analysis of the cell cycle in the DU145, 4-2 and 4-2+PACE4 cell lines revealed the importance of high expression levels of PACE4 to block apoptosis and insure the progression through all the phases of the cell cycle (**figure 6A**). These results also indicated that partial recovery of PACE4 expression blocked DU145 apoptosis but could not reverse the blockade of cells into the G0-G1 phase, suggesting that full PACE4 activity is necessary to sustain the high proliferative state of DU145 cells.

To better understand the role of PACE4 into the proliferation of DU145 cells, we undertook proteomic analyses of DU145 and 4-2 cells derived conditioned media (**figure 6B**

**and 6C**). In the samples issued from 4-2 cells, a fragment derived from the tryptic digestion of TRPS1 (a multitype zinc finger GATA-type transcription factor) was consistently detected [50]. This transcriptional repressor, whose expression is regulated by androgens through the androgen receptor, is normally expressed in the androgen-dependant LNCaP prostate cancer cells [58, 59]. Its lack of expression has been previously reported for the androgen-independent DU145 prostate cancer cells [58], but stable expression of recombinant TRPS1 in those cells allowed the identification of both down-regulated the anti-apoptotic protein POR2 and up-regulated the pro-apoptotic protein KAD2 [60]. Our study now specifically involves PACE4 to this biological system.

Further studies will be needed to better characterize the involvement of PACE4 into the transcriptional control of the pro-apoptotic TRPS1 transcriptional repressor. As mentioned earlier, TRPS1 transcription is known to be regulated by androgens in cells LNCaP prostate cancer cells but not in the DU145 cell line, whose growth is independent of androgens and do not express the androgen receptor [58]. Recently, the bone morphogenic protein 7 (BMP-7), a growth factor belonging to the family of TGF- $\beta$  that necessitates a PC cleavage for its propeptide removal [61], have been identified as an inducer of TRPS1 gene expression in mouse kidneys [62]. Thus, further characterization of the TRPS1 genetic regulation by BMP-7 in DU145, or other members of the TGF- $\beta$  growth factor family that also need furin-like cleavages for their processing [3], would be indicated to better understand the cellular functions of the overexpressed PACE4 in human prostate cancer cells.

The actual genes known to be transcriptionally regulated by

TRPS1 include prostate-specific antigen [59], STAT3 [63], parathyroid hormone-related protein (PTHrP) [64] and osteocalcin [65]. Interestingly, the cleavage of human PTHrP and osteocalcin propeptides, essential for the generation of their mature forms, occurs at typical PCs processing sites (<sub>31</sub>SRRLKR↓AV<sub>38</sub> and <sub>46</sub>VKRPRR↓YL<sub>51</sub>, respectively). Thus the implication PACE4 catalytic activity to the transcriptional control of TRPS1 and DU145 apoptosis triggering could be through these potential substrates. This hypothesis is well supported by our genomic analysis, where the gene PTHLH (coding for PTHrP) was upregulated in 4-2 cells (supplemental **table II**).

In conclusion, this study reports the unique overexpression of PACE4 in prostate cancer tissues. Our results showed that PACE4 plays a major role into the progression of prostate tumors to a higher aggressive status and its inhibition leads to a significant loss of tumorigenicity of the prostate tumor cell line DU145. A better understanding of PACE4 direct and indirect cellular targets contributing to these phenomena remains a great challenge and implicates multiple known oncogenic factors.

#### ACKNOWLEDGMENTS

Authors want to thank Mrs. Xue-Wen Yuan for the helpful technical assistance. This work was funded by a grant from the Canadian Institutes of Health Research (CIHR; to R.D. and J.P.P.), from the Centre National de la Recherche Scientifique, Département de la politique industrielle (to M.S. and I.F.) and from the Ministère de l'Éducation Nationale, de l'Enseignement Supérieur et de la Recherche, Agence National de la recherche (I.F.). Scholarships from

Natural Sciences and Engineering Research Council of Canada (F.D. and S.R.), from Fonds de la Recherche en Santé du Québec (F.D.) and from Région Nord-Pas de Calais (D.B.), were also attributed. J.P.P. holds the Canada Research Chair in Genomic and Catalytic RNA. R.D. and J.P.P. are members of the Centre de Recherche Clinique Etienne-Le Bel (Sherbrooke, Qc, Canada).

#### REFERENCES

1. Hanahan, D. and R.A. Weinberg, *The hallmarks of cancer*. Cell, 2000. **100**(1): p. 57-70.
2. Bassi, D.E., et al., *Proprotein convertases: "master switches" in the regulation of tumor growth and progression*. Mol Carcinog, 2005. **44**(3): p. 151-61.
3. Khatib, A.M., et al., *Proprotein convertases in tumor progression and malignancy: novel targets in cancer therapy*. Am J Pathol, 2002. **160**(6): p. 1921-35.
4. Blanchard, A., et al., *Targeted production of proprotein convertase PC1 enhances mammary development and tumorigenesis in transgenic mice*. Can J Physiol Pharmacol, 2009. **87**(10): p. 831-8.
5. de Cicco, R.L., et al., *Inhibition of proprotein convertases: Approaches to block squamous carcinoma development and progression*. Mol Carcinog, 2007. **46**(8): p. 654-9.
6. Page, R.E., et al., *Increased expression of the pro-protein convertase furin predicts decreased survival in ovarian cancer*. Cell Oncol, 2007. **29**(4): p. 289-99.



7. Scamuffa, N., et al., *Selective inhibition of proprotein convertases represses the metastatic potential of human colorectal tumor cells*. J Clin Invest, 2008. **118**(1): p. 352-63.
8. Bassi, D.E., et al., *Proprotein convertase inhibition results in decreased skin cell proliferation, tumorigenesis, and metastasis*. Neoplasia, 2010. **12**(7): p. 516-26.
9. Fugere, M. and R. Day, *Cutting back on pro-protein convertases: the latest approaches to pharmacological inhibition*. Trends Pharmacol Sci, 2005. **26**(6): p. 294-301.
10. Seidah, N.G., et al., *The activation and physiological functions of the proprotein convertases*. Int J Biochem Cell Biol, 2008. **40**(6-7): p. 1111-25.
11. Clark, D.A., et al., *Protease inhibitors suppress in vitro growth of human small cell lung cancer*. Peptides, 1993. **14**(5): p. 1021-8.
12. Mbikay, M., et al., *Comparative analysis of expression of the proprotein convertases furin, PACE4, PC1 and PC2 in human lung tumours*. Br J Cancer, 1997. **75**(10): p. 1509-14.
13. Cheng, M., et al., *Pro-protein convertase gene expression in human breast cancer*. Int J Cancer, 1997. **71**(6): p. 966-71.
14. Tzimas, G.N., et al., *Abnormal expression and processing of the proprotein convertases PC1 and PC2 in human colorectal liver metastases*. BMC Cancer, 2005. **5**: p. 149.
15. Bassi, D.E., et al., *Elevated furin expression in aggressive human head and neck tumors and tumor cell lines*. Mol Carcinog, 2001. **31**(4): p. 224-32.
16. Fu, Y., et al., *Epigenetic regulation of proprotein convertase PACE4 gene expression in human ovarian cancer cells*. Mol Cancer Res, 2003. **1**(8): p. 569-76.
17. Bassi, D.E., et al., *PACE4 expression in mouse basal keratinocytes results in basement membrane disruption and acceleration of tumor progression*. Cancer Res, 2005. **65**(16): p. 7310-9.
18. Bassi, D.E., et al., *Increased furin activity enhances the malignant phenotype of human head and neck cancer cells*. Am J Pathol, 2003. **162**(2): p. 439-47.
19. Hubbard, F.C., et al., *Expression of PACE4 in chemically induced carcinomas is associated with spindle cell tumor conversion and increased invasive ability*. Cancer Res, 1997. **57**(23): p. 5226-31.
20. Mercapide, J., et al., *Inhibition of furin-mediated processing results in suppression of astrocytoma cell growth and invasiveness*. Clin Cancer Res, 2002. **8**(6): p. 1740-6.
21. Remacle, A.G., et al., *Furin regulates the intracellular activation and the uptake rate of cell surface-associated MT1-MMP*. Oncogene, 2006. **25**(41): p. 5648-55.
22. Coppola, J.M., et al., *A small-molecule furin inhibitor inhibits cancer cell motility and invasiveness*. Neoplasia, 2008. **10**(4): p. 363-70.
23. McColl, B.K., et al., *Proprotein convertases promote processing of VEGF-D, a critical step for binding the angiogenic receptor VEGFR-2*. FASEB J, 2007. **21**(4): p. 1088-98.

24. Basak, A., et al., *A novel enediynyl peptide inhibitor of furin that blocks processing of proPDGF-A, B and proVEGF-C*. PLoS One, 2009. **4**(11): p. e7700.
25. Siegfried, G., et al., *Regulation of the stepwise proteolytic cleavage and secretion of PDGF-B by the proprotein convertases*. Oncogene, 2005. **24**(46): p. 6925-35.
26. Dubois, C.M., et al., *Evidence that furin is an authentic transforming growth factor-beta1-converting enzyme*. Am J Pathol, 2001. **158**(1): p. 305-16.
27. Basak, A., *Inhibitors of proprotein convertases*. J Mol Med, 2005. **83**(11): p. 844-55.
28. Komiyama, T., et al., *Inhibition of furin/proprotein convertase-catalyzed surface and intracellular processing by small molecules*. J Biol Chem, 2009. **284**(23): p. 15729-38.
29. Becker, G.L., et al., *Potent inhibitors of furin and furin-like proprotein convertases containing decarboxylated P1 arginine mimetics*. J Med Chem, 2010. **53**(3): p. 1067-75.
30. Coppola, J.M., et al., *Identification of inhibitors using a cell-based assay for monitoring Golgi-resident protease activity*. Anal Biochem, 2007. **364**(1): p. 19-29.
31. Tafech, A., et al., *Destroying RNA as a therapeutic approach*. Curr Med Chem, 2006. **13**(8): p. 863-81.
32. Chene, L., et al., *Extensive analysis of the 7q31 region in human prostate tumors supports TES as the best candidate tumor suppressor gene*. Int J Cancer, 2004. **111**(5): p. 798-804.
33. Mellinger, G.T., D. Gleason, and J. Bailar, 3rd, *The histology and prognosis of prostatic cancer*. J Urol, 1967. **97**(2): p. 331-7.
34. Schroder, F.H., et al., *The TNM classification of prostate cancer*. Prostate Suppl, 1992. **4**: p. 129-38.
35. Dong, W., et al., *Cellular localization of the prohormone convertases in the hypothalamic paraventricular and supraoptic nuclei: selective regulation of PC1 in corticotrophin-releasing hormone parvocellular neurons mediated by glucocorticoids*. J Neurosci, 1997. **17**(2): p. 563-75.
36. Lucier, J.F., et al., *RiboSubstrates: a web application addressing the cleavage specificities of ribozymes in designated genomes*. BMC Bioinformatics, 2006. **7**: p. 480.
37. D'Anjou, F., et al., *Silencing of SPC2 expression using an engineered delta ribozyme in the mouse betaTC-3 endocrine cell line*. J Biol Chem, 2004. **279**(14): p. 14232-9.
38. Creemers, J.W., et al., *Proprotein processing activity and cleavage site selectivity of the Kex2-like endoprotease PACE4*. FEBS Lett, 1993. **336**(1): p. 65-9.
39. Denault, J.B., et al., *Comparative characterization of two forms of recombinant human SPC1 secreted from Schneider 2 cells*. Protein Expr Purif, 2000. **19**(1): p. 113-24.
40. Seidah, N.G., et al., *cDNA structure, tissue distribution, and chromosomal localization of rat PC7, a novel mammalian proprotein convertase closest to yeast kexin-like proteinases*. Proc Natl Acad Sci U S A, 1996. **93**(8): p. 3388-93.

41. Day, R., et al., *Distribution and regulation of the prohormone convertases PC1 and PC2 in the rat pituitary*. Mol Endocrinol, 1992. **6**(3): p. 485-97.
42. Tusher, V.G., R. Tibshirani, and G. Chu, *Significance analysis of microarrays applied to the ionizing radiation response*. Proc Natl Acad Sci U S A, 2001. **98**(9): p. 5116-21.
43. Tsuji, A., et al., *Engineering of alpha1-antitrypsin variants selective for subtilisin-like proprotein convertases PACE4 and PC6: importance of the P2' residue in stable complex formation of the serpin with proprotein convertase*. Protein Eng Des Sel, 2007. **20**(4): p. 163-70.
44. Bergeron, L.J. and J.P. Perreault, *Target-dependent on/off switch increases ribozyme fidelity*. Nucleic Acids Res, 2005. **33**(4): p. 1240-8.
45. Bergeron, L.J., C. Reymond, and J.P. Perreault, *Functional characterization of the SOFA delta ribozyme*. RNA, 2005. **11**(12): p. 1858-68.
46. Posthaus, H., C.M. Dubois, and E. Muller, *Novel insights into cadherin processing by subtilisin-like convertases*. FEBS Lett, 2003. **536**(1-3): p. 203-8.
47. Nakajima, T., et al., *Prohormone convertase furin has a role in gastric cancer cell proliferation with parathyroid hormone-related peptide in a reciprocal manner*. Dig Dis Sci, 2002. **47**(12): p. 2729-37.
48. Lapointe, J., et al., *Gene expression profiling identifies clinically relevant subtypes of prostate cancer*. Proc Natl Acad Sci U S A, 2004. **101**(3): p. 811-6.
49. Mahloogi, H., D.E. Bassi, and A.J. Klein-Szanto, *Malignant conversion of non-tumorigenic murine skin keratinocytes overexpressing PACE4*. Carcinogenesis, 2002. **23**(4): p. 565-72.
50. Chang, G.T., et al., *Structure and function of GC79/TRPS1, a novel androgen-repressible apoptosis gene*. Apoptosis, 2002. **7**(1): p. 13-21.
51. Malik, T.H., et al., *Transcriptional repression and developmental functions of the atypical vertebrate GATA protein TRPS1*. EMBO J, 2001. **20**(7): p. 1715-25.
52. Uchida, K., et al., *Proprotein convertases regulate activity of prostate epithelial cell differentiation markers and are modulated in human prostate cancer cells*. J Cell Biochem, 2003. **88**(2): p. 394-9.
53. Joshi, M.D., et al., *MUC1 oncoprotein is a druggable target in human prostate cancer cells*. Mol Cancer Ther, 2009. **8**(11): p. 3056-65.
54. Jonckheere, N. and I. Van Seuning, *The membrane-bound mucins: From cell signalling to transcriptional regulation and expression in epithelial cancers*. Biochimie, 2010. **92**(1): p. 1-11.
55. Cozzi, P.J., et al., *MUC1, MUC2, MUC4, MUC5AC and MUC6 expression in the progression of prostate cancer*. Clin Exp Metastasis, 2005. **22**(7): p. 565-73.
56. Kirschenbaum, A., et al., *MUC1 Expression in Prostate Carcinoma: Correlation with Grade and Stage*. Mol Urol, 1999. **3**(3): p. 163-168.
57. O'Connor, J.C., et al., *MUC1 expression in human prostate*

- cancer cell lines and primary tumors. Prostate Cancer Prostatic Dis*, 2005. **8**(1): p. 36-44.
58. Chang, G.T., et al., *The TRPS1 transcription factor: androgenic regulation in prostate cancer and high expression in breast cancer*. *Endocr Relat Cancer*, 2004. **11**(4): p. 815-22.
59. van den Bemd, G.J., et al., *The atypical GATA protein TRPS1 represses androgen-induced prostate-specific antigen expression in LNCaP prostate cancer cells*. *Biochem Biophys Res Commun*, 2003. **312**(3): p. 578-84.
60. Chang, G.T., et al., *Proteomic analysis of proteins regulated by TRPS1 transcription factor in DU145 prostate cancer cells*. *Biochim Biophys Acta*, 2007. **1774**(5): p. 575-82.
61. Swencki-Underwood, B., et al., *Expression and characterization of a human BMP-7 variant with improved biochemical properties*. *Protein Expr Purif*, 2008. **57**(2): p. 312-9.
62. Gai, Z., et al., *Trps1 functions downstream of Bmp7 in kidney development*. *J Am Soc Nephrol*, 2009. **20**(11): p. 2403-11.
63. Suemoto, H., et al., *Trps1 regulates proliferation and apoptosis of chondrocytes through Stat3 signaling*. *Dev Biol*, 2007. **312**(2): p. 572-81.
64. Nishioka, K., et al., *Trps1 deficiency enlarges the proliferative zone of growth plate cartilage by upregulation of Pthrp*. *Bone*, 2008. **43**(1): p. 64-71.
65. Piscopo, D.M., E.B. Johansen, and R. Derynck, *Identification of the GATA factor TRPS1 as a repressor of the osteocalcin promoter*. *J Biol Chem*, 2009. **284**(46): p. 31690-703.

## SUPPLEMENTAL DATA

Supplementary table I. Downregulated genes in 4-2 cells compared to DU145

Accession numbers	Gene symbol	Gene description	Fold change DU145 vs. 4-2	P-value
NM_002456	MUC1	mucin 1	0,080	0,00123
NM_030631	SLC25A21	solute carrier family 25 (mitochondrial oxodicarboxylate carrier), member 21	0,118	0,00037
NM_021034	IFITM3	interferon induced transmembrane protein 3 (1-8U)	0,121	0,00038
NM_006435	IFITM2	interferon induced transmembrane protein 2 (1-8D)	0,142	0,00109
NM_004613	TGM2	transglutaminase 2	0,182	0,00043
NM_013230	CD24	CD24	0,184	0,02952
NM_019601	SUSD2	sushi domain containing 2	0,186	0,01475
NM_004657	SDPR	serum deprivation response	0,209	0,01243
XM_938803	CST6	cystatin E/M	0,282	0,02171
NM_052886	MAL2	mal, T-cell differentiation protein 2	0,302	0,00003
NM_001430	EPAS1	endothelial PAS domain protein 1	0,313	0,00121
NM_000101	CYBA	cytochrome b-245, alpha polypeptide	0,314	0,02607
NM_001038	SCNN1A	sodium channel, nonvoltage-gated 1 alpha	0,317	0,00156
NM_032899	FAM83A	family with sequence similarity 83, member A	0,319	0,00745
NM_000211	ITGB2	integrin, beta 2	0,324	0,00416
NM_001386	DPYSL2	dihydropyrimidinase-like 2	0,325	0,00857
NM_003280	TNNC1	troponin C type 1 (slow)	0,328	0,03211
NM_001823	CKB	creatine kinase, brain	0,347	0,01088
NM_174911	FAM84B	family with sequence similarity 84, member B	0,353	0,00069
NM_001793	CDH3	cadherin 3	0,354	0,00003
NM_007257	PNMA2	paraneoplastic antigen MA2	0,358	0,00459
NM_021135	RPS6KA2	ribosomal protein S6 kinase	0,367	0,00103
NM_020859	SHRM	shroom	0,369	0,00973
NM_199001	MGC59937	Similar to RIKEN cDNA 2310002J15 gene	0,374	0,00505
NM_001003794	MGLL	monoglyceride lipase	0,375	0,00604
NM_003246	THBS1	thrombospondin 1	0,394	0,03839
NM_004360	CDH1	cadherin 1	0,399	0,00005
NM_002754	MAPK13	mitogen-activated protein kinase 13	0,400	0,00182
NM_002354	TACSTD1	tumor-associated calcium signal transducer 1	0,402	0,00197
NM_173584	MGC45840	EF-hand calcium binding domain 4A	0,406	0,00659
AK027147	HS.509165	cDNA: FLJ23494 fis, clone LNG01885	0,407	0,00691
NM_014578	RHOD	ras homolog gene family, member D	0,409	0,00604
NM_182498	C19ORF37	zinc finger protein 428	0,416	0,00178
XM_051862	LOC58489	hypothetical protein from EUROIMAGE 588495	0,416	0,02740
NM_020210	SEMA4B	semaphorin 4B	0,417	0,02468
NM_020770	CGN	cingulin	0,419	0,04416
NM_024329	EFHD2	EF-hand domain family, member D2	0,430	0,00365
NM_019058	DDIT4	DNA-damage-inducible transcript 4	0,431	0,03417
NM_024830	AYTL2	acyltransferase like 2	0,437	0,00391
NM_198282	TMEM173	transmembrane protein 173	0,437	0,00316
NM_005630	SLCO2A1	solute carrier organic anion transporter family, member 2A1	0,437	0,00219
NM_015055	SWAP70	SWAP-70 protein	0,442	0,00778
NM_021101	CLDN1	claudin 1	0,443	0,01301
NM_017917	PPP2R3C	protein phosphatase 2, regulatory subunit B", gamma	0,444	0,00839
NM_005195	CEBPD	CCAAT/enhancer binding protein	0,457	0,03235
NM_018478	C20ORF35	chromosome 20 open reading frame 35	0,461	0,01029
NM_021818	SAV1	salvador homolog 1	0,467	0,00302
NM_000600	IL6	interleukin 6	0,467	0,04991
NM_014220	TM4SF1	transmembrane 4 L six family member 1	0,479	0,00002
BU536065	HS.579631	cDNA clone IMAGE:6563923 5, mRNA sequence	0,485	0,03222
NM_003475	RASSF7	Ras association (RalGDS/AF-6) domain family 7	0,488	0,00515
NM_001013398	IGFBP3	insulin-like growth factor binding protein 3	0,490	0,02959
NM_025151	RAB11FIP1	RAB11 family interacting protein 1 (class I)	0,490	0,04771
NM_033317	ZD52F10	dermokine	0,493	0,00677
NM_001993	F3	coagulation factor III	0,494	0,00051
NM_006169	NNMT	nicotinamide N-methyltransferase	0,495	0,00990
NM_003726	SCAP1	src family associated phosphoprotein 1	0,496	0,00242
XM_944693	ITGB5	integrin, beta 5	0,496	0,00388
NM_018098	ECT2	epithelial cell transforming sequence 2 oncogene	0,498	0,00104

Supplementary table II. Upregulated genes in 4-2 cells compared to DU145

Accession numbers	Gene symbol	Gene description	Fold change DU145 vs. 4-2	P-value
NM_006216	SERPINE2	serpin peptidase inhibitor, clade E, member 2	5,140	0,02748
NM_005242	F2RL1	coagulation factor II (thrombin) receptor-like 1	5,119	0,00063
NM_014689	DOCK10	dedicator of cytokinesis 10	4,418	0,00296
NM_178012	TUBB2B	tubulin, beta 2B	4,330	0,00075
NM_024101	MLPH	melanophilin	4,310	0,00118
NM_003937	KYNU	kynureninase	3,722	0,03011
NM_002281	KRTHB1	keratin 81	3,623	0,01352
NM_003739	AKR1C3	aldo-keto reductase family 1, member C3	3,488	0,00293
NM_177949	ARMCX2	armadillo repeat containing, X-linked 2	3,079	0,00010
NM_002522	NPTX1	neuronal pentraxin I	2,988	0,00034
NM_024016	HOXB8	homeo box B8	2,911	0,00218
NM_000104	CYP1B1	cytochrome P450, family 1, subfamily B, polypeptide 1	2,866	0,00415
NM_002380	MATN2	matrilin 2	2,854	0,00584
NM_000165	GJA1	gap junction protein, alpha 1	2,833	0,00032
NM_020299	AKR1B10	aldo-keto reductase family 1, member B10	2,830	0,00556
NM_153344	C6ORF141	chromosome 6 open reading frame 141	2,727	0,01479
NM_002228	JUN	jun oncogene	2,655	0,00128
NM_004163	RAB27B	clone 25194 mRNA sequence	2,615	0,01647
NM_004398	DDX10	DEAD (Asp-Glu-Ala-Asp) box polypeptide 10	2,593	0,00221
NM_000782	CYP24A1	cytochrome P450, family 24, subfamily A, polypeptide 1	2,504	0,00226
NM_001259	CDK6	cyclin-dependent kinase 6	2,461	0,00582
NM_020919	ALS2	amyotrophic lateral sclerosis 2 (juvenile)	2,439	0,00248
NM_001964	EGR1	early growth response 1	2,438	0,01511
AF131834	HS.4892	clone 24841	2,428	0,00563
NM_003543	HIST1H4H	histone cluster 1, H4h	2,413	0,01138
NM_002639	SERPINB5	serpin peptidase inhibitor, clade B, member 5	2,406	0,00149
NM_012419	RGS17	regulator of G-protein signalling 17	2,351	0,01676
NM_032784	RSPO3	R-spondin 3 homolog	2,348	0,03649
XM_928153	LOC653571	similar to sperm protein associated with the nucleus, X chromosome, family member B1	2,329	0,00901
NM_002737	PRKCA	protein kinase C, alpha	2,326	0,00476
NM_001001390	CD44	CD44 molecule (Indian blood group)	2,306	0,01926
XM_940680	LOC648517	similar to Aldo-keto reductase family 1 member C1	2,269	0,01923
NM_003528	HIST2H2BE	histone cluster 2, H2be	2,264	0,00707
NM_003714	STC2	stanniocalcin 2	2,261	0,02278
NM_012242	DKK1	dickkopf homolog 1	2,249	0,000873
NM_020547	AMHR2	anti-Mullerian hormone receptor, type II	2,213	0,00011
AK091904	HS.202577	cDNA FLJ34585 fis, clone KIDNE2008758	2,202	0,00825
NM_198966	PTH1H	parathyroid hormone-like hormone	2,200	0,00008
NM_015508	TIPARP	TCDD-inducible poly(ADP-ribose) polymerase	2,190	0,02173
NM_001236	CBR3	carbonyl reductase 3	2,175	0,00163
NM_001387	DPYSL3	dihydropyrimidinase-like 3	2,165	0,01420
NM_025145	C10ORF79	chromosome 10 open reading frame 79	2,158	0,00026
NM_001889	CRYZ	crystallin, zeta (quinone reductase)	2,151	0,01170
NM_004403	DFNA5	deafness, autosomal dominant 5	2,141	0,00460
NM_005319	HIST1H1C	histone cluster 1, H1c	2,110	0,00689
NM_001878	CRABP2	cellular retinoic acid binding protein 2	2,094	0,01508
NM_001010924	C10ORF38	chromosome 10 open reading frame 38	2,092	0,00581
NM_000475	NR0B1	nuclear receptor subfamily 0, group B, member 1	2,084	0,00179
NM_004615	TSPAN7	tetraspanin 7	2,084	0,00504
NM_000358	TGFB1	transforming growth factor, beta-induced	2,057	0,02546
NM_005329	HAS3	hyaluronan synthase 3	2,051	0,03194
NM_032168	WDR75	WD repeat domain 75	2,020	0,01148
NM_001008490	KLF6	Kruppel-like factor 6	2,018	0,00732
NM_006120	HLA-DMA	major histocompatibility complex, class II, DM alpha	2,018	0,02250
XM_937100	KRTAP2-1	keratin associated protein 2-1	2,013	0,01978
NM_002178	IGFBP6	insulin-like growth factor binding protein 6	2,009	0,01644
NM_005682	GPR56	G protein-coupled receptor 56	2,009	0,01484
NM_006307	SRPX	sushi-repeat-containing protein, X-linked	2,009	0,00006
NM_002147	HOXB5	homeo box B5	2,006	0,00202
NM_015642	ZBTB20	zinc finger and BTB domain containing 20	2,005	0,01040
NM_020411	XAGE1	X antigen family, member 1D	2,002	0,02842
NM_016352	CPA4	carboxypeptidase A4	2,000	0,03194

## Discussion et perspectives

### 4.1.1 Choix du modèle cellulaire

L'obtention de lignées cellulaires humaines d'adénocarcinome de prostate est particulièrement difficile. Le taux de succès est d'environ 1 % pour les lignées cellulaires susceptibles d'être propagées *in vitro* et d'environ 10 % pour l'obtention de xénogreffes sur l'animal [297]. Habituellement, les cellules tumorales issues de biopsies s'attachent au support plastique et se divisent pendant quelques jours ou quelques semaines avant de mourir par apoptose. Ainsi, seule une dizaine de lignées cellulaires de cancer de prostate humain est disponible (contre plus de 100 pour le cancer du sein, par exemple), ces lignées sont résumées dans le **tableau 16**. Une étude récente a montré que six lignées de cancer de la prostate étaient réellement uniques, tandis que quatre paires de lignées dérivait d'une même origine [298]. Par ailleurs, toutes ces lignées ont été obtenues à partir de métastases et toutes les tentatives d'obtention de lignées à partir d'adénocarcinome de prostate localisé sur l'organe se sont soldées par des échecs.

Lignée	Site métastatique d'origine	Sensibilité aux androgènes	Expression du PSA	Tumorigénicité chez la souris
LNCaP	ganglion	+	+	+ /-
PC3	os	-	-	+
DU145	cerveau	-	-	+
TSU-PR1	ganglion	-	-	+
ARCaP	ascite	+ (faible)	+ (faible)	+
MDA PCa2a	os	+	+	+
MDA PCa2b	os	+	+	+
C4	dérivée de LNCaP	-	+	+
C4-2	dérivée de LNCaP	-	+	+
DuCaP	dure-mère	+	+	+
VCaP	os	+	+	+

Tableau 18 : Lignées cellulaires humaines de cancer de la prostate [299]

La lignée DU145 a été obtenue à partir d'une métastase cérébrale chez un patient atteint d'un cancer de la prostate mais également d'une leucémie [300]. Suivant la situation clinique étudiée, plusieurs modèles cellulaires, animaux ou de xénogreffes humaines existent et permettent de se rapprocher le plus possible des conditions physiologiques pathologiques.

Le **tableau 17** résume les principaux modèles précliniques de cancer de la prostate actuellement utilisés.

<b>Situation clinique</b>	<b>Modèles disponibles</b>
Lésions précancéreuses	-Très peu de modèles -Souris <i>knock-out</i> RXR alpha
Cancer de prostate localisé	-Pas de lignée disponible -Modèles murins orthotopiques (LNCaP et dérivées, PC3, MDA PCa2b, ARCaP)
Rechute biologique isolée	Pas de modèle disponible
Métastases osseuses	-Injections intrafémorales (LNCaP et dérivées, PC3, MDA PCa2b, LuCaP23, LuCaP35) -Injections intracardiaques (PC3 et dérivées, DuCaP, VCaP) -Injections orthotopiques de cellules C4-2B ( <i>LNCaP progression model</i> ) : débattu -Modèle des fragments osseux humains chez la souris SCID -Modèle des fragments de prostate canine à proximité d'un os chez la <i>nude</i> ?
Étude des relations cellules tumorales- ostéoblastes	Modèle de coculture MDA PCa2b et ostéoblastes murins primaires
Formes hormonosensibles	-Lignées LNCaP, MDA PCa2a, MDA PCa2b, DuCaP, VCaP -Xénogreffes LuCaP, LAPC4
Formes hormonoréfractaires	-Lignées PC3, DU145, C4-2, TSU-PR1 -Xénogrefe LAPC3
Étude de la transition vers l'hormonorésistance	-Modèle <i>Dunningrat</i> -Xénogreffes CWR22R, LuCaP23, LAPC4
Étude de l'oncogenèse prostatique	<i>Mouse prostatere constitution</i>
Étude de l'interaction avec le stroma	Lignée C4-2 <i>Mouse prostate reconstitution</i>
Étude du traitement hormonal intermittent	Souris Shionogi, modèle LNCaP

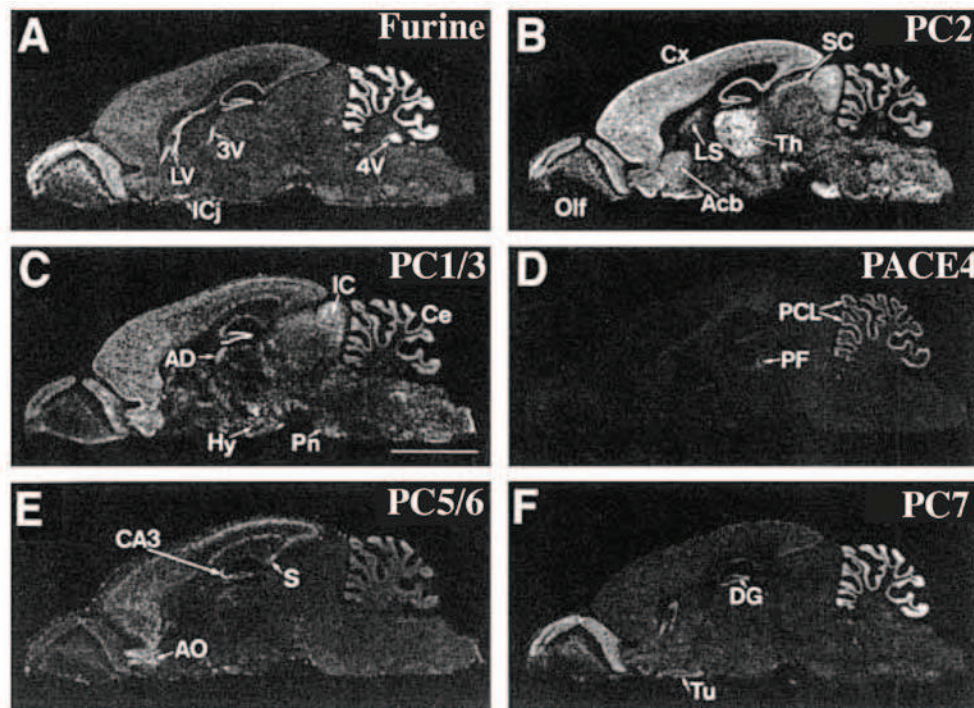
**Tableau 19 : Principaux modèles précliniques de cancer de la prostate actuellement utilisés [299]**



#### 4.1.2 Rôle de PACE4 dans le cancer de la prostate.

Dans notre cas, le modèle cellulaire choisi par le Dr François d'Anjou était les cellules DU145, qui expriment PACE4 de façon endogène. Il a analysé les conséquences d'une diminution de PACE4 sur la prolifération cellulaire en créant une lignée Kock-down pour la PACE4 dans les DU145, qu'il a appelé DU145/4-2. En effet, il a été démontré que les cellules tumorales de prostates androgène-indépendantes se différencient en cellules neuroendocriniennes, dont l'action paracrine de facteurs de croissance sécrétés pourrait contribuer au phénotype plus agressif de ce type de cancer de la prostate, incluant une stimulation de la prolifération, de l'invasion, de l'angiogenèse et une résistance à l'apoptose [301, 302]. Sachant que PACE4 est la seule PC surexprimée dans les tissus tumoraux de prostates analysé, cette cible était indiquée pour cette étude dans les DU145. Ainsi, il a pu démontrer qu'une diminution des niveaux de PACE4 entraînait une diminution de la prolifération des cellules cancéreuses.

Comme détaillé dans le chapitre généralités, les PCs possèdent un large éventail de substrats qui peuvent être impliqués dans de nombreuses pathologies. Un dérèglement de l'expression d'une PC peut donc être lourd de conséquence pour l'homéostasie d'un organe ou même d'un individu. Une question légitime est de savoir si le rôle physiologique de chaque PC est spécifique ou si ces protéases agissent de façon complémentaire (redondance). Au vue de la co-localisation de certaines PCs dans un organe (**Figure 47**), on peut penser que ces dernières agissent ensemble sous la forme d'un « cocktail » de protéase pour aboutir au rôle physiologique voulu. Néanmoins, seule PACE4 est surexprimée dans le cancer de la prostate et nous avons clairement établi que son action est primordiale pour que la cellule cancéreuse maintienne son état pathologique. Ceci laisse penser que suivant le rôle de la protéase et de ces substrats, les PCs possèdent tout de même une action spécifique.



**Figure 47 : Localisation des ARNm de différentes PCs par hybridation in situ dans des coupes sagittales de cerveau de rat adjacentes. Tiré de [303].**

Il est possible d'aboutir à la même conclusion au regard de la difficulté à obtenir des souris knock-out pour une PC. Pour déterminer le rôle biologique de chaque PC, des modèles murins ont été créés (**Tableau 20**). Ceci a permis de mettre l'accent sur la diversité et la complexité des processus impliquant les PCs dans le développement d'un individu. Le concept de redondance suppose que plusieurs PCs pourraient partager certaines fonctions catalytique pour palier l'absence de l'une d'entre elles. Comme pour la furine, PACE4 est exprimé tôt dans le développement. Environ 25% des souris knock-out présentent une létalité prénatale et les souris qui survivent semblent normales malgré certaines malformations pulmonaires ou faciales [304]. La létalité embryonnaire précoce indique un rôle non-redondant des enzymes durant les différentes étapes du développement d'un embryon mais qui n'exclut pas une redondance durant la vie d'adulte.

Convertases/inhibitors	Null gene approach	Null phenotype
Furin	Homologous recombination	Mice die at e10.5 due to hemodynamic insufficiency, ventral closure defect, axial rotation defect, abnormal yolk sac vasculature.
PC1	Homologous recombination	Mice viable, but 60% of normal size with defects in the processing of hormone precursors, hyperproinsulinemia. <b>In humans:</b> severe early obesity, adenocortical insufficiency, and hyperproinsulinemia.
PC2	Neomycin resistance gene	Mice viable, retarded growth, hypoglycemia, defects in the processing of endocrine peptides.
PACE4	Homologous recombination	Mice (25%) die at e10.5 due to situs defects and/or craniofacial malformations.
PC4	Homologous recombination	Mice viable with reduced fertility.
PC5	Homologous recombination	Mice die at e4.5–7.5.
PC7	Homologous recombination	No apparent abnormal phenotype.
SKI-1 (S1P)	Homologous recombination  Inducible Cre recombinase (liver)	Mice die at e4. Disruption of <i>S1P</i> prevents normal epiblast formation and subsequent implantation. Mice with reduced cholesterol and fatty acid synthesis.
PCSK9 (NARC-1)	Homologous recombination	Enhanced cholesterol uptake by the liver. <b>In humans:</b> Mutation-induced gain of function: hypercholesterolemia (autosomal dominant hypercholesterolemia). Mutation-induced loss of function: hypocholesterolemia.

**Tableau 20 : Liste des différents phénotypes de souris knock-out existants pour les PCs. [305]**

Enfin, certains inhibiteurs semblent plus ou moins spécifiques à certaines PCs (**figure 6**) [174]. De plus dans le laboratoire du Pr Robert Day, certains inhibiteurs développés récemment semblent spécifique à une PC (résultats non publiés) montrant ainsi que la séquence de reconnaissance du site catalytique agit sur l'activité des PCs.

Nous pouvons donc en conclure que certains facteurs influent directement sur la spécificité ou la redondance des PCs pour leurs substrats : le site de clivage, le patron d'expression tissulaire et la localisation subcellulaire.

Comme pour la furine, PACE4 est exprimé dans de nombreux tissus où elle mature plusieurs substrats [306]. La maturation de la proPACE4 s'effectue via un autoclivage intramoléculaire du propeptide, ce qui semble être la seule étape limitante de sa sécrétion. Contrairement à la furine, elle ne possède pas de domaine transmembranaire mais un domaine riche en cystéines. Ce domaine est essentiel pour les liaisons de l'enzyme au niveau de la

surface cellulaire où elle semble lier des héparines ou TIMPS-2 [307, 308]. Dernièrement, l'interaction entre PACE4 et ANGPTL3 (Angiopoietin like protein 3) a été décrite comme favorisant le clivage de lipoprotein-lipase (LPL) à la surface des cellules, influençant ainsi le métabolisme des lipoprotéines [309].

TRPS1 (Tricho-rhino-phalangeal syndrome type I ou Zinc finger transcription factor) été détecté dans les cellules avec le knock-down pour la PACE4 et semble sous-exprimé dans la lignée cancéreuse sauvage. Ce répresseur transcriptionnel, dont l'expression semble sensible aux androgènes, est une protéine de 1 281 acides aminés d'environ 160 KDa [310, 311]. Il est connu pour se lier aux séquences d'ADN possédant un motif GATA ou GATA inverse entraînant la répression de l'expression du gène cible [312, 313]. TRPS1 réprime l'expression de POR2 (Voltage dependent anion-selective channel) qui forme un canal au niveau de la membrane mitochondriale permettant la diffusion de petites molécules hydrophiles comme le cytochrome C [314, 315]. POR2 lie et inhibe l'activation de BAK (BCL2-antagonist/killer 1), une protéine pro-apoptotique [316].

L'absence ou la l'expression réduite de TRPS1 dans les lignées cancéreuses peut donc entrainer une expression plus importante de POR2 inhibant ainsi BAK. Ce processus permet donc aux cellules cancéreuses d'acquérir la capacité d'échapper à l'apoptose. PACE4 semble donc jouer un rôle dans ce processus au vu des résultats obtenus par FACS entre les deux lignées cellulaires, mais son niveau d'action et son rôle ne sont pas encore bien établi. Plusieurs hypothèses sont possible dont le rôle clef que pourrait jouer BMP7 (Bone morphogenetic protein 7), un membre de la famille des TGF- $\beta$ , qui semble pouvoir réguler TRPS1 [317]. L'expression de BMP7 est connue pour être diminuée dans les différents stades de développement du cancer de la prostate [318]. De plus, ce facteur de croissance est connu pour être maturé par les proprotéines convertases et pour être sécrété, faisant de lui une cible intéressante pour poursuivre cette étude et mieux comprendre l'implication de PACE4 dans les mécanismes d'apoptose.

L'implication de PACE4 n'est donc pas encore parfaitement établie dans les lignées cancéreuses DU145, comme nous l'avons vu, elle semble être impliquée dans les mécanismes biologiques liés au cancer comme l'apoptose. Néanmoins, il est nécessaire de poursuivre l'étude des différentes protéines rentrant en jeu dans ce mécanisme et de recherche d'autres substrats potentiels de PACE4 par approche en protéomique différentielle directement sur les lignées DU145 et 4-2.

### 4.1.3 Approche protéomique différentielle.

Le plus grand défi demeure l'identification des protéines dont la maturation a été influencée par la réduction des niveaux d'expression de PACE4 au niveau cellulaire. Pour cela, nous avons initié une approche basée sur la technologie SELDI-TOF pour étudier les protéomes des cellules DU145 et DU145/4-2 afin d'identifier les substrats ou des protéines associées à l'action de cette proprotéine convertase. Le SELDI-TOF permet de coupler à la fois la technique de chromatographie et un MALDI-TOF grâce aux surfaces fonctionnalisées des barrettes. Nous avons choisi les barrettes H4 qui sont des phases réverses avec des groupements C16. Nous procédons tout d'abord à une extraction protéique en milieu acide avec une sonde à sonication sur des culots cellulaires de DU145 et de DU145/4-2 en triplicata. Ensuite, nous dosons les protéines via le système BCA qui réagit avec les amines. Nous digérons enfin la même quantité protéique d'échantillon avec de la trypsine pour aboutir aux fragments issus des protéines. Ces digestions sont ensuite appliquées sur la surface fonctionnalisée où il est possible d'éluer l'échantillon pour simplifier le protéome digéré suivant l'hydrophobicité de chaque peptide. Ainsi en comparant les peptides de digestion des triplicatas de chaque condition, nous avons pu mettre en évidence des m/z spécifiques à la lignée cancéreuse sauvage, et spécifiques à la lignée « knock-down » comme le montre la **figure 48**. Une liste de m/z sur ou sous-régulés a pu être élaborée et nous nous intéressons aujourd'hui à l'identification de ces substrats potentiels de PACE4 à l'aide du « Lucid proteomics system » élaboré en collaboration entre Biorad et Bruker. Ce système permet d'analyser les barrettes SELDI au sein d'un appareil Bruker de type Ultraflex et donc d'aboutir à la ms<sup>2</sup> des ions d'intérêt via le mode LIFT [31]. Nous travaillons aujourd'hui à l'uniformisation d'un protocole commun au SELDI-TOF et au MALDI-TOF-TOF pour pouvoir travailler sur les mêmes barrettes en parallèle entre la France et le Québec.

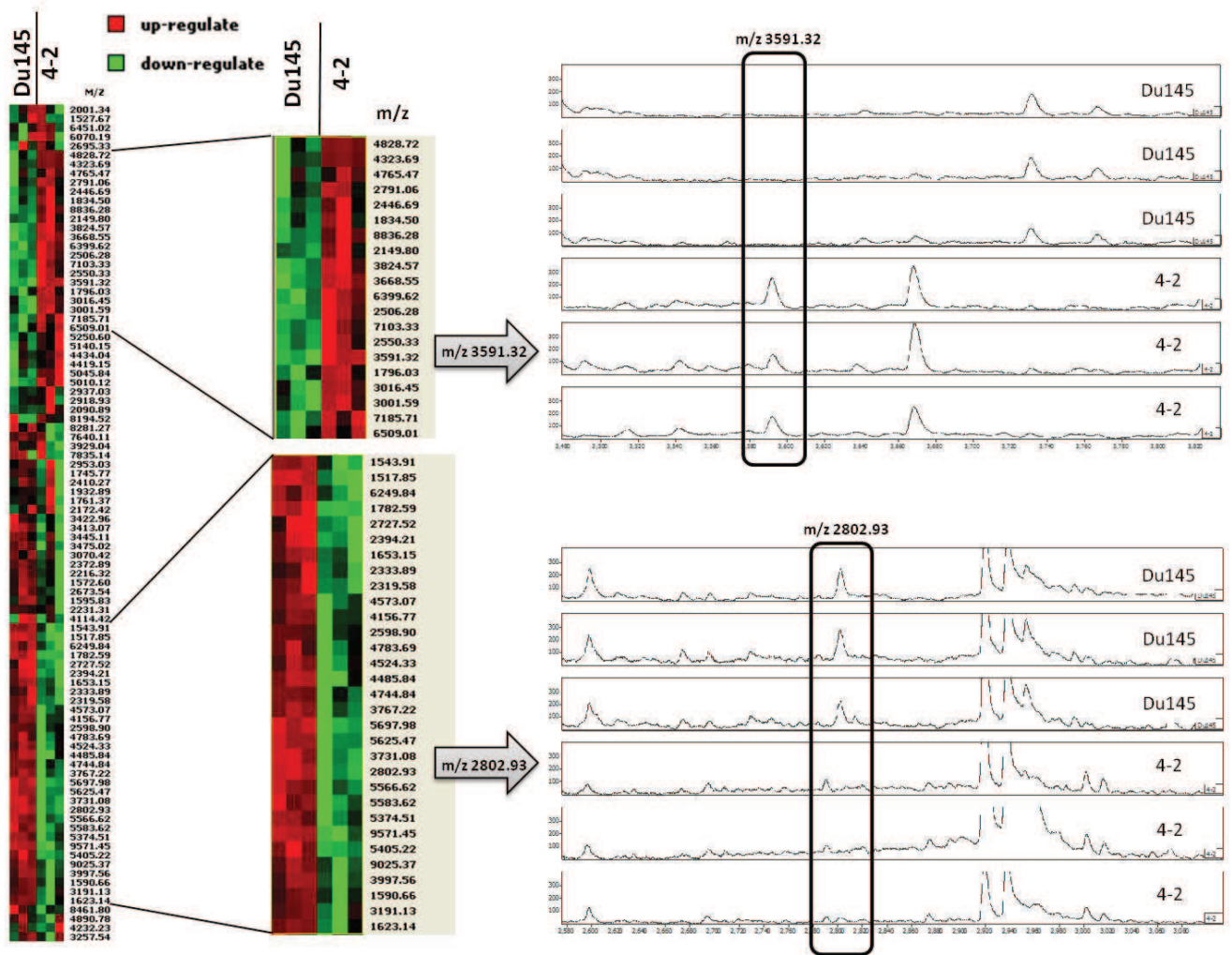


Figure 48 : Analyse SELDI-TOF des peptides de digestion différentiels entre la lignée DU145 et la lignée 4-2. Les ions m/z 3591.32 et m/z 2802.93 sont des exemples des variations détectables entre les deux conditions.

Très récemment, Overall et al. ont mis au point une autre technique de protéomique quantitative basée sur les marquages isobariques et la sélection négative spécialement pour l'étude des substrats des protéases [319]. Cette méthode, appelée iTRAQ-TAILS (isobaric tag for relative and absolute quantitation - terminal amine isotopic labeling of substrates), combine la technique de marquage iTRAQ [320] et TAILS [321] et a été décrite dernièrement pour l'étude des N-terminomes. La **figure 49** détaille l'approche générale. Différents échantillons ayant subi des niveaux différents de protéolyse sont marqués par les labels iTRAQ (qui donnent des ions rapporteurs à 114, 115, 116 et 117 en  $m/z$ ) sur leurs N-termini libres et sur leurs résidus lysine. Le mélange de tous les échantillons marqués est soumis à une digestion trypsique pour générer des peptides de digestion avec un N-terminal marqué ou possédant une modification (acétylation par exemple) ou encore des peptides internes ou C-

terminal ayant un N-terminal libre. Une sélection négative permet ensuite de lier les peptides avec un N-terminal libre sur un polymère dendritique de polyaldehyde. Ce polymère et ses peptides liés sont éliminés par ultracentrifugation. Avec une analyse  $ms^2$  sur une chaîne nanoLC-ESI-Orbitrappe par exemple, il est possible de quantifier les peptides ayant le label iTRAQ et d'obtenir leur identification via l'interrogation des banques de données à partir des spectres de fragmentation. La comparaison des ratios des peptides va permettre de mettre en évidence les fragments dus à la protéolyse de l'enzyme d'intérêt de comparer chaque condition. Cette toute nouvelle méthodologie semble donc parfaitement adaptée à notre thématique puisqu'elle permet à la fois la quantification, l'étude différentielle et l'identification des protéines d'intérêt. Ce qui pourrait permettre de compléter nos approches actuelles sur SELDI-TOF et de valider les protéines qui seront identifiées comme étant présentes ou absentes dans notre lignée DU145/4-2 (ayant le knock-down pour PACE4) et la lignée DU145.

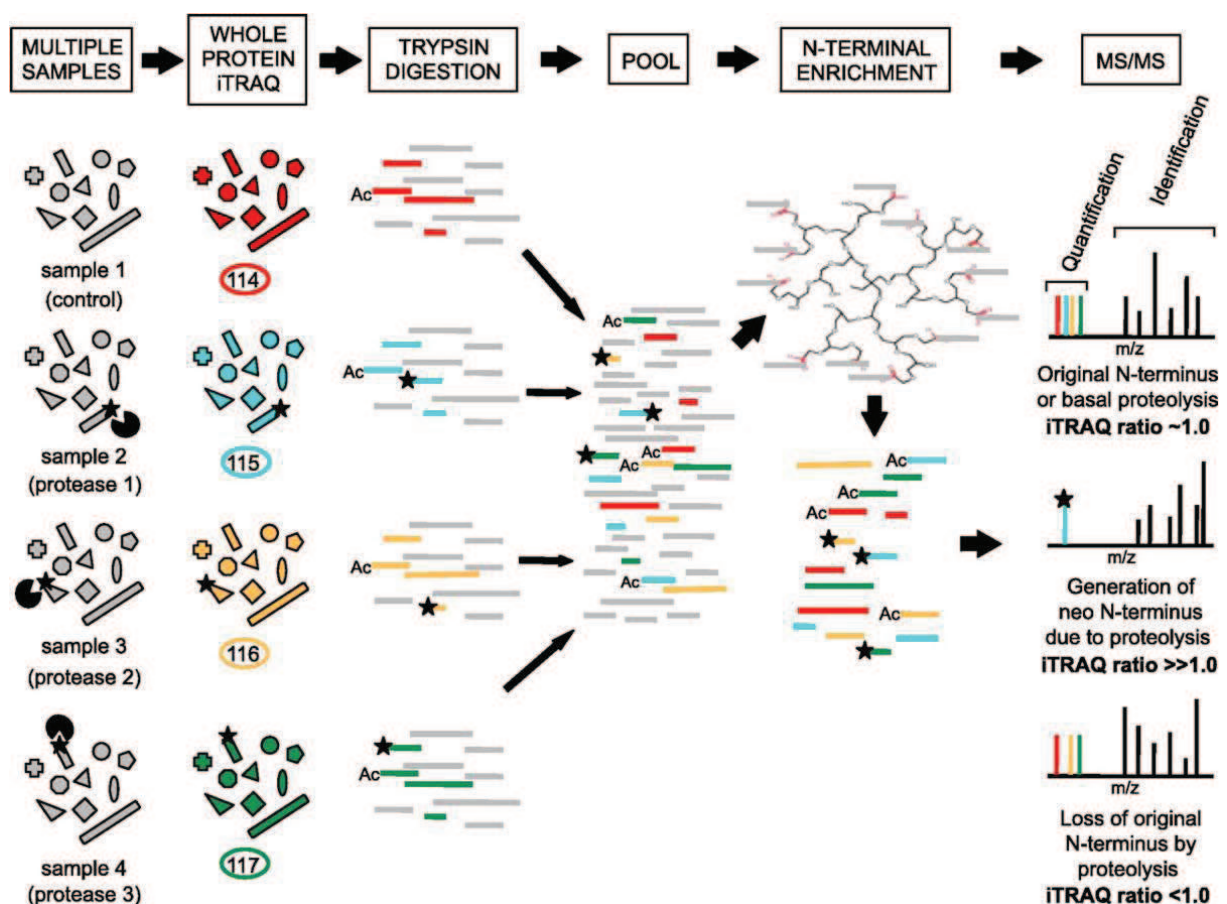


Figure 49 : Principe de la technique iTRAQ-TAILS pour la recherche des substrats de protéases [319].

## Conclusion

La proprotéine convertase PACE4 possède donc des fonctions spécifiques dans les lignées de cancer de prostate DU145. Elle joue un rôle dans la capacité de prolifération de ces cellules et dans les mécanismes de protection contre l'apoptose. Ces fonctions uniques peuvent s'afficher au niveau de la nature des substrats produits par leur action catalytique, mais également au niveau des conséquences cellulaires résultant de l'action de ces produits maturés (comme nous le pensons pour TRPS1). La recherche des substrats des PCs en général et de PACE4 plus particulièrement reste une perspective très intéressante pour la compréhension des fonctions et des mécanismes impliquant ces protéases. Ces validations permettront de valider définitivement PACE4 comme une cible thérapeutique pour le cancer de prostate qui pourra être régulée grâce aux différents inhibiteurs pharmacologiques existants ou en développement.



## Chapitre 5 : Conclusion générale et perspectives

---

Comme nous l'avons vu, la protéomique représente aujourd'hui un domaine qui ne cesse de se développer. Les protéines sont les acteurs directs influençant l'état physiologique de la cellule aux travers de leurs rôles effecteurs, activateurs ou répresseurs, dans toutes les voies de signalisation et dans toutes les interactions intermoléculaires. Les peptides et les protéines semblent donc être des cibles thérapeutiques et des marqueurs pathologiques parfaits puisqu'une modification de leur expression peut directement être liée à un état ou à un stade pathologique. D'un point de vue plus fondamental, la découverte des processus impliquant ces molécules permet de mieux comprendre les mécanismes mis en jeu dans les pathologies.

Dans ce contexte, l'apport de nouvelles techniques et technologies pour l'étude des protéines n'a cessé de croître. Au premier rang desquelles, la spectrométrie de masse s'est imposée comme la technologie la plus efficace pour la séparation, la détection et l'identification des protéines dans des échantillons biologiques complexes [3]. En effet, l'apport des méthodes off-gels basées sur la MS ou la MS/MS a permis de se libérer des problèmes de reproductibilité des méthodes in-gels même si de gros progrès ont été faits. Ainsi les méthodes de marquage comme l'iTRAQ, l'ICAT, l'aqua ou le TAILS ou encore les méthodes couplant la chromatographie liquide et la MS comme le SELDI permettent aujourd'hui la séparation, la détection, la quantification et l'identification de molécules cibles entre différentes conditions de manière efficace mais sans conserver l'information de localisation des molécules d'intérêt.

C'est pourquoi l'imagerie par spectrométrie de masse MALDI (**chapitre 2**) a connu un essor très rapide en un peu plus de dix ans et semble aujourd'hui une technologie de plus en plus adaptée en protéomique clinique [7]. En effet, depuis la première image MALDI de 1997[62], de très nombreux développements ont été réalisés à plusieurs niveaux. Aujourd'hui, les protocoles de préparation tissulaire sont bien définis selon le type de molécules que l'on étudie [85, 88, 90, 93, 108, 111, 112]. Cette préparation comprend à la fois des lavages de l'échantillon et/ou l'application de matrices spécifiques à la détection d'ions d'intérêt (lipides, protéines, peptides). En plus de ces développements, de gros efforts ont été apportés par les sociétés commercialisant les spectromètres de masse et les équipes spécialisées en MSI pour optimiser les techniques d'imagerie. Ces efforts ont principalement porté sur :

- les systèmes laser et leur fréquence permettant la réalisation d'images de 10 000 points en moins d'une heure avec les derniers appareils.
- la résolution spatiale toujours améliorée par les progrès sur la focalisation des faisceaux laser
- les détecteurs dont la sensibilité est améliorée et qui permettant donc une gamme de détection plus importante avec une meilleure résolution spectrale.
- les couplages de la source MALDI avec d'autres types d'analyseurs que le TOF.
- les logiciels d'acquisition et de retraitement dont les caractéristiques permettent une analyse poussée des données et offrent de plus en plus d'études possibles suivant le choix de l'utilisateur [135].

De même, il faut noter que contrairement aux premières analyses MSI effectuées au début des années 2000, il existe maintenant des méthodes automatisées et robotisées pour les traitements des coupes de tissus et pour le dépôt de la matrice, ceci permettant alors d'aboutir à des protocoles reproductibles, répétables et donc standardisés [90].

Néanmoins, en raison de la grande quantité de spectres obtenus suite à l'acquisition MSI, il nous était difficile d'exploiter pleinement la totalité des informations avant l'application et l'utilisation des analyses multivariées qui ont permis de palier à ce problème. En effet, comme nous l'avons vu dans le **chapitre 3**, les analyses PCA permettent de simplifier la totalité des données MSI en les représentant dans un espace multidimensionnel dans lequel les coordonnées de chaque spectre sont déterminées par les composantes principales. Ainsi, la comparaison des spectres les uns par rapport aux autres devient possible et nous permet de mettre en évidence très rapidement les variables ( $m/z$ ) propres à chaque groupe de spectres. La classification hiérarchique des spectres issus d'une image MSI nous a permis de mettre en évidence des profils peptidiques et/ou protéiques d'une zone histologique précise dans d'un tissu. En effet, si certains spectres se classent sous la forme d'un cluster (possible de visualiser par superposition sur l'image optique) et coïncident avec une zone histologique précise, ceci nous montre que des molécules spécifiques à chaque zone existent et permettent donc de discriminer les groupes de spectres. Nous avons également pu démontrer, grâce à ces classifications, qu'aucune délocalisation n'était visible après les différentes étapes de préparation d'échantillon que notre laboratoire a mis au point. Ainsi, nous avons pu valider nos démarches sur tissus frais et tissus FFPE pour la recherche de biomarqueurs.

En présence d'une coupe de tissu avec plusieurs zones histologiques, il restait très difficile de détecter les ions propres aux zones les plus petites. En effet, la taille des zones histologiques d'un tissu est importante, puisqu'au plus une zone sera petite, au moins elle possédera de spectres acquis. Ces spectres seront donc « noyés » dans la quantité d'informations recueillies sur la totalité de la coupe et il est très difficile de pointer les  $m/z$  spécifiques des zones plus petites. C'est pourquoi nous avons développé l'approche PCA-SDA qui combine une première analyse des données d'une image MSI par classification hiérarchique pour ainsi mettre en évidence toutes les zones d'un tissu (étape pouvant être confirmée par coloration de la coupe ayant servie pour la MSI). Une fois les différentes zones détectées à l'aide des clusters, les spectres de chaque groupe d'intérêt sont exportés et une analyse PCA est effectuée. Il est alors possible d'analyser très finement les données générées par MSI sans risquer de perdre d'information.

La principale limite actuelle de ces approches réside dans les logiciels commerciaux comme ClinproTools qui demandent énormément de ressource informatique et qui ne permettent pas l'analyse d'image avec un très grand nombre de points. C'est pourquoi, notre laboratoire s'est associé avec le Pr Edouardo Macagno à San Diego (USA) et la doctorante Jocelyne Bruand pour développer un logiciel gérant le traitement d'un nombre de données illimitées et qui nous permettra de retraiter la totalité de nos images comme sur animaux entiers (whole body) qui génèrent, en raison de leur taille, un grand nombre de spectres [322]. En effet, l'utilisation d'une approche par PCA-SDA nécessite une résolution spatiale suffisante pour distinguer les différentes zones histologiques. Comme il est possible d'analyser par MSI des coupes dont la superficie est de plus en plus importante en raison des progrès techniques, et avec l'amélioration croissante de la résolution spatiale, il est impératif d'utiliser un logiciel permettant d'analyser rapidement un grand nombre de spectres par PCA.

Une perspective importante de la classification hiérarchique est la modélisation des groupes de spectres comme une possible nouvelle méthode de diagnostic. En effet, comme décrit dans la discussion du **chapitre 3**, cette méthode permettrait d'identifier des régions histologiques pathologiques à partir de leurs profils peptidiques et/ou protéiques, et non plus en fonction d'un marqueur unique. Le profil de l'ensemble des molécules détectées à l'aide d'un protocole standardisé et reproductible de MSI, permettrait de distinguer une zone en fonction de l'abondance relative de tous les ions détectés. La mise en place de cette méthodologie demanderait cependant du temps en raison des différentes étapes de validation du modèle mais serait extrêmement prometteuse à l'heure où les premières utilisations de la MSI ont été réalisées *in vivo*.

C'est donc en raison des nombreux avantages que nous venons de voir concernant l'imagerie MALDI, que notre laboratoire a initié une étude en protéomique clinique basée sur la MSI à partir de biopsies de cancer d'ovaire. Cette étude réalisée en collaboration avec l'hôpital Jeanne de Flandre est décrite dans le **chapitre 3**. Ainsi en couplant les protocoles de préparation d'échantillon élaborés au laboratoire [85, 88, 90, 109] avec les analyses multivariées et la classification hiérarchique sur les données de MSI, nous avons été en mesure de détecter et d'identifier plusieurs biomarqueurs potentiels de cette pathologie. Parmi ceux-ci, nous avons retrouvé Reg- $\alpha$  qui correspond au fragment C-terminal de l'activateur de l'immunoproteasome 11S dont une première identification avait été effectuée par Rémi Lemaire au sein du laboratoire. L'observation de la dégradation d'une sous-unité de l'immunoprotéasome, dont nous n'avons pas encore défini les effecteurs, pourrait être une capacité acquise pour échapper au système immunitaire en entraînant une dérégulation du processus de présentation des antigènes. Ce phénomène déjà décrit dans les mécanismes d'infection virale pourrait expliquer comment les cellules cancéreuses échappent à la surveillance du système immunitaire. L'identification de la surexpression de la mucine 9, qui n'avait pas encore été décrite dans le cadre du cancer de l'ovaire, pourrait s'avérer un excellent marqueur de cette pathologie. En effet, les mucines ont la capacité de servir de récepteurs de surface cellulaire et de capteurs de signaux en réponse à des stimuli externes pour une variété de réponses cellulaires : prolifération et croissance cellulaires, différenciation et apoptose. Ces données suggèrent donc que l'expression aberrante des mucines pourrait être impliquée dans le développement et la progression du cancer de l'ovaire. Enfin, L'identification de la surexpression de protéines comme la profilin-1, la cofilin-1, la vimentin et la cytokératine 19 nous montre que des phénomènes importants doivent se passer au niveau de la matrice extracellulaire et du cytosquelette. Or, on parle souvent de transition épithélio-mésenchymateuse au niveau du cancer lors du passage d'un épithélium, dont les cellules polarisées sont liées par des jonctions serrées, des jonctions gap, des jonctions adhérentes, et des desmosomes, à un mésenchyme où les cellules ne sont plus liées entre elles mais constituent un tissu de soutien relié par la matrice extracellulaire et où les cellules sont capables de migration. Plusieurs événements cellulaires sont nécessaires à la transition épithélio-mésenchymateuse : la réorganisation de l'adhérence intercellulaire et de l'adhérence matricielle, le remodelage du cytosquelette et le changement de polarité cellulaire.

Ces observations rejoignent les connaissances actuelles impliquant les proprotéines convertases dans la tumorigénèse comme décrit dans le **chapitre 2**. C'est pourquoi nous envisageons d'étudier leurs implications précises dans le cancer de l'ovaire avec, dans un

premier temps, la création de modèle cellulaire knock-down pour les PCs dans le cancer de l'ovaire (à partir de la lignée SKOV3). Ces travaux, amorcés par Rémi Longuespée (doctorant), permettront d'une part de constater l'implication de la sous-expression de PCs sur la prolifération, l'adhésion et l'invasion de ces cellules modèles pour ensuite rechercher les différents substrats des PCs touchés par les knock-down, et d'autre part de tenter de mieux comprendre l'implication des protéines mises en évidence par imagerie MALDI dans le cancer de l'ovaire. Il serait également intéressant de poursuivre l'hypothèse d'une étiologie virale, pouvant également impliquer les PCs, du cancer de l'ovaire par une étude génomique ou protéomique sur tissu et/ou cellules.

Il n'existe aujourd'hui aucun doute sur l'implication des PCs dans les phénomènes de tumorigénèse, comme le Dr François D'Anjou a pu le démontrer durant sa thèse avec l'identification de la surexpression de PACE4 dans le cancer de la prostate (**chapitre 4**). En effet, PACE4 s'est avéré être la seule PC surexprimée dans cette pathologie suite à l'étude différentielle effectuée sur des biopsies humaines. C'est pourquoi nous nous sommes intéressés à l'étude des substrats potentiels de cette protéase par des analyses protéomiques différentielles entre une lignée wildtype du cancer de la prostate (DU145) et une lignée Knock-down pour PACE4 (DU145/4-2). Ainsi, à partir de ces modèles *in vitro*, nous avons pu identifier le facteur de transcription TRPS1 comme étant surexprimé dans la lignée DU145/4-2. Ceci nous a permis de mettre en évidence une modulation de l'apoptose entre nos deux lignées. En effet, suite à l'expression réduite de PACE4, nous avons constaté que les lignées knock-down présentaient un profil apoptotique revenu à la « normale » comparé à la lignée DU145. Cette dernière semble posséder un système d'évasion apoptotique caractéristique des lignées cancéreuses. TRPS1 n'étant pas un substrat des PCs, il serait intéressant de poursuivre l'étude jusqu'à l'identification de la protéine maturée reliant PACE4 et TRPS1. Pour cela, nous avons démarré une étude protéomique différentielle à partir de notre système SELDI-TOF qui nous a permis de détecter des fragments de digestion différentiels issus de protéines qui pourraient être des substrats de PACE4. C'est pourquoi, nous souhaitons essayer le système Lucid qui permettrait d'identifier par  $m/z$  ces ions d'intérêt à partir de notre Ultraflex II. Il serait également possible de croiser cette technique avec de nouvelles approches par marquage comme l'ITRAQ/TAILS qui semble convenir pour notre problématique.

On peut également se demander quel serait l'impact d'un inhibiteur spécifique à PACE4 sur le phénotype cancéreux. Ce type d'inhibiteur pharmacologique est actuellement en test au sein du laboratoire du Pr Day [323], et il sera très intéressant d'évaluer dans un

premier temps son impact *in vitro* ou sur des xénogreffes de cancer de prostate chez la souris pour définitivement valider PACE4 comme étant une cible thérapeutique potentielle de choix pour le cancer de la prostate.

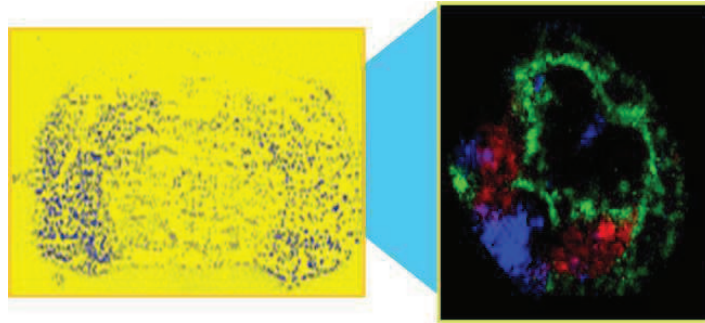
L'imagerie par spectrométrie de masse rentre donc dans une nouvelle ère, celle de la protéomique clinique au chevet du patient d'un côté et de l'autre celle des DMPK (Drug Metabolism and Pharmacokinetics) à haut débit pour les petites molécules ou les anticorps thérapeutiques. Ces deux voies nécessitent encore des développements pour rentrer à l'hôpital aux côtés du TEP (tomographie par émission de positons) et de l'IRM (imagerie par résonance magnétique). La MSI suit donc le chemin de sa grande sœur l'IRM qui en 50 ans est désormais devenue une technique incontournable en médecine. Qui aurait parié en 1969 que la RMN puisse devenir un outil médical qui permettrait la détection de tumeurs, hormis le Pr. Damadian. Il a fallu attendre 1977 pour que les premières images d'un corps humain vivant apparaissent (**Figure 50**).



**Figure 50 : Réprésentation des Prs Damadian et Lauterbur avec leur première image IRM obtenue sur un corps humain en 1977.**

L'imagerie MALDI fait donc ce même périple. En 10 ans en regardant les premières images présenter à l'ASMS en 1999, personne ne pouvait s'imaginer que l'imagerie en serait désormais à ce type de résolution (10 $\mu$ m) et d'avancée (**Figure 51**). Il est indéniable que ces cinq prochaines années vont permettre à cette technologie de s'ancrer dans cette problématique clinique avec une perspective également dans le cadre de l'environnement comme outil permettant de faire l'interface entre santé et environnement. Elle sera donc au

côté de pathologiste pour son diagnostic sur des biopsies extemporanée, et peut être dans les mains du chirurgien pour définir les zones cancéreuses des zones en évolutions grâce à une miniaturisation. L'imagerie par spectrométrie de masse est donc par ce type d'étude en train de réaliser ses classes pour devenir un nouvel outil de diagnostic de demain.



**Figure 51 : Evolution de l'imagerie MALDI depuis la première image en 1999 sur des cerveaux murins [324].**





## Références bibliographiques

---

1. *Finishing the euchromatic sequence of the human genome*. Nature, 2004. **431**(7011): p. 931-45.
2. Sikaroodi, M., Y. Galachiantz, and A. Baranova, *Tumor markers: the potential of "omics" approach*. Curr Mol Med, 2010. **10**(2): p. 249-57.
3. Hudler, P., M. Gorsic, and R. Komel, *Proteomic strategies and challenges in tumor metastasis research*. Clin Exp Metastasis, 2010. **27**(6): p. 441-51.
4. Vegvari, A. and G. Marko-Varga, *Clinical protein science and bioanalytical mass spectrometry with an emphasis on lung cancer*. Chem Rev, 2010. **110**(5): p. 3278-98.
5. Micallef, J., et al., *Applying mass spectrometry based proteomic technology to advance the understanding of multiple myeloma*. J Hematol Oncol, 2010. **3**: p. 13.
6. Wang, P., J.R. Whiteaker, and A.G. Paulovich, *The evolving role of mass spectrometry in cancer biomarker discovery*. Cancer Biol Ther, 2009. **8**(12): p. 1083-94.
7. Franck, J., et al., *MALDI imaging mass spectrometry: state of the art technology in clinical proteomics*. Mol Cell Proteomics, 2009. **8**(9): p. 2023-33.
8. Tanaka, H., et al., *[Imaging mass spectrometry: principle and application]*. Tanpakushitsu Kakusan Koso, 2009. **54**(3): p. 224-9.
9. Walch, A., et al., *MALDI imaging mass spectrometry for direct tissue analysis: a new frontier for molecular histology*. Histochem Cell Biol, 2008. **130**(3): p. 421-34.
10. Fournier, I., M. Wisztorski, and M. Salzet, *Tissue imaging using MALDI-MS: a new frontier of histopathology proteomics*. Expert Rev Proteomics, 2008. **5**(3): p. 413-24.
11. Wisztorski, M., et al., *[MALDI imaging: a new technology to discover and validate new biomarkers]*. Med Sci (Paris), 2007. **23 Spec No 1**: p. 31-6.
12. Wisztorski, M., et al., *MALDI direct analysis and imaging of frozen versus FFPE tissues: what strategy for which sample?* Methods Mol Biol, 2010. **656**: p. 303-22.
13. Kaletas, B.K., et al., *Sample preparation issues for tissue imaging by imaging MS*. Proteomics, 2009. **9**(10): p. 2622-33.
14. James, P., *Protein identification in the post-genome era: the rapid rise of proteomics*. Q Rev Biophys, 1997. **30**(4): p. 279-331.
15. Lander, E.S., et al., *Initial sequencing and analysis of the human genome*. Nature, 2001. **409**(6822): p. 860-921.
16. O'Farrell, *High resolution two-dimensional electrophoresis of proteins*. J. Biol. Chem, 1975. **250**: p. 4007-4021.
17. Kenrick, K.G. and J. Margolis, *Isoelectric focusing and gradient gel electrophoresis: a two-dimensional technique*. Anal Biochem, 1970. **33**(1): p. 204-7.
18. Mikkers, F., F. Everaerts, and T. Verheggen, *High-performance zone electrophoresis*. J Chromatogr 1979(169): p. 11-20.
19. Jorgenson, J.W. and K.D. Lukacs, *Capillary zone electrophoresis*. Science, 1983. **222**(4621): p. 266-72.
20. Ding, J. and P. Vouros, *Advances in CE/MS*. Anal Chem, 1999. **71**(11): p. 378A-385A.
21. Zamfir, A. and J. Peter-Katalinic, *Capillary electrophoresis-mass spectrometry for glycoscreening in biomedical research*. Electrophoresis, 2004. **25**(13): p. 1949-63.
22. Unger, M., *Capillary electrophoresis of natural products: current applications and recent advances*. Planta Med, 2009. **75**(7): p. 735-45.
23. Hutchens, T.W. and T. Tai-Tung Yip, *New desorption strategies for the mass spectrometric analysis of macromolecules*. Rapid Communications in Mass Spectrometry, 1993. **7**: p. 576-580.

24. Xu, G., et al., *Application of SELDI-TOF-MS to identify serum biomarkers for renal cell carcinoma*. *Cancer Lett*, 2009. **282**(2): p. 205-13.
25. van Winden, A.W., et al., *Validation of previously identified serum biomarkers for breast cancer with SELDI-TOF MS: a case control study*. *BMC Med Genomics*, 2009. **2**: p. 4.
26. Chen, L., et al., *SELDI protein chip technology for the detection of serum biomarkers for liver disease*. *Protein Pept Lett*, 2009. **16**(5): p. 467-72.
27. Zhang, X., et al., *Potential biomarkers of acute cerebral infarction detected by SELDI-TOF-MS*. *Am J Clin Pathol*, 2008. **130**(2): p. 299-304.
28. Wang, S., et al., *Detection of serum proteomic changes and discovery of serum biomarkers for Kashin-Beck disease using surface-enhanced laser desorption ionization mass spectrometry (SELDI-TOF MS)*. *J Bone Miner Metab*, 2008. **26**(4): p. 385-93.
29. Liu, W., et al., *Using SELDI-TOF MS to identify serum biomarkers of rheumatoid arthritis*. *Scand J Rheumatol*, 2008. **37**(2): p. 94-102.
30. Langbein, S., *Identification of disease biomarkers by profiling of serum proteins using SELDI-TOF mass spectrometry*. *Methods Mol Biol*, 2008. **439**: p. 191-7.
31. Jourdain, S., A. Bulman, and E. Dalmaso, *The Lucid Proteomics System for top-down biomarker research*. *Arch Physiol Biochem*.
32. Posthumus, M.A., et al., *Laser desorption-mass spectrometry of polar nonvolatile bio-organic molecules*. *Anal Chem*, 1978. **50**(7): p. 985-991.
33. Karas, M. and F. Hillenkamp, *Laser desorption ionization of proteins with molecular masses exceeding 10,000 daltons*. *Anal Chem*, 1988. **60**(20): p. 2299-301.
34. Tanaka, K., 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**(8): p. 151-153.
35. Hillenkamp, F., et al., *Matrix-assisted laser desorption/ionization mass spectrometry of biopolymers*. *Anal Chem*, 1991. **63**(24): p. 1193A-1203A.
36. Hillenkamp, F. and M. Karas, *Mass spectrometry of peptides and proteins by matrix-assisted ultraviolet laser desorption/ionization*. *Methods Enzymol*, 1990. **193**: p. 280-95.
37. Dole, M., et al., *Molecular beams of macroions*. *J. Chem. Phys.*, 1968. **49**: p. 2240-2249.
38. Fenn, J.B., et al., *Electrospray ionization for mass spectrometry of large biomolecules*. *Science*, 1989. **246**(4926): p. 64-71.
39. Fenn, M.Y.J., *Electrospray ion source. Another variation on the Free-jet theme*. *J. Phys. Chem.*, 1984. **88**(4451-4460).
40. Yamashita, M. and J. Fenn, *Negative Ion Production with the electrospray Ion Source*. *J. Phys. Chem.*, 1984. **88**: p. 4671-4675.
41. Ikonomou, M.G., A.T. Blades, and P. Kebarle, *Electrospray-Ion spray: A comparison of mechanisms and performance*. *Anal. Chem.*, 1991. **63**: p. 1989-98.
42. Blades, A.T., M.G. Ikonomou, and P. Kebarle, *Mechanism of electrospray mass spectrometry. Electrospray as an electrolysis cell*. *Anal Chem*, 1991. **63**: p. 2109 - 2114.
43. Kebarle, P. and U.H. Verkerk, *Electrospray: from ions in solution to ions in the gas phase, what we know now*. *Mass Spectrom Rev*, 2009. **28**(6): p. 898-917.
44. Iribarne, J.V. and B.A. Thomson, *On the evaporation of small ions from charged droplets*. *J. Chem. Phys.*, 1976. **64**: p. 2287-2294.
45. Nguyen, S. and J.B. Fenn, *Gas-phase ions of solute species from charged droplets of solutions*. *Proc Natl Acad Sci U S A*, 2007. **104**(4): p. 1111-7.
46. Kebarle, P., *A brief overview of the present status of the mechanisms involved in electrospray mass spectrometry*. *J Mass Spectrom*, 2000. **35**(7): p. 804-17.

47. Ikonomidou, M., A.T. Blades, and P. Kebarle, *Investigations of the Electrospray Interface for Liquid Chromatography Mass Spectrometry*. Anal Chem, 1990. **62**: p. 957-967.
48. Covey, T., B.I. Shushan, and R. Bonner, *Practical advantages of Ionspray over Electrospray*. Sciex, 1990.
49. Cole, R.B., *Some tenets pertaining to electrospray ionization mass spectrometry*. J Mass Spectrom, 2000. **35**(7): p. 763-72.
50. Wilm, M. and M. Mann, *Analytical properties of the nanoelectrospray ion source*. Anal Chem, 1996. **68**(1): p. 1-8.
51. Fligge, T.A., et al., *Direct monitoring of protein-chemical reactions utilising nanoelectrospray mass spectrometry*. J Am Soc Mass Spectrom, 1999. **10**(2): p. 112-8.
52. Shevchenko, A., et al., *Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels*. Anal Chem, 1996. **68**(5): p. 850-8.
53. Glish, G.L. and R.W. Vachet, *The basics of mass spectrometry in the twenty-first century*. Nat Rev Drug Discov, 2003. **2**(2): p. 140-50.
54. Henzel, W.J., et al., *Identifying proteins from two-dimensional gels by molecular mass searching of peptide fragments in protein sequence databases*. Proc Natl Acad Sci U S A, 1993. **90**(11): p. 5011-5.
55. James, P., et al., *Protein identification by mass profile fingerprinting*. Biochem Biophys Res Commun, 1993. **195**(1): p. 58-64.
56. Yates, J.R., 3rd, et al., *Peptide mass maps: a highly informative approach to protein identification*. Anal Biochem, 1993. **214**(2): p. 397-408.
57. Roepstorff, P. and J. Fohlman, *Proposal for a common nomenclature for sequence ions in mass spectra of peptides*. Biomed Mass Spectrom, 1984. **11**(11): p. 601.
58. Biemann, K., *Appendix 5. Nomenclature for peptide fragment ions (positive ions)*. Methods Enzymol, 1990. **193**: p. 886-7.
59. Fournier, I., R. Day, and M. Salzet, *Direct analysis of neuropeptides by in situ MALDI-TOF mass spectrometry in the rat brain*. Neuro Endocrinol Lett, 2003. **24**(1-2): p. 9-14.
60. Castaing, R. and G. Slodzian, *J Microsc*, 1962. **1**: p. 395-410.
61. Kaufmann, R., F. Hillenkamp, and R. Wechsung, *The laser microprobe mass analyzer (LAMMA): a new instrument for biomedical microprobe analysis*. Med Prog Technol, 1979. **6**(3): p. 109-21.
62. Caprioli, R.M., T.B. Farmer, and J. Gile, *Molecular imaging of biological samples: localization of peptides and proteins using MALDI-TOF MS*. Anal Chem, 1997. **69**(23): p. 4751-60.
63. Chaurand, P., M. Stoeckli, and R.M. Caprioli, *Direct profiling of proteins in biological tissue sections by MALDI mass spectrometry*. Anal Chem, 1999. **71**(23): p. 5263-70.
64. Takats, Z., et al., *Mass spectrometry sampling under ambient conditions with desorption electrospray ionization*. Science, 2004. **306**(5695): p. 471-3.
65. Wiseman, J.M., et al., *Tissue imaging at atmospheric pressure using desorption electrospray ionization (DESI) mass spectrometry*. Angew Chem Int Ed Engl, 2006. **45**(43): p. 7188-92.
66. McDonnell, L.A. and R.M. Heeren, *Imaging mass spectrometry*. Mass Spectrom Rev, 2007. **26**(4): p. 606-43.
67. Wiseman, J.M., et al., *Mass spectrometric profiling of intact biological tissue by using desorption electrospray ionization*. Angew Chem Int Ed Engl, 2005. **44**(43): p. 7094-7.

68. Kruse, R.A., et al., *Direct assay of Aplysia tissues and cells with laser desorption/ionization mass spectrometry on porous silicon*. J Mass Spectrom, 2001. **36**(12): p. 1317-22.
69. May, T.W., et al., *A Lead Isotope Distribution Study in Swine Tissue Using ICP-MS*. Atom. Spectrosc., 1999. **20**(6): p. 199-211.
70. Luxembourg, S.L., et al., *High-spatial resolution mass spectrometric imaging of peptide and protein distributions on a surface*. Anal Chem, 2004. **76**(18): p. 5339-44.
71. Zenobi, R. and V.V. Deckert, *Scanning Near-Field Optical Microscopy and Spectroscopy as a Tool for Chemical Analysis*. Angew Chem Int Ed Engl, 2000. **39**(10): p. 1746-1756.
72. Spengler, B. and M. Hubert, *Scanning microprobe matrix-assisted laser desorption ionization (SMALDI) mass spectrometry: instrumentation for sub-micrometer resolved LDI and MALDI surface analysis*. J Am Soc Mass Spectrom, 2002. **13**(6): p. 735-48.
73. Kindness, A., C.N. Sekaran, and J. Feldmann, *Two-dimensional mapping of copper and zinc in liver sections by laser ablation-inductively coupled plasma mass spectrometry*. Clin Chem, 2003. **49**(11): p. 1916-23.
74. Ford, M.J. and G.J. Van Berkel, *An improved thin-layer chromatography/mass spectrometry coupling using a surface sampling probe electrospray ion trap system*. Rapid Commun Mass Spectrom, 2004. **18**(12): p. 1303-9.
75. Novikov, A., et al., *Matrix-implanted laser desorption/ionization mass spectrometry*. Anal Chem, 2004. **76**(24): p. 7288-93.
76. Takats, Z., et al. *Jet Desorption Ionization - A new ionization method for tissue analysis*. in *54th ASMS Conference on Mass Spectrometry 2006*. Seattle, Washington.
77. Huang, M.Z., et al., *Direct protein detection from biological media through electrospray-assisted laser desorption ionization/mass spectrometry*. J Proteome Res, 2006. **5**(5): p. 1107-16.
78. Balog, J., et al., *Identification of Biological Tissues by Rapid Evaporative Ionization Mass Spectrometry*. Anal Chem, 2010.
79. Jimenez, C.R., et al., *Neuropeptide expression and processing as revealed by direct matrix-assisted laser desorption ionization mass spectrometry of single neurons*. J Neurochem, 1994. **62**(1): p. 404-7.
80. Li, L., et al., *In situ sequencing of peptides from biological tissues and single cells using MALDI-PSD/CID analysis*. Anal Chem, 1999. **71**(24): p. 5451-8.
81. Li, L., R.W. Garden, and J.V. Sweedler, *Single-cell MALDI: a new tool for direct peptide profiling*. Trends Biotechnol, 2000. **18**(4): p. 151-60.
82. Rubakhin, S.S., et al., *Measuring the peptides in individual organelles with mass spectrometry*. Nat Biotechnol, 2000. **18**(2): p. 172-5.
83. Stoeckli, M., T.B. Farmer, and R.M. Caprioli, *Automated mass spectrometry imaging with a matrix-assisted laser desorption ionization time-of-flight instrument*. J Am Soc Mass Spectrom, 1999. **10**(1): p. 67-71.
84. Lemaire, R., et al. *Direct analysis and MALDI imaging on formalin-fixed paraffin-embedded (FFPE) tissues : Application to Parkinson disease*. in *54th ASMS Conference on Mass Spectrometry*. 2006. Seattle, Washington.
85. Lemaire, R., et al., *Direct analysis and MALDI imaging of formalin-fixed, paraffin-embedded tissue sections*. J Proteome Res, 2007. **6**(4): p. 1295-305.
86. Svensson, M., et al., *Heat stabilization of the tissue proteome: a new technology for improved proteomics*. J Proteome Res, 2009. **8**(2): p. 974-81.
87. Schwartz, S.A., M.L. Reyzer, and R.M. Caprioli, *Direct tissue analysis using matrix-assisted laser desorption/ionization mass spectrometry: practical aspects of sample preparation*. J Mass Spectrom, 2003. **38**(7): p. 699-708.

88. Lemaire, R., et al., *MALDI-MS direct tissue analysis of proteins: Improving signal sensitivity using organic treatments*. Anal Chem, 2006. **78**(20): p. 7145-53.
89. Seeley, E.H., et al., *Enhancement of protein sensitivity for MALDI imaging mass spectrometry after chemical treatment of tissue sections*. J Am Soc Mass Spectrom, 2008. **19**(8): p. 1069-77.
90. Franck, J., et al., *Improving tissue preparation for matrix-assisted laser desorption ionization mass spectrometry imaging. Part 1: using microspotting*. Anal Chem, 2009. **81**(19): p. 8193-202.
91. Grey, A.C., et al., *MALDI imaging mass spectrometry of integral membrane proteins from ocular lens and retinal tissue*. J Proteome Res, 2009. **8**(7): p. 3278-83.
92. Leinweber, B.D., et al., *Improved MALDI-TOF imaging yields increased protein signals at high molecular mass*. J Am Soc Mass Spectrom, 2009. **20**(1): p. 89-95.
93. Franck, J., et al., *MALDI mass spectrometry imaging of proteins exceeding 30000 Da*. Med Sci Monit, 2010: p. In press.
94. van Remoortere, A., et al., *MALDI imaging and profiling MS of higher mass proteins from tissue*. JASMS, 2010. In press.
95. El Ayed, M., et al., *MALDI imaging mass spectrometry in ovarian cancer for tracking, identifying, and validating biomarkers*. Med Sci Monit, 2010. **16**(8): p. BR233-45.
96. Kocher, T., A. Engstrom, and R.A. Zubarev, *Fragmentation of peptides in MALDI in-source decay mediated by hydrogen radicals*. Anal Chem, 2005. **77**(1): p. 172-7.
97. Schnaible, V., et al., *Screening for disulfide bonds in proteins by MALDI in-source decay and LIFT-TOF/TOF-MS*. Anal Chem, 2002. **74**(19): p. 4980-8.
98. Gao, J., et al., *A programmable fragmentation analysis of proteins by in-source decay in MALDI-TOF mass spectrometry*. Anal Chem, 2002. **74**(6): p. 1449-57.
99. Lennon, J.J. and K.A. Walsh, *Locating and identifying posttranslational modifications by in-source decay during MALDI-TOF mass spectrometry*. Protein Sci, 1999. **8**(11): p. 2487-93.
100. Calligaris, D., et al., *MALDI in-source decay of high mass protein isoforms: application to alpha- and beta-tubulin variants*. Anal Chem, 2010. **82**(14): p. 6176-84.
101. Debois, D., et al., *MALDI-in source decay applied to mass spectrometry imaging: a new tool for protein identification*. Anal Chem, 2010. **82**(10): p. 4036-45.
102. Groseclose, M.R., et al., *Identification of proteins directly from tissue: in situ tryptic digestions coupled with imaging mass spectrometry*. J Mass Spectrom, 2007. **42**(2): p. 254-62.
103. D'Amico, F., E. Skarmoutsou, and F. Stivala, *State of the art in antigen retrieval for immunohistochemistry*. J Immunol Methods, 2009. **341**(1-2): p. 1-18.
104. Gustafsson, J.O., et al., *Citric Acid Antigen Retrieval (CAAR) for tryptic peptide imaging directly on archived formalin-fixed paraffin-embedded tissue*. J Proteome Res, 2010.
105. Ronci, M., et al., *Protein unlocking procedures of formalin-fixed paraffin-embedded tissues: application to MALDI-TOF imaging MS investigations*. Proteomics, 2008. **8**(18): p. 3702-14.
106. Tian, Y. and H. Zhang, *Isolation of proteins by heat-induced extraction from formalin-fixed, paraffin-embedded tissue and preparation of tryptic peptides for mass spectrometric analysis*. Curr Protoc Mol Biol, 2010. **Chapter 10**: p. Unit 10 26 1-7.
107. Yamashita, S., *Heat-induced antigen retrieval: mechanisms and application to histochemistry*. Prog Histochem Cytochem, 2007. **41**(3): p. 141-200.
108. Franck, J., et al., *On tissue protein identification improvement by N-terminal peptide derivatization*. Methods Mol Biol, 2010. **656**: p. 323-38.

109. Franck, J., et al., *On-tissue N-terminal peptide derivatizations for enhancing protein identification in MALDI mass spectrometric imaging strategies*. Anal Chem, 2009. **81**(20): p. 8305-17.
110. Beavis, R.C. and B.T. Chait, *Matrix-assisted laser desorption ionization mass-spectrometry of proteins*. Methods Enzymol, 1996. **270**: p. 519-51.
111. Lemaire, R., et al., *Solid ionic matrixes for direct tissue analysis and MALDI imaging*. Anal Chem, 2006. **78**(3): p. 809-19.
112. Meriaux, C., et al., *Liquid ionic matrixes for MALDI mass spectrometry imaging of lipids*. J Proteomics, 2010. **73**(6): p. 1204-18.
113. Chan, K., et al., *MALDI mass spectrometry imaging of gangliosides in mouse brain using ionic liquid matrix*. Anal Chim Acta, 2009. **639**(1-2): p. 57-61.
114. Rujoi, M., R. Estrada, and M.C. Yappert, *In situ MALDI-TOF MS regional analysis of neutral phospholipids in lens tissue*. Anal Chem, 2004. **76**(6): p. 1657-63.
115. Jackson, S.N., H.Y. Wang, and A.S. Woods, *In situ structural characterization of phosphatidylcholines in brain tissue using MALDI-MS/MS*. J Am Soc Mass Spectrom, 2005. **16**(12): p. 2052-6.
116. Jackson, S.N., H.Y. Wang, and A.S. Woods, *Direct profiling of lipid distribution in brain tissue using MALDI-TOFMS*. Anal Chem, 2005. **77**(14): p. 4523-7.
117. Jackson, S.N., et al., *Direct tissue analysis of phospholipids in rat brain using MALDI-TOFMS and MALDI-ion mobility-TOFMS*. J Am Soc Mass Spectrom, 2005. **16**(2): p. 133-8.
118. Cha, S. and E.S. Yeung, *Colloidal graphite-assisted laser desorption/ionization mass spectrometry and MS<sub>n</sub> of small molecules. 1. Imaging of cerebroside directly from rat brain tissue*. Anal Chem, 2007. **79**(6): p. 2373-85.
119. Zhang, H., S. Cha, and E.S. Yeung, *Colloidal graphite-assisted laser desorption/ionization MS and MS<sub>n</sub> of small molecules. 2. Direct profiling and MS imaging of small metabolites from fruits*. Anal Chem, 2007. **79**(17): p. 6575-84.
120. Cha, S., et al., *Direct profiling and imaging of plant metabolites in intact tissues by using colloidal graphite-assisted laser desorption ionization mass spectrometry*. Plant J, 2008. **55**(2): p. 348-60.
121. Astigarraga, E., et al., *Profiling and imaging of lipids on brain and liver tissue by matrix-assisted laser desorption/ionization mass spectrometry using 2-mercaptopbenzothiazole as a matrix*. Anal Chem, 2008. **80**(23): p. 9105-14.
122. Benabdellah, F., et al., *In situ primary metabolites localization on a rat brain section by chemical mass spectrometry imaging*. Anal Chem, 2009. **81**(13): p. 5557-60.
123. Lorkiewicz, P. and M.C. Yappert, *Titanium microparticles and nanoparticles as matrixes for in vitro and in situ analysis of small molecules by MALDI-MS*. Anal Chem, 2009. **81**(16): p. 6596-603.
124. Aerni, H.R., D.S. Cornett, and R.M. Caprioli, *Automated acoustic matrix deposition for MALDI sample preparation*. Anal Chem, 2006. **78**(3): p. 827-34.
125. Baluya, D.L., T.J. Garrett, and R.A. Yost, *Automated MALDI matrix deposition method with inkjet printing for imaging mass spectrometry*. Anal Chem, 2007. **79**(17): p. 6862-7.
126. Chen, Y., et al., *Imaging MALDI mass spectrometry using an oscillating capillary nebulizer matrix coating system and its application to analysis of lipids in brain from a mouse model of Tay-Sachs/Sandhoff disease*. Anal Chem, 2008. **80**(8): p. 2780-8.
127. Hankin, J.A., R.M. Barkley, and R.C. Murphy, *Sublimation as a method of matrix application for mass spectrometric imaging*. J Am Soc Mass Spectrom, 2007. **18**(9): p. 1646-52.

128. Puolitaival, S.M., et al., *Solvent-free matrix dry-coating for MALDI imaging of phospholipids*. J Am Soc Mass Spectrom, 2008. **19**(6): p. 882-6.
129. Goodwin, R.J., et al., *Use of a solvent-free dry matrix coating for quantitative matrix-assisted laser desorption ionization imaging of 4-bromophenyl-1,4-diazabicyclo(3.2.2)nonane-4-carboxylate in rat brain and quantitative analysis of the drug from laser microdissected tissue regions*. Anal Chem, 2010. **82**(9): p. 3868-73.
130. Holle, A., et al., *Optimizing UV laser focus profiles for improved MALDI performance*. J Mass Spectrom, 2006. **41**(6): p. 705-16.
131. Fournier, I., et al., *Irradiation effects in MALDI, ablation, ion production, and surface modifications. Part II. 2,5-dihydroxybenzoic acid monocrystals*. J Am Soc Mass Spectrom, 2003. **14**(8): p. 893-9.
132. Trim, P.J., et al., *Introduction of a 20 kHz Nd:YVO4 laser into a hybrid quadrupole time-of-flight mass spectrometer for MALDI-MS imaging*. Anal Bioanal Chem, 2010. **397**(8): p. 3409-19.
133. Bouschen, W., et al., *Matrix vapor deposition/recrystallization and dedicated spray preparation for high-resolution scanning microprobe matrix-assisted laser desorption/ionization imaging mass spectrometry (SMALDI-MS) of tissue and single cells*. Rapid Commun Mass Spectrom, 2010. **24**(3): p. 355-64.
134. Trim, P.J., et al., *Introduction of a 20 kHz Nd:YVO4 laser into a hybrid quadrupole time-of-flight mass spectrometer for MALDI-MS imaging*. Anal Bioanal Chem.
135. Jardin-Mathe, O., et al., *MITICS (MALDI Imaging Team Imaging Computing System): a new open source mass spectrometry imaging software*. J Proteomics, 2008. **71**(3): p. 332-45.
136. Schwamborn, K., et al., *Identifying prostate carcinoma by MALDI-Imaging*. Int J Mol Med, 2007. **20**(2): p. 155-9.
137. Masumori, N., et al., *A probasin-large T antigen transgenic mouse line develops prostate adenocarcinoma and neuroendocrine carcinoma with metastatic potential*. Cancer Res, 2001. **61**(5): p. 2239-49.
138. Lemaire, R., et al., *Specific MALDI imaging and profiling for biomarker hunting and validation: fragment of the 11S proteasome activator complex, Reg alpha fragment, is a new potential ovary cancer biomarker*. J Proteome Res, 2007. **6**(11): p. 4127-34.
139. Rauser, S., et al., *Classification of HER2 receptor status in breast cancer tissues by MALDI imaging mass spectrometry*. J Proteome Res. **9**(4): p. 1854-63.
140. Pevsner, P.H., et al., *Mass spectrometry MALDI imaging of colon cancer biomarkers: a new diagnostic paradigm*. Biomark Med, 2009. **3**(1): p. 55-69.
141. Bhattacharya, S.H., A.A. Gal, and K.K. Murray, *Laser capture microdissection MALDI for direct analysis of archival tissue*. J Proteome Res, 2003. **2**(1): p. 95-8.
142. Sinha, T.K., et al., *Integrating spatially resolved three-dimensional MALDI IMS with in vivo magnetic resonance imaging*. Nat Methods, 2008. **5**(1): p. 57-9.
143. Schwartz, S.A., et al., *Protein profiling in brain tumors using mass spectrometry: feasibility of a new technique for the analysis of protein expression*. Clin Cancer Res, 2004. **10**(3): p. 981-7.
144. Morrow, P.K., F. Zambrana, and F.J. Esteva, *Recent advances in systemic therapy: Advances in systemic therapy for HER2-positive metastatic breast cancer*. Breast Cancer Res, 2009. **11**(4): p. 207.
145. Han, X., et al., *Substantial sulfatide deficiency and ceramide elevation in very early Alzheimer's disease: potential role in disease pathogenesis*. J Neurochem, 2002. **82**(4): p. 809-18.
146. Murphy, E.J., et al., *Phospholipid composition and levels are altered in Down syndrome brain*. Brain Res, 2000. **867**(1-2): p. 9-18.

147. Han, X., et al., *Alterations in myocardial cardiolipin content and composition occur at the very earliest stages of diabetes: a shotgun lipidomics study*. *Biochemistry*, 2007. **46**(21): p. 6417-28.
148. Shimma, S., et al., *MALDI-based imaging mass spectrometry revealed abnormal distribution of phospholipids in colon cancer liver metastasis*. *J Chromatogr B Analyt Technol Biomed Life Sci*, 2007. **855**(1): p. 98-103.
149. Liu, Y., et al., *Elevation of sulfatides in ovarian cancer: An integrated transcriptomic and lipidomic analysis including tissue-imaging mass spectrometry*. *Mol Cancer*. **9**(1): p. 186.
150. Thomas, R.L., Jr., et al., *Mass spectrometric identification of increased C16 ceramide levels during apoptosis*. *J Biol Chem*, 1999. **274**(43): p. 30580-8.
151. Fuchs, B., J. Schiller, and M.A. Cross, *Apoptosis-associated changes in the glycerophospholipid composition of hematopoietic progenitor cells monitored by <sup>31</sup>P NMR spectroscopy and MALDI-TOF mass spectrometry*. *Chem Phys Lipids*, 2007. **150**(2): p. 229-38.
152. Fahy, E., et al., *Update of the LIPID MAPS comprehensive classification system for lipids*. *J Lipid Res*, 2009. **50 Suppl**: p. S9-14.
153. Vidova, V., et al., *Visualizing spatial lipid distribution in porcine lens by MALDI imaging high-resolution mass spectrometry*. *J Lipid Res*.
154. Wang, H.Y., et al., *A Minimalist Approach to MALDI Imaging of Glycerophospholipids and Sphingolipids in Rat Brain Sections*. *Int J Mass Spectrom*, 2008. **278**(2-3): p. 143-149.
155. Vrkoslav, V., et al., *MALDI imaging of neutral cuticular lipids in insects and plants*. *J Am Soc Mass Spectrom*. **21**(2): p. 220-31.
156. Sugiura, Y., et al., *Visualization of the cell-selective distribution of PUFA-containing phosphatidylcholines in mouse brain by imaging mass spectrometry*. *J Lipid Res*, 2009. **50**(9): p. 1776-88.
157. Meriaux, C., et al., *Liquid ionic matrixes for MALDI mass spectrometry imaging of lipids*. *J Proteomics*. **73**(6): p. 1204-18.
158. Solon, E.G., et al., *Autoradiography, MALDI-MS, and SIMS-MS imaging in pharmaceutical discovery and development*. *AAPS J*. **12**(1): p. 11-26.
159. Khatib-Shahidi, S., et al., *Direct molecular analysis of whole-body animal tissue sections by imaging MALDI mass spectrometry*. *Anal Chem*, 2006. **78**(18): p. 6448-56.
160. Pacholski, M.L. and N. Winograd, *Imaging with mass spectrometry*. *Chem Rev*, 1999. **99**(10): p. 2977-3006.
161. Todd, P.J., et al., *Organic ion imaging of biological tissue with secondary ion mass spectrometry and matrix-assisted laser desorption/ionization*. *J Mass Spectrom*, 2001. **36**(4): p. 355-69.
162. Brunelle, A., D. Touboul, and O. Laprevote, *Biological tissue imaging with time-of-flight secondary ion mass spectrometry and cluster ion sources*. *J Mass Spectrom*, 2005. **40**(8): p. 985-99.
163. Guerquin-Kern, J.L., et al., *Progress in analytical imaging of the cell by dynamic secondary ion mass spectrometry (SIMS microscopy)*. *Biochim Biophys Acta*, 2005. **1724**(3): p. 228-38.
164. Ifa, D.R., et al., *Desorption electrospray ionization and other ambient ionization methods: current progress and preview*. *Analyst*. **135**(4): p. 669-81.
165. Girod, M., et al., *Desorption electrospray ionization imaging mass spectrometry of lipids in rat spinal cord*. *J Am Soc Mass Spectrom*. **21**(7): p. 1177-89.



166. Crecelius, A.C., et al., *Three-dimensional visualization of protein expression in mouse brain structures using imaging mass spectrometry*. J Am Soc Mass Spectrom, 2005. **16**(7): p. 1093-9.
167. Puente, X.S., et al., *Human and mouse proteases: a comparative genomic approach*. Nat Rev Genet, 2003. **4**(7): p. 544-58.
168. Overall, C.M. and C.P. Blobel, *In search of partners: linking extracellular proteases to substrates*. Nat Rev Mol Cell Biol, 2007. **8**(3): p. 245-57.
169. Steiner, D.F., et al., *Insulin biosynthesis: evidence for a precursor*. Science, 1967. **157**(789): p. 697-700.
170. Roebroek, A.J., et al., *Characterization of human c-fes/fps reveals a new transcription unit (fur) in the immediately upstream region of the proto-oncogene*. Mol Biol Rep, 1986. **11**(2): p. 117-25.
171. Fuller, R.S., A.J. Brake, and J. Thorner, *Intracellular targeting and structural conservation of a prohormone-processing endoprotease*. Science, 1989. **246**(4929): p. 482-6.
172. Rouille, Y., et al., *Proteolytic processing mechanisms in the biosynthesis of neuroendocrine peptides: the subtilisin-like proprotein convertases*. Front Neuroendocrinol, 1995. **16**(4): p. 322-61.
173. Seidah, N.G., A.M. Khatib, and A. Prat, *The proprotein convertases and their implication in sterol and/or lipid metabolism*. Biol Chem, 2006. **387**(7): p. 871-7.
174. Fugere, M. and R. Day, *Cutting back on pro-protein convertases: the latest approaches to pharmacological inhibition*. Trends Pharmacol Sci, 2005. **26**(6): p. 294-301.
175. Barrett, A.J., N.D. Rawlings, and E.A. O'Brien, *The MEROPS database as a protease information system*. J Struct Biol, 2001. **134**(2-3): p. 95-102.
176. Seidah, N.G., M. Chretien, and R. Day, *The family of subtilisin/kexin like pro-protein and pro-hormone convertases: divergent or shared functions*. Biochimie, 1994. **76**(3-4): p. 197-209.
177. Zhou, A., et al., *Proteolytic processing in the secretory pathway*. J Biol Chem, 1999. **274**(30): p. 20745-8.
178. Poirier, S., et al., *Dissection of the endogenous cellular pathways of PCSK9-induced low density lipoprotein receptor degradation: evidence for an intracellular route*. J Biol Chem, 2009. **284**(42): p. 28856-64.
179. Nassoury, N., et al., *The cellular trafficking of the secretory proprotein convertase PCSK9 and its dependence on the LDLR*. Traffic, 2007. **8**(6): p. 718-32.
180. Benjannet, S., et al., *NARC-1/PCSK9 and its natural mutants: zymogen cleavage and effects on the low density lipoprotein (LDL) receptor and LDL cholesterol*. J Biol Chem, 2004. **279**(47): p. 48865-75.
181. Khatib, A.M., et al., *Proprotein convertases in tumor progression and malignancy: novel targets in cancer therapy*. Am J Pathol, 2002. **160**(6): p. 1921-35.
182. Aaronson, S.A., *Growth factors and cancer*. Science, 1991. **254**(5035): p. 1146-53.
183. Duguay, S.J., J. Lai-Zhang, and D.F. Steiner, *Mutational analysis of the insulin-like growth factor I prohormone processing site*. J Biol Chem, 1995. **270**(29): p. 17566-74.
184. Duguay, S.J., et al., *Post-translational processing of the insulin-like growth factor-2 precursor. Analysis of O-glycosylation and endoproteolysis*. J Biol Chem, 1998. **273**(29): p. 18443-51.
185. Kayo, T., et al., *Proprotein-processing endoprotease furin controls growth of pancreatic beta-cells*. Diabetes, 1997. **46**(8): p. 1296-304.
186. Siegfried, G., et al., *The proteolytic processing of pro-platelet-derived growth factor-A at RRKR(86) by members of the proprotein convertase family is functionally*

- correlated to platelet-derived growth factor-A-induced functions and tumorigenicity.* Cancer Res, 2003. **63**(7): p. 1458-63.
187. Leitlein, J., et al., *Processing of immunosuppressive pro-TGF-beta 1,2 by human glioblastoma cells involves cytoplasmic and secreted furin-like proteases.* J Immunol, 2001. **166**(12): p. 7238-43.
  188. Siegfried, G., et al., *The secretory proprotein convertases furin, PC5, and PC7 activate VEGF-C to induce tumorigenesis.* J Clin Invest, 2003. **111**(11): p. 1723-32.
  189. Khatib, A.M., *Regulation of Carcinogenesis, Angiogenesis and Metastasis by the Proprotein Convertases (PCs) 2006*, Netherlands: Springer.
  190. Komada, M., et al., *Proteolytic processing of the hepatocyte growth factor/scatter factor receptor by furin.* FEBS Lett, 1993. **328**(1-2): p. 25-9.
  191. Papa, V., et al., *Elevated insulin receptor content in human breast cancer.* J Clin Invest, 1990. **86**(5): p. 1503-10.
  192. Khatib, A.M., et al., *Inhibition of proprotein convertases is associated with loss of growth and tumorigenicity of HT-29 human colon carcinoma cells: importance of insulin-like growth factor-1 (IGF-1) receptor processing in IGF-1-mediated functions.* J Biol Chem, 2001. **276**(33): p. 30686-93.
  193. Hynes, R.O., *Integrins: versatility, modulation, and signaling in cell adhesion.* Cell, 1992. **69**(1): p. 11-25.
  194. Hynes, R.O., *Cell adhesion: old and new questions.* Trends Cell Biol, 1999. **9**(12): p. M33-7.
  195. Brakebusch, C. and R. Fassler, *The integrin-actin connection, an eternal love affair.* EMBO J, 2003. **22**(10): p. 2324-33.
  196. Juliano, R.L. and S. Haskill, *Signal transduction from the extracellular matrix.* J Cell Biol, 1993. **120**(3): p. 577-85.
  197. Giancotti, F.G. and F. Mainiero, *Integrin-mediated adhesion and signaling in tumorigenesis.* Biochim Biophys Acta, 1994. **1198**(1): p. 47-64.
  198. Lissitzky, J.C., et al., *Endoproteolytic processing of integrin pro-alpha subunits involves the redundant function of furin and proprotein convertase (PC) 5A, but not paired basic amino acid converting enzyme (PACE) 4, PC5B or PC7.* Biochem J, 2000. **346 Pt 1**: p. 133-8.
  199. Berthet, V., et al., *Role of endoproteolytic processing in the adhesive and signaling functions of alphavbeta5 integrin.* J Biol Chem, 2000. **275**(43): p. 33308-13.
  200. Daemi, N., et al., *Anti-beta4 integrin antibodies enhance migratory and invasive abilities of human colon adenocarcinoma cells and their MMP-2 expression.* Int J Cancer, 2000. **85**(6): p. 850-6.
  201. Khatib, A.M., et al., *Regulation of urokinase plasminogen activator/plasmin-mediated invasion of melanoma cells by the integrin vitronectin receptor alphaVbeta3.* Int J Cancer, 2001. **91**(3): p. 300-8.
  202. Lehmann, M., et al., *Lack of integrin alpha-chain endoproteolytic cleavage in furin-deficient human colon adenocarcinoma cells LoVo.* Biochem J, 1996. **317 ( Pt 3)**: p. 803-9.
  203. Heil, M., et al., *Vascular endothelial growth factor (VEGF) stimulates monocyte migration through endothelial monolayers via increased integrin expression.* Eur J Cell Biol, 2000. **79**(11): p. 850-7.
  204. Mantovani, A., et al., *Regulation of endothelial cell function by pro- and anti-inflammatory cytokines.* Transplant Proc, 1998. **30**(8): p. 4239-43.
  205. Poher, J.S. and R.S. Cotran, *The role of endothelial cells in inflammation.* Transplantation, 1990. **50**(4): p. 537-44.

206. Balaram, S.K., D.K. Agrawal, and J.D. Edwards, *Insulin like growth factor-1 activates nuclear factor-kappaB and increases transcription of the intercellular adhesion molecule-1 gene in endothelial cells*. Cardiovasc Surg, 1999. **7**(1): p. 91-7.
207. Ishizuka, T., et al., *Endothelin-1 enhances vascular cell adhesion molecule-1 expression in tumor necrosis factor alpha-stimulated vascular endothelial cells*. Eur J Pharmacol, 1999. **369**(2): p. 237-45.
208. Denault, J.B., et al., *Processing of proendothelin-1 by human furin convertase*. FEBS Lett, 1995. **362**(3): p. 276-80.
209. Black, R.A. and J.M. White, *ADAMs: focus on the protease domain*. Curr Opin Cell Biol, 1998. **10**(5): p. 654-9.
210. Lunn, C.A., et al., *Purification of ADAM 10 from bovine spleen as a TNFalpha convertase*. FEBS Lett, 1997. **400**(3): p. 333-5.
211. Christofori, G. and H. Semb, *The role of the cell-adhesion molecule E-cadherin as a tumour-suppressor gene*. Trends Biochem Sci, 1999. **24**(2): p. 73-6.
212. Kinsella, A.R., et al., *The role of the cell-cell adhesion molecule E-cadherin in large bowel tumour cell invasion and metastasis*. Br J Cancer, 1993. **67**(5): p. 904-9.
213. Birchmeier, W. and J. Behrens, *Cadherin expression in carcinomas: role in the formation of cell junctions and the prevention of invasiveness*. Biochim Biophys Acta, 1994. **1198**(1): p. 11-26.
214. Mareel, M., et al., *E-cadherin/catenin/cytoskeleton complex: a regulator of cancer invasion*. J Cell Physiol, 1997. **173**(2): p. 271-4.
215. De Wever, O., et al., *Critical role of N-cadherin in myofibroblast invasion and migration in vitro stimulated by colon-cancer-cell-derived TGF-beta or wounding*. J Cell Sci, 2004. **117**(Pt 20): p. 4691-703.
216. Shimazui, T., et al., *The level of cadherin-6 mRNA in peripheral blood is associated with the site of metastasis and with the subsequent occurrence of metastases in renal cell carcinoma*. Cancer, 2004. **101**(5): p. 963-8.
217. Lee, S.W., *H-cadherin, a novel cadherin with growth inhibitory functions and diminished expression in human breast cancer*. Nat Med, 1996. **2**(7): p. 776-82.
218. Stetler-Stevenson, W.G. and A.E. Yu, *Proteases in invasion: matrix metalloproteinases*. Semin Cancer Biol, 2001. **11**(2): p. 143-52.
219. Baker, A.H., D.R. Edwards, and G. Murphy, *Metalloproteinase inhibitors: biological actions and therapeutic opportunities*. J Cell Sci, 2002. **115**(Pt 19): p. 3719-27.
220. Noda, M., et al., *RECK: a novel suppressor of malignancy linking oncogenic signaling to extracellular matrix remodeling*. Cancer Metastasis Rev, 2003. **22**(2-3): p. 167-75.
221. Liotta, L.A., et al., *Metastatic potential correlates with enzymatic degradation of basement membrane collagen*. Nature, 1980. **284**(5751): p. 67-8.
222. Basset, P., et al., *Matrix metalloproteinases as stromal effectors of human carcinoma progression: therapeutic implications*. Matrix Biol, 1997. **15**(8-9): p. 535-41.
223. Westermarck, J. and V.M. Kahari, *Regulation of matrix metalloproteinase expression in tumor invasion*. FASEB J, 1999. **13**(8): p. 781-92.
224. Nelson, A.R., et al., *Matrix metalloproteinases: biologic activity and clinical implications*. J Clin Oncol, 2000. **18**(5): p. 1135-49.
225. Sato, H., et al., *A matrix metalloproteinase expressed on the surface of invasive tumour cells*. Nature, 1994. **370**(6484): p. 61-5.
226. Polette, M. and P. Birembaut, *Membrane-type metalloproteinases in tumor invasion*. Int J Biochem Cell Biol, 1998. **30**(11): p. 1195-202.
227. Endo, K., et al., *Cleavage of syndecan-1 by membrane type matrix metalloproteinase-1 stimulates cell migration*. J Biol Chem, 2003. **278**(42): p. 40764-70.

228. Collier, I.E., et al., *H-ras oncogene-transformed human bronchial epithelial cells (TBE-1) secrete a single metalloprotease capable of degrading basement membrane collagen*. J Biol Chem, 1988. **263**(14): p. 6579-87.
229. Li, Y., et al., *Cleavage of lumican by membrane-type matrix metalloproteinase-1 abrogates this proteoglycan-mediated suppression of tumor cell colony formation in soft agar*. Cancer Res, 2004. **64**(19): p. 7058-64.
230. Koshikawa, N., et al., *Role of cell surface metalloprotease MT1-MMP in epithelial cell migration over laminin-5*. J Cell Biol, 2000. **148**(3): p. 615-24.
231. Ratnikov, B.I., et al., *An alternative processing of integrin alpha(v) subunit in tumor cells by membrane type-1 matrix metalloproteinase*. J Biol Chem, 2002. **277**(9): p. 7377-85.
232. Nagase, H. and J.F. Woessner, Jr., *Matrix metalloproteinases*. J Biol Chem, 1999. **274**(31): p. 21491-4.
233. Guo, C., et al., *Paradigmatic identification of MMP-2 and MT1-MMP activation systems in cardiac fibroblasts cultured as a monolayer*. J Cell Biochem, 2005. **94**(3): p. 446-59.
234. Sato, T., et al., *Furin-independent pathway of membrane type 1-matrix metalloproteinase activation in rabbit dermal fibroblasts*. J Biol Chem, 1999. **274**(52): p. 37280-4.
235. Sardinha, T.C., et al., *Membrane-type 1 matrix metalloproteinase mRNA expression in colorectal cancer*. Dis Colon Rectum, 2000. **43**(3): p. 389-95.
236. Yana, I. and S.J. Weiss, *Regulation of membrane type-1 matrix metalloproteinase activation by proprotein convertases*. Mol Biol Cell, 2000. **11**(7): p. 2387-401.
237. Rozanov, D.V. and A.Y. Strongin, *Membrane type-1 matrix metalloproteinase functions as a proprotein self-convertase. Expression of the latent zymogen in Pichia pastoris, autolytic activation, and the peptide sequence of the cleavage forms*. J Biol Chem, 2003. **278**(10): p. 8257-60.
238. Primakoff, P. and D.G. Myles, *The ADAM gene family: surface proteins with adhesion and protease activity*. Trends Genet, 2000. **16**(2): p. 83-7.
239. Loechel, F., et al., *Human ADAM 12 (meltrin alpha) is an active metalloprotease*. J Biol Chem, 1998. **273**(27): p. 16993-7.
240. Wu, E., P.I. Croucher, and N. McKie, *Expression of members of the novel membrane linked metalloproteinase family ADAM in cells derived from a range of haematological malignancies*. Biochem Biophys Res Commun, 1997. **235**(2): p. 437-42.
241. Schlondorff, J. and C.P. Blobel, *Metalloprotease-disintegrins: modular proteins capable of promoting cell-cell interactions and triggering signals by protein-ectodomain shedding*. J Cell Sci, 1999. **112 ( Pt 21)**: p. 3603-17.
242. Wolfsberg, T.G. and J.M. White, *ADAMs in fertilization and development*. Dev Biol, 1996. **180**(2): p. 389-401.
243. Koike, H., et al., *Membrane-anchored metalloprotease MDC9 has an alpha-secretase activity responsible for processing the amyloid precursor protein*. Biochem J, 1999. **343 Pt 2**: p. 371-5.
244. Nath, D., et al., *Interaction of metargidin (ADAM-15) with alphavbeta3 and alpha5beta1 integrins on different haemopoietic cells*. J Cell Sci, 1999. **112 ( Pt 4)**: p. 579-87.
245. Lopez-Perez, E., et al., *Constitutive alpha-secretase cleavage of the beta-amyloid precursor protein in the furin-deficient LoVo cell line: involvement of the pro-hormone convertase 7 and the disintegrin metalloprotease ADAM10*. J Neurochem, 2001. **76**(5): p. 1532-9.

246. Chen, M.S., et al., *Evidence that distinct states of the integrin alpha6beta1 interact with laminin and an ADAM*. J Cell Biol, 1999. **144**(3): p. 549-61.
247. Eto, K., et al., *RGD-independent binding of integrin alpha9beta1 to the ADAM-12 and -15 disintegrin domains mediates cell-cell interaction*. J Biol Chem, 2000. **275**(45): p. 34922-30.
248. Kuno, K., et al., *Molecular cloning of a gene encoding a new type of metalloproteinase-disintegrin family protein with thrombospondin motifs as an inflammation associated gene*. J Biol Chem, 1997. **272**(1): p. 556-62.
249. Kuno, K., Y. Terashima, and K. Matsushima, *ADAMTS-1 is an active metalloproteinase associated with the extracellular matrix*. J Biol Chem, 1999. **274**(26): p. 18821-6.
250. Basak, A. and C. Lazure, *Synthetic peptides derived from the prosegments of proprotein convertase 1/3 and furin are potent inhibitors of both enzymes*. Biochem J, 2003. **373**(Pt 1): p. 231-9.
251. Becker, G.L., et al., *Potent inhibitors of furin and furin-like proprotein convertases containing decarboxylated P1 arginine mimetics*. J Med Chem, 2010. **53**(3): p. 1067-75.
252. Fugere, M., et al., *Short polybasic peptide sequences are potent inhibitors of PC5/6 and PC7: Use of positional scanning-synthetic peptide combinatorial libraries as a tool for the optimization of inhibitory sequences*. Mol Pharmacol, 2007. **71**(1): p. 323-32.
253. Basak, A., *Inhibitors of proprotein convertases*. J Mol Med, 2005. **83**(11): p. 844-55.
254. Bassi, D.E., et al., *Proprotein convertase inhibition results in decreased skin cell proliferation, tumorigenesis, and metastasis*. Neoplasia, 2010. **12**(7): p. 516-26.
255. McCombie, G., et al., *Spatial and spectral correlations in MALDI mass spectrometry images by clustering and multivariate analysis*. Anal Chem, 2005. **77**(19): p. 6118-24.
256. Maddalo, G., et al., *Analytical assessment of MALDI-TOF Imaging Mass Spectrometry on thin histological samples. An insight in proteome investigation*. Clin Chim Acta, 2005. **357**(2): p. 210-8.
257. Van de Plas, R., et al., *Prospective exploration of biochemical tissue composition via imaging mass spectrometry guided by principal component analysis*. Pac Symp Biocomput, 2007: p. 458-69.
258. Trim, P.J., et al., *Matrix-assisted laser desorption/ionisation mass spectrometry imaging of lipids in rat brain tissue with integrated unsupervised and supervised multivariate statistical analysis*. Rapid Commun Mass Spectrom, 2008. **22**(10): p. 1503-9.
259. McDonnell, L.A., et al., *Mass spectrometry image correlation: quantifying colocalization*. J Proteome Res, 2008. **7**(8): p. 3619-27.
260. Deininger, S.O., et al., *MALDI Imaging Combined with Hierarchical Clustering as a New Tool for the Interpretation of Complex Human Cancers*. J Proteome Res, 2008.
261. Rauser, S., et al., *Classification of HER2 receptor status in breast cancer tissues by MALDI imaging mass spectrometry*. J Proteome Res, 2010. **9**(4): p. 1854-63.
262. Rauser, S., H. Hofler, and A. Walch, *[MALDI imaging mass spectrometry for direct tissue analysis]*. Pathologe, 2009. **30 Suppl 2**: p. 140-5.
263. Ruschoff, J., et al., *HER2 diagnostics in gastric cancer-guideline validation and development of standardized immunohistochemical testing*. Virchows Arch.
264. Bast, R.C., Jr., et al., *Reactivity of a monoclonal antibody with human ovarian carcinoma*. J Clin Invest, 1981. **68**(5): p. 1331-7.

265. Moore, R.G., et al., *The use of multiple novel tumor biomarkers for the detection of ovarian carcinoma in patients with a pelvic mass*. Gynecol Oncol, 2008. **108**(2): p. 402-8.
266. Skates, S.J., et al., *Calculation of the risk of ovarian cancer from serial CA-125 values for preclinical detection in postmenopausal women*. J Clin Oncol, 2003. **21**(10 Suppl): p. 206s-210s.
267. Cazares, L.H., et al., *Imaging mass spectrometry of a specific fragment of mitogen-activated protein kinase/extracellular signal-regulated kinase kinase 2 discriminates cancer from uninvolved prostate tissue*. Clin Cancer Res, 2009. **15**(17): p. 5541-51.
268. Samsi, S.S., et al., *Imaging mass spectrometry analysis for follicular lymphoma grading*. Conf Proc IEEE Eng Med Biol Soc, 2009. **2009**: p. 6969-72.
269. Djidja, M.C., et al., *MALDI-ion mobility separation-mass spectrometry imaging of glucose-regulated protein 78 kDa (Grp78) in human formalin-fixed, paraffin-embedded pancreatic adenocarcinoma tissue sections*. J Proteome Res, 2009. **8**(10): p. 4876-84.
270. McNaught, K.S., et al., *Failure of the ubiquitin-proteasome system in Parkinson's disease*. Nat Rev Neurosci, 2001. **2**(8): p. 589-94.
271. Sompallae, R., et al., *Epstein-barr virus encodes three bona fide ubiquitin-specific proteases*. J Virol, 2008. **82**(21): p. 10477-86.
272. Ozols, R.F., *Future directions in the treatment of ovarian cancer*. Semin Oncol, 2002. **29**(1 Suppl 1): p. 32-42.
273. Alexander-Sefre, F., U. Menon, and I.J. Jacobs, *Ovarian cancer screening*. Hosp Med, 2002. **63**(4): p. 210-3.
274. Hollingsworth, M.A. and B.J. Swanson, *Mucins in cancer: protection and control of the cell surface*. Nat Rev Cancer, 2004. **4**(1): p. 45-60.
275. Moniaux, N., et al., *Structural organization and classification of the human mucin genes*. Front Biosci, 2001. **6**: p. D1192-206.
276. Maines-Bandiera, S., et al., *Oviductal glycoprotein (OVGP1, MUC9): a differentiation-based mucin present in serum of women with ovarian cancer*. Int J Gynecol Cancer, 2010. **20**(1): p. 16-22.
277. Chauhan, S.C., D. Kumar, and M. Jaggi, *Mucins in ovarian cancer diagnosis and therapy*. J Ovarian Res, 2009. **2**: p. 21.
278. Lansac, G., et al., *Lipopolysaccharide mediated regulation of neuroendocrine associated proprotein convertases and neuropeptide precursor processing in the rat spleen*. J Neuroimmunol, 2006. **171**(1-2): p. 57-71.
279. Sorem, J., T.S. Jardetzky, and R. Longnecker, *Cleavage and secretion of Epstein-Barr virus glycoprotein 42 promote membrane fusion with B lymphocytes*. J Virol, 2009. **83**(13): p. 6664-72.
280. Apte, S. and D.A. Sanders, *Effects of retroviral envelope-protein cleavage upon trafficking, incorporation, and membrane fusion*. Virology, 2010. **405**(1): p. 214-24.
281. Nakayama, K., *Furin: a mammalian subtilisin/Kex2p-like endoprotease involved in processing of a wide variety of precursor proteins*. Biochem J, 1997. **327** ( Pt 3): p. 625-35.
282. Xiao, Y., et al., *Cell-surface processing of extracellular human immunodeficiency virus type 1 Vpr by proprotein convertases*. Virology, 2008. **372**(2): p. 384-97.
283. Seidah, N.G., *Unexpected similarity between the cytosolic West Nile virus NS3 and the secretory furin-like serine proteinases*. Biochem J, 2006. **393**(Pt 2): p. e1-3.
284. Greenlee, R.T., et al., *Cancer statistics, 2001*. CA Cancer J Clin, 2001. **51**(1): p. 15-36.

285. Bolla, M., et al., *Long-term results with immediate androgen suppression and external irradiation in patients with locally advanced prostate cancer (an EORTC study): a phase III randomised trial*. Lancet, 2002. **360**(9327): p. 103-6.
286. Pilepich, M.V., et al., *Phase III radiation therapy oncology group (RTOG) trial 86-10 of androgen deprivation adjuvant to definitive radiotherapy in locally advanced carcinoma of the prostate*. Int J Radiat Oncol Biol Phys, 2001. **50**(5): p. 1243-52.
287. Messing, E.M., et al., *Immediate hormonal therapy compared with observation after radical prostatectomy and pelvic lymphadenectomy in men with node-positive prostate cancer*. N Engl J Med, 1999. **341**(24): p. 1781-8.
288. Walsh, P.C., *Immediate versus deferred treatment for advanced prostatic cancer: initial results of the Medical Research Council trial*. The Medical Research Council Prostate Cancer Working Party Investigators Group. J Urol, 1997. **158**(4): p. 1623-4.
289. Petrylak, D.P., et al., *Docetaxel and estramustine compared with mitoxantrone and prednisone for advanced refractory prostate cancer*. N Engl J Med, 2004. **351**(15): p. 1513-20.
290. Wang, M.C., et al., *Purification of a human prostate specific antigen*. Invest Urol, 1979. **17**(2): p. 159-63.
291. Riegman, P.H., et al., *Characterization of the prostate-specific antigen gene: a novel human kallikrein-like gene*. Biochem Biophys Res Commun, 1989. **159**(1): p. 95-102.
292. Lee, C., et al., *Demonstration of the role of prostate-specific antigen in semen liquefaction by two-dimensional electrophoresis*. J Androl, 1989. **10**(6): p. 432-8.
293. Iwakiri, J., et al., *An analysis of urinary prostate specific antigen before and after radical prostatectomy: evidence for secretion of prostate specific antigen by the periurethral glands*. J Urol, 1993. **149**(4): p. 783-6.
294. Thompson, I.M., et al., *Prevalence of prostate cancer among men with a prostate-specific antigen level < or =4.0 ng per milliliter*. N Engl J Med, 2004. **350**(22): p. 2239-46.
295. Raaijmakers, R., et al., *Prostate cancer detection in the prostate specific antigen range of 2.0 to 3.9 ng/ml: value of percent free prostate specific antigen on tumor detection and tumor aggressiveness*. J Urol, 2004. **171**(6 Pt 1): p. 2245-9.
296. Uchida, K., et al., *Proprotein convertases regulate activity of prostate epithelial cell differentiation markers and are modulated in human prostate cancer cells*. J Cell Biochem, 2003. **88**(2): p. 394-9.
297. Navone, N.M., et al., *Model systems of prostate cancer: uses and limitations*. Cancer Metastasis Rev, 1998. **17**(4): p. 361-71.
298. van Bokhoven, A., et al., *Molecular characterization of human prostate carcinoma cell lines*. Prostate, 2003. **57**(3): p. 205-25.
299. Fizaki, K. and N.M. Navone, *Modèles précliniques de cancer de la prostate*. Bull Cancer.2005. **92**: p. 129-41.
300. Stone, K.R., et al., *Isolation of a human prostate carcinoma cell line (DU 145)*. Int J Cancer, 1978. **21**(3): p. 274-81.
301. Hu, Y., et al., *Molecular characterization of a metastatic neuroendocrine cell cancer arising in the prostates of transgenic mice*. J Biol Chem, 2002. **277**(46): p. 44462-74.
302. Huang, J., et al., *Function and molecular mechanisms of neuroendocrine cells in prostate cancer*. Anal Quant Cytol Histol, 2007. **29**(3): p. 128-38.
303. Bergeron, F., R. Leduc, and R. Day, *Subtilase-like pro-protein convertases: from molecular specificity to therapeutic applications*. J Mol Endocrinol, 2000. **24**(1): p. 1-22.
304. Constam, D.B. and E.J. Robertson, *SPC4/PACE4 regulates a TGFbeta signaling network during axis formation*. Genes Dev, 2000. **14**(9): p. 1146-55.

305. Scamuffa, N., et al., *Proprotein convertases: lessons from knockouts*. FASEB J, 2006. **20**(12): p. 1954-63.
306. Taniguchi, T., et al., *A critical role for the carboxy terminal region of the proprotein convertase, PACE4A, in the regulation of its autocatalytic activation coupled with secretion*. Biochem Biophys Res Commun, 2002. **290**(2): p. 878-84.
307. Nour, N., et al., *The cysteine-rich domain of the secreted proprotein convertases PC5A and PACE4 functions as a cell surface anchor and interacts with tissue inhibitors of metalloproteinases*. Mol Biol Cell, 2005. **16**(11): p. 5215-26.
308. Tsuji, A., et al., *Secretory proprotein convertases PACE4 and PC6A are heparin-binding proteins which are localized in the extracellular matrix. Potential role of PACE4 in the activation of proproteins in the extracellular matrix*. Biochim Biophys Acta, 2003. **1645**(1): p. 95-104.
309. Liu, J., et al., *Angiopoietin-like protein 3 inhibits lipoprotein lipase activity through enhancing its cleavage by proprotein convertases*. J Biol Chem, 2010.
310. Momeni, P., et al., *Mutations in a new gene, encoding a zinc-finger protein, cause tricho-rhino-phalangeal syndrome type I*. Nat Genet, 2000. **24**(1): p. 71-4.
311. Chang, G.T., et al., *Characterization of a zinc-finger protein and its association with apoptosis in prostate cancer cells*. J Natl Cancer Inst, 2000. **92**(17): p. 1414-21.
312. Chang, G.T., et al., *Structure and function of GC79/TRPS1, a novel androgen-repressible apoptosis gene*. Apoptosis, 2002. **7**(1): p. 13-21.
313. van den Bemd, G.J., et al., *The atypical GATA protein TRPS1 represses androgen-induced prostate-specific antigen expression in LNCaP prostate cancer cells*. Biochem Biophys Res Commun, 2003. **312**(3): p. 578-84.
314. Chang, G.T., et al., *Proteomic analysis of proteins regulated by TRPS1 transcription factor in DU145 prostate cancer cells*. Biochim Biophys Acta, 2007. **1774**(5): p. 575-82.
315. De Pinto, V., et al., *New functions of an old protein: the eukaryotic porin or voltage dependent anion selective channel (VDAC)*. Ital J Biochem, 2003. **52**(1): p. 17-24.
316. Cheng, E.H., et al., *VDAC2 inhibits BAK activation and mitochondrial apoptosis*. Science, 2003. **301**(5632): p. 513-7.
317. Gai, Z., et al., *Trps1 functions downstream of Bmp7 in kidney development*. J Am Soc Nephrol, 2009. **20**(11): p. 2403-11.
318. Masuda, H., et al., *Expression of bone morphogenetic protein-7 (BMP-7) in human prostate*. Prostate, 2004. **59**(1): p. 101-6.
319. Prudova, A., et al., *Multiplex N-terminome analysis of MMP-2 and MMP-9 substrate degradomes by iTRAQ-TAILS quantitative proteomics*. Mol Cell Proteomics. **9**(5): p. 894-911.
320. Ross, P.L., et al., *Multiplexed protein quantitation in Saccharomyces cerevisiae using amine-reactive isobaric tagging reagents*. Mol Cell Proteomics, 2004. **3**(12): p. 1154-69.
321. Kleifeld, O., et al., *Isotopic labeling of terminal amines in complex samples identifies protein N-termini and protease cleavage products*. Nat Biotechnol. **28**(3): p. 281-8.
322. Bruand, J., et al., *Automated Querying and Identification of Novel Peptides using MALDI Mass Spectrometric Imaging*. Mol Cell Proteomics, 2010. **In Press**.
323. Day, R., M. Fugère, and W. Neugebauer, *MULTI-LEU PEPTIDES AND ANALOGUES THEREOF AS SELECTIVE PACE4 INHIBITORS AND EFFECTIVE ANTIPROLIFERATIVE AGENTS*. 2010.
324. Stoeckli, M., P. Chaurand, and R.M. Caprioli, *Applications of MALDI MS Imaging to Biological Samples*, in ASMS. 1999.



**Annexe I** : MALDI imaging mass spectrometry: state of the art technology in clinical proteomics. *Mol Cell Proteomics*. 2009 Sep;8(9):2023-33.

**Annexe II** : MITICS (MALDI Imaging Team Imaging Computing System): a new open source mass spectrometry imaging software. *J Proteomics*. 2008 Aug 21;71(3):332-45.

**Annexe III** : MALDI imaging of formalin-fixed paraffin-embedded tissues: application to model animals of Parkinson disease for biomarker hunting. *J Proteome Res*. 2008 Mar;7(3):969-78.

**Annexe IV** : Contribution of MALDI tissue profiling and MALDI-Imaging for the study of saponins in Cuvierian tubules of *Holothuria forskali*. *PLOS ONE*. In Press

**Annexe I : MALDI imaging mass spectrometry: state of the art technology in clinical proteomics.**

# MALDI Imaging Mass Spectrometry

STATE OF THE ART TECHNOLOGY IN CLINICAL PROTEOMICS\*

Julien Franck‡§, Karim Arafah‡§, Mohamed Elayed‡, David Bonnel‡, Daniele Vergara‡, Amélie Jacquet‡, Denis Vinatier¶, Maxence Wisztorski‡, Robert Day||, Isabelle Fournier‡\*\*, and Michel Salzet‡ ††

A decade after its inception, MALDI imaging mass spectrometry has become a unique technique in the proteomics arsenal for biomarker hunting in a variety of diseases. At this stage of development, it is important to ask whether we can consider this technique to be sufficiently developed for routine use in a clinical setting or an indispensable technology used in translational research. In this report, we consider the contributions of MALDI imaging mass spectrometry and profiling technologies to clinical studies. In addition, we outline new directions that are required to align these technologies with the objectives of clinical proteomics, including: 1) diagnosis based on profile signatures that complement histopathology, 2) early detection of disease, 3) selection of therapeutic combinations based on the individual patient's entire disease-specific protein network, 4) real time assessment of therapeutic efficacy and toxicity, 5) rational redirection of therapy based on changes in the diseased protein network that are associated with drug resistance, and 6) combinatorial therapy in which the signaling pathway itself is viewed as the target rather than any single "node" in the pathway. *Molecular & Cellular Proteomics* 8:2023–2033, 2009.

MS has become a versatile tool that we are familiar with in large part due to important electronic and informatics advancements. The ability to obtain the molecular weight is one of the first steps in the identification of a molecule. With the addition of primary structural information mass spectrometry has become a useful technique to identify molecules within complex mixtures.

Biological specimens, such as tissues, urine, or plasma, are complex and highly heterogeneous, which makes them inherently difficult to analyze. Further research and developments are necessary to achieve reliable biological models for understanding and studying pathologies. Therefore, it is of primary

importance to identify the constituents of these systems and subsequently understand how they function within the framework of the tissue. With regard to clinical proteomics, there is the added dimension of disease, and therefore, the main goal is to characterize the cellular circuitry with a focus on the impact of the disease and/or therapy on these cellular networks.

Mass spectrometry has become a centerpiece technology predominantly in the field of proteomics. Nonetheless a more comprehensive understanding of the constituents of biological systems will be aided by determining the constituent distribution. This anatomical dimension has been added through mass spectrometry imaging (MSI)<sup>1</sup> especially using MALDI-MSI.

MALDI is an ion source that is well compatible with the introduction of raw materials and surfaces. Shortly after its introduction, MALDI was used for direct tissue profiling. The first applications were neurobiological studies on dissected organs from the mollusk *Lymnaea stagnalis* (1–8), crustaceans (9), and other mollusks (10, 11). More recently, MALDI was used to generate profiles from tissue sections and ion images using a scanning method to analyze the surface (12) (Fig. 1). This led to the first MALDI MS tissue section imaging micrographs in 1997 (13–15). These studies were followed by 10 years of intense efforts to improve the sensitivity, reproducibility, data processing, tissue preservation, and preparation treatments to fully characterize the proteome leading to a clear improvement of molecular images (16–39) (Fig. 2).

These developments led to clinical studies using MALDI-MSI technology. Clinical proteomics has many objectives including 1) diagnosis based on signatures as a complement to histopathology, 2) early disease detection, 3) individualized selection of therapeutic combinations that best target the patient's entire disease-specific protein network, 4) real time assessment of therapeutic efficacy and toxicity, 5) rational redirection of therapy based on changes in the diseased protein network that are associated with drug resistance, and 6) combinatorial therapy in which the signaling pathway itself is viewed as the target rather than any single "node" in the pathway.

From the ‡MALDI Imaging Team, Laboratoire de Neuroimmunologie des Annélides, IFR 147, CNR-FRE 2933, University of Lille1, 59655 Villeneuve d'Ascq, France, ¶Service de gynécologie, Hôpital Jeanne de Flandre, 59037 Lille, France, and ||Département de Pharmacologie, Faculté de médecine, Université de Sherbrooke, Sherbrooke, Québec J1H 5N4, Canada

Received, December 8, 2008, and in revised form, May 5, 2009

Published, MCP Papers in Press, May 18, 2009, DOI 10.1074/mcp.R800016-MCP200

<sup>1</sup> The abbreviations used are: MSI, mass spectrometry imaging; PCA, principal component analysis; XRT, x-ray therapy; EGFR, epidermal growth factor receptor; MRI, magnetic resonance imaging.

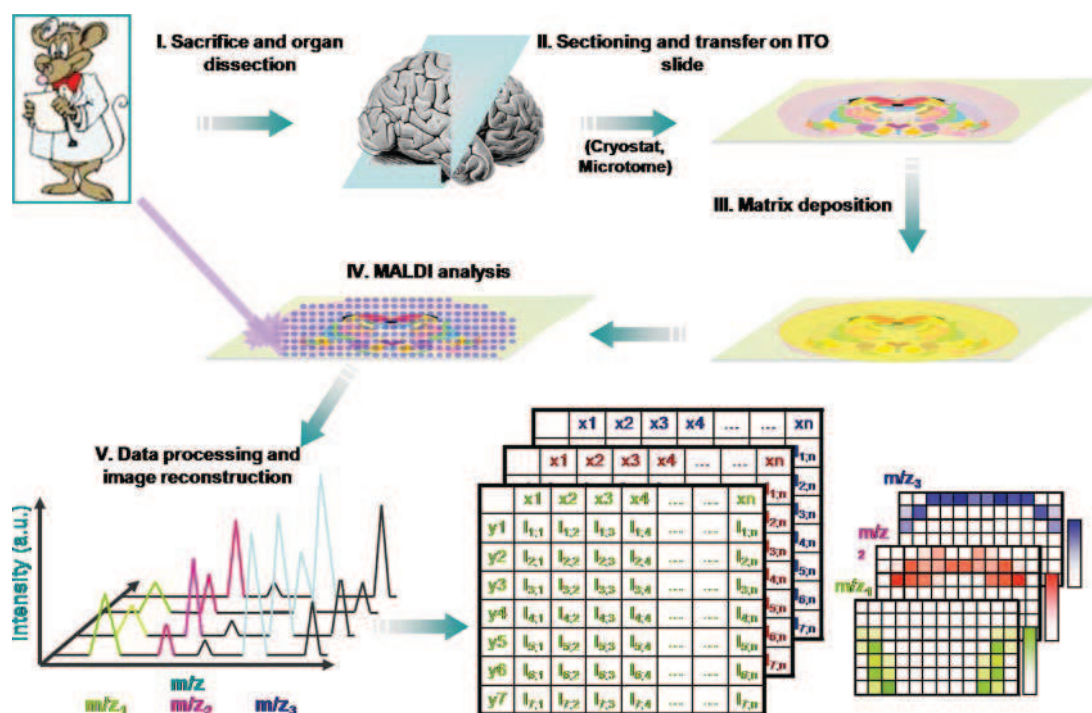


FIG. 1. Schematic representation of the MALDI-MSI work flow. After tissue sectioning and transfer onto a conductive and transparent sample plate, the MALDI matrix is deposited, and data are acquired by recording mass spectra according to a raster of points covering the surface to be analyzed. Mass spectra recorded with their coordinates on the tissue are processed, and molecular images of the localization of molecules can be reconstructed. *a.u.*, arbitrary units; *ITO*, indium tin oxide.

Based on these key objectives, can we consider MALDI-MSI a mature technology for use in clinical studies? What is the potential impact of this technology in anatomy/pathology and disease? By reviewing each objective, do we have sufficient evidence that MALDI-MSI satisfies the criteria imposed by clinical proteomics? We will now specifically address each of these key points.

#### DIAGNOSIS BASED ON SIGNATURES AS A COMPLEMENT TO HISTOPATHOLOGY

In some cases, diagnosis or tissue classification cannot be easily achieved through standard histological staining. Further refinements based on molecular signatures and statistical data, which are currently missing, are crucial for improved diagnostics. The development of rapid and reliable screening of human tissues for diagnostics (e.g. biopsies or smears) has been improved with modern proteomics. By using MALDI-MSI, a molecular diagnosis could be done on tissue directly in the environment of the tumors. MALDI-MSI could help to detect the tumor boundary or infiltration of adjacent normal tissue that presents a normal histology. It could also help to detect the early stage of pathology that presents no histological modifications and to prevent tumor recurrence at the site of surgical resection. One of the major advances of MSI is the correlation of the MALDI images with histological information. MALDI-MSI software (for a review, see Ref. 40) superimposes the MALDI images over a macroscopic or microscopic optical

image of the sample taken before MALDI measurement. Although the primary macroscopic optical image is sufficient to recognize the outline of the tissue and define the measurement area, it is not usually possible to observe histological features in the image (in contrast to microscopic images). For a histological interpretation, it is necessary to use stained tissue sections. Two approaches have been used to correlate histology with MALDI-MSI results: performing MALDI-MSI and histological staining on consecutive sections (41, 42) or staining the sample after MALDI measurement (43). The latter technique has been successfully used by pathologists (Fig. 3) (44), which suggests that combining MALDI-MSI and classic histological staining provides pathologists with more information to make better diagnoses. The next step is not only to perform a diagnosis based on  $m/z$  signatures but also on molecular data generated from identification of specific biomarkers that have been characterized as pathological signatures.

However, another challenge for pathologists is tissue classification, which is required to catalogue tumors or benign tissues. The major technological improvement that MALDI-MSI provides is the direct identification of novel markers within an *in situ* context from fixed sections/biopsy embedded in paraffin (e.g. archived material) (42). Several studies on cancer and neurodegenerative diseases have demonstrated that MALDI-MSI is a key technology for identifying biomark-

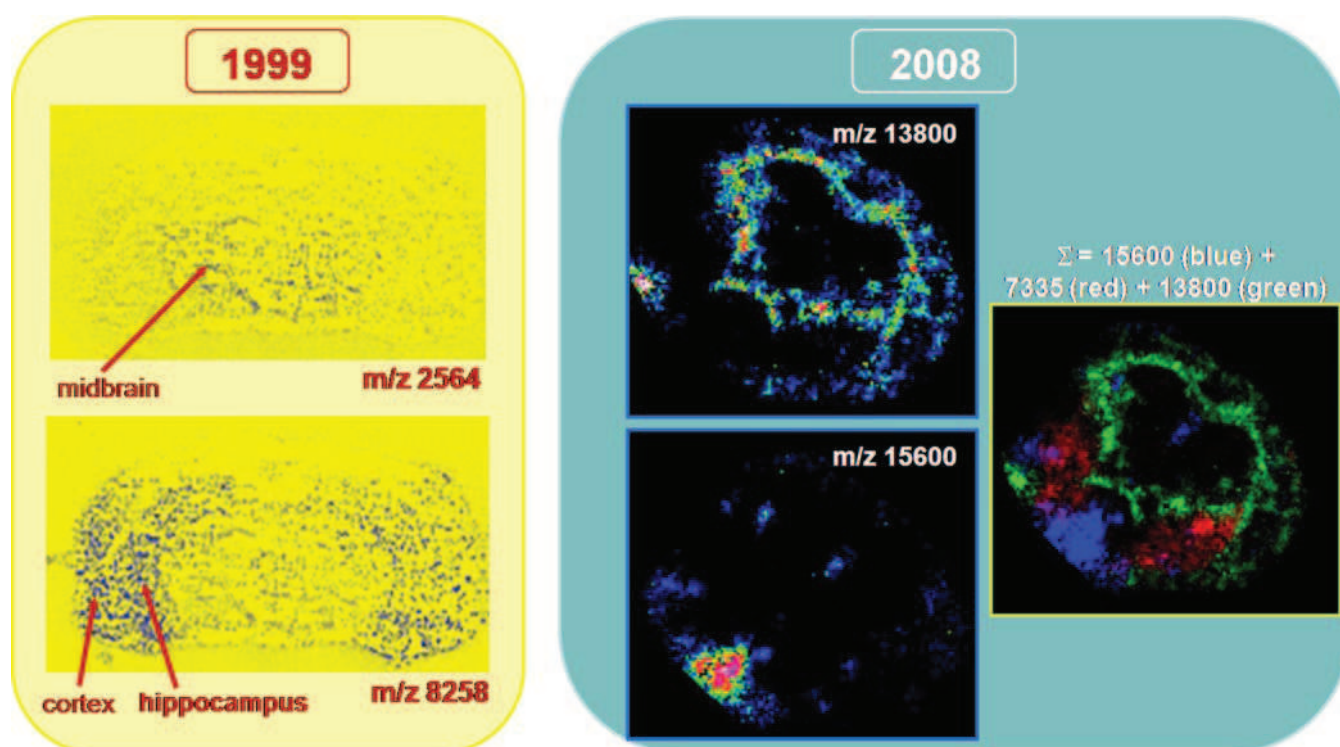


FIG. 2. Ten years' evolution from one of the first MALDI images presented in 1999 at the 47th ASMS Conference on Mass Spectrometry and Allied Topics (*left*) (reprinted with permission of Caprioli and co-workers (84)) and molecular images obtained by our group for mouse stem cells injected in brain tissue sections (*right*) (M. Wisztorski, C. Meriaux, M. Salzert, and I. Fournier, unpublished results).

ers, assessing their localization, and cross-validation (29, 45–51). The use of archived, formalin-fixed, paraffin-embedded tissue from hospital pathology departments represents a “gold mine” of existing data (42, 52–54). The application of MALDI MS imaging to archived materials could lead to the creation of an international disease marker database that would facilitate the development of early diagnostics for various pathologies as well as for follow-up examination of disease progression.

Therefore, the addition of statistical analysis will be very important for the comparison of the different tissue components (e.g. tumor *versus* benign or healthy). Each tissue type depends upon the nature of its composition of cells. Thus, biocomputational methods are absolutely necessary to identify individualized molecular patterns to aid in diagnosis and prognosis.

The advantage of MALDI-MSI is the ability to obtain a large collection of mass spectra spread out over a tissue section while retaining the absolute spatial location of these measurements for subsequent analysis and imaging. One of the statistical techniques to reduce the complexity of the information in multidimensional data sets in MALDI-MSI is principal component analysis (PCA) (55). PCA is a multivariate preanalysis tool that allows for the correlation and identification of the major spatial and mass-related trends in the data that guide further downstream analysis (56). PCA reduces the

dimensionality of the data set but does not classify the spectra. This is a transformation of the original coordinate system defined by peak intensities to a coordinate system that better explains the variance within the data set. This has been recently used in a prostate cancer study (43).

The next required step is the hierarchical clustering of the tissue based on PCA statistical analyses that reflect the most important variance of ions within the tissue (57). Dendrograms can be constructed, and each branch represents ions present in the same group of cells (e.g. epithelial cancer cells *versus* benign cells). Thus, this representation provides access to huge numbers of individual spectra and reduces the complexity of the data set. It can also be correlated with histology as previously used for mouse kidney (Fig. 4) (44), gastric (58), and ovarian cancer (Fig. 5).

#### EARLY DISEASE DETECTION

Based on MALDI MS profiling (Fig. 6a) and imaging strategies (Fig. 6, b and c), several biomarkers have been identified in various cancer studies. In stage III and IV ovarian cancer, a highly prevalent (80%) biomarker has been identified using MALDI MS and nano-LC-nano-ESI MS using MS and MS/MS after separation by reverse phase HPLC and trypsin enzymatic digestion. This marker with an  $m/z$  of 9744 corresponds to an 84-amino acid fragment from the 11 S proteasome activator complex (PA28  $\alpha$  or REG- $\alpha$ ) (33). This biomarker was

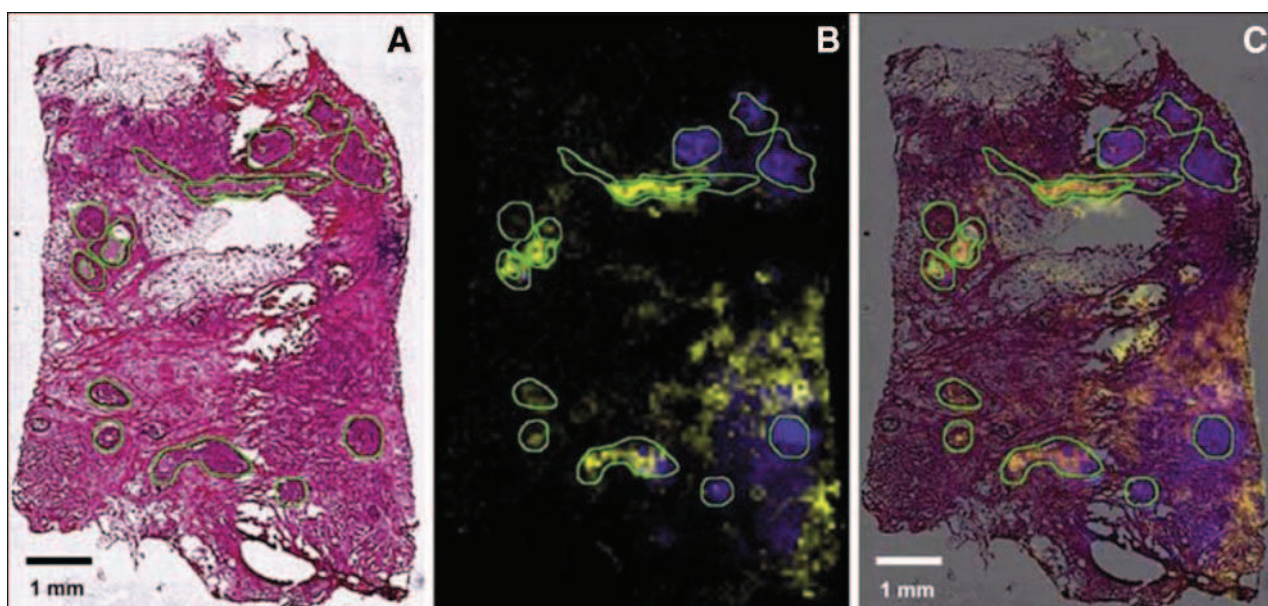
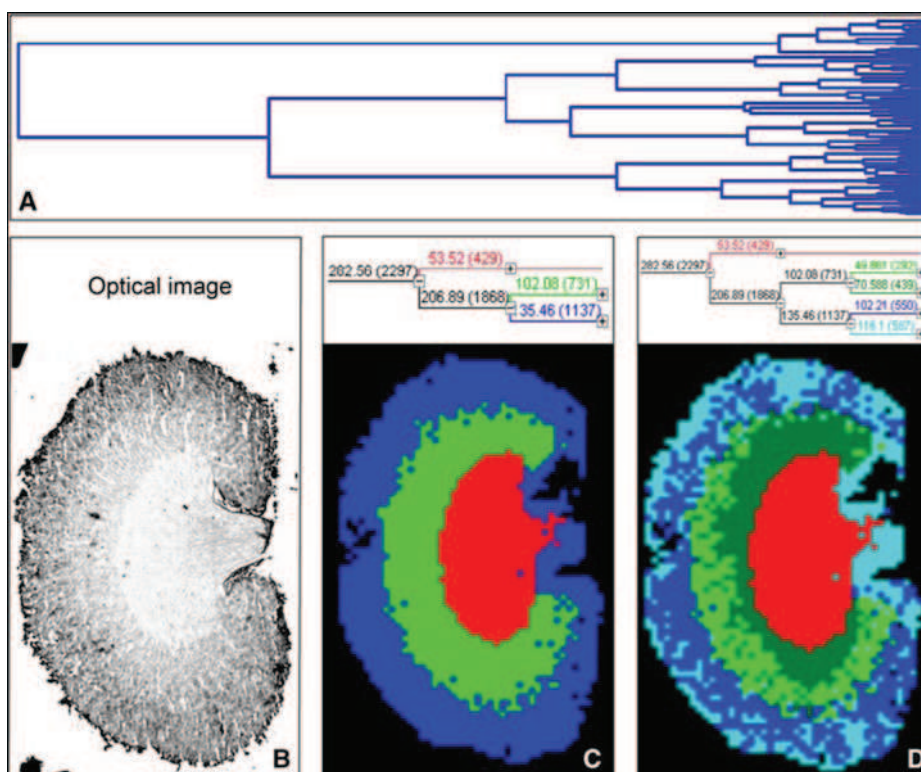


FIG. 3. Mass species representing molecular features of preinvasive and invasive lesions of the breast. A, optical microscopic image of hematoxylin- and eosin-stained tissue section showing several carcinomatous *in situ* regions (outlined in green). The staining was done after MALDI measurement of the tissue section. This allows an unambiguous correlation with MALDI imaging results. B, visualization of ion density images of two selected masses ( $m/z$  9750 shown in yellow and  $m/z$  4519 shown in blue). C, overlay of hematoxylin and eosin staining and MALDI molecular image. The distribution of these two masses suggests a divergent clonal evolution of the preinvasive lesions. These two masses are also present in the invasive cancer cells surrounding some carcinomas *in situ* (right site). Scanning resolution, 80  $\mu\text{m}$ . Scale bar, 1 mm. (Reprinted with permission of Walch *et al.* (44).)

FIG. 4. Hierarchical clustering of a mouse kidney data set achieved by MALDI-MSI. A, full dendrogram of all spectra in a mouse kidney data set. B, optical image of the mouse kidney analyzed by MALDI-MSI. C and D, reconstruction of selected dendrogram branches and corresponding images.

The three main branches reflect the renal cortex (blue), medulla (green), and pelvis (red). C, the medulla branch separates into two distinct areas, whereas the cortex branch further differentiates into fat and connective tissue of the renal capsule and hilus and the actual cortex (D). (Reprinted with permission of Walch *et al.* (44).)



validated using MALDI MSI (Fig. 6, b and c), classic immunohistochemistry with an antibody raised against the C-terminal part of the protein containing the fragment of interest (Fig. 6,

d, e, and f), specific MALDI-MSI using the Tag-mass concept (Fig. 6g) (59), quantitative PCR, and Western blot (Fig. 6h). Recently, we confirmed the REG- $\alpha$  expression in an ovarian

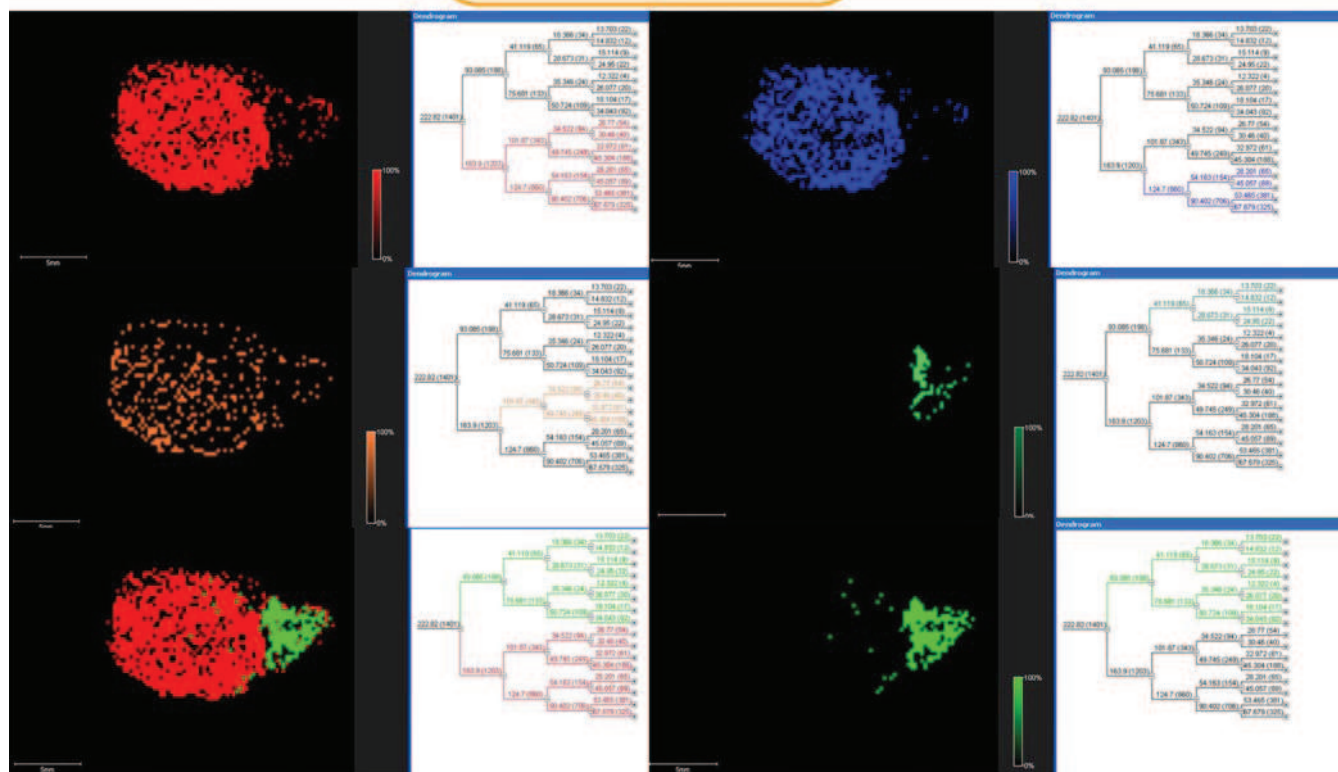
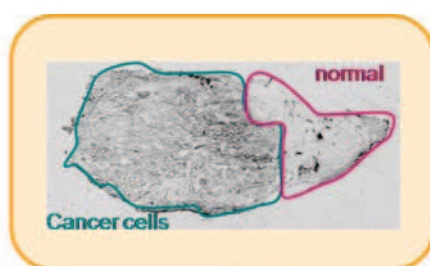


FIG. 5. Hierarchical clustering using the ClinProt tool (Bruker Daltonics) after PCA of a stage 4 mucous ovarian carcinoma section covered with ionic matrix using the Shimadzu CHIP 1000 microspotter. *a*, optical image of the ovarian carcinoma section. *b–f*, reconstructed selected dendrograms and corresponding images. The two main branches reflect the carcinoma (*red*; *a*) and the healthy (*green*; *f*) parts in the section. *b* and *c* are two carcinoma subclasses, and *d* is a subclass of the healthy part. *e* represents a merge of the two branches.

epithelial cell line by quantitative PCR.<sup>2</sup> Immunohistochemistry confirmed the epithelial expression of this fragment with a nuclear localization in benign epithelial cells and a cytoplasmic localization in carcinoma cells (Fig. 6, *d* and *e*). This localization pattern indicates that this antibody can be used to discriminate borderline tumor cases, which are the most difficult to diagnose. Thus, a specific antibody that discriminates between cells transitioning from benign to malignant will be an asset for early diagnosis. Taken together, these studies indicate that direct tissue analysis and specific MALDI-MSI strategies facilitate biomarker identification and validation.

In addition, data can be obtained from fundamental studies by analyzing the ontogeny of protein expression during mor-

phogenesis and tumorigenesis, and proteins that could potentially serve as biomarkers for diagnosing diseases can be identified as demonstrated by an MSI study on murine prostate cancer development (60). Murine prostate during development (1–5 weeks of age), at sexual maturation (6 weeks of age), and in adulthood (at 10, 15, or 40 weeks of age) was compared with prostate tumors from 15-week-old mice genetically engineered to express the large T antigen gene under the control of the prostate-specific probasin promoter (LPB-Tag mice). This approach identified proteins that were differentially expressed at specific time points during prostate development. The expression of probasin and spermine-binding protein, which are associated with prostate maturation, decreased during prostate tumor formation (60). This study was the first use of MALDI-MSI to follow ontogeny to tumorigenesis (60).

<sup>2</sup> M. El Ayel, D. Bonnel, I. Fournier, and M. Salzet, unpublished data.

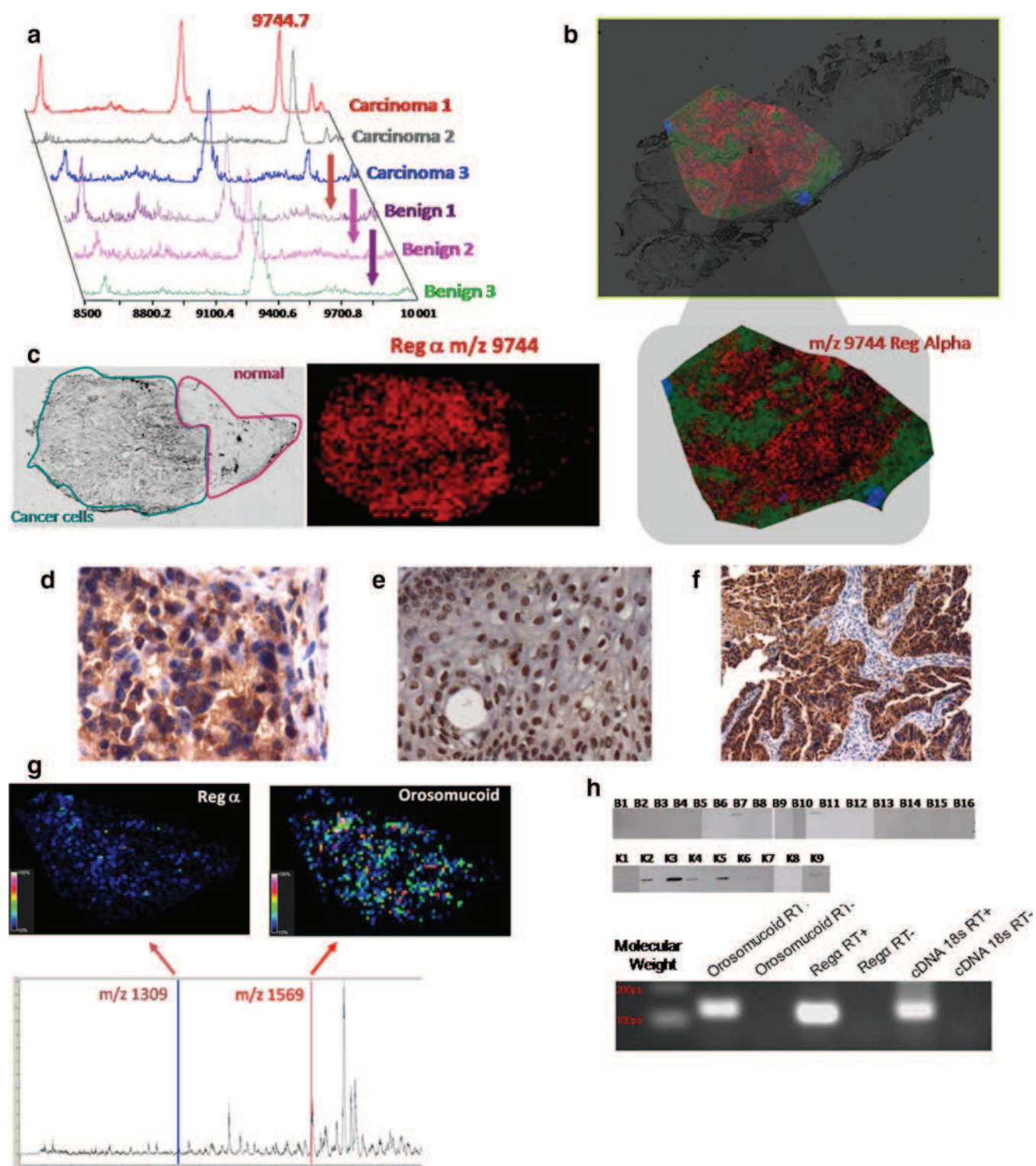


FIG. 6. Validation of C-terminal fragment of the immunoproteasome REG- $\alpha$  and orosomuroid as ovarian biomarkers. *a*, MALDI-MS profiles of three ovarian carcinomas versus benign tumor samples. *b*, MALDI-MSI molecular image of REG- $\alpha$  fragment ( $m/z$  9744) at a resolution of 50  $\mu\text{m}$  from an ovarian carcinoma tissue section. *c*, optical image of the tissue section with the region of interest defined (cancerous versus healthy part) and MALDI-MSI of the REG- $\alpha$  fragment showing its presence exclusively in the cancer part. *d-f*, immunocytochemical data obtained after the antigen retrieval technique and H&E coloration with the anti-C-terminal REG- $\alpha$  antibody: *d*, cytoplasm localization of the anti-C-terminal REG- $\alpha$  labeling in ovary carcinoma; *e*, nucleus localization of the anti-C-terminal REG- $\alpha$  labeling in ovary benign tumor; *f*, epithelial cells labeled with the anti-C-terminal REG- $\alpha$  in ovarian carcinoma. *g*, specific MALDI imaging analysis using the



There is no doubt of the usefulness of MALDI-MSI in biomarker development for early diagnosis. However, MALDI-MSI is still not being routinely used in a clinical setting and has not yet been adjusted to conform to clinical proteomics procedures. Only a limited number of international groups have used this technology effectively in clinical settings; however, the number of clinical studies applying MSI has dramatically increased in the past 2 years (43, 44, 54, 61–66).

INDIVIDUALIZED SELECTION OF THERAPEUTIC COMBINATIONS THAT BEST TARGET THE PATIENT'S ENTIRE DISEASE-SPECIFIC PROTEIN NETWORK

MALDI-MSI is highly advantageous for *in situ* drug tracking. In fact, it enables the detection of both endogenous and exogenous compounds present in tissues with molecular specificity and preserves their spatial orientation. This unique combination coupled with excellent sensitivity and rapid analysis presents potential advantages for a wide range of applications in diverse biological fields. As described previously, recent advances have demonstrated that the technique can be applied to cancer research, neuroscience, and pharmaceutical development (67). MALDI-MSI can be used in clinical studies to provide a molecular *ex vivo* view of resected organs. This allows for the label-free tracking of both endogenous and exogenous compounds with spatial resolution and molecular specificity (67–77). Several examples support the idea that MALDI-MSI technology will become a key tool in drug development (67–73), including novel drug design through the ability to analyze metabolic pathways directly in tissues (e.g. through *in situ* multiplex metabolite analysis), as well as in the elucidation of secondary effects and unexpected feedback loops (78). Currently MSI of biomolecules and chemical compounds in cell-based assays and highly complex tissue sections is used in parallel with classic mass spectrometry ionization techniques to identify chemical compounds interfering with enzymatic function, receptor-ligand binding, or molecules modulating a protein-protein interaction.

There is evidence supporting MSI as a key technique that can be used in combination with other therapeutic technologies. Recently the efficacy of combining radiation (XRT) with a dual epidermal growth factor receptor (EGFR)/vascular endothelial growth factor receptor inhibitor, AEE788, in prostate cancer models with different levels of EGFR expression was analyzed using Doppler sonography, tumor blood vessel destruction (visualized by immunohistochemistry), and MSI (76). Tumor xenografts established from DU145 or PC-3 prostate cancer cell lines inoculated into the hind limbs of athymic nude mice were assigned to four treatment groups: 1) control,

2) AEE788, 3) XRT, and 4) AEE788 and XRT. AEE788 had a radiosensitization effect in human umbilical vein endothelial cells and increased their susceptibility to apoptosis. Therefore, concurrent AEE788/XRT treatment compared with either treatment alone led to a significant delay in tumor growth in animals bearing DU145 tumors. Conversely there was no effect on the growth of PC-3 tumors with combination therapy. In DU145 tumors, there was a significant decrease in tumor blood flow with combination therapy as assessed by using Doppler sonography and tumor blood vessel destruction. MSI demonstrated that AEE788 is bioavailable and heterogeneously distributed in DU145 tumors receiving therapy, supporting the efficacy of the combination of AEE788 and XRT *in vitro* and *in vivo* in DU145-based models. In contrast, in PC-3-based models, the tumors were adequately treated with XRT alone without any added benefit from combination therapy. These findings correlated with differences in EGFR expression. Overall this study demonstrated the effects of therapeutics on both tumor cell proliferation and vascular destruction using complementary technologies, including MALDI-MSI in a clinical proteomics protocol.

REAL TIME ASSESSMENT OF THERAPEUTIC EFFICACY AND TOXICITY

MSI technology will also help significantly advance the analysis of novel therapeutics and may provide deeper insight into therapeutic and toxicological processes, revealing the mechanism of efficacy or side effects at the molecular level (79). A study by Atkinson *et al.* (80) using AQ4N (baxoxatrone) (1,4-bis-5,8-dihydroxyanthracene-9,10-dione) as a prodrug demonstrated that MSI can be used for both drug and clinical development. In hypoxic cells, AQ4N is reduced to AQ4 (cytotoxic form), which is a topoisomerase II inhibitor. By inhibiting topoisomerase II within these hypoxic areas, AQ4N sensitizes tumors to existing chemo- and radiotherapy treatments. The distribution of AQ4N and AQ4 in treated H460 human tumor xenografts has been examined by MALDI-MSI, and images of the distribution of AQ4N and AQ4 show little overlap (80). The distribution of ATP in the tumor xenografts was studied as an endogenous marker of hypoxia because concentrations of ATP are known to decrease with hypoxia. The ATP distribution was similar to that of AQ4N, suggesting that in regions with abundant ATP expression (*i.e.* normoxic tissue) there was no evidence of conversion of AQ4N to AQ4. This indicates that the cytotoxic metabolite AQ4 is confined to hypoxic regions of the tumor (80).

Tag-mass concept with anti-C-terminal REG- $\alpha$  and an anti-human IgG tag (reporter *m/z* 1309) and anti-orosomuroid and anti-human monoclonal antibody (reporter *m/z* 1569). *h, top*, Western blot analyses with the anti-C-terminal REG- $\alpha$  (immunoproteasome 11 S) of the 16 benign tumors and nine carcinomas (33); *bottom*, quantitative PCR validation of REG- $\alpha$  and orosomuroid from the SKVO3 ovarian cancer epithelial cell line.

### RATIONAL REDIRECTION OF THERAPY BASED ON CHANGES IN THE DISEASED PROTEIN NETWORK THAT ARE ASSOCIATED WITH DRUG RESISTANCE

There are no studies that have used MALDI-MSI to redirect therapy. However, this objective will be the next challenge for this research field.

### COMBINATORIAL THERAPY IN WHICH THE SIGNALING PATHWAY ITSELF IS VIEWED AS THE TARGET RATHER THAN ANY SINGLE NODE IN THE PATHWAY

Similar to the fifth objective, there are no studies on combinatorial therapy that focus on a signaling pathway as a whole because of the infancy of this technology.

### CONCLUSION

MALDI MSI emerged only 10 years ago, so it is still a young technology that is continuously evolving. Further developments are still needed to establish this technology in a clinical setting. For example, standardization of protocols must be undertaken between tissue collection, storage, preparation, and data acquisition. Moreover improvements in resolution will be sought because the minimum area that can be examined is currently in the range of a few cells. The actual resolution that can be routinely achieved while keeping good sensitivity (*i.e.* sufficient ion yields) is  $50 \times 50 \mu\text{m}^2$ . To achieve increased resolution it is necessary to reduce the surface area irradiated by the laser beam, which most obviously can be achieved by decreasing the laser beam diameter. However, under 30–40- $\mu\text{m}$  laser beam diameter a decrease in ion production resulting in decreased sensitivity is observed. Thus, active research to decrease the size of the irradiated area without loss of sensitivity or attempts to better understand and optimize the ionization processes involved in MALDI are underway. From our own work, we have developed masks that achieve a resolution of  $30 \times 30 \mu\text{m}$  without reducing the production of ions and result in an increased sensitivity of 2–3-fold (81). A new generation of masks at  $10 \times 10 \mu\text{m}$  is now being developed. Nonetheless the present levels of sensitivity allow the detection of a small group of cells but are not sufficient to detect discrete modification at a single cell level. However, detection limits are the same as those obtained for the analysis of a complex mixture using classical MALDI-TOF procedures. Contrary to the predictions that the most abundant proteins are the only ones to be detected with MALDI, it has been observed that ionization efficiency is an important parameter and that a low abundance molecule that is well ionized can also be detected. The detection of low abundance proteins can also be improved with the development of statistical software allowing the treatment of large cohorts of patient with large data sets for biomarkers, tissue classification, and stage of disease development. Other limitations include the detectable mass range, which is typically between  $m/z$  400 and 30,000. The lower limit is due to the use of matrix that masks the analysis below this  $m/z$ . The upper limitation is not understood, and various protocols are being tested to

overcome this limitation, such as new tissue treatments, new matrices, or development of new ion sources to generate multicharged ions. One difficulty that has been overcome is the direct identification of biomarkers on tissues. Bottom-up strategies using on-tissue trypsin digestion have been developed for frozen (31) and formalin-fixed, paraffin-embedded (42, 54) tissues. Ideally the ability to perform on-tissue top-down protein characterization is one of our future objectives. MALDI imaging will ultimately provide high resolution molecular imaging but will also result in direct biomarker identification with statistical validation such that it will become an essential proteomics tool in clinical histopathology.

Additional needed developments will be three-dimensional reconstruction to obtain tumor maps (82). MALDI-MSI will improve tissue classification necessary to perform retrospective studies, will assist clinical studies from the bench to bedside, and will provide a remarkable follow-up procedure. Improved tissue classification using MALDI-MSI on the same tissues used by pathologists for diagnosis will speed up the process of molecular diagnosis. Molecular tissue classification after MALDI-MSI based on known biomarkers or using unsupervised multivariable analyses can positively affect patient treatment. For example, borderline ovarian cancers are difficult to detect clinically until they are advanced in size or stage. The most common presenting symptoms are abdominal pain, increasing girth or abdominal distension, and abdominal mass. Approximately 23% of patients are asymptomatic. With such tumors, correct diagnosis is difficult to reach, and the molecular profiles provided by MALDI-MSI may facilitate classification and aid the development of a treatment strategy. Moreover depending on the nature of the malignancy (*e.g.* serous or mucous with or without cell infiltration), the therapeutic strategy is different. Should MALDI-MSI tumor classification libraries be created, these could permit clinicians to individually tailor patient treatments in a practical manner. Based on the tissue biomarkers identified by MSI, it could be important to follow up the evolution of the malignancy during treatment, for example after cisplatin treatment to define whether resistance to the treatment may appear. Furthermore in cases where traditional biomarkers cannot be clearly detected in biopsies, MALDI-MSI could become critical to the outcome. This points to the importance of establishing tumor MSI libraries to facilitate multicenter studies and the creation of MSI classification maps.

At this point, it is obvious that further clinical studies using MALDI-MSI technology are required. Nonetheless MALDI-MSI has opened the door to molecular tissue classification, which could be of great use to pathologists with regard to diagnosis but also in drug development and diagnosis coupled with magnetic resonance imaging (MRI) technology. A major advance for MALDI-MSI will be its coupling with positron emission tomography, x-ray, computed tomography instrumentation, and MRI for both preclinical and clinical research. The complementarities between non-invasive tech-

niques and molecular data obtained from MALDI MS imaging will result in a more precise diagnosis. In clinical studies, the need for information on the spatial localization of pathologically gene-encoded products has become more pressing. The three-dimensional volume reconstructions generated by MALDI-MSI data (83) now offer the possibility to compare the molecular data with data obtained using positron emission tomography or MRI. These associations will enhance the use of MALDI-MSI. Ultimately comparing the MRI image of a tumor and the image generated by MALDI-MSI at a molecular level will provide a comprehensive data set for diagnosis and treatment selection.

\* This work was supported by grants from the CNRS Département de la politique industrielle (to M. S. and I. F.), Ministère de L'Education Nationale, de L'Enseignement Supérieur et de la Recherche, Institut National du Cancer (to I. F.), Agence National de la recherche (to I. F.), the Institut National du Cancer (to I. F.), and the Canadian Institutes of Health Research (to R. D. and M. S.).

§ Co-authors.

\*\* To whom correspondence may be addressed. E-mail: isabelle.fournier@univ-lille1.fr.

‡‡ To whom correspondence may be addressed. E-mail: michel.salzet@univ-lille1.fr.

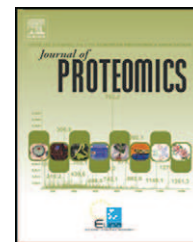
#### REFERENCES

- van Veelen, P. A., Jiménez, C. R., Li, K. W., Wildering, W. C., Geraerts, W. P., Tjaden, U. R., and van der Greef, J. (1993) Direct peptide profiling of single neurons by matrix-assisted laser desorption/ionization mass spectrometry. *Org. Mass Spectrom.* **28**, 1542–1546
- Jiménez, C. R., van Veelen, P. A., Li, K. W., Wildering, W. C., Geraerts, W. P., Tjaden, U. R., and van der Greef, J. (1994) Neuropeptide expression and processing as revealed by direct matrix-assisted laser desorption/ionization mass spectrometry of single neurons. *J. Neurochem.* **62**, 404–407
- Li, K. W., Hoek, R. M., Smith, F., Jiménez, C. R., van der Schors, R. C., van Veelen, P. A., Chen, S., van der Greef, J., Parish, D. C., and Benjamin, P. R. (1994) Direct peptide profiling by mass spectrometry of single identified neurons reveals complex neuropeptide-processing pattern. *J. Biol. Chem.* **269**, 30288–30292
- Li, K. W., Jiménez, C. R., Van Veelen, P. A., and Geraerts, W. P. (1994) Processing and targeting of a molluscan egg-laying peptide prohormone as revealed by mass spectrometric peptide fingerprinting and peptide sequencing. *Endocrinology* **134**, 1812–1819
- Li, K. W., van Golen, F. A., van Minnen, J., van Veelen, P. A., van der Greef, J., and Geraerts, W. P. (1994) Structural identification, neuronal synthesis, and role in male copulation of myomodulin-A of Lymnaea: a study involving direct peptide profiling of nervous tissue by mass spectrometry. *Brain Res. Mol. Brain Res.* **25**, 355–358
- Dreisewerd, K., Kingston, R., Geraerts, W. P., and Li, K. W. (1997) Direct mass spectrometric peptide profiling and sequencing of nervous tissues to identify peptides involved in male copulatory behavior in Lymnaea stagnalis. *Int. J. Mass Spectrom.* **169**, 291–299
- Jiménez, C. R., and Burlingame, A. L. (1998) Ultramicroanalysis of peptide profiles in biological samples using MALDI mass spectrometry. *Exp. Nephrol.* **6**, 421–428
- Jiménez, C. R., Li, K. W., Dreisewerd, K., Spijker, S., Kingston, R., Bateman, R. H., Burlingame, A. L., Smit, A. B., van Minnen, J., and Geraerts, W. P. (1998) Direct mass spectrometric peptide profiling and sequencing of single neurons reveals differential peptide patterns in a small neuronal network. *Biochemistry* **37**, 2070–2076
- Redeker, V., Toullec, J. Y., Vinh, J., Rossier, J., and Soyey, D. (1998) Combination of peptide profiling by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry and immunodetection on single glands or cells. *Anal. Chem.* **70**, 1805–1811
- Rubakhin, S. S., Li, L., Moroz, T. P., and Sweedler, J. V. (1999) Characterization of the Aplysia californica cerebral ganglion F cluster. *J. Neurophysiol.* **81**, 1251–1260
- Sweedler, J. V., Li, L., Floyd, P., and Gilly, W. (2000) Mass spectrometric survey of peptides in cephalopods with an emphasis on the FMRFamide-related peptides. *J. Exp. Biol.* **203**, 3565–3573
- Caprioli, R. M., Farmer, T. B., and Gile, J. (1997) Molecular imaging of biological samples: localization of peptides and proteins using MALDI-TOF MS. *Anal. Chem.* **69**, 4751–4760
- Stoeckli, M., Chaurand, P., Hallahan, D. E., and Caprioli, R. M. (2001) Imaging mass spectrometry: a new technology for the analysis of protein expression in mammalian tissues. *Nat. Med.* **7**, 493–496
- Chaurand, P., Schwartz, S. A., and Caprioli, R. M. (2002) Imaging mass spectrometry: a new tool to investigate the spatial organization of peptides and proteins in mammalian tissue sections. *Curr. Opin. Chem. Biol.* **6**, 676–681
- Fournier, I., Day, R., and Salzet, M. (2003) Direct analysis of neuropeptides by in situ MALDI-TOF mass spectrometry in the rat brain. *Neuro Endocrinol. Lett.* **24**, 9–14
- Caldwell, R. L., and Caprioli, R. M. (2005) Tissue profiling by mass spectrometry: a review of methodology and applications. *Mol. Cell. Proteomics* **4**, 394–401
- Jurchen, J. C., Rubakhin, S. S., and Sweedler, J. V. (2005) MALDI-MS imaging of features smaller than the size of the laser beam. *J. Am. Soc. Mass Spectrom.* **16**, 1654–1659
- Maddalo, G., Petrucci, F., Iezzi, M., Pannellini, T., Del Boccio, P., Ciavardelli, D., Biroccio, A., Forli, F., Di Ilio, C., Ballone, E., Urbani, A., and Federici, G. (2005) Analytical assessment of MALDI-TOF imaging mass spectrometry on thin histological samples. An insight in proteome investigation. *Clin. Chim. Acta* **357**, 210–218
- Aerni, H. R., Cornett, D. S., and Caprioli, R. M. (2006) Automated acoustic matrix deposition for MALDI sample preparation. *Anal. Chem.* **78**, 827–834
- Altelaar, A. F., Klinkert, I., Jalink, K., de Lange, R. P., Adan, R. A., Heeren, R. M., and Piersma, S. R. (2006) Gold-enhanced biomolecular surface imaging of cells and tissue by SIMS and MALDI mass spectrometry. *Anal. Chem.* **78**, 734–742
- Crossman, L., McHugh, N. A., Hsieh, Y., Korfmacher, W. A., and Chen, J. (2006) Investigation of the profiling depth in matrix-assisted laser desorption/ionization imaging mass spectrometry. *Rapid Commun. Mass Spectrom.* **20**, 284–290
- Lemaire, R., Tabet, J. C., Ducoroy, P., Hendra, J. B., Salzet, M., and Fournier, I. (2006) Solid ionic matrixes for direct tissue analysis and MALDI imaging. *Anal. Chem.* **78**, 809–819
- Lemaire, R., Wisztorski, M., Desmons, A., Tabet, J. C., Day, R., Salzet, M., and Fournier, I. (2006) MALDI-MS direct tissue analysis of proteins: improving signal sensitivity using organic treatments. *Anal. Chem.* **78**, 7145–7153
- Stauber, J., Lemaire, R., Wisztorski, M., Ait-Menguellet, S., Lucot, J. P., Vinatier, D., Desmons, A., Deschamps, M., Proess, G., Rudolf, I., Salzet, M., and Fournier, I. (2006) New developments in MALDI imaging mass spectrometry for pathological proteomic studies: introduction to a novel concept, the specific MALDI imaging. *Mol. Cell. Proteomics* **5**, S247–249
- Wiseman, J. M., Ifa, D. R., Song, Q., and Cooks, R. G. (2006) Tissue imaging at atmospheric pressure using desorption electrospray ionization (DESI) mass spectrometry. *Angew. Chem. Int. Ed. Engl.* **45**, 7188–7192
- Wisztorski, M., Brunet, L., Dreisewerd, K., Hillenkamp, F., Berkenkamp, S., Salzet, M., and Fournier, I. (2006) Effect of metals coating for UV MALDI-TOF mass spectrometry imaging (MALDI IMS) and direct tissue analysis in UV/IR MALDI-o-TOF mass spectrometry, in *Proceedings of the 54th ASMS Conference on Mass Spectrometry and Allied Topics, Seattle, May 30–June 2, 2006*, Abstract No. ThP 328, American Society for Mass Spectrometry, Santa Fe, NM
- Chaurand, P., Schriver, K. E., and Caprioli, R. M. (2007) Instrument design and characterization for high resolution MALDI-MS imaging of tissue sections. *J. Mass Spectrom.* **42**, 476–489
- Cornett, D. S., Reyzer, M. L., Chaurand, P., and Caprioli, R. M. (2007) MALDI imaging mass spectrometry: molecular snapshots of biochemical systems. *Nat. Methods* **4**, 828–833
- Dreisewerd, K., Lemaire, R., Pohlentz, G., Salzet, M., Wisztorski, M., Berkenkamp, S., and Fournier, I. (2007) Molecular profiling of native and

- matrix-coated tissue slices from rat brain by infrared and ultraviolet laser desorption/ionization orthogonal time-of-flight mass spectrometry. *Anal. Chem.* **79**, 2463–2471
30. Garrett, T. J., Prieto-Conaway, M. C., Kovtoun, V., Bui, H., Izgarian, N., Stafford, G., and Yost, R. A. (2007) Imaging of small molecules in tissue sections with a new intermediate-pressure MALDI linear ion trap mass spectrometer. *Int. J. Mass Spectrom.* **260**, 166–176
  31. Groseclose, M. R., Andersson, M., Hardesty, W. M., and Caprioli, R. M. (2007) Identification of proteins directly from tissue: in situ tryptic digestions coupled with imaging mass spectrometry. *J. Mass Spectrom.* **42**, 254–262
  32. Hankin, J. A., Barkley, R. M., and Murphy, R. C. (2007) Sublimation as a method of matrix application for mass spectrometric imaging. *J. Am. Soc. Mass Spectrom.* **18**, 1646–1652
  33. Lemaire, R., Menguellet, S. A., Stauber, J., Marchaudon, V., Lucot, J. P., Collinet, P., Farine, M. O., Vinatier, D., Day, R., Ducoroy, P., Salzet, M., and Fournier, I. (2007) Specific MALDI imaging and profiling for biomarker hunting and validation: fragment of the 11S proteasome activator complex, Reg alpha fragment, is a new potential ovary cancer biomarker. *J. Proteome Res.* **6**, 4127–4134
  34. McLean, J. A., Ridenour, W. B., and Caprioli, R. M. (2007) Profiling and imaging of tissues by imaging ion mobility-mass spectrometry. *J. Mass Spectrom.* **42**, 1099–1105
  35. Taban, I. M., Altelaar, A. F., van der Burg, Y. E., McDonnell, L. A., Heeren, R. M., Fuchser, J., and Baykut, G. (2007) Imaging of peptides in the rat brain using MALDI-FTICR mass spectrometry. *J. Am. Soc. Mass Spectrom.* **18**, 145–151
  36. Chen, Y., Allegood, J., Liu, Y., Wang, E., Cachón-Gonzalez, B., Cox, T. M., Merrill, A. H., Jr., and Sullards, M. C. (2008) Imaging MALDI mass spectrometry using an oscillating capillary nebulizer matrix coating system and its application to analysis of lipids in brain from a mouse model of Tay-Sachs/Sandhoff disease. *Anal. Chem.* **80**, 2780–2788
  37. Seeley, E. H., Oppenheimer, S. R., Mi, D., Chaurand, P., and Caprioli, R. M. (2008) Enhancement of protein sensitivity for MALDI imaging mass spectrometry after chemical treatment of tissue sections. *J. Am. Soc. Mass Spectrom.* **19**, 1069–1077
  38. Taira, S., Sugiura, Y., Moritake, S., Shimma, S., Ichiyanagi, Y., and Setou, M. (2008) Nanoparticle-assisted laser desorption/ionization based mass imaging with cellular resolution. *Anal. Chem.* **80**, 4761–4766
  39. Trim, P. J., Atkinson, S. J., Princivalle, A. P., Marshall, P. S., West, A., and Clench, M. R. (2008) Matrix-assisted laser desorption/ionisation mass spectrometry imaging of lipids in rat brain tissue with integrated unsupervised and supervised multivariate statistical analysis. *Rapid Commun. Mass Spectrom.* **22**, 1503–1509
  40. Jardin-Mathé, O., Bonnel, D., Franck, J., Wisztorski, M., Macagno, E., Fournier, I., and Salzet, M. (2008) MITICS (MALDI Imaging Team Imaging Computing System): a new open source mass spectrometry imaging software. *J. Proteomics* **71**, 332–345
  41. Chaurand, P., Schwartz, S. A., Billheimer, D., Xu, B. J., Crecelius, A., and Caprioli, R. M. (2004) Integrating histology and imaging mass spectrometry. *Anal. Chem.* **76**, 1145–1155
  42. Lemaire, R., Desmons, A., Tabet, J. C., Day, R., Salzet, M., and Fournier, I. (2007) Direct analysis and MALDI imaging of formalin-fixed, paraffin-embedded tissue sections. *J. Proteome Res.* **6**, 1295–1305
  43. Schwamborn, K., Krieg, R. C., Reska, M., Jakse, G., Kneuchel, R., and Wellmann, A. (2007) Identifying prostate carcinoma by MALDI-Imaging. *Int. J. Mol. Med.* **20**, 155–159
  44. Walch, A., Rauser, S., Deininger, S. O., and Höfler, H. (2008) MALDI imaging mass spectrometry for direct tissue analysis: a new frontier for molecular histology. *Histochem. Cell Biol.* **130**, 421–434
  45. Brown, L. M., Helmke, S. M., Hunsucker, S. W., Netea-Maier, R. T., Chiang, S. A., Heinz, D. E., Shroyer, K. R., Duncan, M. W., and Haugen, B. R. (2006) Quantitative and qualitative differences in protein expression between papillary thyroid carcinoma and normal thyroid tissue. *Mol. Carcinog.* **45**, 613–626
  46. Chaurand, P., Sanders, M. E., Jensen, R. A., and Caprioli, R. M. (2004) Proteomics in diagnostic pathology: profiling and imaging proteins directly in tissue sections. *Am. J. Pathol.* **165**, 1057–1068
  47. Johnson, M. D., Floyd, J. L., and Caprioli, R. M. (2006) Proteomics in diagnostic neuropathology. *J. Neuropathol. Exp. Neurol.* **65**, 837–845
  48. Marko-Varga, G., Lindberg, H., Löfdahl, C. G., Jönsson, P., Hansson, L., Dahlbäck, M., Lindquist, E., Johansson, L., Foster, M., and Fehniger, T. E. (2005) Discovery of biomarker candidates within disease by protein profiling: principles and concepts. *J. Proteome Res.* **4**, 1200–1212
  49. Meistermann, H., Norris, J. L., Aerni, H. R., Cornett, D. S., Friedlein, A., Erskine, A. R., Augustin, A., De Vera Mudry, M. C., Ruepp, S., Suter, L., Langen, H., Caprioli, R. M., and Ducret, A. (2006) Biomarker discovery by imaging mass spectrometry: transthyretin is a biomarker for gentamicin-induced nephrotoxicity in rat. *Mol. Cell. Proteomics* **5**, 1876–1886
  50. Sköld, K., Svensson, M., Nilsson, A., Zhang, X., Nydahl, K., Caprioli, R. M., Svenningsson, P., and Andrén, P. E. (2006) Decreased striatal levels of PEP-19 following MPTP lesion in the mouse. *J. Proteome Res.* **5**, 262–269
  51. Stoeckli, M., Knochenmuss, R., McCombie, G., Mueller, D., Rohner, T., Staab, D., and Wiederhold, K. H. (2006) MALDI MS imaging of amyloid. *Methods Enzymol.* **412**, 94–106
  52. Groseclose, M. R., Massion, P. P., Chaurand, P., and Caprioli, R. M. (2008) High-throughput proteomic analysis of formalin-fixed paraffin-embedded tissue microarrays using MALDI imaging mass spectrometry. *Proteomics* **8**, 3715–3724
  53. Ronci, M., Bonanno, E., Colantoni, A., Pieroni, L., Di Ilio, C., Spagnoli, L. G., Federici, G., and Urbani, A. (2008) Protein unlocking procedures of formalin-fixed paraffin-embedded tissues: application to MALDI-TOF imaging MS investigations. *Proteomics* **8**, 3702–3714
  54. Stauber, J., Lemaire, R., Franck, J., Bonnel, D., Croix, D., Day, R., Wisztorski, M., Fournier, I., and Salzet, M. (2008) MALDI imaging of formalin-fixed paraffin-embedded tissues: application to model animals of Parkinson disease for biomarker hunting. *J. Proteome Res.* **7**, 969–978
  55. Van de Plas, R., Ojeda, F., Dewil, M., Van Den Bosch, L., De Moor, B., and Waelkens, E. (2007) Prospective exploration of biochemical tissue composition via imaging mass spectrometry guided by principal component analysis. *Pac. Symp. Biocomput.* 458–469
  56. Dijkstra, M. C., Carolan, V., Loadman, P. M., and Clench, M. R. (2008) Method development for protein profiling in biological tissues by matrix-assisted laser desorption/ionisation mass spectrometry imaging. *Rapid Commun. Mass Spectrom.* **22**, 1615–1618
  57. McCombie, G., Staab, D., Stoeckli, M., and Knochenmuss, R. (2005) Spatial and spectral correlations in MALDI mass spectrometry images by clustering and multivariate analysis. *Anal. Chem.* **77**, 6118–6124
  58. Deininger, S. O., Ebert, M. P., Fütterer, A., Gerhard, M., and Röcken, C. (2008) MALDI imaging combined with hierarchical clustering as a new tool for the interpretation of complex human cancers. *J. Proteome Res.* **7**, 5230–5236
  59. Lemaire, R., Stauber, J., Wisztorski, M., Van Camp, C., Desmons, A., Deschamps, M., Proess, G., Rudlof, I., Woods, A. S., Day, R., Salzet, M., and Fournier, I. (2007) Tag-mass: specific molecular imaging of transcriptome and proteome by mass spectrometry based on photocleavable tag. *J. Proteome Res.* **6**, 2057–2067
  60. Chaurand, P., Rahman, M. A., Hunt, T., Mobley, J. A., Gu, G., Latham, J. C., Caprioli, R. M., and Kasper, S. (2008) Monitoring mouse prostate development by profiling and imaging mass spectrometry. *Mol. Cell. Proteomics* **7**, 411–423
  61. Din, S., Lennon, A. M., Arnott, I. D., Hupp, T., and Satsangi, J. (2007) Technology insight: the application of proteomics in gastrointestinal disease. *Nat. Clin. Pract. Gastroenterol. Hepatol.* **4**, 372–385
  62. Fournier, I., Wisztorski, M., and Salzet, M. (2008) Tissue imaging using MALDI-MS: a new frontier of histopathology proteomics. *Expert Rev. Proteomics* **5**, 413–424
  63. Francesc, S., Dani, F. R., Traldi, P., Mastrobuoni, G., Pieraccini, G., and Moneti, G. (2009) MALDI mass spectrometry imaging, from its origins up to today: the state of the art. *Comb. Chem. High Throughput Screen.* **12**, 156–174
  64. Norris, J. L., Cornett, D. S., Mobley, J. A., Andersson, M., Seeley, E. H., Chaurand, P., and Caprioli, R. M. (2007) Processing MALDI mass spectra to improve mass spectral direct tissue analysis. *Int. J. Mass Spectrom.* **260**, 212–221
  65. Tilz, G. P., Wiltgen, M., Demel, U., Faschinger, C., Schmidinger, H., and Hermetter, A. (2007) Insights into molecular medicine: development of new diagnostic and prognostic parameters. *Wien. Med. Wochenschr.* **157**, 122–129
  66. Wisztorski, M., Lemaire, R., Stauber, J., Menguellet, S. A., Croix, D., Mathé, O. J., Day, R., Salzet, M., and Fournier, I. (2007) New developments in

- MALDI imaging for pathology proteomic studies. *Curr. Pharm. Des.* **13**, 3317–3324
67. Reyzer, M. L., and Caprioli, R. M. (2007) MALDI-MS-based imaging of small molecules and proteins in tissues. *Curr. Opin. Chem. Biol.* **11**, 29–35
68. Hsieh, Y., Casale, R., Fukuda, E., Chen, J., Knemeyer, I., Wingate, J., Morrison, R., and Korfmacher, W. (2006) Matrix-assisted laser desorption/ionization imaging mass spectrometry for direct measurement of clozapine in rat brain tissue. *Rapid Commun. Mass Spectrom.* **20**, 965–972
69. Hsieh, Y., Chen, J., and Korfmacher, W. A. (2007) Mapping pharmaceuticals in tissues using MALDI imaging mass spectrometry. *J. Pharmacol. Toxicol. Methods* **55**, 193–200
70. O'Brien, E., Dedova, I., Duffy, L., Cordwell, S., Karl, T., and Matsumoto, I. (2006) Effects of chronic risperidone treatment on the striatal protein profiles in rats. *Brain Res.* **1113**, 24–32
71. Reyzer, M. L., Hsieh, Y., Ng, K., Korfmacher, W. A., and Caprioli, R. M. (2003) Direct analysis of drug candidates in tissue by matrix-assisted laser desorption/ionization mass spectrometry. *J. Mass Spectrom.* **38**, 1081–1092
72. Rubakhin, S. S., Jurchen, J. C., Monroe, E. B., and Sweedler, J. V. (2005) Imaging mass spectrometry: fundamentals and applications to drug discovery. *Drug Discov. Today* **10**, 823–837
73. Wang, H. Y., Jackson, S. N., McEuen, J., and Woods, A. S. (2005) Localization and analyses of small drug molecules in rat brain tissue sections. *Anal. Chem.* **77**, 6682–6686
74. Dekker, L. J., van Kampen, J. J., Reedijk, M. L., Burgers, P. C., Gruters, R. A., Osterhaus, A. D., and Luiders, T. M. (2009) A mass spectrometry based imaging method developed for the intracellular detection of HIV protease inhibitors. *Rapid Commun. Mass Spectrom.* **23**, 1183–1188
75. Hopfgartner, G., Varesio, E., and Stoeckli, M. (2009) Matrix-assisted laser desorption/ionization mass spectrometric imaging of complete rat sections using a triple quadrupole linear ion trap. *Rapid Commun. Mass Spectrom.* **23**, 733–736
76. Huamani, J., Willey, C., Thotala, D., Niermann, K. J., Reyzer, M., Leavitt, L., Jones, C., Fleishcher, A., Caprioli, R., Hallahan, D. E., and Kim, D. W. (2008) Differential efficacy of combined therapy with radiation and AEE788 in high and low EGFR-expressing androgen-independent prostate tumor models. *Int. J. Radiat. Oncol. Biol. Phys.* **71**, 237–246
77. Trim, P. J., Henson, C. M., Avery, J. L., McEwen, A., Snel, M. F., Claude, E., Marshall, P. S., West, A., Princivalle, A. P., and Clench, M. R. (2008) Matrix-assisted laser desorption/ionization-ion mobility separation-mass spectrometry imaging of vinblastine in whole body tissue sections. *Anal. Chem.* **80**, 8628–8634
78. Reyzer, M. L., Caldwell, R. L., Dugger, T. C., Forbes, J. T., Ritter, C. A., Guix, M., Arteaga, C. L., and Caprioli, R. M. (2004) Early changes in protein expression detected by mass spectrometry predict tumor response to molecular therapeutics. *Cancer Res.* **64**, 9093–9100
79. Khatib-Shahidi, S., Andersson, M., Herman, J. L., Gillespie, T. A., and Caprioli, R. M. (2006) Direct molecular analysis of whole-body animal tissue sections by imaging MALDI mass spectrometry. *Anal. Chem.* **78**, 6448–6456
80. Atkinson, S. J., Loadman, P. M., Sutton, C., Patterson, L. H., and Clench, M. R. (2007) Examination of the distribution of the bioreductive drug AQ4N and its active metabolite AQ4 in solid tumours by imaging matrix-assisted laser desorption/ionisation mass spectrometry. *Rapid Commun. Mass Spectrom.* **21**, 1271–1276
81. Wisztorski, M., Verplanck, N., Thomy, V., Stauber, J., Camart, J. C., Salzet, M., and Fournier, I. (2007) Use of Masks in MALDI-IMS: an easy tool for increasing spatial resolution of images by decreasing irradiated area, in *Proceedings of the 55th ASMS Conference on Mass Spectrometry and Allied Topics, Indianapolis, June 4–7, 2007*, Abstract No. WPD-066, American Society for Mass Spectrometry, Santa Fe, NM
82. Sinha, T. K., Khatib-Shahidi, S., Yankeelov, T. E., Mapara, K., Ehteshami, M., Cornett, D. S., Dawant, B. M., Caprioli, R. M., and Gore, J. C. (2008) Integrating spatially resolved three-dimensional MALDI IMS with in vivo magnetic resonance imaging. *Nat. Methods* **5**, 57–59
83. Andersson, M., Groseclose, M. R., Deutch, A. Y., and Caprioli, R. M. (2008) Imaging mass spectrometry of proteins and peptides: 3D volume reconstruction. *Nat. Methods* **5**, 101–108
84. Stoeckli, M., Chaurand, P., and Caprioli, R. M. (1999) Applications of MALDI MS imaging to biological samples, in *Proceedings of the 47th ASMS Conference on Mass Spectrometry and Allied Topics, Dallas, June 13–17, 1999*, Abstract No. TPG-264, American Society for Mass Spectrometry, Santa Fe, NM

**Annexe II: MITICS (MALDI Imaging Team Imaging Computing System): a new open source mass spectrometry imaging software.**

available at [www.sciencedirect.com](http://www.sciencedirect.com)[www.elsevier.com/locate/jprot](http://www.elsevier.com/locate/jprot)

## MITICS (MALDI Imaging Team Imaging Computing System): A new open source mass spectrometry imaging software

Olivia Jardin-Mathé<sup>a</sup>, David Bonnel<sup>a</sup>, Julien Franck<sup>a</sup>, Maxence Wisztorski<sup>a</sup>,  
Eduardo Macagno<sup>b</sup>, Isabelle Fournier<sup>a,\*</sup>, Michel Salzet<sup>a,\*</sup>

<sup>a</sup>University of Lille1, CNRS FRE 2933, MALDI Imaging team, Laboratoire de Neuroimmunologie des Annélides, Villeneuve d'Ascq, France

<sup>b</sup>Section of Cell and Developmental Biology, Division of Biological Sciences, University of California San Diego, La Jolla, CA 92093, USA

### ARTICLE INFO

#### Keywords:

MALDI imaging  
Mass spectrometry  
Software

### ABSTRACT

MITICS is a new software developed for MALDI imaging. We tried to render this software compatible with all types of instruments. MITICS is divided in two parts: MITICS control for data acquisition and MITICS Image for data processing and images reconstruction. MITICS control is available for Applied BioSystems MALDI-TOF instruments and MITICS Image for both Applied BioSystems and Bruker Daltonics ones. MITICS Control provides an interface to the user for setting the acquisition parameters for the imaging sequence, namely set instruments acquisition parameters, create the raster of acquisition and control post-acquisition data processing, and provide this settings to the automatic acquisition software of the MALDI instrument. MITICS Image ensures image reconstruction, files are first converted to XML files before being loaded in a database. In MITICS image we have chosen to implement different data representations and calculations for image reconstruction. MITICS Image uses three different representations that have shown to ease extraction of information from the whole data set. It also offers image reconstruction base either on the maximum peak intensity or the peak area. Image reconstruction is possible for single ions but also by summing signals of different ions. MITICS was validated on biological cases.

© 2008 Elsevier B.V. All rights reserved.

## 1. Introduction

Mass spectrometry analyzes proteins based on the mass-to-charge ( $m/z$ ) ratio of ionized molecules. A typical mass spectrometer consists of an ionization source, a mass analyzer, and a detector for counting the number of analytes at each  $m/z$  ratio. In addition, mass spectrometers are commonly coupled with separation devices, such as liquid chromatography instrumentation. Aside from providing rapid data on various proteins present in biological samples, mass spectrometry also offers specific information about post-translational modifications associated with, for example, a particular disease state. Recent instrument developments and electronic improvements now place mass spectrometry at the

forefront of many biological applications. In particular, this technology has been crucial in efforts to identify the human proteome. In combination with classical biochemical methods, mass spectrometry has been largely responsible for revealing an expanded biological diversity at the protein level, due to alternative splicing and post-translational modifications that could not be predicted by genome sequence analysis alone.

Although mass spectrometry technologies are efficient in the identification of novel proteins, until recently they did not allow the determination of chemicals distribution or the regulation of peptide/proteins expressed in tissues or single cells. With new emerging mass spectrometry technologies, it has now been demonstrated that direct tissue analysis is

\* Corresponding authors.

E-mail addresses: [Isabelle.fournier@univ-lille1.fr](mailto:Isabelle.fournier@univ-lille1.fr) (I. Fournier), [michel.salzet@univ-lille1.fr](mailto:michel.salzet@univ-lille1.fr) (M. Salzet).

feasible using matrix-assisted laser desorption/ionization (MALDI) sources. A major advantage of direct MALDI analysis is to avoid time-consuming extraction, purification or separation steps, which also have the potential for producing artifacts. Therefore, direct MALDI analysis in tissue sections enables the acquisition of cellular expression profiles while maintaining cellular and molecular integrity [1–7]. With automation and the ability to reconstruct complex spectral data with imaging software, it is now possible to produce multiplex imaging maps of selected bio-molecules within tissue sections [8,9]. Thus, direct MALDI analysis obtained from tissue sections is convertible into imaging maps, a method now known as MALDI imaging. MALDI imaging combines the power of mass spectrometry, namely exquisite sensitivity, with unequivocal positional information within a normal morphological context [8,9]. Every MALDI imaging experiment is comprised of two steps: (a) sample preparation and treatment, and (b) data acquisition and processing. Given the nature and magnitude of the physical measurements, computationally powerful and transparent data processing and displaying tools are a critical requirement for the successful analysis of the large number of spectra obtained in any MALDI imaging experiment.

Currently, four principal MALDI imaging software packages are available, some commercially (Table 1). The original one, MMSIT, was written by M. Stoeckli for use on Applied Biosystems MALDI-TOF and MALDI-TOF/TOF instruments [8]. This software is available at <http://www.maldi-msi.org>. For Applied Biosystems, BioMap and more recently Tissue view 1.0 have been developed for 4700 and 4800 MALDI-TOF/TOF™ Analyzers, QStar Pulsar System, QSTAR® XL Hybrid LC/MS/MS System and the QSTAR. Bruker, Shimadzu and Thermo Fisher have developed their own software's namely FlexImaging, Axima2 Analyze and ImageQuest. The latest ones, Create Target and Analyze this! are two new software assisting MALDI-IMS on Bruker Reflex IV and Ultraflex II instruments [10]. We have developed MITICS to have a software that is as possible workable with all type of data whatever the instrument used for both data acquisition when necessary and for data processing. Data processing for MALDI imaging is a difficult part that is always a compromise between finding the processing that better extracts real biological data and processing that could generate non relevant biological information as translate into image representations. Here, we have been interested by adding new features to the software for trying to extract as many data as possible. We have also aim in releasing software that is open and implementable by who ever wants to. MITICS is composed of two parts namely MITICS Control and MITICS Image which are respectively dedicated to data acquisition and to data processing and image reconstruction. MITICS Control is actually available for MALDI-TOF Applied BioSystems instruments. MITICS Image is actually interfaced with Applied Biosystems and Bruker Daltonics data files format. MITICS converts all reading files, such as .dat files in Data Explorer for Applied Biosystems, to .xml files, the format used by FlexAnalysis for Bruker Daltonics. MITICS uses two different interfaces, one to load all the spectra from each run into a single database that is readily available for repeated analysis, and the second to reconstruct and display interactively the molecular images of interest. The

software also offers two alternative computational tools to create molecular maps, using the maximum intensity value of a peak to show where the sample is present, or use the integrated area of the peak above a set threshold. This presents an interest for MALDI-TOF data since peaks are broadenings and decreasing in intensity with increasing  $m/z$  values. Thus as a general trend, maximum intensity of a peak shows to be of interest for smaller  $m/z$  values whereas peak area is better fitting for higher  $m/z$  ratios. Threshold values for minimum and maximum have also been included to take into account for signal/noise ratio or intensity variations between experiments. This allows for defining what is the minimum intensity or peak area to be taken into account and considered as a real signal and what is the maximum intensity of a peak excluding intensity variations between collected data at different coordinates. MITICS offers also three different representations of the data collected, the classical averaged spectrum as well as two new representations that are the sum of all the collected mass spectra as well as a spectrum displaying for each  $m/z$  the maximum intensity as found in the data set. These different representations have shown to enlarge the number of signals that have a biological significance, i.e. better exploitation of data can be taken out of using these three representations while avoiding to browse data set  $m/z$  by  $m/z$ . MITICS is presented here on two different biological applications for validation of the software running. MITICS was applied to a case of biomarker following in brain tissues of rats treated with 6-OHDA, fixed in formalin and embedded for the past 12 years in paraffin, and to peptide mapping in dissected leech embryo whole mounts.

## 2. Experimental

### 2.1. Materials

$\alpha$ -cyano-4-hydroxycinnamic acid (HCCA), ammonium bicarbonate, trisma base, toluene, ethanol (HPLC grade), acetonitrile (HPLC grade), methanol (HPLC grade), Trifluoroacetic acid, aniline, Angiotensin II, Des-Arg-Bradykinin, Substance P, ACTH 18–39, ACTH 7–38 and bovine Insulin were obtained from Sigma–Aldrich and used as described. Trypsin was obtained from Promega.

### 2.2. Animal surgery and treatment

Wistar male rats (180–200 g) were anesthetized with pentobarbital (50 mg/kg, i.p.) and infused over 8 min with 6-OHDA (8 mg in 4 ml of 0.05% ascorbic acid in saline) at coordinates A 5 23.8 mm, L 5 1.5 mm, H 5 28.5 mm [11].

Leech embryos were obtained from a laboratory colony. Prior to dissection, they were anesthetized in 8% Ethanol Ringer's, then opened dorsally and pinned flat in a Sylgard-covered dish. After removal of the yolk, the embryos were briefly submerged in Methanol, transferred to metal-coated (ITO) slides and gently dried on a hot plate at ~45 °C.

### 2.3. Tissue fixation

Rats were sacrificed 4 h after the last infusion of 6-OHDA. Their brains were removed and fixed for 40 min at 48 °C in 4%



Table 1 – Comparison table between all imaging mass spectrometry software already known

Software	Type	Different choices for calculations	Input format	Output format	Spectrometers compatibility	Use of databank	Requirements	Maximum data points	Processing Time	Availability	Free	Authors	Code
MMSIT	Acquisition	No	Instrument format .dat	Analyze 7.5	AB MALDI/TOF	Yes	Aquris DP211 digitizer board	Unknown	Few seconds	<a href="http://www.maldi-si.org">www.maldi-si.org</a>	Yes	M. Stoeckli	VB6 and assembler C#
4700/4800 Imaging Create Target AND Analyze This!	Acquisition	No	Instrument format	Analyze 7.5	AB MALDI-TOF/TOF	Unknown	Instruments software's	32 768 per spectrum	–	<a href="http://www.maldi-msi.org">www.maldi-msi.org</a>	Yes	M. Stoeckli	C#
Axima2 Analyze oMALDI Server 5.1	Acquisition/data conversion	No	Binary	Analyze 7.5	BD MALDI/TOF and MALDI-TOF/TOF	No	Instruments software's and Biomap	Theoretically 1 million per picture	25 mn for .img conversion	<a href="http://www.maldi-msi.org">www.maldi-msi.org</a>	Yes	C. Ceuppens et al.	Visual Basic 6
Flex Imaging	Acquisition/data conversion	No	Instrument format	Analyze 7.5	SZU	Unknown	Windows	Unknown	–	<a href="http://www.maldi-msi.org">www.maldi-msi.org</a>	Yes	M. Setou et al.	Unknown
Tissue view 1.0	Acquisition/Image reconstruction	No	Unknown	Analyze 7.5	AB/MDS SCIEX oMALDI QSTAR Hybrid LC/MS/MS	Unknown	Instruments software's	Unknown	Unknown	AB	No	AB/MDS SCIEX	C++
Image Quest	Acquisition/Image reconstruction	No	Instrument format .img files	bmp, jpeg, jpg, tif, tiff and MALDI-TOF/TOF	BD MALDI/TOF and MALDI-TOF/TOF	Unknown	Instruments software's	Unknown	Few seconds	BD	No	BD	Unknown
BIOMAP	Images reconstruction	No	Instrument format	tiff, jpeg, Analyze 7.5	AB/MDS SCIEX MALDI-TOF/TOF and QSTAR	Yes	Instruments software's	Unknown	Few seconds	AB	No	AB/MDS SCIEX	IDL
MITICS	Acquisition/Image reconstruction	No	Instrument format .xml	Unknown	TS MALDI LTQ XL, LTQ Orbitrap hybrid series	Unknown	Instruments software's	Unknown	Unknown	TS	No	TS	Unknown
			Analyze 7.5, .img, tiff, .dicom, .png	tiff, jpeg, Analyze 7.5 and 4700 and 4800 MALDI-TOF/TOF, MALDI Micro MX(tm) (Waters)	AB MALDI-TOF and 4700 and 4800 MALDI-TOF/TOF, MALDI Micro MX(tm) (Waters)	Yes	IDL RT 6.1 or higher	32 768 per spectrum for older versions	Few seconds	<a href="http://www.maldi-msi.org">www.maldi-msi.org</a>	Yes	M. Rausch with additions by M. Stoeckli	IDL
			bmp, jpg, jp2, gif, pict, png, tif	bmp, jpg, Analyze 7.5 and 4700 and 4800 MALDI-TOF/TOF, MALDI Micro MX(tm) (Waters)	AB MALDI/TOF, BD MALDI/TOF, TOF/TOF	Yes, MySQL	Instruments software's, IDL VN and Data miner	Theoretically 1 million per picture	45 min/16000 points	Upon email	Yes	O. Jardin-Mathé et al.	C++, IDL, Visual Basic 6

Table has been created from data collected via different sources and may not be exhaustive or may contain uncompleted data and could be submitted to changes.  
 AB: Applied Biosystems; BD: Bruker Daltonics; SZU: Shimadzu; TS: Thermo Scientific.

paraformaldehyde. The accuracy of the 6-OHDA lesion was checked by immunocytochemistry with an anti-tyrosine hydroxylase antibody according to a previously published procedure [11]. These brain tissue blocks were then stored in a box at room temperature 12 years before being used in this experimental plan procedure.

#### 2.4. Tissue dewaxing

Rat tissue sections of 10  $\mu\text{m}$  thickness were obtained using a microtome (E. Leitz, Westlars, Germany) and mounted on ITO (Indium Teen Oxided) one side coated conductive glass slides. Paraffin was removed by 2 baths of 5 min in toluene and lightly rehydrated with graded ethanol (100°, 96°, 70° and 30°) baths before drying at room temperature.

#### 2.5. In situ trypsin digestion for MALDI imaging of rat brain tissue sections

Trypsin deposition (20  $\mu\text{g}/\text{mL}$  in water) was performed using an automatic vibration vaporization system (ImagePrep, Bruker Daltonics, Bremen, Germany). The quantity of trypsin was set, after optimization, to 1 mL over 30 cycles in order to cover the whole surface of the tissue. The system settings were: Spray power 30%, Spray modulation 40%, time spray on 2 s, incubation time: 30 s and dry time 30 s. HCCA/Aniline solid ionic matrix was used for these experiments [12]. The matrix preparation was described previously. Briefly, just prior to use, 4.8  $\mu\text{L}$  of Aniline (1 eq.) was added to the matrix solution (HCCA 10 mg/mL in ACN/TFA 0.1%, 6:4, v/v) and sonicated for 15 min. HCCA/Aniline was then deposited on the tissue section using the vibration vaporization system according to the same parameters as described above.

#### 2.6. Mass spectrometry

##### 2.6.1. MALDI MS direct analysis of tissue sections after in situ automatic trypsin digestion

Trypsin fingerprint experiments on FFPE tissue sections after in situ digestion of the whole tissue section using the vibration sprayer system were performed on an Ultraflex II TOF-TOF instrument (Bruker Daltonics, Bremen, Germany) equipped with LIFT III cell and a Smartbeam laser with a repetition rate up to 200 Hz. For MS/MS experiments, parameters were set as follow: laser repetition rate was 100 Hz with 33% attenuation, ion source voltages were respectively 8 kV and 7.3 kV on a MALDI sample plate and 1st electrode; LIFT cell was pulse from ground for electrode 1 and 2 to 19 kV and in the last step electrode 3 was decrease to 3.2 kV; reflector end voltage was set to 29.5 kV and mid-grid to 13.85 kV. Trypsin fingerprint protein identification in databanks were performed using the Biotools 3.0 software (Bruker Daltonics, Bremmen, Germany) connected to Mascot search engine and interrogating the NCBI, Swissprot and ESTs databanks with Oxydation as Variable modification, Monoisotopic as mass values, a peptide mass tolerance of  $\pm 0.1$  Da with a peptide charge state of 1, a number of maximum missed cleavages of 2 and number of queries of 300.

#### 2.7. MALDI MSI

MALDI-MSI of rat brain tissue sections was performed on an Ultraflex II TOF-TOF (Bruker Daltonics, Bremen, DE). After dewaxing, images were obtained in positive reflector mode after trypsin digestion and matrix deposition. 1 mL of the mix was applied onto the tissue using the ImagePrep sprayer as explained above and dried at room temperature. Acquisition was performed using a Smartbeam laser; with a repetition rate of 100 Hz and a spatial resolution of  $100 \times 100 \mu\text{m}$ . Images were reconstructed using either MITICS or FlexImaging v2.0 software (Bruker Daltonics, Bremen, Germany). Average mass spectra were recorded in positive mode for each of 12919 spots, covering the whole tissue section with an average of 500 laser shots per position.

MALDI-MSI was also performed on several dissected leech embryos (see above) at the E12 stage of development, mounted either with the inside of the embryo or the epidermis facing the laser beam. In one case the endodermal layer was removed during dissection. For reference, histological images were recorded prior to covering the tissues with a layer of ionic matrix.

---

### 3. Results and discussion

#### 3.1. Software development

The new MITICS MS imaging software is divided in two parts, namely “MITICS Control” and “MITICS Image” (Fig. 1). MITICS Control is dedicated to create an interface for data acquisition on the instrument and thus requires that the instruments acquisition software be run in parallel, while MITICS Imaging is dedicated to post-acquisition data processing and image reconstruction. MITICS Control has been programmed in C++ for faster execution, whereas MITICS Image has been coded under IDL (Interactive Data Language) v 6.0 (ITT VIS, Paris, France), since IDL offers a powerful computing environment for data analysis and visualization. MITICS needs the IDL Virtual machine and the IDL Dataminer module to run. A temporary licence to test it can be obtained by emailing RSI support.

#### 3.2. Data acquisition using MITICS: MITICS control

Almost all MALDI instruments now provide an automatic acquisition mode for high throughput analysis of samples in standard proteomics workflows. In the automatic acquisition mode, the software uses a table that defines the positions of the spots to be analyzed. MITICS provides a simple and easy interface with the automatic acquisition software of the instrument for defining the image raster and the acquisition parameters of the instrument. For MALDI-TOF Voyager series Applied Biosystems instruments, MITICS provides to the acquisition software (Voyager Instrument Control Panel) all parameters that are required in the MALDI-IMS sequence. For this purpose, the interface creates three files: .bic, .plt and .seq that are given to the instrument for data acquisition. The .bic file of MITICS Control substitutes itself for the acquisition parameters normally administered by the Voyager Instrument

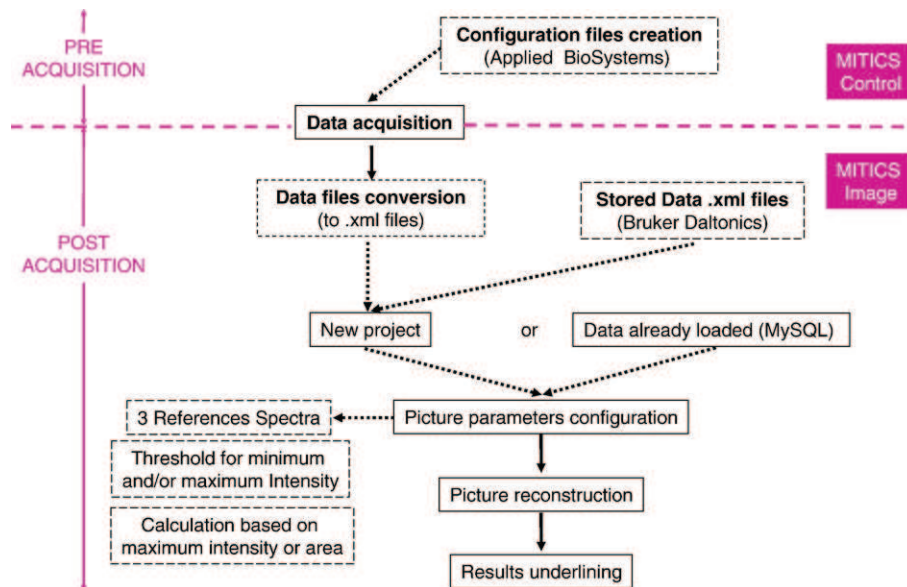


Fig. 1 – Scheme of the MITICS software global workflow including MITICS Control and MITICS Image parts.

acquisition software. The analysis parameters *i.e.* laser intensity, number of shots to be acquired per position, acceleration voltage on target and extraction grids, guide wire voltage, extraction delay time, mass scale, and low mass threshold are define and registered in the .bic file at the tab “BIC file” of the graphic interface (GUI). By contrast, the .plt and .sql files replace the instrument software for adjustment automatic mode values, *i.e.* the definition of the spots that will be sampled by the laser. Based on user-provided information, the interface calculates a table that is stored in the .plt files and given to the instrument during the automatic acquisition. For raster creation, the user defines Start and End coordinates and either a step size between spots or the number of spots per line or per column. The software then calculates automatically either the number of points or the step size, respectively. The X and Y dimensions have been decoupled, so that interval size can be set to a different value for each dimension. The third file (.sql) contains all the file information used by MITICS Control for data acquisition by providing these information’s to the automatic acquisition software, namely the name and the location of the .bic and .plt files, and the recorded data files. MITICS Control also allows the application of post-acquisition macros to the recorded data. For example, the user can define an automatic baseline correction to be performed on each spectrum, or apply a smoothing or noise removal function to the data. It is also possible to enter the name of the suitable function in the text area provided for defining such manipulations prior to acquiring the spectra. A similar interface is currently under development for Bruker Daltonics MALDI instruments.

### 3.3. Data processing and image reconstruction: MITICS Image

Both Biomap and FlexImaging softwares provide sets of analytical tools for image reconstruction and data or image processing, including display an image of an ion of interest,

review of spectra behind each pixel of an image, definition of regions of interest (ROI) for comparison and statistical analysis, or co-registration of images with tissue section pictures. However, the first step for an external reconstruction tool is to import and convert the data collected during the IMS experiments. Due to the differences between the formats of the recorded data depending on the instrument, different GUIs are required. For MITICS, in order to speed up the calculation and optimize the browsing of the spectra, we have chosen to import the data files from the instruments, convert them (if necessary) to XML file format, use XML parsing to collect data in the files and store them in a databank that is powered by MySQL. Spectra in XML format files are easily obtained thanks to existing command lines that run in FlexAnalysis (SaxBasic coding) for Bruker Daltonics and to a new macro that has been written in DataExplorer (Visual Basic) for Applied Biosystems devices. MySQL has been chosen for the database mainly because of its accuracy but also for its speed and ease of use.

The initial loading of IMS data requires XML files to be written and data to be transferred to the database and stored. At the second loading of the data, the data already stored in the database are extracted, which increase the speed of processing. This allows the user to store data in order to rebuild pictures later, facilitates search and interrogation of the spectra and enables the delocalization of the data bank to another computer on the network. Actually, it is not necessary that the database reside on the computer where the software is running. Moreover, since it is possible to link several databases to MITICS, as long as the ODBC sources are correctly configured, the data can be recorded and stored on several computers. This is a great advantage, knowing a picture composed of about 7000 points with a mass range from 600 to 24000 Da (a whole acquired data folder of 4,33 Go) requires a database size of ~1.2 GB. Conversely, if MITICS is running on several computers on a network, several pictures can be reconstructed at the same time from a single data set of the

database, allowing for high throughput data processing and image reconstruction. Because MALDI-IMS data contains up to hundreds of different peaks, i.e., potential images, this can be a great advantage.

### 3.3.1. XML files

MITICS is designed to operate on .xml files, but in some instruments the XML files may not be automatically created during spectra acquisition. To solve this potential difficulty, a macro can be added to MITICS control. This function will be executed for every spectrum during the data acquisition, so as to get all the files available in XML as soon as the acquisition processes completed. If a picture has to be created from spectra previously acquired, XML files can still be produced by running a batch process in *FlexAnalysis* for an analysis realised on a Bruker Daltonics device, or by executing a visual basic macro in *DataExplorer* for an analysis performed on an Applied Biosystems instrument. In MITICS macro is available that allows the selection of an appropriate macro and its application to all the files of a folder in order to transform .dat file to .xml files. Fig. 2 presents a typical XML file as obtained from Bruker Daltonics instruments after data exportation. Each collected XML file corresponds to a collected mass spectrum of the image data. The file name includes the (X,Y) coordinates at which this

average spectrum was collected. The data itself is composed of blocks that give all relevant information for image reconstruction on a precise peak defined by its *m/z* value. An image sequence is then represented by a collection of XML files. Originally XML files were found to be the easiest way to easily obtain compatible data from various instruments. Although, recently on the initiative of HUPO a new international standard format for imaging data “i-mzML” as been proposed (<http://www.i-mzml.org/>) as an extension of the “mzML” format also proposed as international standard for MS data in general. This means that spectra in this format should be available for all instruments soon. “i-mzML” format are composed of two files one is an XML file that contains informations relative to the experiment. The second file is a binary file containing of the recorded mass spectra. We propose in the future to replace our XML conversion and develop the software to use these standard file formats since this will give more simplicity and unity to the software while allowing to easily getting through data of all instruments.

### 3.3.2. Processing time

Tests were performed on a computer running Windows XP with an Intel Pentium 4 (3.2 GHz) processor and 1.00Go of RAM. In these conditions, it took 45 min to first load and store in the

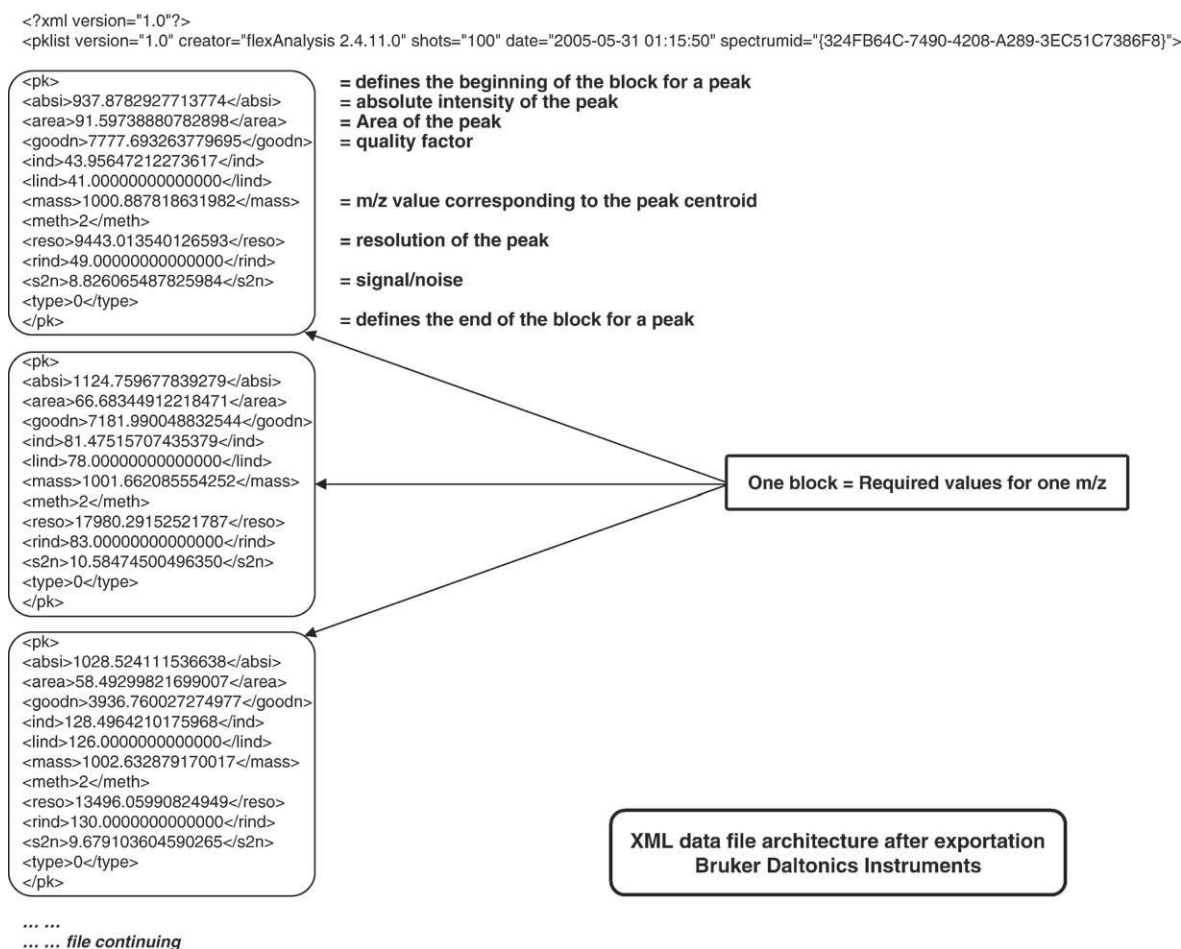
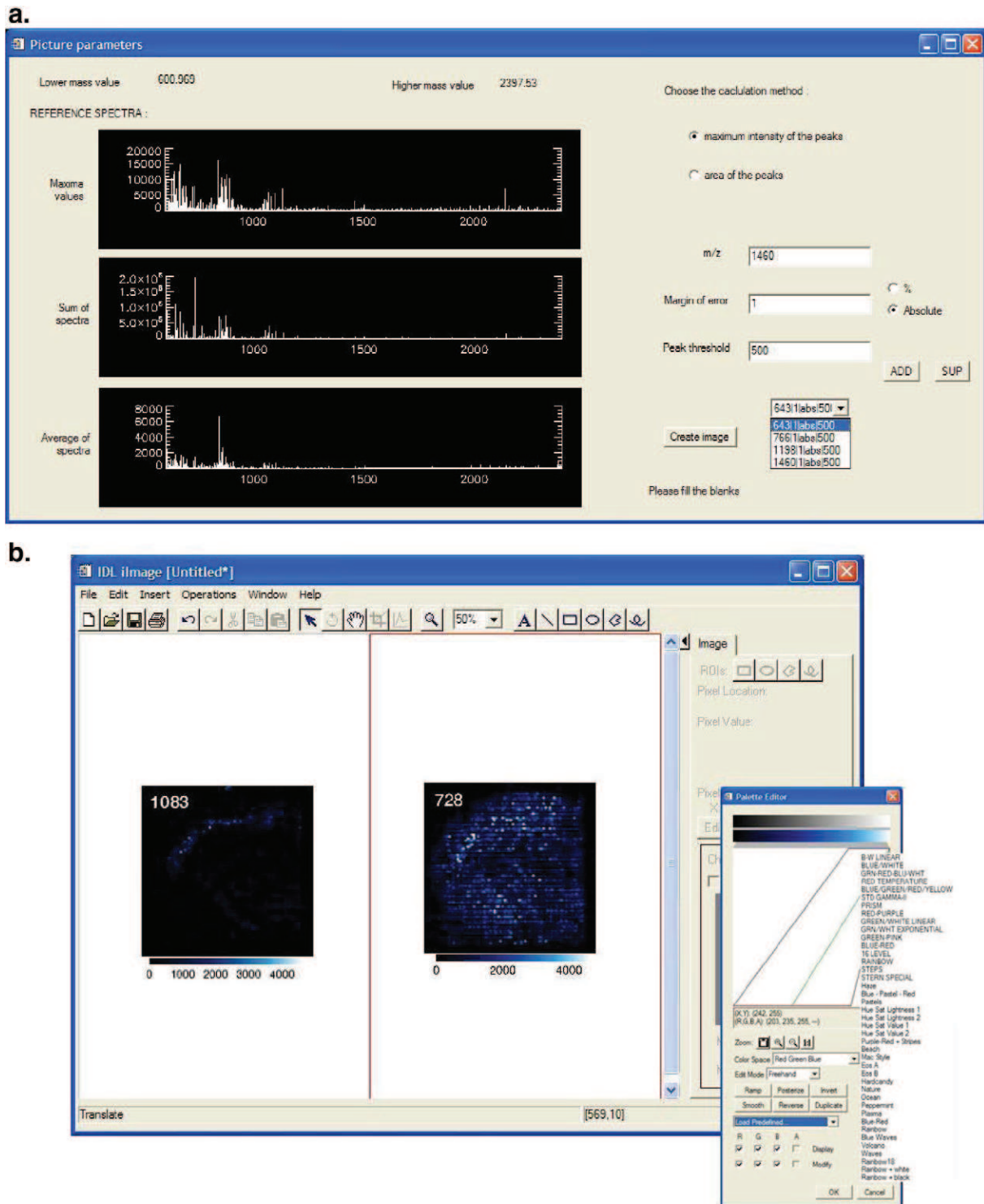


Fig. 2 – Example of the architecture of an XML file as obtained after exportation in batched process from a Bruker Daltonics instrument and prior to its loading into the Database.

database a 16000 point IMS data set. Then, reloading of the data directly from the databank took less than 1 min. Image reconstruction from firstly opened or reloaded data took also less than 1 min.

### 3.3.3. Data loading and databank storage

When data are first loaded on MITICS, the user has as first choice to define the MS system used for acquisition, e.g., an Applied Voyager DE-STR MALDI-TOF button or a Bruker



**Fig. 3** – Snapshots presenting the interface proposed by MITICS Image for data processing and Image reconstruction. a) Presents the interface for image parameters setting including the three different representations of the image data as well as the choice for using maximum peak intensity or peak area for image reconstruction. b. Presents the graphic interface after image reconstruction and the palette editor with the predefined color scales.

UltraFlex MALDI-TOF-TOF button. For data originating from MITICS control the .plt file (containing the coordinates of the picture's pixels) has to be provided. For all data, the location of the IMS data must be given. While processing the data, a continuously updated message appears at the bottom of the window to inform the user about the progress in the analysis. At the end of processing, a new window appears asking the user to define the parameters for image reconstruction. The XML files are loaded into the database the first time. The database itself is composed of three tables that respectively include the name of the project, the acquisition position (image rastering) and the spectral information's ( $m/z$ , intensity, area,...). These three tables are linked together and associate to a project a specific set of positions and for each of these positions spectral data are linked. After loading in the database the data are immediately available for image reconstruction. They are kept there for faster reloading. This allows also as mentioned to perform image reconstruction from the network on different computers and process several

imaging sequences at the same time from delocalized computers.

One part of the software is dedicated to the configuration of the databank used. The software can check automatically on the computer where the available ODBC sources are located in the Administration Tools of the Control Panel. In this manner, when a new project is created, collected spectral data can be loaded into the right database.

### 3.3.4. Image reconstruction

Fig. 3 presents the user interface of MITICS for image reconstruction. Image reconstruction is the most difficult part to be settled to find what the best representation of the data is without inducing false representation. For image reconstruction we have decided to implement new representations of the MS data collection and 2 different types of calculations for image reconstruction. For image reconstruction, the user must provide an  $m/z$  value within a margin error. Threshold values (minimum and maximum) can also be defined for taking into

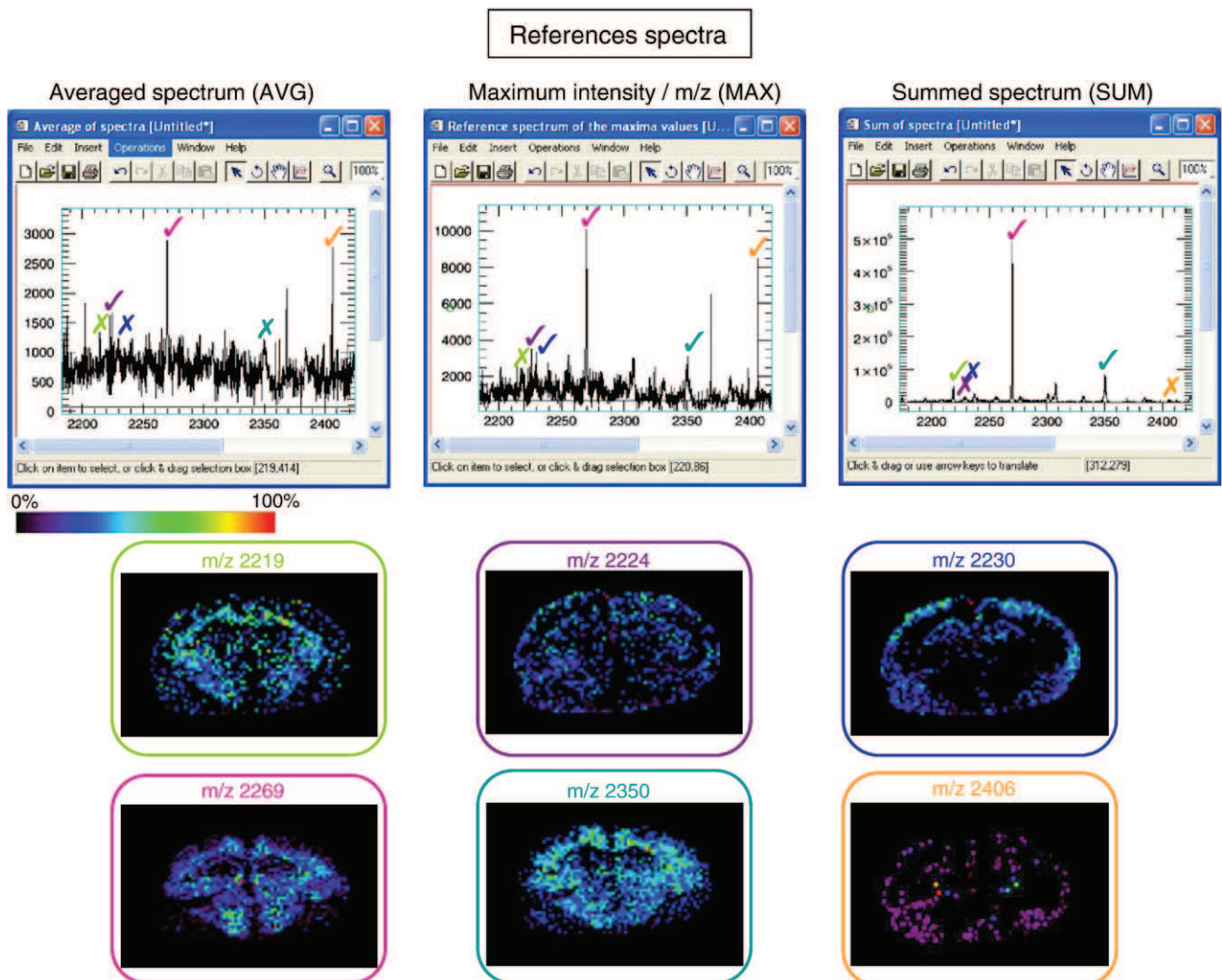


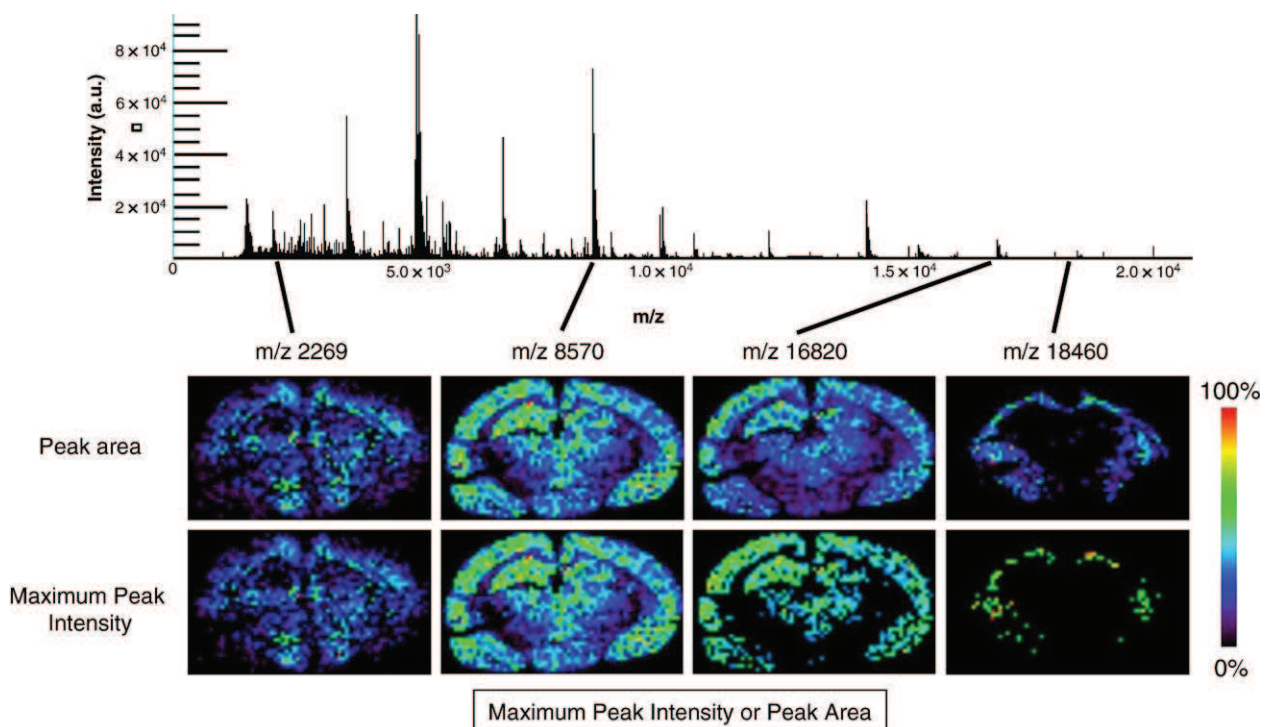
Fig. 4– Presentation of the three representations of image data proposed by MITICS and their interest for image reconstruction. Each spectrum correspond to one of the different representation (average, sum or maximum intensity) and images derived from the peaks observed in each of the representation according to the color code/ It is clear that all peaks are not displayed by all representations.

account signal variability (according to the position on the tissue). Minimum intensity threshold allows for defining the minimum value for intensity or area to consider that the signal is a peak. If the value is under the threshold then the value is set to 0 (no signal). The maximum value algorithm considers signal variability from one spot to another and decides that above a certain threshold values will be set to 100% intensity. After parameter adjustments, clicking the “add” button includes the selected ion in a list of ions that have been analyzed. Regarding the error margin, it can be defined either in percent of the total mass range or with an absolute value to add and subtract from the  $m/z$  value to obtain the boundaries of the accepted interval.

For data representation we have decided to had to the classical reconstructed averaged mass spectrum of all data (AVG), a mass spectrum corresponding to the sum of all mass spectra from the data collection (SUM) and a representative mass spectrum obtained by plotting for each  $m/z$  the maximum intensity as found in the data collection (MAX). This was motivated by the fact that significant biological information can be lost in the classical averaged spectrum representation of the data collection. This is for example the case for molecules that are either intense but much localized and for which the signal will be diluted by averaging or for molecules which are present in most of the tissue but at a quite low level of intensity. This well illustrated on Fig. 4 on the image reconstruction using MITICS for data obtained on a rat brain tissue section. The top part of the figure displayed the three different representations (AVG, SUM and MAX) of the same data set and the bottom part of the figure the

reconstructed molecular images using MITICS. It is clear from the different spectrum representations that peaks observables are different in each of the representation. Logically, the AVG representation gives the lowest signal intensity and the SUM representation reaches the highest level which as the convenience of increasing signal/noise ratio. To summarize, it can be shown that a peak that is quite intense with a large representation on the tissue section will be found in the three representations (e.g.  $m/z$  2269, pink). On the contrary a peak with a large representation on the tissue section but with a weak intensity is better depicted by the SUM or MAX spectra (e.g.  $m/z$  2350, pale blue). Peaks with weak intensity and more localized representation on the tissue section are only found in the MAX spectrum. This clearly illustrates the interest of adding different representation of the image data set for faster access to relevant information while avoiding browsing the entire  $m/z$  range to find hidden data in the averaged representation. All of the displayed representation allow for zooming in the corresponding window for ease of visualization.

We also choose to propose two methods of calculations either measuring the peak maximum intensity or its area. Because on MALDI-TOF instruments the resolution is not constant over the whole mass range (TOF peaks become broader with increasing  $m/z$  because of variations in initial positions and velocities), the maximum peak intensity does not always accurately describe the data. For this reason, measurement of the peak area instead of the maximum intensity can be displayed in MITICS Imaging. The user can then choose the calculation method. Peak intensity is better



**Fig. 5** – Molecular Image reconstructed with MITICS of 4 peptides/proteins at respectively  $m/z$  2269, 8570, 16 820 and 18 460 compared for the two calculations methods (maximum peak intensity or peak area). Data collected after an imaging sequence in positive mode on a rat brain tissue section using HCCA/Aniline as matrix.





for lower  $m/z$  ratio or for peaks very closed in  $m/z$  almost overlapping and peak area give better description for higher  $m/z$  signals if no overlapping of peaks is observed. Fig. 5 proposed the comparison for an imaging sequence recorded from a rat brain tissue section (same data collection as Fig. 4) of the two calculations methods. As expected the molecular images are almost identical by the two calculation methods for lower  $m/z$  ratio (e.g.  $m/z$  2269 or  $m/z$  8570). However, for higher  $m/z$  reconstructed images based on the peak area give better description of the compound localization and better fit the morphology of the tissue (ions at  $m/z$  16820 and  $m/z$  18460). Image reconstruction based on the peak area is not always an absolute requirement but can benefit to images for higher mass proteins.

Finally, the software can reconstruct several images at once if the  $m/z$ , error margins and threshold values have been settled. The software also gives the possibility of adding the signal of several ions and displayed the sum as one unique image. This was added in order to take into account the possibility to sum several ions corresponding to a same molecule i.e. for taking into account the presence of protonated or cationized ions of a same compound, different charge states for higher mass compounds or to sum all signals of different peptides generated by enzymatic digestion but belonging to the same original protein. Using this way we aimed in increasing the signal level (i.e. images with higher contrast) as well as averaging spot to spot variations due for example to a higher abundance of salts in some area leading to increase of cationized ion signals at the expense of protonated ions or to differences in enzyme activity depending the molecular environment at a precise location.

After image reconstruction, the software allows defining regions of interest, moving or turning the picture, resizing it, or

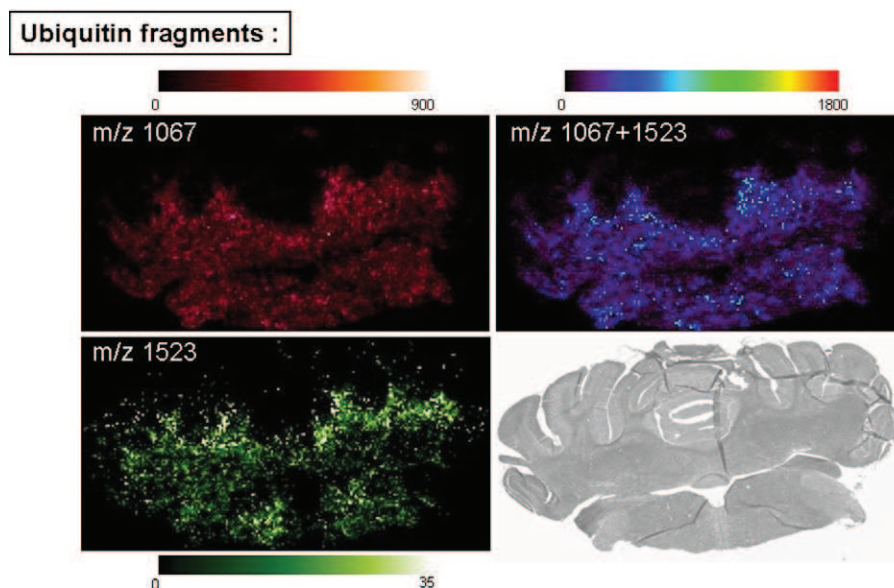
adding some geometrical shapes on it. It is also possible to save the image or print it, or to open a saved file or a new empty one. Some copy and/or paste or zoom functions are available as well. Another picture can be opened in the same window, making comparisons easier to do and some operations are available to work on the image, like filters for example, which permits getting a smoother picture compared to the initial simple hard pixel display. It is possible, moreover, to mark the results with a wide range of colors and gradations of basic tints or pre-specified multichromatic sets which are adaptable according to the wanted intensities. Lastly, the picture can be reprocessed after it has been displayed without the need to initiate another computation from the beginning — for example, maximum and minimum thresholds can be changed easily thanks to the Window Level Control in the Image Panel.

### 3.4. Application and validation of MITICS on biological cases

#### 3.4.1. Applications to biomarkers using Parkinson's model animals

As previously demonstrated [13,14], brain FFPE tissue blocks of ten 6-OHDA treated versus ten untreated rats were systematically studied and compared for biomarker hunting using trypsin digestion followed by nanoLC-ion-trap identification. Several biomarkers have been identified by this strategy to be either up or down regulated between control and injected animals especially Ubiquitin, Alpha enolase, CRMP2, Peroxidoxin 6 and F1 ATPase were found to be up-regulated. Some of these proteins and others (Profilin, Ferritin Heavy Chain, HSP27, Apolipoprotein E, Vimentin) were also previously found by other proteomics approaches [15] or direct tissue analysis [16,17] to be up-regulated.

To access the efficiency of MITICS software on a real biological case we have used it to reconstruct molecular



**Fig. 6**—Molecular images reconstructed using MITICS software of  $m/z$  1067 and 1523 ions and the composite image summing these two ions that correspond to two trypsin digestion fragments of Ubiquitin proteins as found by MALDI imaging in the reflectron positive mode on a FFPE rat 6-OHDA rat brain tissue section after tissue dewaxing and in situ trypsin digestion. Last panel presents photography of the tissue section prior to the experiment.

images from FFPE 6-OHDA rat brain tissue sections. We have performed, after dewaxing and rehydration, on tissue enzymatic digestion with trypsin of the tissue sections using a micro-spotting device and searched for fragments of the already known biomarkers by comparison with *in silico* digestions of these biomarkers (Table 2). Several digestions peptides of the markers of interest were found in the MALDI mass after enzymatic digestion and are given Table 2. These fragments were used to obtain molecular images of their repartition in the 6-OHDA injected rat brain sections. Fig. 6 presents the reconstructed images of two fragments ( $m/z$  1067 and  $m/z$  1523) of ubiquitin. The images show that the two fragments do present the same localization in the medulla oblongata whereas no signal is found in the cerebellum part of the section by comparison to the tissue section photography taken before the experiment where morphology is well apparent. The composite image of the sum of these two fragments is presented in the last panel. As previously mentioned possibility to sum different ions corresponding to a same molecule leads to a more contrasted image with higher signal intensity and fewer variations inter spots. Molecular image for  $m/z$  1523 ion presents a more diffuse image than those of  $m/z$  1067. This is not due to the software treatment but is probably due to the presence of a second smaller peak under the envelope of the peak corresponding to the ubiquitin digestion fragment that do not exactly present the same distribution as our fragment. Such effect is removed by using the composite image visualization, showing again the interest of using images based on several ion signals. Apolipoprotein E fragments were also found in the images data. The molecular images for two different digestion peptides of Apolipoprotein E respectively  $m/z$  726 and 1523 are shown Fig. 7. Here, molecular images reconstructed either on the peak maximum

intensity or area, are compared. Again we observed that in some cases area value give better description of the data than the peak maximum intensity.

3.4.2. *Applications to peptides/proteins localization in the leech*  
A second application was performed using MITICS for validation by studying peptides and proteins located in the CNS of the leech comparing leech embryos and adults. For adults, a metamer of the animal was removed and open on top of the dorsal part to obtain a strip of a metamer. An imaging sequence was performed on this leech segment to image the dorso-ventral orientation of molecules in the animal. Fig. 8 shows the difference for the  $m/z$  range 3200–3600 of the two representations AVG and MAX. In this  $m/z$  range only few and weak intensity peaks are observed in the AVG representation whereas more intense and numerous peaks are depicted in the MAX spectrum. By selecting on of these peaks ( $m/z$  3507 $\pm$ 0.5) it is possible to reconstruct the molecular image for this specific protein. It appears that it is highly localized in the nervous system (here a nervous ganglia). This validates again the interest of representation image data under several representations to extract interesting biological features. Whole mounted embryos were also studied using MALDI imaging. From the collected data set molecular images were reconstructed for two peptides at  $m/z$  3507 and 6443 using MITICS but also using Fleximaging and Biomap software's (Fig. 9). For ion at  $m/z$  3507 the three software's give very similar results especially for Biomap and MITICS that exactly provide the same color scale. Slight differences for Fleximaging are due to the color scale and the normalization of the data. For ion at  $m/z$  6443 more differences are observed. The signal of this ion is quite weak and the ion localization is hardly observed in the nephridia of the animal. It is a bit clearer on the molecular

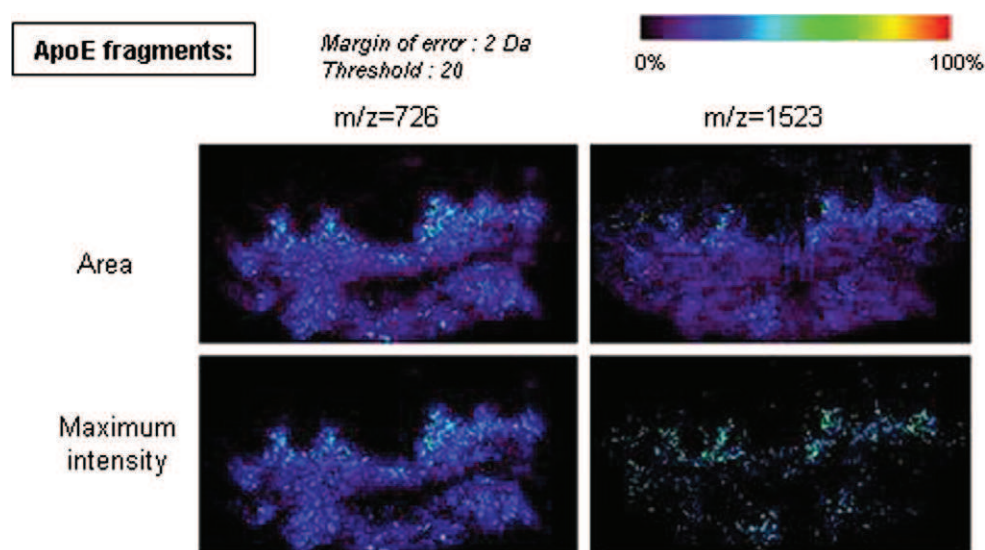
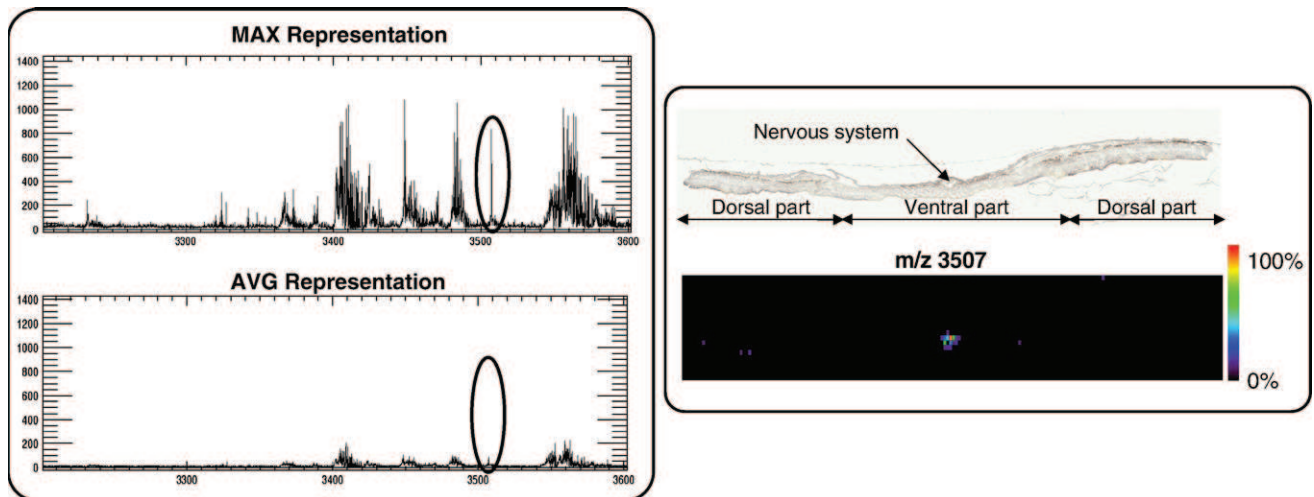


Fig. 7 – Comparison of the images reconstructed with MITICS based on Apolipoprotein E (Apo E) trypsin specific peptides  $m/z$  726 and 1523 using either the peak maximum intensity or the peak with a threshold of 20 and a margin of error of 2 Da for comparison of the two calculation methods. Data have been obtained after an imaging sequence in the positive reflectron mode on a FFPE 6-OHDA rat brain tissue section after tissue dewaxing and *in situ* trypsin digestion.



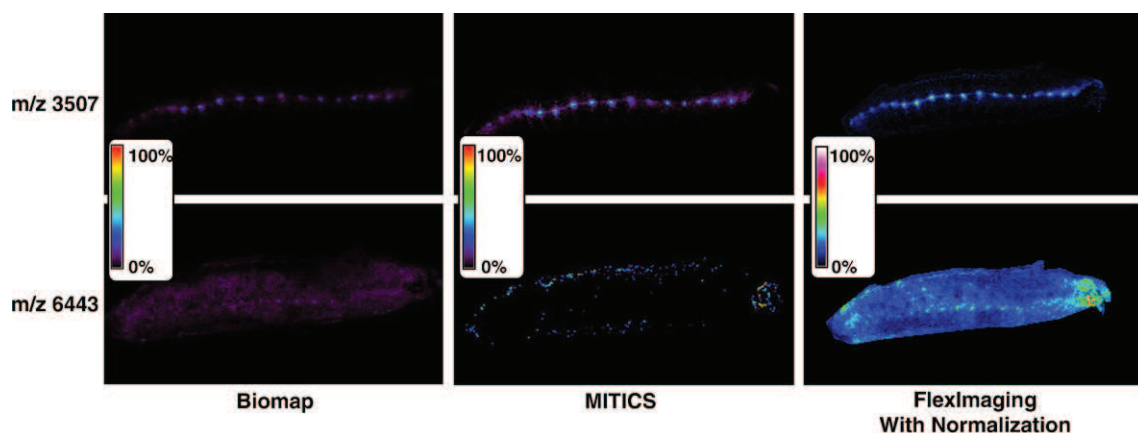
**Fig. 8**—Comparison of two representations proposed by MITICS (average and maximum of intensity per  $m/z$ ) for a MALDI imaging sequence performed on a leech segment. Data show that images of the ion at  $m/z$  16 620 are equivalents even if the peak of interest is only observable from the representation of maximum intensity per  $m/z$  showing the interest of displaying several representations. Data have been obtained by an imaging sequence in the linear positive mode after dissection of one leech segment and its opening on top of the dorsal part to obtain a dorso-ventral leech strip where the nervous ganglia of this peculiar segment is observable in the middle of the ventral part. HCCA/aniline was used as matrix for this imaging sequence.

image from Fleximaging due to normalization of the data. As observed here normalization of data find interest by enhancing signal weakest signal but as a tendency to increase images noise. However, this is something not yet implemented in MITICS that would be required in the next future to increase performances of the software.

#### 4. Conclusion

MITICS is novel integrated software for IMS that is designed for compatibility with all MALDI instruments. MITICS provide for the moment completely developed solutions for processing data from Applied Biosystems and Bruker Daltonics MALDI-TOF instruments, but it is aimed to evolve to be

compatible with SCIEX or Shimadzu or even other instruments. MITICS includes two parts, namely MITICS Control for data acquisition on the instrument and MITICS Image for data processing and image reconstruction. Data are processed following XML conversion and data extraction using an XML parser before storage in a database. Addition of databanks to the software allows for simultaneous reconstruction of numerous images in the goal of high throughput imaging to be developed. MITICS provides different views of data than other available software, namely different spectral representations for ions of interest, and offers the flexibility to use peak area instead of maximum intensity as the measurement basis for image reconstruction, which has proved to be better in describing biological informations for higher  $m/z$  ratio. Finally, MITICS tends to be universal software, easily modified for



**Fig. 9**—MALDI molecular images reconstructed with Biomap, MITICS and FLExImaging for validation of MITICS software of two ions  $m/z$  3507 and 6443 found after a MALDI imaging sequence of dissected stage E12 specimens leech embryos in whole mount. MALDI images were reconstructed from the mass spectra collected in the linear positive mode from 20 000 locations covering the embryos completely in a rectangular raster of points  $60 \mu\text{m}$  apart using HCCA/aniline as matrix.

integration of new tools and is available on request on our web site (<http://www.maldi-imaging.com>).

## Acknowledgements

Supported by grants from Centre National de la Recherche Scientifique (CNRS), Ministère de L'Éducation Nationale, de L'Enseignement Supérieur et de la Recherche (ACI Jeunes Chercheurs ACI JC4074 to I. Fournier), ANR PCV (to I. Fournier). Also supported by collaboration Bruker Daltonics and the laboratory for the MALDI imaging technology development.

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.jprot.2008.07.004](https://doi.org/10.1016/j.jprot.2008.07.004).

## REFERENCES

- [1] Fournier I, Day R, Salzet M. Direct analysis of neuropeptides by in situ MALDI-TOF mass spectrometry in the rat brain. *Neuro Endocrinol Lett* 2003;24(1–2):9–14.
- [2] Li L, Garden RW, Romanova EV, Sweedler JV. In situ sequencing of peptides from biological tissues and single cells using MALDI-PSD/CID analysis. *Anal Chem* 1999;71(24):5451–8.
- [3] Chaurand P, Stoeckli M, Caprioli RM. Direct profiling of proteins in biological tissue sections by MALDI mass spectrometry. *Anal Chem* 1999;71(23):5263–70.
- [4] Redeker V, Toullec JY, Vinh J, Rossier J, Soyez D. Combination of peptide profiling by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry and immunodetection on single glands or cells. *Anal Chem* 1998;70(9):1805–11.
- [5] Jimenez CR, Li KW, Dreisewerd K, Spijker S, Kingston R, Bateman RH, et al. Direct mass spectrometric peptide profiling and sequencing of single neurons reveals differential peptide patterns in a small neuronal network. *Biochemistry* 1998;37(7):2070–6.
- [6] Dreisewerd KKR, Geraerts WPM, Li KW. Direct mass spectrometric peptide profiling and sequencing of nervous tissues to identify peptides involved in male copulatory behavior in *Lymnaea stagnalis*. *Int J Mass Spectrom Ion Process* 1997;169/170:291–9.
- [7] Li KW, Hoek RM, Smith F, Jimenez CR, van der Schors RC, van Veelen PA, et al. Direct peptide profiling by mass spectrometry of single identified neurons reveals complex neuropeptide-processing pattern. *J Biol Chem* 1994;269(48):30288–92.
- [8] Stoeckli M, Farmer TB, Caprioli RM. Automated mass spectrometry imaging with a matrix-assisted laser desorption ionization time-of-flight instrument. *J Am Soc Mass Spectrom* 1999;10(1):67–71.
- [9] Caprioli RM, Farmer TB, Gile J. Molecular imaging of biological samples: localization of peptides and proteins using MALDI-TOF MS. *Anal Chem* 1997;69(23):4751–60.
- [10] Clerens S, Ceuppens R, Arckens L. CreateTarget and Analyze This!: new software assisting imaging mass spectrometry on Bruker Reflex IV and Ultraflex II instruments. *Rapid Commun Mass Spectrom* 2006;20(20):3061–6.
- [11] Tonnaer JA, Lammers AJ, Wieringa JH, Steinbusch HW. Immunohistochemical evidence for degeneration of cholinergic neurons in the forebrain of the rat following injection of AF64A-picrylsulfonate into the dorsal hippocampus. *Brain Res* 1986;370(1):200–3.
- [12] Lemaire R, Tabet JC, Ducoroy P, Hendra JB, Salzet M, Fournier I. Solid ionic matrixes for direct tissue analysis and MALDI imaging. *Anal Chem* 2006;78(3):809–19.
- [13] Stauber J, Lemaire R, Franck J, Bonnel D, Croix D, Day R, et al. MALDI imaging of formalin-fixed paraffin-embedded tissues: application to model animals of Parkinson disease for biomarker hunting. *J Proteome Res* 2008;7(3):969–78.
- [14] Lemaire R, Desmons A, Tabet JC, Day R, Salzet M, Fournier I. Direct analysis and MALDI imaging of formalin-fixed, paraffin-embedded tissue sections. *J Proteome Res* 2007;6(4):1295–305.
- [15] Basso M, Giraud S, Corpillo D, Bergamasco B, Lopiano L, Fasano M. Proteome analysis of human substantia nigra in Parkinson's disease. *Proteomics* 2004;4(12):3943–52.
- [16] Pierson J, Norris JL, Aerni HR, Svenningsson P, Caprioli RM, Andren PE. Molecular profiling of experimental Parkinson's disease: direct analysis of peptides and proteins on brain tissue sections by MALDI mass spectrometry. *J Proteome Res* 2004;3(2):289–95.
- [17] Pierson J, Svenningsson P, Caprioli RM, Andren PE. Increased levels of ubiquitin in the 6-OHDA-lesioned striatum of rats. *J Proteome Res* 2005;4(2):223–6.

**Annexe III: MALDI imaging of formalin-fixed paraffin-embedded tissues: application to model animals of Parkinson disease for biomarker hunting.**

## MALDI Imaging of Formalin-Fixed Paraffin-Embedded Tissues: Application to Model Animals of Parkinson Disease for Biomarker Hunting

J. Stauber,<sup>§,†</sup> R. Lemaire,<sup>§,†</sup> J. Franck,<sup>†</sup> D. Bonnel,<sup>†</sup> D. Croix,<sup>†</sup> R. Day,<sup>‡</sup> M. Wisztorski,<sup>†</sup>  
 I. Fournier,<sup>\*,†</sup> and M. Salzet<sup>\*,†</sup>

Université des Sciences et Technologies de Lille, FRE-CNRS 2933, MALDI Imaging Team, Laboratoire de Neuroimmunologie des Annelides, IFR 147, Cité Scientifique, 59655 Villeneuve d'Ascq Cedex, France, and Université de Sherbrooke, Département de Pharmacologie, Faculté de médecine, Sherbrooke, Québec, J1H 5N4, Canada

Received July 24, 2007

A common technique for the long-term storage of tissues in hospitals and clinical laboratories is preservation in formalin-fixed paraffin-embedded (FFPE) blocks. Such tissues stored for more than five years have not been useful for proteomic studies focused on biomarker discovery. Recently, MS-based proteomic analyses of FFPE showed positive results on blocks stored for less than 2 days. However, most samples are stored for more than one year, and thus our objective was to establish a novel strategy using as a model system 6-hydroxydopamine (6-OHDA) treated rat brain tissues stored in FFPE blocks for more than 9 years. We examined MALDI tissue profiling combining the use of automatic spotting of the MALDI matrix with in situ tissue enzymatic digestion. On adjacent sections, the identification of compounds is carried out by tissue digestion followed by nanoLC/MS-MS analysis. The combination of these approaches provides MALDI direct analysis, MALDI/MS imaging, as well as the localization of a large number of proteins. This method is validated since the analyses confirmed that ubiquitin, *trans*-elongation factor 1, hexokinase, and the Neurofilament M are down-regulated as previously shown in human or Parkinson animal models. In contrast, peroxiredoxin 6, F1 ATPase, and  $\alpha$ -enolase are up-regulated. In addition, we uncovered three novel putative biomarkers, the *trans*-elongation factor 1 (eEF1) and the collapsin response mediator 1 and 2 from protein libraries. Finally, we validate the CRMP-2 protein using immunocytochemistry and MALDI imaging based on the different ions from tryptic digestion of the protein. The access to archived FFPE tissue using MALDI profiling and imaging opens a whole new area in clinical studies and biomarker discovery from hospital biopsy libraries.

**Keywords:** MALDI direct analysis • MALDI imaging • FFPE tissues • Parkinson • biomarkers

### Introduction

Parkinson's disease (PD) is the second most common neurological disease after Alzheimer's disease. PD is characterized by a selective degeneration of dopaminergic neurons in the *substantia nigra pars compacta* and by cytoplasmic inclusions (Lewy Bodies) where specific proteins are stored like the  $\alpha$ -synuclein.<sup>1</sup> The clinical symptoms are severe motor dysfunctions, including rigidity, postural imbalance, slowness of movements, and uncontrollable tremor. Mutations in genes encoding  $\alpha$ -synuclein, parkin, and ubiquitin carboxy-terminal hydrolase L1 (UCH L1) have been identified in sporadic familial forms of PD.<sup>2</sup> Several proteomic studies using 2D gel analyses in animal models and humans have identified biomarkers implicated in this pathology<sup>3–7</sup> (Table 1). In Parkin knockdown models, pyruvate dehydrogenase, NADH ubiquinone oxyreductase 30

kDa, cytochrome c oxidase, peroxiredoxin 1, 2, and 6, lactoylglutathione lyase, vacuolar protein sorting 29, crystalline chain b, and heterogeneous nuclear ribonucleoprotein 1 are all down regulated.<sup>4</sup> Similarly, in 6-OHDA animal models,<sup>8</sup> as well as in humans,  $\alpha$ -enolase,  $\beta$ -Actin, Lasp-1, and neurofilament triplet L and M are also down-regulated.<sup>3–6</sup> In contrast, human peroxiredoxin 2, complexin I, fatty acid binding protein, L type calcium channel d subunit, mitochondrial complex III, and ATP synthase D chain are up-regulated.<sup>5</sup> These results have been confirmed by Strey et al.<sup>7</sup> in SOD1 gene studies, where  $\alpha$ -enolase as well as HSP25, HSP27, phosphatidylinositol transfer protein, apolipoprotein E, and ferritin heavy chain are up-regulated. These contradictory studies for peroxiredoxin 2 or profilin show the difficulty to get real biomarkers from these classical techniques. Molecular profiling of proteins and peptides can be performed directly on or near the surface of brain tissue sections with high specificity and sensitivity by utilizing matrix-assisted laser desorption/ionization mass spectrometry (MALDI MS). This technique was first introduced by Caprioli et al.<sup>9</sup> and adapted in other studies.<sup>10</sup> The first tissue profiling

\* Corresponding authors. E-mail: isabelle.fournier@univ-lille1.fr; michel.salzet@univ-lille1.fr.

§ Equal contribution.

† Université des Sciences et Technologies de Lille.

‡ Université de Sherbrooke.

**Table 1.** Comparison of Parkinson Biomarkers Identified Using Different Proteomics Approaches on Different Models

model	proteomic approach	protein	modification	references
Parkin -/-	2D Gel	Pyruvate Deshydrogenase	Down-regulated	4
		NADH Ubiquinone Oxyreductase 24 kDa	Up-regulated	
		NADH Ubiquinone Oxyreductase 30 kDa	Down-regulated	
		Cytochrome c Oxydase	Down-regulated	
		Peroxiredoxin 1	Down-regulated	
		Peroxiredoxin 2	Down-regulated	
		Peroxiredoxin 6	Down-regulated	
		Lactoylglutathione Lyase	Down-regulated	
		Profilin	Down-regulated	
		Vacuolor Protein Sorting 29	Down-regulated	
		$\alpha$ -Crystallin Chain b	Down-regulated	
		Heterogeneous Nuclear Ribonucleoprotein 1	Down-regulated	
		Lasp-1	Down-regulated	
		$\alpha$ -enolase	Up-regulated	
		$\beta$ Actin	Down-regulated	
6-OHDA	MDA <sup>a</sup>	Calmodulin	Down-regulated	11
		Cytochrome C	Down-regulated	
		Cytochrome C oxidase	Down-regulated	
		Ubiquitin	Up-regulated	
Human	Blood	Serum Creatine Kinase	Up-regulated	31
		Neurofilament Triplet L	Down-regulated	
		Neurofilament Triplet M	Down-regulated	
		Peroxiredoxin 2	Up-regulated	
		Mitochondrial Complex III	Up-regulated	
	2D Gel	ATP Synthase D chain	Up-regulated	
		Profilin	Up-regulated	
		L type Calcium Chanel d Subunit	Up-regulated	
		Fattu Acid Binding Protein	Up-regulated	
		PEP 19	Up-regulated	
MTP G93 A (gene SOD1)	MDA 2D Gel	$\alpha$ -Enolase	Down-regulated	32
		HSP 25	Up-regulated	
		HSP 27	Up-regulated	7
		Phosphatidylinositol Transger Protein a	Up-regulated	
		Apolipoprotein E	Up-regulated	
		Ferritin Heavy Chain	Up-regulated	

<sup>a</sup> MDA: MALDI Direct Analysis MS.

studies on the 6-OHDA Parkinson model have been performed by the Per Andr n group.<sup>11</sup> From their studies, calmodulin, cytochrome c, and cytochrome c oxidase appear to be down-regulated with the exception of ubiquitin.<sup>11</sup>

In this study, we looked into Parkinson's biomarkers using 6-OHDA injected rats as model animals. Because brains available for this study had been stored after formalin fixation and paraffin embedding (FFPE) over 9 years, a more direct strategy was to perform in situ tissue enzymatic digestion prior to analysis. Indeed, formalin fixation induces protein cross-linking inside the tissue by formation of methylene bridges. If such samples are extremely stable over time and allow for high preservation of tissue integrity at the ultracellular level, they raise analytical difficulties by the formation of a global protein network. Thus, enzymatic digestion gives the opportunity for retrieving small protein pieces that are possible to be analyzed by conventional technologies. Protein identification and localization were obtained using two different and complementary strategies: MALDI MS direct profiling of tissues and imaging with nanoLC-nanoESI MS. The first strategy allows for quickly observing variation in protein abundances looking to their digestion peptides by comparison of model animal brains with control brains, whereas the second strategy allows high-throughput identification of proteins including proteins of interest. Such a combined approach allowed us to confirm that hexokinase and neurofilament M proteins are down-regulated, whereas by contrast, collapsin response mediator proteins 1 and 2 (CRMP1 and CRMP2), peroxidoredoxine 6, F1 ATPase,

ubiquitin, and  $\alpha$ -enolase are up-regulated. Moreover, we also uncovered that the *trans*-elongation factor 1 (eEF1) is down-regulated.

### Experimental

**Material.**  $\alpha$ -Cyano-4-hydroxycinnamic acid (HCCA), 2,5-dihydroxybenzoic acid (2,5-DHB), ammonium bicarbonate, trisma base, toluene, ethanol, Angiotensin II, Des-Arg-Bradykinin, substance P, ACTH 18-39, ACTH 7-38, and bovine insulin were obtained from Sigma-Aldrich and used as provided. Trypsin was from Promega. Trifluoroacetic acid (TFA) was purchased from Applied Biosystems. Acetonitrile p.a. and methanol p.a. were from J.T. Baker.

**Animal Surgery and Treatment.** Wistar male rats (180–200 g) were anesthetized with pentobarbital (50 mg/kg, i.p.) and infused over 8 min with 6-OHDA (8 mg in 4 mL of 0.05% ascorbic acid in saline) at coordinates A 5 23.8 mm, L 5 1.5 mm, and H 5 28.5 mm.<sup>12</sup>

**Tissue Fixation.** Rats were sacrificed 4 h after the last drug injection. Their brains were removed and fixed for 40 min at 48 °C in 4% paraformaldehyde. The accuracy of the 6-OHDA lesion was checked by immunocytochemistry with an antityrosine hydroxylase antibody according to a previously published procedure.<sup>12</sup> These brain tissue blocks were then stored in a box at room temperature 9 years before use in this experimental plan procedure.

**Tissue Dewaxing.** Tissue sections of 10  $\mu$ m were obtained using a microtome (E. Leitz, Westlard, Germany) and applied

**MALDI Imaging of FFPE Tissues**

onto ITO (Indium Teen Oxide) one-side coated, conductive glass slides. Paraffin was removed by 2 baths of 5 min of toluene and lightly rehydrated with graded ethanol (100°, 96°, 70°, and 30°) baths before drying at room temperature.

**In Situ Trypsin Digestion. For MALDI direct analysis.** 2  $\mu\text{L}$  of enzyme (trypsin 0.033  $\mu\text{g}/\mu\text{L}$  in 25 mM Tris buffer, pH 7.4) was added at different spots on the tissue to obtain representative protein/peptide profiles. The digest is performed at room temperature, and the tissue is covered with a cap to decrease liquid evaporation. Each 10 min, enzyme solution is added on the same spots. Digestion is stopped by rinsing the tissue with 80% cold ethanol ( $-20^\circ\text{C}$ ) to remove salts. The sections are allowed to dry at room temperature. 30  $\mu\text{L}$  of matrix solution (HCCA, 20 mg in ACN:H<sub>2</sub>O, TFA 0.1% (7:3, v/v)) is then applied on the tissue using a micropipette to cover the whole tissue section and dried at room temperature.

**For MALDI Imaging.** Trypsin microspotting: spots of enzyme (trypsin at 0.05  $\mu\text{g}/\mu\text{L}$  in water) were performed using a high accurate position automatic Chemical Inkjet Printer (CHIP-1000, Shimadzu Biotech, Kyoto, Japan). Thus, the whole tissue section was microspotted with enzyme following a regular raster of spots of  $\sim 200\ \mu\text{m}$  size. The quantity of trypsin was, after optimization, set to 20 nL by 40 cycles of 500 pL on each spot position to cover the surface of the tissue. 40 nL of ionic matrix was then spotted on the same position as that for the enzyme using 40 cycles of 500 pL. The matrix preparation was described previously. Briefly, just prior to use, 4.8  $\mu\text{L}$  of Aniline (1 equiv) was added to the matrix solution (HCCA, 10 mg/mL in ACN/TFA 0.1%, 6:4, v/v) and sonicated during 2 min.

**For nanoLC-ESI MS Analysis.** On a section of 2  $\text{cm}^2$ , in situ enzymatic digestion is performed by adding 15  $\mu\text{L}$  of trypsin enzyme (0.033  $\mu\text{g}/\mu\text{L}$  in 25 mM Tris buffer, pH 7.4) for 1 h at room temperature. After enzymatic digestion, purification of the resulting digestion peptides was achieved by using reverse phase C<sub>8</sub> coated silica magnetic beads (ClinProts, Bruker Daltonics, Germany) according to the protocol of the manufacturer modified for tissues. For this, 15  $\mu\text{L}$  of binding solution was directly applied onto the tissue during 1 min, then 15  $\mu\text{L}$  of magnetic bead was added on the section. Extraction occurred during 10 min. During this step, beads and digested products were mixed 3 times using a micropipette directly onto the tissue. The digestion solution and beads were, then, deposited in a polypropylene tube and washed 3 times using 500  $\mu\text{L}$  of H<sub>2</sub>O/TFA 0.1%. Peptides were eluted from the beads with 30  $\mu\text{L}$  of ACN/H<sub>2</sub>O (1:1, v/v), and solution was dried by vacuum centrifugation. For nanoLC-MS/MS identification, peptides were redissolved in H<sub>2</sub>O/MEOH 0.1% formic acid (9:1 v/v) after elution and evaporation.

**Mass Spectrometry. MALDI MS Direct Analysis of Tissue Sections.** Spectra were acquired on a Voyager-DE STR mass spectrometer (Applied Biosystems, Framingham, MA, USA) with delayed extraction (DE) and a 337 nm pulsed nitrogen laser with a repetition of 3 Hz. External calibration was performed using a solution of standard peptides (bradykinin 1.6  $\mu\text{M}$ , substance P 1.6  $\mu\text{M}$ , ACTH 18-39 1.6  $\mu\text{M}$ , ACTH 7-38 3.2  $\mu\text{M}$ , bovine insulin 4.8  $\mu\text{M}$ , and bovine ubiquitin 4.8  $\mu\text{M}$  in H<sub>2</sub>O). Slices were visualized in the mass spectrometer using a color CCD camera (SONY). Each spectrum is the result of the average of 200 laser shots of the area of interest.

**MALDI MS Direct Analysis of Tissue Sections after in Situ Automatic Trypsin Digestion.** Trypsin fingerprint experiments of FFPE tissue sections after in situ digestion of the whole tissue section were performed on an Ultraflex II TOF-TOF instrument (Bruker Daltonics, Bremen, DE) equipped with

a LIFT III cell and Smartbeam laser with a repetition rate up to 200 Hz. For MS/MS experiments, parameters were set as follows: laser repetition rate was 100 Hz with 33% attenuation; ion source voltages were, respectively, 8 and 7.3 kV on the MALDI sample plate and first electrode; LIFT cell was pulsed from ground for electrodes 1 and 2 to 19 kV, and in the last step electrode 3 was decreased to 3.2 kV; reflector end voltage was set to 29.5 kV and midgrid to 13.85 kV. Trypsin fingerprint protein identifications in databanks were performed using the Biotoools 3.0 software (Bruker Daltonics, Bremen, DE) connected to the Mascot search engine and interrogating the NCBI, Swissprot, and ESTs databanks with oxydation as variable modification, monoisotopic as mass values, a peptide mass tolerance  $\pm 0.1$  Da with a peptide charge state of 1+, a maximum of missed cleavages of 2, and a number of queries of 300.

**MALDI MSI.** For MALDI-MSI of fixed and paraffin embedded tissues stored for 9 years, images were performed on an Ultraflex II TOF-TOF (Bruker Daltonics, Bremen, DE). After dewaxing, images were obtained in positive reflector mode. After trypsin digestion and matrix deposition, 20 nL of the mix was applied onto the tissue using microspotting as explained above and dried at room temperature. Acquisition was realized using a Smartbeam laser, with a repetition rate of 100 Hz and a spatial resolution of  $100 \times 100\ \mu\text{m}$ . For image reconstruction, the FlexImaging v2.0 software (Bruker Daltonics, Bremen, DE) was used. For the positive mode, 12 919 spots covering the whole slice with 500 laser shots per position were scanned. From each position, the software measures an average mass spectrum with its coordinates on the slice.

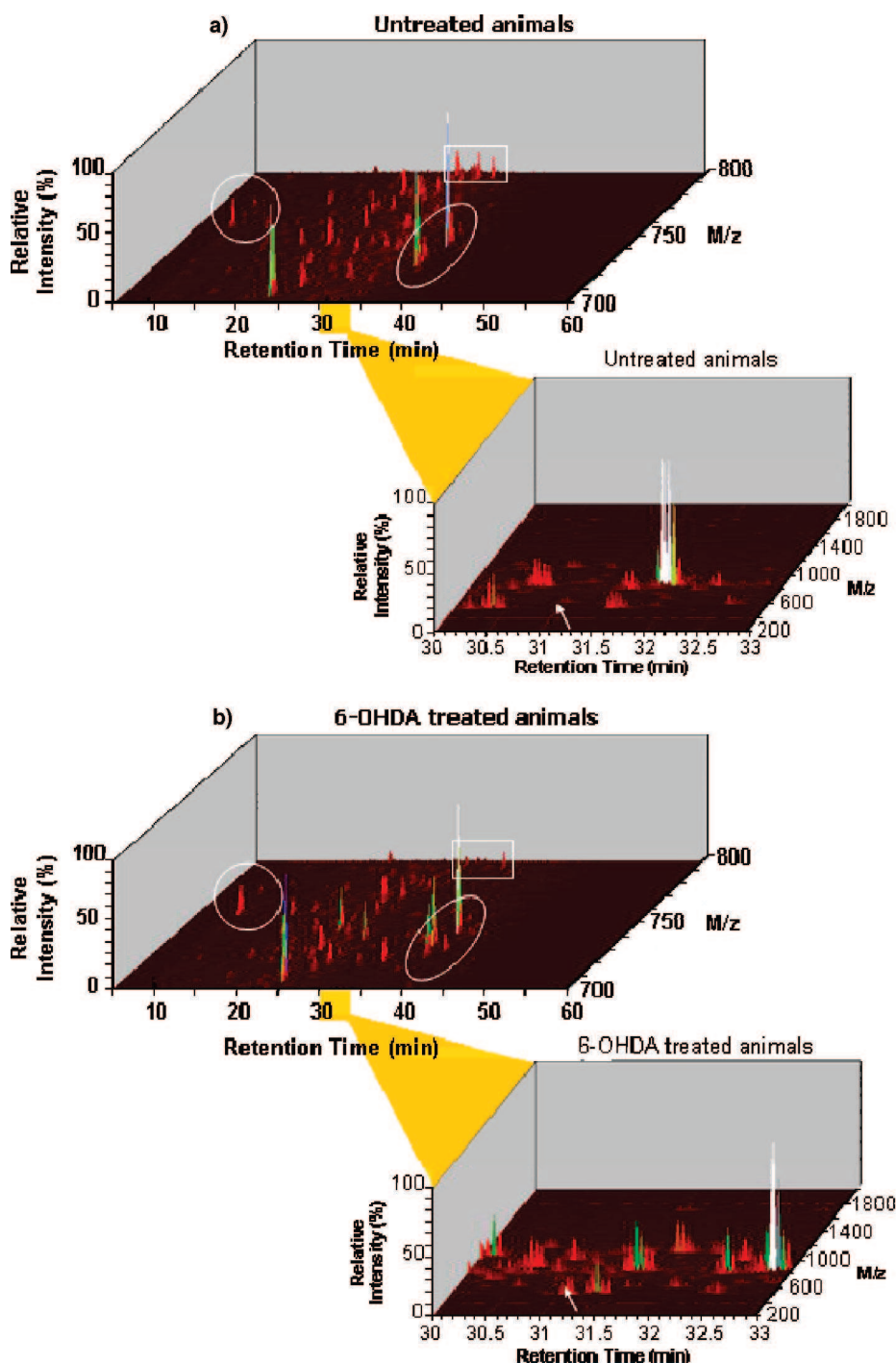
**nanoLC-nanoESI-IT MS and MS/MS.** Analyses were performed on an ion trap mass spectrometer (LCQ deca XP plus, Thermo electron, Manchester, UK) equipped with a nanoESI ion source and online coupled to a nanoHPLC system. 0.5  $\mu\text{L}$  of digest was injected with a Switchos Autosampler (Dionex corporation), and separation was performed on a C18 silica bonded stationary phase (75  $\mu\text{m}$  i.d., 150 mm long, 3  $\mu\text{m}$  100 Å pore size, Dionex). Samples were washed for 2 min at 10  $\mu\text{L}/\text{min}$  with 100% mobile phase A (95% H<sub>2</sub>O, 5% ACN, 0.1% formic acid). Peptides were then eluted using a linear gradient of 1%/min mobile phase B (ACN 80%, H<sub>2</sub>O 20%, formic acid 0.08%) for 70 min at a flow rate of 0.2  $\mu\text{L}/\text{min}$ . The LCQ was operated in a data-dependent MS/MS mode in which one MS full scan was followed by one MS/MS scan on the most abundant peptide ion. Collision energy was set to 35%. The heated capillary temperature and electrospray voltage were 160 °C and 1.5 kV, respectively.

Protein identification was performed under the MASCOT sequence query search program using the SwissProt database filtered for the taxonomy "*Rattus norvegicus*". A tolerance of 2 Da for peptide and 0.8 Da for MS/MS was set. Only protein sequences with a MOWSE score higher than 36 (indicating significant homology or identity) and identified in several samples representing at least two significant MS/MS were considered. Methionine oxidation and acetylation of N-terminals were defined as variable modification.

**Results and Discussion**

Brain tissue blocks of five 6-OHDA treated versus five untreated rats were systematically studied and compared for biomarkers hunting. The FFPE brains had been stored for 9 years before use in the present report. Formalin fixation provokes the formation of protein–nucleic acid and protein–





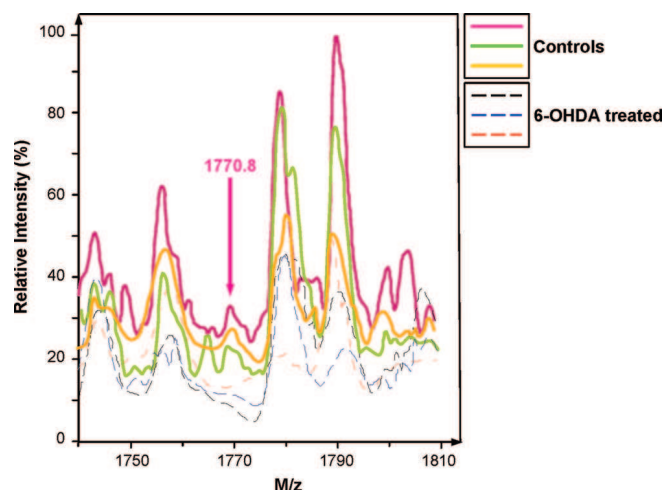
**Figure 1.** Comparison of 3D intensity vs *m/z* vs retention time plots reconstructed from the data collected on the nanoLC-nanoESI-IT MS experiments after in situ trypsin digestion and extraction of peptides for (a) control animals and (b) 6-OHDA treated model animals. The first 3D view corresponds to retention time of 0–60 min vs *m/z* 700–800, whereas the second view corresponds to a retention time of 30–33 min vs *m/z* 200–1800.

protein cross-linking in the intracellular environment arising from the reactivity of formaldehyde with the side-chains of lysyl, arginyl, tyrosyl, aspartyl, histidyl, and seryl residues.<sup>13</sup> The methylene bonds are difficult to break without destroying the peptide backbone. To solve this problem, we used endopeptidases directly on tissue sections to break down the protein network and retrieve digestion peptides that can also be used for protein identification.

In a first strategy, digestion of the whole tissue section followed by extraction of the resulting solution from the tissue,

purification, separation, and analysis on a nanoLC-nanoESI-IT system is used. By plotting intensity vs retention time vs *m/z*, differences between our two physiological states can easily be observed. Comparison of nanoLC profiles from Parkinson animal models to controls reflects the presence of numerous proteins, among which most are common, and a few of them are specific to each condition, exemplified by the plots presented Figure 1. For each peak observed on the TIC chromatogram, an MS is performed followed by MS/MS on the most abundant ions. More than 100 proteins can be identified for

## MALDI Imaging of FFPE Tissues



**Figure 2.** Superposition of six mass spectra obtained by MALDI direct analysis in the same region of FFPE tissue sections after in situ trypsin digestion for three 6-OHDA treated animals vs three control animals. Arrows indicate an ion at  $m/z$  1770.3 identified as a digestion fragment of the neurofilament triplet M protein which is a potential biomarker.

each type of sample using this method. Among all the identified proteins, some of the proteins are known to be present at low concentration. Interestingly, such FFPE tissues seem to bring an easier access to lowly abundant proteins than samples stored frozen without treatment, probably because the protein cross-linking network renders some proteins more accessible than they are in the normal tissue.

In a second step, we compared the results of the classical proteomic “Bottom-up” approach with the direct analysis by MALDI. MALDI direct analysis of the tissue after digestion allows for identifying several proteins by peptide mass fingerprint and confirmation of the modulation of the previously identified proteins. As given for example in Figure 2, the neurofilament triplet M protein is clearly observed in the control samples, whereas this protein is almost not observable in the 6-OHDA treated samples. Table 2 gives examples of proteins that were identified in the MALDI direct analysis approach. Considering the subcellular localization of these compounds, we can confirm that MALDI direct analysis after in situ digestion allows the detection of proteins from different compartments such as cytosol, the nuclear envelope, and the cytoskeleton and not just from the cell surface. Various functional proteins were detected including enzymes or signal transduction molecules or regulators. The technique allows the detection of very high mass compounds directly from the tissue like vimentin, neuromodulin, and neurogenesis.

Considering the data obtained by MALDI and ESI, in the case of the animal treated with 6-OHDA, the results obtained (Table 3) are in agreement with (Table 4) those obtained from 6-OHDA rat brain frozen tissue section by Pierson et al.<sup>11</sup> (Table 1). 6-OHDA is known to inhibit the mitochondrial transport chain, and this, along with the resulting production of reactive oxygen species, contributes to neuronal death.<sup>14</sup> This agent appears to induce neuronal death by activating transcription as well as DNA repair enzymes.<sup>15</sup> When activated, the signal transcription pathways initiated by PERK and IRE1a induce a characteristic set of genes encoding ER chaperones and nuclear transcription factors that lead to reduction of ER stress or to death.<sup>16</sup>

Based on these data, comparison of the proteins present or absent in both conditions revealed the presence of specific

proteins which have been identified in the Parkinson model, such as hexokinase (2-7-1-1), neurofilament M protein, peroxidoxin 6, F1 ATPase, and  $\alpha$ -enolase (Tables 1 and 3).

The neurofilament M protein has been described as down-regulated,<sup>17,5</sup> and its down regulation is correlated with decreased mRNA during the disease and is dependent upon disease severity.<sup>17</sup> Our results using direct analysis of FFPE tissues support these findings.

Due to 6-OHDA treatment provoking oxidative stress, several molecules have been found to be up-regulated under these conditions, such as peroxydioxin 6, known as an antioxidative protein. Similar results were found Strey et al.<sup>7</sup> using classical proteomic studies. In contrast, peroxydioxin 2 is described as being down-regulated in the parkin knockout model but is up-regulated in human Parkinson’s disease studies. These data reaffirm the need to consider multiplex biomarkers for pathology.  $\alpha$ -Enolase was also increased in our study as previously observed elsewhere in the case of Parkinson’s or Alzheimer’s disease.<sup>3,7</sup> This molecule is known to be the target of specific oxidation or nitroization.<sup>3</sup> We also observed up-regulation of F1 ATPase as previously demonstrated by Seo et al.<sup>18</sup>

The ubiquitin complex was observed to be up-regulated in our study on FFPE tissue as shown by Pierson et al.<sup>11</sup> It is known that high levels of ubiquitin and ubiquitinated proteins are present in Lewy bodies indicating that protein degradation is impaired in PD.<sup>2</sup> Proteins conjugated with a chain of ubiquitin moieties are targeted to the ubiquitin–proteasome system complex, where they undergo proteolytic degradation. Genetic studies of PD have identified mutations in the genes coding for proteins involved in the ubiquitin–proteasome degradation pathway.<sup>19,20</sup>

Interestingly, results obtained from 6-OHDA treated animals after trypsin digestion and nano-LC-nanoESI-IT MS and MS/MS experiments revealed a novel biomarker, the collapsin response mediator protein 2. Figure 3 presents the results of identification for CRMP2 proteins. Several digestion fragments of this protein are found in nanoESI (Figure 3a), and databank interrogation using the MS/MS experiments confirms this identification with attribution of the MS/MS data (as exemplified for ion at  $m/z$  508.59) that is clearly consistent with ion fragment series expected for such instruments (Figure 3b). MALDI direct analysis experiments give good agreement with the previous identification even if digestion fragments observed in MALDI are different from those observed in nanoESI execution of one peptide as highlighted in the CRMP2 sequence (Figure 3c). MS/MS performed directly on the tissue section in MALDI confirmed that the observed fragments would be consistent with digestion peptides of the CRMP2 protein. Globally, all the data combined together give identification to this up-regulated protein as the CRMP2 splice variant B of the CRMP2 protein (*Rattus norvegicus*, P47492).

Two splice variants have been recently found.<sup>21</sup> Indeed, the CRMP2A is the long N-terminal isoform (75 kDa) and induces oriented microtubule patterns in cultured fibroblasts, a pattern also observed in axons. Conversely, CRMP2B, the shortest variant (64 kDa), induces disorientation of microtubule patterns in cultured fibroblasts and reduces axon length when overexpressed in retinal explants.<sup>21</sup> In an adult brain, it is known that expression of CRMPs is dramatically down-regulated.<sup>22</sup> In our study, an increase in of CRMP2B in 6-OHDA treated animals has been observed which is in agreement with previous molecular data.<sup>23</sup> To go deeper inside the understanding of the role CRMP2, we have consecutively performed MALDI imaging of CRMP2 in the

**Table 2.** Examples of Proteins Detected by MALDI Direct Analysis MS from the FFPE Rat Brain Tissues after in Situ Enzymatic Digestion, with the Corresponding Score, Sequence Coverage, and Matched Peptides in the Protein Sequence

protein	match score	mass (Da)	datbank	sequence coverage (%)	peptides (bold) matched in the protein sequence
Chordin-Like Protein 1 Precursor (Neuralin-1): Neurogenesis-1	39	50526	Swissprot	48	MEGIKYIASL VFFFVLEAS <b>KTEPVKHSET</b> <b>YCMFQDKKYR</b> VGEKWHPLYE PYGLVVCVNC ICSNGNVLC SRVRCPTLHC <b>LSPVHIPHLC</b> <b>CPRCPDSLPP</b> MNSKVTSKSC EYNGTTYQH ELFIAEGLFQ NRQPNQCSQC <b>SCSEGNVYCG</b> <b>LKTCPKLTCA</b> FVPVSPDSCC RVCRRGDGELS WEHSDADIFR QPANREARHS <b>YLRSPYDPPP</b> <b>SRQAGGLPRF</b> AGSRSHRGAV <b>IDSQQASGTI</b> <b>VQIVINNKHK</b> HGQVCVSN GK <b>TYSHGESWHP</b> <b>NLRAFGIVEC</b> VLCTCNVTKQ ECKKIHCNPR YPCKYPQKLD GKCKKVCPEE PPSQNFDSKG <b>SFCGEETMPV</b> YEAVLVEDGE <b>TARKVALETE</b> KPPQVEVHVW TIRKILQHF HIEKISKEMF GGLHHFKLVT <b>RTTMNQWKIF</b> <b>AEGEAQLSQM</b> <b>CSSRVCRTTEL</b> EDLVQVLYLE RPEKDHCR MLCCMRRRTKQ VEKNDQDKI EQDGVKPEDK <b>AHKAATKIQA</b> SFRGHITRKK <b>LKDEKKGDAP</b> <b>AAEAIAKEKD</b> <b>DAPVADGVEK</b> <b>KEGDGSATTD</b> <b>AAPATSPKAE</b> <b>EPSKAGDAPS</b> EEKKGEGDAA PSEKAGSAE TESAATTD <b>NSPSSKAEDG</b> PAKEPKQAD VPAAVTDAAA TTPAAEDAAK AAQPPTETAE SSQAEEKEA VDEAKPKESA RQDEGKEDPE ADQEHA
Neuromodulin (Axonal Membrane Protein GAP-43)	24	23589	Swissprot	47	<b>MQPASDSRWR</b> VTGLQSEPT <b>QTLAHSGLS</b> <b>EGWRGRGKAW</b> <b>PSQGVSPTPV</b> PERKRGLRRA KMAAAAAAAAA <b>ATEQQGNSNP</b> VKKSMREKAV ERRNVNKEHN SNFKAGYPI DEDRLHKTGL RGRKGNLAIC VIVLLFILAV INLLITLVIW <b>AVIRIGPNGC</b> <b>DSLEFHESGL</b> <b>LRFKQVSDMG</b> <b>IHPLYKSTV</b> GGRRNENLVI TGNNQPIVFQ QGTTKLSVEK NKTSITSDIG MQFFDPRTON ILFSTDYETH EFHLPVGVKS LNVQKASTER <b>ITSNATSDLN</b> <b>IKVDGRAIVR</b> <b>GNEGVFIMGK</b> TIEFHMRGDV ELKAENSIL NGTVMVSPTR LPSSSGGDQS GSGDWVRYKL CMCADGTLFK VQVTSNMGCG QVSDNPGCNT H STRSVSSSSY <b>RRMFGSGS</b> <b>SRPSSNRSYV</b> TTSTRYSLG SALRPSTSR LYSSSPGGAY VTRSSAVRLR SSMPPGVRLLQ <b>DSVDFSLADA</b> <b>INTEFKNTRT</b> NEKVELQELN DRFANYIDKV RFLEQQNKIL LAELEQLKGQ <b>GKSRLGDLYE</b> <b>EEMRELRRQV</b> <b>DQLTNDKARV</b> EVERDNLAED <b>IMRLREKLQE</b> <b>EMLQREEAES</b> <b>TLQSFQRDQV</b> <b>NASLARLDLE</b> RKVESLQEEI AFLKKLHDEE IQELQAQIQE QHVQIDVDVS KPDLLAALRD <b>VRQQYESVAA</b> <b>KNLQEAEEWY</b> <b>KSKFADLSEA</b> ANRNDALRQ AKQESNEYRR QVQSLTCEVD ALKGTNESLE <b>RQMRMEENF</b> <b>ALEAANYQDT</b> <b>IGRLQDEIQN</b> MKEEMARHLR EYQDLLNVKM <b>ALDIEIATYR</b> <b>KLEGEESRI</b> <b>SLPLPNFSSL</b> NLRETNLESL <b>PLVDTHSKRT</b> LLIKTIVETRD GQVINETSQH HDDLE <b>RSPSRQSW</b> _M RMPCHRSKA <b>TTLCAWCSPA</b> <b>ATRTWTLCEL</b> <b>CGPQEISRHR</b> <b>WTLQWPSMTC</b> <b>LWSWTF</b> _ILS <b>TRKPPCGSWT</b> <b>CAPLCCPKLR</b> <b>NFCRASM RAM</b> SRLGVPPSS_ <b>SCSGSYPSL</b> <b>TFWQPRPLWV</b> <b>WTSAERRGCI</b> <b>SVDSASSSR</b> <b>ASVASSRAGR</b> _V_AA
Dystrophin-Associated Glycoprotein) (43DAG)	50	41643	NCBI	47	NGTMSPAITS CPHVKRAQW PVENPQRWSM PEPLGRRKIA MRSAPWCGTS VLRLSSAMC PPSGVSPVGT GKSLES LAQT PTPSTGYRS GA_DPHGGAA PAWPTERSFH NVPRMLSPA SHADRASHST WCLLVAFCHI RNPLKKEKKK KK <b>SAEMVDSLQT</b> <b>SPKSQWCAEP</b> <b>R</b> _PEVVFCLS HLEGCIDVVW NSCPLXWTGL PD_SHMG_SG LHGPSSCWPA GCPSSCSPS_ FARWSVS_TL <b>PPARSCQNGW</b> <b>LCHVASAGAS</b> <b>PRRTHPAPCP</b> <b>HRGSIPVRLG</b> <b>R</b> _MYTEHLSG PSEGADCGF_ TQTRRRQPSV PSQSCPSWHR SQ_RCQSDPR PRYSGPQWAQ HVRG_VSSLS SGTSAFRWG_ MVPQSQASW_ SAAVRPAVEL CSCRGVADGC RTCGWCNPPG PRS
Vimentin	54	53569	NCBI	42	
UI-R-BJ2-bqp-d-02-0-UI.s1 UI-R-BJ2 <i>Rattus norvegicus</i> cDNA Clone	58	17842	EST	83	
UI-R-DZ1-cnf-g-02-0-UI.s1 NCI_CGAP_DZ1 <i>Rattus norvegicus</i> cDNA clone	54	16280	EST	78	
AGENCOURT_28540227 NIH_MGC_249 <i>Rattus norvegicus</i> cDNA clone	52	26005	EST	57	

6-OHDA treated samples to obtain the localization of the protein. Figure 3 presents the MALDI molecular images obtained from

such samples and reconstructed using different ions corresponding to different digestion fragments of the protein. All these ions

**Table 3.** Differential Protein Found after Trypsic Digestion by Comparing 6-OHDA FFPE Tissues to Control FFPE Tissues of Rat Brains Using either MALDI-MS Direct Analysis or NanoLC-nanoESI-IT MS on Analogue Zones

Markers	protein	m/zMDA <sup>a</sup>	m/z ESI	mascot score (%)	detected on healthy zone (number)	detected on damaged zone (number)	known in Parkinson's disease literature
Down-regulated	<i>trans</i> -elongation Factor 1 (eEF1)	1025.5	1025.8	66	3/5	0/5	
	Hexokinase (2-7-1-1)	1399.9	1400.5	76	3/5	0/5	33
	Neurofilament M Protein	1770.6	1770.8	57	4/5	0/5	5
Up-regulated	Peroxidoxin 6	1499.9	1498.8	65	0/5	4/5	5
	F1 ATPase		2007.3	81	0/5	4/5	18
	$\alpha$ -Enolase	3062.9	3063.6	62	0/5	3/5	34
	Ubc-Ubiquitin	647.4		78	0/5	4/5	11
	CRMP2	1083	1083.6	70	0/5	4/5	

<sup>a</sup> MDA: MALDI Direct Analysis MS.

**Table 4.** Data for the Identification of the CRMP2 Protein (MW<sub>avg</sub> = 62270.62 u, score 245) in 6-OHDA Injected Rat Brain FFPE Tissues after Trypsin Digestion and Nano-LC-nanoESI-IT MS and MS/MS Experiments: (a) Peptide Mass Fingerprint Identification Data, (b) Peptide Sequence Tag (MS/MS) Identification Data for the m/z 508.59 Ion Assigned as Fragment SAAEVIAQAR of CRMP2, and (c) CRMP2 Protein (*Rattus norvegicus*, P47492) Sequence<sup>a</sup>

a)	Observed M/z ([M+H] <sup>+</sup> , [M+2H] <sup>2+</sup> or [M+3H] <sup>3+</sup> )	Experimental Mw monoisotopic	Calculated Mw monoisotopic	Peptide Position	Peptide
	508.590	1015.166	1014.546	259-268	SAAEVIAQAR
	543.064	1084.114	1083.629	441-451	GSPLVVISQGK
	562.215	1683.623	1681.863	452-467	IVLEDGTLHVTEGSGR
	571.065	1140.116	1139.601	472-480	KPFPDFVYK
	571.066	1710.178	1710.913	468-480	YIPRKPFPDFVYK
	648.020	1294.025	1293.686	64-75	MVIPGGIDVHTR
	675.321	674.314	674.360	566-572	ANITSLG
	834.047	2499.119	2498.160	271-293	GTVVYGEPIASLGTGDSHYWSK
	863.482	1724.950	1724.804	375-390	MDENQFVAVTSTNAAK
	896.971	1791.928	1791.827	346-361	DNFTLIPEGTNGTEER
	956.396	1910.777	1910.015	174-189	FQLTDSQIYEVLSVIR
	959.045	1916.076	1914.957	401-418	ISVGSADLVWDPDSVK

b)	Peptide Sequence SAAEVIAQAR M/z observed 508.590 ([M+H] <sup>+</sup> ) Ion Score 63	
	Matched ions M/z	Ion assignment
1	229.129	[y <sub>2</sub> -17] <sup>+</sup>
2	230.113	b <sub>3</sub> <sup>+</sup>
3	246.156	y <sub>2</sub> <sup>+</sup>
4	341.146	[b <sub>3</sub> -18] <sup>+</sup>
5	359.156	b <sub>4</sub> <sup>+</sup>
6	374.215	y <sub>3</sub> <sup>+</sup>
7	440.214	[b <sub>5</sub> -18] <sup>+</sup>
8	445.252	y <sub>4</sub> <sup>+</sup>
9	458.224	b <sub>5</sub> <sup>+</sup>
10	553.298	[b <sub>6</sub> -18] <sup>+</sup>
11	558.336	y <sub>5</sub> <sup>+</sup>
12	571.309	b <sub>6</sub> <sup>+</sup>
13	642.346	b <sub>7</sub> <sup>+</sup>
14	657.404	y <sub>6</sub> <sup>+</sup>
15	752.394	[b <sub>8</sub> -18] <sup>+</sup>
16	769.420	[y <sub>7</sub> -17] <sup>+</sup>
17	770.404	b <sub>8</sub> <sup>+</sup>
18	840.457	[y <sub>8</sub> -17] <sup>+</sup>
19	841.441	b <sub>9</sub> <sup>+</sup>
20	857.484	y <sub>8</sub> <sup>+</sup>

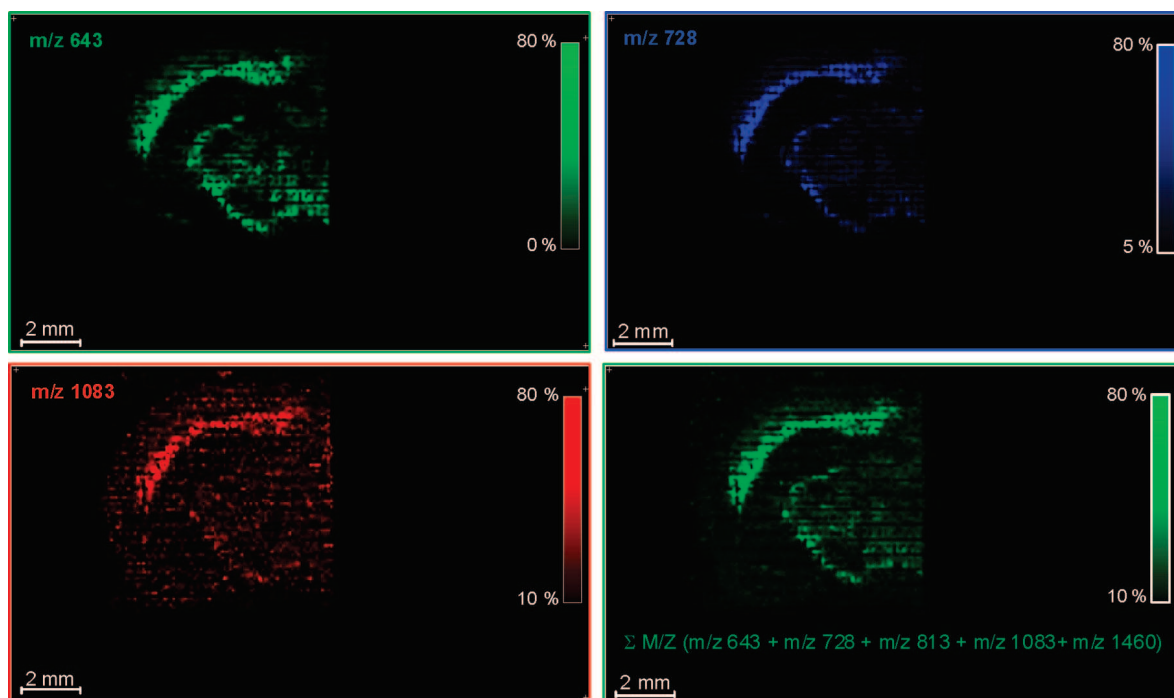
  

c)	MS/MS Sequence SAAEVIAQAR			
1	MSYQGKKNIP	RITSDRLLIK	GGKIVNDQDS	FYADIYMEDG
41	LIKQIGENLI	VPGGVKTIEA	HSRMVIPGGI	DVHTRFQMPD
81	QGMTSADDF	QGTKAALAGG	TTMIIDHVVP	EPGTSLLAAF
121	DQWREWADSK	SCCDYSLHVD	ITEWHKGIQE	EMEALVKDHG
161	VNSFLVYMAF	KDRFQLTDSQ	IYEVLSVIRD	IGAIQVHAE
201	NGDIAEEQQ	RILDLGITGP	EGHVLSRPEE	VEAEAVNRSI
241	TIANQTNCPL	YVTKVMSKSA	AEVIAQARKK	GTVVYGEPI
281	ASLGTGDSHY	WSKNWAKAAA	FVTSPPLSPD	PTTPDFLNSL
321	LSCGDLQVTG	SAHCTFNNTAQ	KAVGKDNFTL	IPEGTNGTEE
361	RMSVIWDKAV	VTGKMDENQF	VAVTSTNAAK	VFNLVPRKGR
401	ISVGSADLV	IWDPDSVKTI	SAKTHNSALE	YNI FEGMECR
441	GSPLVVISQG	KIVLEDGTLH	VTEGSGRYIP	RKFPDFVYK
481	RIKARSRLAE	LRGVPRGLYD	GPVCEVSVTP	KTVTPASSAK
521	TSPAKQAPP	VRNLHQSGFS	LSGAQIDDNI	PRRTTQRIVA
561	PPGGRANITSLG			

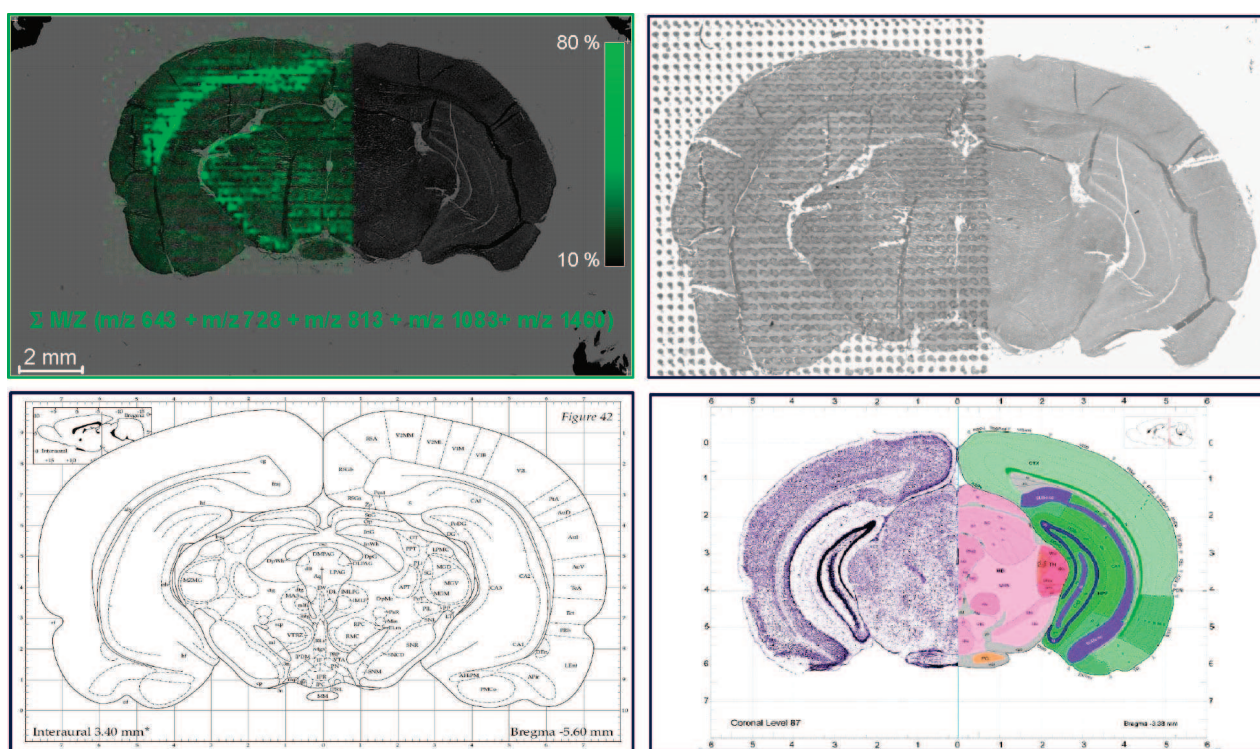
<sup>a</sup> Red indicates parts of the sequence found in the nanoLC-nanoESI-IT MS experiments. Blue indicates those obtained during MALDI direct analysis. Purple indicates the common ones.

globally present an equivalent localization in the rat brain section, even if some of them present more contrasting images. The last image is a composite image taking into account the signal of all ions corresponding to identify digestion fragments of CRMP2 for a global view. The most striking feature is the important localization of the protein in very specific regions of the brain with a highly contrasted signal. This is especially true for the *Corpus callosum* where ions are always found to be very intense. Figure

4 presents the map of the corresponding section and compares the MALDI images from the 6-OHDA treated animals to the expression of the CRMP2 mRNA in a normal adult mouse. As found in the literature for the rat, this variant of CRMP2 is normally not present in the *Corpus callosum*. It is located predominantly in dendrites of specific neuronal populations, such as cortical pyramidal neurons, hippocampal CA1 pyramidal cells, or Purkinje cerebellar cells.<sup>24</sup> Thus, localization in the *Corpus*



**Figure 3.** MALDI molecular images reconstructed from the data collected after acquisition on a 6-OHDA treated rat brain section originally conserved after formalin fixation and paraffin embedding over 9 years after in situ automatic trypsin digestion. (a) Image of ion at *m/z* 643, (b) *m/z* 728, and (c) *m/z* 1083 and (d) composite image of all ions corresponding to digestion fragments of the CRMP2 protein and detected in the MALDI experiment.



**Figure 4.** (a) Composite MALDI molecular image reconstructed using all detected digestion fragments of CRMP2 protein on a 6-OHDA treated FFPE tissue rat brain section after paraffin removal and in situ trypsin digestion. (b) Optical image of the tissue section after automatic spotting of trypsin and solid ionic matrix HCCA/ANI. (c) Map of the corresponding tissue section with a different region indicated (Bregma Index). (d) Expression of *Dpysl2* mRNA (coding for CRMP2 protein) in the mouse brain (data from Allen Brain Atlas, <http://www.brain-map.org>).

*callosum* is in line with an involvement in neurodegenerative diseases. In fact, *Corpus callosum* is known to be a brain area implicated in dementia in lot of neurodegenerative diseases.<sup>25</sup>

Thus, in PD, the CRMP (intracellular protein mediating Semaphorin3A) and the mitochondrial stress protein HSP60 mRNA encoding these proteins are up-regulated. Antibodies

directed against collapsin-1 (Sema3A, that functions in signaling growth cone collapse, chemorepulsion, and neuronal apoptosis during early development of the central nervous system) provided marked and prolonged protection of several neuronal cell types from dopamine-induced apoptosis.<sup>23</sup> It has been hypothesized that these proteins are positive mediators of DA-induced neuronal apoptosis in PD. In PD, nigral neuronal death could be due to excessive oxidative stress generated by auto and enzymatic oxidation of the endogenous neurotransmitter dopamine (DA), the formation of neuromelanin, and the presence of high concentrations of iron. DA toxicity is mediated through its oxidative metabolites, whereas thiol-containing antioxidants provided marked protection against DA toxicity and ascorbic acid accelerated DA-induced death. In Alzheimer's disease, the CRMP-2 is known to be implicated in neurite degeneration, acting on the assembly and polymerization of microtubules.<sup>26</sup> Accumulation of Sema3A overlaps the appearance of phosphorylated MAP1B and tau in many neurons, suggesting that Sema3A signaling at some level may be coupled to these previously identified cytoskeletal markers of neurodegeneration.<sup>27</sup> The hippocampus of patients with AD express phosphorylated MAP1B, collapsin-response mediator protein 2 (CRMP-2), Plexins A1 and A2, and a processed form of Sema3A.<sup>27</sup>

Thus, we speculate that CRMP factors are good biomarkers for neurodegenerative diseases like PD or AD.

While our methodology is validated by confirming previously discovered molecules implicated in PD, we have also uncovered previously undescribed changes, such as the down-regulation of *trans*-elongation factor 1 (eEF1) in 6-OHDA samples. Studies on regulatory factors of heat shock factor (HSF-1) modulation and their involvement in the age-associated attenuation have demonstrated an attenuated response to stress which is characteristic of senescence. Heat shock (HS), a significant form of stress, is delayed and reduced in aging organisms. The proteins I-HSF [HSF inhibitor], eEF1, and a large noncoding RNA (HSR) participate in regulation and activation of HSF-1 in early stages of heat shock gene transcription. It is proposed that structural changes in any one or combination of these factors in response to heat shock may contribute to the age-associated attenuation in response to stress.<sup>28</sup> These results are in agreement with the fact that 6-OHDA treatment provokes a down-regulation of the eEF1. Moreover, 6-OHDA provokes an oxidative stress stimulating HSP expression as demonstrated by Strey et al.,<sup>7</sup> and it is down-regulated in these conditions. Finally, eEF1 is present in the dendrites of neurons that exhibit synaptic plasticity, and its translation is locally regulated.<sup>29</sup> In the hippocampus, the dendritic mRNAs is highly expressed in cell bodies and is also concentrated in the zone of termination of commissural/associational afferents in the inner molecular layer, suggesting that mRNA localization is in some way related to the distribution of different types of synapses.<sup>30</sup> In the case of the 6-OHDA model, this treatment provokes a decrease of synaptic plasticity and thus a down-regulation of eEF1 levels.

In conclusion, using 9 year old FFPE tissues, we confirmed that molecules acting in oxidative stress are regulated as already demonstrated in various proteomic studies on human or fresh frozen tissues in a 6-OHDA animal model. New biomarkers were found such as eEF1 and CRMP. This confirms the high interest to use these FFPE tissues stored in hospital tissue libraries for biomarker hunting or validation in the case of clinical studies.

**Acknowledgment.** Supported by grants from Centre National de la Recherche Scientifique (CNRS), Ministère de L'Education Nationale, de L'Enseignement Supérieur et de la Recherche (ACI Jeunes Chercheurs ACI JC4074 to I. Fournier), and the Canadian Institutes of Health Research (CIHR) to R. Day and M. Salzet). Also supported by a collaboration agreement between Shimadzu and the Laboratoire de Neuro-immunologie des Annélides. The authors would also like to thank A. Page and C. Rolando of the Proteomic Platform of USTL.

## References

- (1) Beal, M. F.; Hantraye, P. Novel therapies in the search for a cure for Huntington's disease. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98* (1), 3–4.
- (2) Giasson, B. I.; Lee, V. M. Are ubiquitination pathways central to Parkinson's disease. *Cell* **2003**, *114* (1), 1–8.
- (3) De Iulius, A.; Grigoletto, J.; Recchia, A.; Giusti, P.; Arslan, P. A proteomic approach in the study of an animal model of Parkinson's disease. *Clin. Chim. Acta* **2005**, *357* (2), 202–9.
- (4) Palacino, J. J.; Sagi, D.; Goldberg, M. S.; Krauss, S.; Motz, C.; Wacker, M.; Klose, J.; Shen, J. Mitochondrial dysfunction and oxidative damage in parkin-deficient mice. *J. Biol. Chem.* **2004**, *279* (18), 18614–22.
- (5) Basso, M.; Giraudo, S.; Corpillo, D.; Bergamasco, B.; Lopiano, L.; Fasano, M. Proteome analysis of human substantia nigra in Parkinson's disease. *Proteomics* **2004**, *4* (12), 3943–52.
- (6) Basso, M.; Giraudo, S.; Lopiano, L.; Bergamasco, B.; Bosticco, E.; Cinquepalmi, A.; Fasano, M. Proteome analysis of mesencephalic tissues: evidence for Parkinson's disease. *Neurol. Sci.* **2003**, *24* (3), 155–6.
- (7) Strey, C. W.; Spellman, D.; Stieber, A.; Gonatas, J. O.; Wang, X.; Lambris, J. D.; Gonatas, N. K. Dysregulation of stathmin, a microtubule-destabilizing protein, and up-regulation of Hsp25, Hsp27, and the antioxidant peroxiredoxin 6 in a mouse model of familial amyotrophic lateral sclerosis. *Am. J. Pathol.* **2004**, *165* (5), 1701–18.
- (8) Ungerstedt, U.; Ljungberg, T.; Steg, G. Behavioral, physiological, and neurochemical changes after 6-hydroxydopamine-induced degeneration of the nigro-striatal dopamine neurons. *Adv. Neurol.* **1974**, *5*, 421–6.
- (9) Caprioli, R. M.; Farmer, T. B.; Gile, J. Molecular imaging of biological samples: localization of peptides and proteins using MALDI-TOF MS. *Anal. Chem.* **1997**, *69* (23), 4751–60.
- (10) Fournier, I.; Day, R.; Salzet, M. Direct analysis of neuropeptides by in situ MALDI-TOF mass spectrometry in the rat brain. *Neuro. Endocrinol. Lett.* **2003**, *24* (1–2), 9–14.
- (11) Pierson, J.; Norris, J. L.; Aerni, H. R.; Svenningsson, P.; Caprioli, R. M.; Andren, P. E. Molecular profiling of experimental Parkinson's disease: direct analysis of peptides and proteins on brain tissue sections by MALDI mass spectrometry. *J. Proteome Res.* **2004**, *3* (2), 289–95.
- (12) Tonnaer, J. A.; Lammers, A. J.; Wieringa, J. H.; Steinbusch, H. W. Immunohistochemical evidence for degeneration of cholinergic neurons in the forebrain of the rat following injection of AF64A-pyridylsulfonate into the dorsal hippocampus. *Brain Res.* **1986**, *370* (1), 200–3.
- (13) Shi, S. R.; Imam, S. A.; Young, L.; Cote, R. J.; Taylor, C. R. Antigen retrieval immunohistochemistry under the influence of pH using monoclonal antibodies. *J. Histochem. Cytochem.* **1995**, *43* (2), 193–201.
- (14) Grunblatt, E.; Mandel, S.; Youdim, M. B. Neuroprotective strategies in Parkinson's disease using the models of 6-hydroxydopamine and MPTP. *Ann. N.Y. Acad. Sci.* **2000**, *899*, 262–73.
- (15) Herceg, Z.; Wang, Z. Q. Functions of poly(ADP-ribose) polymerase (PARP) in DNA repair, genomic integrity and cell death. *Mutat. Res.* **2001**, *477* (1–2), 97–110.
- (16) Mori, K. Tripartite management of unfolded proteins in the endoplasmic reticulum. *Cell* **2000**, *101* (5), 451–4.
- (17) Liu, Q.; Xie, F.; Siedlak, S. L.; Nunomura, A.; Honda, K.; Moreira, P. I.; Zhua, X.; Smith, M. A.; Perry, G. Neurofilament proteins in neurodegenerative diseases. *Cell. Mol. Life Sci.* **2004**, *61* (24), 3057–75.
- (18) Seo, B. B.; Nakamaru-Ogiso, E.; Cruz, P.; Flotte, T. R.; Yagi, T.; Matsuno-Yagi, A. Functional expression of the single subunit NADH dehydrogenase in mitochondria in vivo: a potential therapy for complex I deficiencies. *Hum. Gene Ther.* **2004**, *15* (9), 887–95.

- (19) Leroy, E.; Boyer, R.; Auburger, G.; Leube, B.; Ulm, G.; Mezey, E.; Harta, G.; Brownstein, M. J.; Jonnalagada, S.; Chernova, T.; Dehejia, A.; Lavedan, C.; Gasser, T.; Steinbach, P. J.; Wilkinson, K. D.; Polymeropoulos, M. H. The ubiquitin pathway in Parkinson's disease. *Nature* **1998**, 395 (6701), 451–2.
- (20) Leroy, E.; Boyer, R.; Polymeropoulos, M. H. Intron-exon structure of ubiquitin c-terminal hydrolase-L1. *DNA Res.* **1998**, 5 (6), 397–400.
- (21) Yuasa-Kawada, J.; Suzuki, R.; Kano, F.; Ohkawara, T.; Murata, M.; Noda, M. Axonal morphogenesis controlled by antagonistic roles of two CRMP subtypes in microtubule organization. *Eur. J. Neurosci.* **2003**, 17 (11), 2329–43.
- (22) Ricard, D.; Stankoff, B.; Bagnard, D.; Aguera, M.; Rogemond, V.; Antoine, J. C.; Spassky, N.; Zalc, B.; Lubetzki, C.; Belin, M. F.; Honnorat, J. Differential expression of collapsin response mediator proteins (CRMP/ULIP) in subsets of oligodendrocytes in the postnatal rodent brain. *Mol. Cell. Neurosci.* **2000**, 16 (4), 324–37.
- (23) Barzilai, A.; Zilkha-Falb, R.; Daily, D.; Stern, N.; Offen, D.; Ziv, I.; Melamed, E.; Shirvan, A. The molecular mechanism of dopamine-induced apoptosis: identification and characterization of genes that mediate dopamine toxicity. *J. Neural. Transm. Suppl.* **2000**, (60), 59–76.
- (24) Bretin, S.; Reibel, S.; Charrier, E.; Maus-Moatti, M.; Auvergnon, N.; Thevenoux, A.; Glowinski, J.; Rogemond, V.; Premont, J.; Honnorat, J.; Gauchy, C. Differential expression of CRMP1, CRMP2A, CRMP2B, and CRMP5 in axons or dendrites of distinct neurons in the mouse brain. *J. Comp. Neurol.* **2005**, 486 (1), 1–17.
- (25) Charrier, E.; Reibel, S.; Rogemond, V.; Aguera, M.; Thomasset, N.; Honnorat, J. Collapsin response mediator proteins (CRMPs): involvement in nervous system development and adult neurodegenerative disorders. *Mol. Neurobiol.* **2003**, 28 (1), 51–64.
- (26) Gu, Y.; Hamajima, N.; Ihara, Y. Neurofibrillary tangle-associated collapsin response mediator protein-2 (CRMP-2) is highly phosphorylated on Thr-509, Ser-518, and Ser-522. *Biochemistry* **2000**, 39 (15), 4267–75.
- (27) Good, P. F.; Alapat, D.; Hsu, A.; Chu, C.; Perl, D.; Wen, X.; Burstein, D. E.; Kohtz, D. S. A role for semaphorin 3A signaling in the degeneration of hippocampal neurons during Alzheimer's disease. *J. Neurochem.* **2004**, 91 (3), 716–36.
- (28) Shamovsky, I.; Gershon, D. Novel regulatory factors of HSF-1 activation: facts and perspectives regarding their involvement in the age-associated attenuation of the heat shock response. *Mech. Ageing Dev.* **2004**, 125 (10–11), 767–75.
- (29) Huang, F.; Chotiner, J. K.; Steward, O. The mRNA for elongation factor 1alpha is localized in dendrites and translated in response to treatments that induce long-term depression. *J. Neurosci.* **2005**, 25 (31), 7199–209.
- (30) Zhong, J.; Zhang, T.; Bloch, L. M. Dendritic mRNAs encode diversified functionalities in hippocampal pyramidal neurons. *BMC Neurosci.* **2006**, 7, 17.
- (31) Takubo, H.; Shimoda-Matsubayashi, S.; Mizuno, Y., Serum creatine kinase is elevated in patients with Parkinson's disease: a case controlled study. *Parkinsonism Relat. Disord* **2003**, 9 Suppl 1, S43–6.
- (32) Skold, K.; Svensson, M.; Nilsson, A.; Zhang, X.; Nydahl, K.; Caprioli, R. M.; Svenningsson, P.; Andren, P. E. Decreased striatal levels of PEP-19 following MPTP lesion in the mouse. *J. Proteome Res.* **2006**, 5 (2), 262–9.
- (33) Pastoris, O.; Dossena, M.; Foppa, P.; Catapano, M.; Ferrari, R.; Dagani, F. Biochemical evaluations in skeletal muscles of primates with MPTP Parkinson-like syndrome. *Pharmacol. Res.* **1995**, 31 (6), 361–9.
- (34) Poon, H. F.; Frasier, M.; Shreve, N.; Calabrese, V.; Wolozin, B.; Butterfield, D. A. Mitochondrial associated metabolic proteins are selectively oxidized in A30P alpha-synuclein transgenic mice--a model of familial Parkinson's disease. *Neurobiol. Dis.* **2005**, 18 (3), 492–8.

PR070464X

**Annexe IV: Contribution of MALDI tissue profiling and MALDI-  
Imaging for the study of saponins in Cuvierian tubules of  
Holothuria forskali**



# Localization of secondary metabolites in marine invertebrates: Contribution of MALDI MSI for the study of saponins in Cuvierian tubules of *H. forskali*

Séverine Van Dyck<sup>1</sup>, Céline Meriaux<sup>2</sup>, David Bonnel<sup>2</sup>, Michel Salzet<sup>2</sup>, Isabelle Fournier<sup>2</sup>, Patrick Flammang<sup>1§</sup> and Maxence Wisztorski<sup>2§</sup>

<sup>1</sup>Université de Mons, Laboratoire de Biologie marine, 6 Avenue du Champ de Mars, B-7000 Mons, Belgium

<sup>2</sup>Université Lille Nord de France, CNRS FRE 3249, Laboratoire de Neuroimmunologie et Neurochimie Evolutives, MALDI Imaging Team, Université de Lille 1, 59655 Villeneuve d'Ascq, France.

§Co-corresponding authors

Email addresses:

SVD: [severinevandyck@hotmail.com](mailto:severinevandyck@hotmail.com)

CM: [celine1.meriaux@etudiant.univ-lille1.fr](mailto:celine1.meriaux@etudiant.univ-lille1.fr)

DB: [D.Bonnel@etudiant.univ-lille1.fr](mailto:D.Bonnel@etudiant.univ-lille1.fr)

MS: [michel.salzet@univ-lille1.fr](mailto:michel.salzet@univ-lille1.fr)

IF: [isabelle.fournier@univ-lille1.fr](mailto:isabelle.fournier@univ-lille1.fr)

PF: [Patrick.Flammang@umons.ac.be](mailto:Patrick.Flammang@umons.ac.be)

MW: [maxence.wisztorski@univ-lille1.fr](mailto:maxence.wisztorski@univ-lille1.fr)

## **Abstract**

### **Background**

Several species of sea cucumbers of the family Holothuriidae possess a particular mechanical defense system called the Cuvierian tubules. This system consists of multiple tubules that can be ejected outside the animal towards a predator in response to an attack. In holothuriids, triterpene glycosides (saponins) appear to be particularly concentrated in the Cuvierian tubules. In the present study, the precise localization of saponins in the Cuvierian tubules of *Holothuria forskali* is investigated. Classical histochemical labeling using lectin was firstly performed but did not generate any conclusive results. Thus, MALDI mass spectrometry (MALDI-MS) was directly applied for direct tissue profiling and completed by MALDI mass spectrometry Imaging (MALDI-MSI) added to statistical multivariate tests using the Principal Component Analysis (PCA) method. A comparison between the tubules of relaxed and stressed animals was realized.

### **Results**

These analyses allowed the detection of three groups of ions, corresponded to the isomeric saponins of the tubules of *H. forskali* already described in literature. Saponins detected at  $m/z$  1287 and 1303 were the most abundant and were apparently localized in the connective tissue of the tubules of both relaxed and stressed individuals. Saponins at  $m/z$  1125 and 1141 were detected in lower amount and were present mainly in tissues of relaxed animals without any precise localization. Finally, saponin ions at 1433, 1449, 1463 and 1479 were mainly observed in some Cuvierian tubules of stressed holothuroids. Except for  $m/z$  1449 which showed a diffuse signal, these saponins were localized in the outer part of the connective tissue. The saponin group  $m/z$  14xx seems therefore to be stress-specific and could come from modifications to the saponins of one the ones with  $m/z$  of 12xx. All together, these data show that, thanks to MALDI-MSI, peculiar saponins, the ones related to the group with  $m/z$  14xx, are specific for stress signals compared to the two other groups.

### **Conclusions**

The results of the present study allow, for the first time to precisely visualize secondary metabolites that are not possible to observe by conventional techniques.

Moreover, this reflects that not all saponins participate to the immediate stress responses and may participate to secondary response of the holothuriids as chemosignaling molecules. Finally, this present report demonstrates that MALDI-MSI opens the door of the chemical ecology domain and will allow investigating environmental changes at the biodiversity level through ecological molecular imaging investigation.

## Introduction

Sea cucumbers seem to be vulnerable animals regarding their numerous predators referenced in the literature [1]. However, authors consider predation on adult holothuroids to be infrequent [2,3]. Amongst the numerous anti-predator mechanisms developed by these animals, the toxicity of the body wall and the Cuvierian tubules seems to be the most effective against non-specialist predators [4,5]. Cuvierian tubules are little caeca located in the posterior part of the animal, that can be ejected toward a predator in response to an aggression. Expelled tubules lengthen into sticky white threads susceptible to entangle the predator [6,7]. Although only some species of holothuroids from the family Holothuriidae, including *Holothuria forskali*, possess Cuvierian tubules, all the sea cucumbers contain saponins in their body wall and viscera. Saponins of sea cucumbers are secondary metabolites having a triterpene glycoside structure. They have long been suggested playing a role in the defense of these animals as a toxin [4,8]. Indeed, due to their membranolytic action [9,10], saponins have a wide range of pharmacological effects such as cytotoxicity to tumor cells, antifungal and hemolytic actions [11,12]. Triterpene glycosides are particularly concentrated in the Cuvierian tubules of holothuroids [13,14] and specific congener mixtures are usually associated with the defensive function of this system [11,15].

Recent studies demonstrated that mass spectrometry (MS) procedures represent very valuable techniques for the detection and identification of saponins [16,17]. However, classical techniques require the complete pounding of the tissue to allow the extraction of its content. In consequence, the determination of the precise localization of saponins in the tissue is impossible. For 10 years, MALDI-MSI (Matrix-Assisted Laser Desorption/Ionization-Mass Spectrometry Imaging), also called MALDI-Imaging, has undergone many developments [18,19] and has achieved a certain maturity, allowing it to be now used in many domains like clinical proteomics [20] or pharmaceutical fields [21]. This procedure permits to detect and localize ions of interest directly on tissue section and nearly without any preparation. The benefit of this technique is the high sensitivity and the use of a soft ionization technique, like MALDI-MS, which allows a label free molecular imaging of a biological tissue section. The great added value of using MALDI-MS direct tissue analysis is to detect hundred of molecules in one spatially-resolved analysis. By this way, MALDI-MSI or MALDI profiling will open the door of chemical ecology domain and will permit to

investigate the impact of environmental changes on the biodiversity at the molecular level. MALDI-MSI can be used to see interaction between pathogen-host, exogenous chemicals (pesticides, fungicides etc..), chemical mediators present in abiotic interactions in low amount in complex medium. In the present study, we investigate the triterpene glycosides in *H. forskali* and tempted to localize saponins in their Cuvierian tubules. We balanced classical histochemical labeling to MALDI-MS direct tissue profiling and MALDI-MSI in order to detect and perform the spatial localization of saponins in the Cuvierian tubules added to statistical multivariate test using the Principal Component Analysis (PCA) method. PCA allows highlighting groups of close ions from dense and complex data sets and was conducted to compare the molecular data of Cuvierian tubules from stressed and relaxed holothuroids.

## **Results and Discussion**

### **Use of lectins to localize saponins on tissue sections**

As no antibodies to saponins are available, lectins were considered as a tool to detect these molecules on Cuvierian tubule sections. Lectins are proteins or glycoproteins of non-immune origin that are able to bind specific carbohydrate motifs in a way similar to the formation of the antibody-antigen complex. Each lectin recognizes specifically one oligosaccharidic chain but it may also bind to similar oligo- or mono-saccharides, although with a lower affinity. To the best of our knowledge, no commercially available lectin is specific of the saponin carbohydrate moiety which encloses glucose (Glc), 3-*O*-methylglucose (MeGlc), quinovose (Qui) and xylose (Xyl) residues. However, mannose-binding lectins like Con A, LCA and PSA also show an affinity for glucose [22] and could therefore label saponins.

To confirm that glucose-binding lectins are able to recognize the saccharidic chain of the saponins of *H. forskali*, we used a lectin-binding assay based on the technique developed by Smith [23] to detect cell surface glycolipids of human and bovine erythrocytes on thin layer chromatograms (TLC). These lectin-binding assays were realized directly on saponin extracts spotted onto nitrocellulose membrane and TLC plate strips. The development of the dark blue staining at the level of the saponin spots varied with the lectin used except for the positive control for which the reaction was always strong and fast (data not show). Although the three lectins labeled the saponin spots, Con A presented the strongest and fastest reaction compared to LCA and PSA which were both characterized by a much lighter staining. These results indicated that the considered lectins probably bind to the saccharidic chain of saponins and thus could also be used on Cuvierian tubule sections.

### **Histology and histochemistry**

Most of the ultrastructural information available on Cuvierian tubules comes from VandenSpiegel and Jangoux [24] and VandenSpiegel et al. [25], who clarified much of the structure and function of these organs in the species *H. forskali*. Quiescent Cuvierian tubules are hollow organs consisting of a narrow central lumen surrounded by a thick wall. The tubule wall is made up of an outer mesothelium and an inner epithelium encompassing a thick connective tissue sheath (Fig. 1A). This sheath includes longitudinal and circular muscle fibers that separate it into a thick, collagen-

rich inner layer and a much thinner outer layer. The other cell types present in the connective tissue sheath are the vacuolar cells, always present in the vicinity of the muscular layer, and the neurosecretory-like cells whose processes form an extensive network in the inner connective tissue layer. The mesothelium is the tissue layer involved in the adhesive process. It is a pseudostratified epithelium made up of two superposed cell layers, an outer layer of adluminal cells and an inner layer of granular cells which is highly folded along the long axis of the tubule. The inner epithelium also consists of two cell types: inner epithelial cells and spherulocytes.

Before the lectin experiments, the PAS method was used on Cuvierian tubule sections for staining structures containing a high proportion of carbohydrate-containing molecules and macromolecules (Fig. 1A). PAS-positive molecules are present on the whole section but mainly concentrated in the mesothelium (especially in adluminal cells), in the muscular fibers, and in the inner epithelium (especially in spherulocytes). Three lectins specific to glucose- and mannose-containing oligosaccharides (Con A, LCA and PSA) were then used to label the Cuvierian tubules. Investigations were performed on paraffin sections and on cryo-sections. The pattern of labeling was very similar on both types of sections and with the three lectins tested (Figs. 1B-D). The major difference laid in the labeling intensity, reactivity with Con A being always stronger than with the other two lectins (Fig. 1C). In the mesothelium, only adluminal cells were intensely labeled at the level of large intra-cytoplasmic vesicles. Muscle cells also presented important lectin reactivity. In the inner epithelium, both the inner epithelial cells and the spherulocytes were lightly labeled, the former with Con A and the latter with LCA and PSA (Figs. 1B-D). Finally, two types of lectin reactive structures were also observed in the connective tissue sheath: a few scattered vacuolar cells and a network of narrow cell processes.

Although several cell types were reactive with the three lectins tested, this does not mean that all these cells contain saponins. Indeed these three lectins certainly label other molecules like tissue glycoproteins. It was postulated that comparison between paraffin and cryo-sections could give a clue about saponin localization as the ethanol solutions used to prepare the former should have extracted the saponins from the tissues. However, no difference was found between paraffin and cryo-sections, so either saponins have been extracted in the two types of sections (saponins are also soluble in water) or they have not been extracted (possibly because of the fixatives

used). In both cases, the use of lectins only cannot give any precise information about the localization of saponins in the Cuvierian tubules.

### **Detection of saponins directly on Cuvierian tubules**

After lectin experiments, a direct analysis of a cryo-section through a bundle of a dozen Cuvierian tubule was done using a MALDI-Imaging mass spectrometer. Contrary to the usual procedure [26,27], rinsing steps were not possible due to the high solubility of saponins and the risk of their delocalization or their loss. After matrix deposition, a mass spectrum was averaged by randomly acquired data across the entire tissue section surface (Fig. 2A). This spectrum displays, in the mass range  $m/z$  1100 to 1500, eight major peaks at  $m/z$  1125, 1141, 1287, 1303, 1433, 1449, 1463 and 1479 (marked with an asterisk in the Figure 2A). In view of their  $m/z$ , these peaks might correspond to triterpene glycosides. A comparison was done between the MALDI-MS direct tissue profiling spectrum and a mass spectrum from a classical MALDI-tof analysis following chemical extraction of the saponins (Fig. 2). The pattern of the spectra was very similar for both types of analyses and presents the same  $m/z$  ratios with a conservation of the relative abundance of the ions for the eight major peaks.

In order to obtain more information on the compounds detected by MALDI-MS direct tissue profiling and to confirm that these peaks correspond to saponins, tandem mass spectrometry (MS/MS) was undertaken directly on tissue section. As an example, the MALDI-MS/MS mass spectrum for the ions detected at  $m/z$  1287, which present the most intense signal, is depicted in Figure 3. MS/MS spectra of saponins in the Cuvierian tubules were characterized by the presence of two typical signals at  $m/z$  507 and 523, corresponding to the oligosaccharidic chains [MeGlc-Glc-Qui + Na<sup>+</sup>] and [MeGlc-Glc-Glc + Na<sup>+</sup>], respectively (Fig. 3). This permitted to highlight the presence of isomers. In our case, these two ions were observed in the MALDI-MS/MS spectrum realized directly from the tissue section and confirmed that the ions detected at  $m/z$  1287 correspond to the isomeric saponins holothurinosides E and E1. Similar fragmentation patterns were obtained for the 8 peaks observed upon MALDI-MS direct profiling of the Cuvierian tubules, each one presenting  $m/z$  507 and 523 ions as the key decomposition products. These results confirmed the detection of the 16 principal saponins in Cuvierian tubules (Table 1). Like in the extraction experiments holothurinosides E and E1 (detected at  $m/z$  1287) were characterized by



the most intense signal (Fig. 2B) and consequently seemed, in first approximation, to be the most abundant saponins in the Cuvierian tubules. On the other hand, holothurinosides C and C1 (at  $m/z$  1125) presented the least intense signal.

This first approach allows showing that the MALDI-MS direct tissue analysis makes possible the detection of saponins directly on tissue, with results similar to those obtained by conventional techniques, but without any step of extraction and by consuming only a small part of tissue sample.

### **Tissue profiling of Cuvierian tubules at different state of stress**

The saponin content of Cuvierian tubule sections was investigated by MALDI profiling in order to compare their composition depending to the stress level of holothuroids. For an in-depth comparison, 3 sections in different parts of a group of tubules were realized for both relaxed and stressed individuals of *H. forskali*. These tissue sections were randomly analyzed by selecting arbitrary positions on the sample. This type of analysis is defined by the fact that the specific spot positions are selected for the data acquisition and that it does not result in a full molecular image. Approximately 150 spots were selected for each condition. For the statistical analyses, the mass spectra were internally recalibrated and normalized on the total ion count. A virtual gel view of individual spectra is presented in Figure 4A. Some differences can be observed from a visual inspection of the spectra, especially in the mass range of saponins. Peaks at  $m/z$  1125 and 1141 seemed to be present exclusively in the relaxed state whereas ions at  $m/z$  1433, 1449, 1463 and 1479 appeared to be more abundant in a stressed specimen.

To confirm this visual feeling, a statistical analysis was done on the data set. One of the techniques used to reduce the complexity of the information is the PCA [28]. The 3D plots of the first three principal components are shown in Figure 4. When all the ions of the MALDI spectra are considered for PCA analysis, the data were quite scattered (Fig. 4B) but the stressed state data (green dots) were clearly distinguishable from the relaxed group (red dots). Considering only the saponin peaks for PCA analysis (Fig. 4C), the data were more concentrated in a specific region of the principal component space. The Figure 4D showed the result of a 3D loading plot of the first three loadings. Regarding to the loading axis, three different groups having a high variance could be observed. The first group was composed of the saponins detected at  $m/z$  1125 and 1141 (red oval) and seemed to be more specific to the

relaxed data set. Another group (green oval), comprising peaks at  $m/z$  1433, 1449, 1463 and 1479, was characteristic of the stressed state. A third group, between the two others and combining the saponins at  $m/z$  1287 and 1303, may also be obtained. The multivariate analysis used here thus demonstrated the capacity of identifying differential aspects of the profile of ionic species from relaxed or stressed animals. The point to emphasize is that a correct grouping can be achieved by using only the peaks of saponins. Our findings suggest that saponins are the most variable compounds during a stress response.

With FlexImaging software, it is possible to visualize, using a color gradient, the distribution of the signal obtained for each group of saponins previously defined after PCA on the direct profiling of Cuvierian tubule sections. Results are shown in Figure 5. As expected from the PCA results, the three groups have different localizations. The group 11xx was more conspicuous in the data of the relaxed state whereas it corresponded to a low signal in the stressed state (Fig. 5B). In comparison, the group 14xx was more specifically correlated to the stressed state. Finally the group of  $m/z$  1287 and 1303 was present in the same proportion whatever the physiological state of the animal. Results are summarized in Table 2.

In view of these results, some hypotheses can be made concerning the kinetic of saponin use in the Cuvierian tubules during a stress. Holothurinosides C/C1 and desholothurins A/A1 ( $m/z$  1125 and 1141) seem to be rapidly used or released outside of the Cuvierian tubules. On contrary, holothurinosides F/F1, G/G1, H/H1 and I/I1 ( $m/z$  1433, 1449, 1463 and 1479), appear to be abundantly produced in a state of stress compared to their low level in the relaxed condition. In view of the structures of these different saponins, it seems likely that, in case of a prolonged stress situation, holothurinosides C/C1 ( $m/z$  1125) are converted to holothurinosides F/F1 and H/H1 ( $m/z$  1433 and 1463, respectively), and desholothurins A/A1 ( $m/z$  1141) to holothurinosides G/G1 and I/I1 ( $m/z$  1449 and 1479, respectively). This is simply done by the addition of a disaccharide: either Qui-Glc or MeGlc-Glc. These modifications would make the saponins more hydrophilic (i.e. more soluble in sea water) and more membranolytic (i.e. more toxic) [29]. Concerning the last group including the holothurinosides A/A1 and E/E1 ( $m/z$  1303 and 1287) their expression level look to be the same in the relaxed or stressed states.

### **MALDI-Imaging of Cuvierian tubules**

MALDI-MS direct tissue analysis is able to detect saponins on Cuvierian tubules and to study the changes during a stress response. Our purpose was to establish if MALDI-MSI could be performed on Cuvierian tubules, not only to detect saponins, but also to visualize their spatial distribution. To achieve a molecular image, a cross section of a dozen of tubules was prepared and a measurement region was defined. Tubules are very small, i.e. a cross section presents a diameter of 500  $\mu\text{m}$ . To localize the saponins as precisely as possible, the distance between each point was set to 60  $\mu\text{m}$ . In total, a cross section of 15 tubules was analyzed at a spatial resolution of 60 x 60  $\mu\text{m}^2$ .

Figure 6 shows the molecular images for each  $m/z$  ratio (corresponding to a pair of isomeric saponins). Ions detected at  $m/z$  843 were used as a counterstained due to their specific localization in the mesothelium, thereby defining the tubules. The data clearly provided evidences that there is a differential distribution of saponins within a cross section of Cuvierian tubules. Ions at  $m/z$  1125 and 1141 were slightly detected and without any specific location. This is in agreement with the average spectrum in Figure 2, in which these saponins presented a very low intensity. Concerning ions at  $m/z$  1287 and 1303, the most intense in the average spectrum, they had internal localization concentrated in what appears to be the connective tissue. Saponins group at  $m/z$  14xx showed a distinct distribution, different from the previous one, except for  $m/z$  1449 which presented a more diffuse signal. This group of saponins was only present in some tubules while saponins at  $m/z$  1287 and 1303 were distributed in all the tubules imaged. Moreover the distribution was not uniform inside the connective tissue and seemed to concentrate on the outer edges.

The fact that saponins at  $m/z$  1433, 1463 and 1479 are only present in some tubules could be correlated to the fact that this saponin group would be specific to stressed Cuvierian tubules. They could be specifically expressed in the tubules which are ready to be expelled. Indeed, although individuals of *H. forskali* possess between 200 and 600 tubules, only 10 to 20 of them are fired at a time 25, presumably in a predefined sequence. Alternatively, the difference of localization of the saponins at  $m/z$  14xx could correspond to a mix of fully functional tubules and regenerating tubules. Indeed, to maintain the line of defense, *H. forskali* is capable to regenerate its Cuvierian tubules [25]. During the collect of the tubules directly after dissection, no distinction was possible between complete tubules and regenerating ones.

## Conclusions

In conclusion, we have shown that it is possible to use direct tissue mass spectrometry profiling and MALDI-Imaging to study saponins contained in the Cuvierian tubules of holothuroids. These molecules are soluble in most solvents and their localization is therefore impossible through classical cytochemistry methods. Moreover, detection by histochemistry with lectins is difficult to perform due to their weak specificity against saponins. In this context, the use of MALDI tissue analysis (profiling or imaging) seemed a good alternative. We have demonstrated that, by using MALDI profiling on Cuvierian tubules, it was possible to detect signals of the 16 principal saponins of the Cuvierian tubules of *H. forskali*. The comparison obtained by a statistical analysis of data from relaxed and stressed holothuroids indicates that there is a differential composition in saponins depending on the condition of the animal. Finally, we demonstrated that MALDI-Imaging can be used to localize saponins inside Cuvierian tubule sections. An important future extension might be the use of MALDI-Imaging at higher resolution to check if saponins are specific to certain cell types. Indeed, the current resolution is too low for this purpose. However, comparing the saponin-specific MALDI-MSI images with the high-resolution but non-specific lectin-labeled images suggests that saponins would be produced at the level of the connective tissue layer, either in vacuolar cells, in neurosecretory-like cells, or in both cell types.

All together, this study reflects that MALDI profiling and MALDI-MSI can be used for chemical ecological studies to track some specific chemical signaling molecules like terpenes or pheromones. This study represents the first one performed in the domain of ecology.

## Methods

### Chemicals and materials

$\alpha$ -cyano-4-hydroxycinnamic acid (HCCA), trifluoroacetic acid (TFA), acetonitrile (ACN), aniline (ANI), were purchased from Sigma-Aldrich (Saint-Quentin Fallavier, France).

### Sampling

Individuals of *H. forskali* were collected at depths ranging from 10 to 30 m by scuba diving at the Biological Station of Banyuls-sur-Mer (Pyrénées-Orientales, France). They were transported to the Marine Biology Laboratory of the University of Mons, where they were kept in a marine aquarium with closed circulation (13 °C, 33‰

salinity). Two different physiological states were investigated: relaxed and stressed individuals (only the former were used, however, for the histochemical approach). Relaxed animals were anesthetized 2 h in a 0.1% solution of 1-phenoxy-2-propanol (Sigma-Aldrich, Germany) in sea water. Stressed individuals were mechanically disturbed for 4 h by repetitive hitting using a specific device. The device consisted of three loads hanging from a horizontal rotating bar and immersed in the tank containing the animal. Loads were embedded in parafilm in order to avoid injuries. Animals stimulated in this way contracted but never expelled their Cuvierian tubules. After each treatment, the holothuroids were dissected in order to collect their Cuvierian tubules. A fraction of the tubules was stored in 70% ethanol for saponin extraction, another fraction was fixed for histology and histochemistry, and a last fraction was quickly frozen in liquid nitrogen for MALDI-MSI.

#### **Lectin assays**

Lectins are proteins or glycoproteins of non-immune origin that are able to bind carbohydrates without chemically modifying them[30]. Three glucose- and mannose-specific, biotinylated lectins purchased from Vector Laboratories (Burlingame, CA) were tested on saponin extracts from Cuvierian tubules: concanavaline A (Con A), *Lens culinaris* agglutinin (LCA) and *Pisum sativum* agglutinin (PSA). Saponins were extracted. The final butanol fraction was spotted onto strips from two different supports: nitrocellulose membranes (BA 85, Schleicher & Schuell, Dassel, Germany) and thin layer chromatography plates (TLC plates, silica gel on aluminium, Sigma-Aldrich, St. Louis, MO). A diluted aqueous solution of commercial milk powder was used as a positive control. The strips were pre-incubated for 30 min in phosphate-buffered saline containing 0.05% of Tween 20 (MP Biomedicals, Solon, OH) and 0.1% bovine serum albumin (Sigma, St. Louis, MO) (PBS-Tween-BSA), rinsed in PBS, and were then incubated for 2 h at room temperature in the biotinylated lectins diluted at 0.25 µg/ml in PBS-Tween-BSA. After several washes in PBS, the strips were incubated for 1 h in a solution of streptavidine-conjugated alkaline phosphatase (Vector Laboratories, Burlingame, CA) diluted 1:500 in PBS-Tween-BSA, then washed again in PBS-Tween-BSA. Finally, the strips were incubated for 10 min in a 2% solution of NBT/BCIP (Roche, Mannheim, Germany) in the revelation buffer (0.1 M Tris-HCl, pH 9.5, 0.1 M NaCl, 0.05 M MgCl<sub>2</sub>). The reaction was stopped in deionized water.

### **Histology and histochemistry**

Cuvierian tubules were fixed in Bouin's fluid, in 10% formalin in sea water, or in 4% paraformaldehyde in PBS. Bouin's fluid-fixed tubules were dehydrated in graded ethanol, embedded in paraffin wax using a routine method, and sectioned at 7  $\mu\text{m}$  using a rotary microtome (Microm HM340E). Formalin- and paraformaldehyde-fixed tubules were embedded in a commercial tissue freezing medium (TBS from Triangle Biomedical Sciences, Durham, N.C.) in order to realize 10  $\mu\text{m}$  cryo-sections using a Leica Cryocut 1800 cryostat operated at  $-25\text{ }^{\circ}\text{C}$ . Some sections (paraffin- and cryo-sections) were stained with Heidenhain's azan trichrome stain, with Masson's trichrome stain or with the periodic acid-Schiff (PAS) method. Other sections were subjected to an indirect histochemical labeling method.

For lectin histochemistry, non-specific background staining was blocked by a 30 min pre-incubation of the sections in PBS containing 0.05% of Tween 20 and 3% BSA (PBS-Tween-BSA2). Biotinylated lectins, diluted at 2.5 or 0.25  $\mu\text{g}/\text{ml}$  in PBS-Tween-BSA2, were then applied for 2 h at room temperature. After several washes in PBS, the sections were incubated for 30 min in the ABCComplex solution (Dako, Denmark) consisting of streptavidin and biotinylated peroxidase. Following a final wash in PBS, lectin-carbohydrate complexes were revealed by incubation of the sections with a 0.05% solution of DAB (3,3'-diaminobenzidine tetrahydrochloride; Sigma, St. Louis, MO) in PBS containing 0.02%  $\text{H}_2\text{O}_2$ . The reaction was controlled under microscope and, when the optimal contrast was obtained, the revelation was stopped in deionized water. Finally, the sections were lightly counterstained with hemalum and luxol blue and observed with a Zeiss Axioscope A1 microscope (Carl Zeiss MicroImaging, Göttingen, Germany) equipped with a Zeiss Icc-3 digital camera.

### **Tissue preparation for mass spectrometry**

Thin 12  $\mu\text{m}$  tissue sections were obtained from a frozen bundle of Cuvierian tubules using a cryostat CM1510S (Leica Microsystems, Nanterre, France) and applied onto an indium-tin oxide (ITO)-coated conductive glass slides (Bruker Daltonics, Bremen, Germany).

### **Matrix deposition**

A solid ionic HCCA/ANI matrix was used to analyze the saponins on tissue sections and was prepared following a previously established procedure [31]. Briefly 7.2  $\mu\text{l}$  of ANI (1.5 equivalent) was added to a solution containing 10 mg/ml of HCCA (1 equivalent) in ACN/aqueous TFA 0.1% (6:4, v/v). The matrix solution was then deposited using an automatic vibration vaporization system (ImagePrep, Bruker Daltonics, Bremen, Germany) to cover the whole surface of the tissue section. The ImagePrep method for HCCA/ANI deposition is based on a modified HCCA method included in the ImagePrep.

#### **MALDI-MS direct tissue analysis**

Data were acquired at different position on the tissue section using an UltraFlex II MALDI-TOF/TOF mass spectrometer (Bruker Daltonics, Bremen, Germany) equipped with a Smartbeam<sup>TM</sup> laser. Direct profiling data were performed in positive reflectron mode, for a mass range from  $m/z$  800 to 3000. A total of 300 spectra were acquired at each spot at a laser frequency of 100 Hz. FlexImaging<sup>TM</sup> 2.1 software was used for automatic data acquisition of 150 spots for each condition. Standards for spectral calibration consisted of a mixed solution of peptides ranging between  $m/z$  900 to 3500 Da.

#### **PCA statistical analysis**

Statistical analyses were performed using ClinProTools 2.2 software (Bruker Daltonics, Bremen, Germany). After loading data of each condition, mass spectra were internally recalibrated and normalized. An average spectrum created from all single spectra was used for peak picking. A first PCA analysis was done using all peaks detected in the average mass spectrum. A second PCA was completed using only peaks corresponding to the saponins.

#### **MALDI-Imaging**

The images were acquired using an UltraFlex II MALDI-TOF/TOF mass spectrometer (Bruker Daltonics, Bremen, Germany) equipped with a Smartbeam<sup>TM</sup> laser. Images data were performed in positive reflectron mode, and MALDI-MS spectra were acquired in the  $m/z$  range from  $m/z$  800 to 3000. A total of 300 spectra were acquired at each spot at a laser frequency of 100 Hz. Spatial resolution was set to 60 x 60  $\mu\text{m}^2$ . FlexImaging<sup>TM</sup> 2.1 software was used for molecular image

reconstruction. Standards for spectral calibration consisted of a mixed solution of peptides ranging between m/z 900 to 3500 Da.

## Acknowledgements

The authors are grateful to A. Desmons for her help to tissue sections and to M. Bohem for her contribution to histochemical analyses.

## References

1. Francour P (1997) Predation on Holothurians: A Literature Review. *Invertebrate Biology* 116: 52-60.
2. Da Silva J, Cameron L, Frankboner PV (1986) Movement and orientation patterns in the commercial sea cucumber *Parastichopus californicus*. *Marine Behavior and Physiology* 12: 133-147.
3. Massin C, Jangoux M (1976) Observations écologiques sur *Holothuria tubulosa*, *H. poli* et *H. forskali* (Echinodermata, Holothuroidea) et comportement alimentaire de *H. tubulosa*. *Cahiers de Biologie Marine* 18: 45-59.
4. Bakus GJ (1968) Defensive mechanisms and ecology of some tropical holothurians. *Marine Biology*: 23-32.
5. Bingham BL, Braithwaite LF (1986) Defense adaptations of the dendrochirote holothurian *Psolus chitonoides* Clark. *Journal of experimental marine biology and ecology*: 311-322.
6. Flammang P, Ribesse J, Jangoux M (2002) Biomechanics of adhesion in sea cucumber Cuvierian tubules. *Integr Comp Biol*: 1107-1115.
7. Hamel J, Mercier A (2000) Cuvierian tubules in tropical holothurians: usefulness and efficiency as a defence mechanism. *Marine and Freshwater Behaviour and Physiology*: 115-139.
8. Kalyani GA, Kakrani HKN, Hukkeri VI (1988) Holothurin - A review. *Indian J Natural Products* 4: 3.
9. Verbist JE (1993) Pharmacological effects of compounds from Echinoderms. *Echinoderm Studies* 4: 111-186.
10. Popov AM (2002) Comparative study of cytotoxic and hemolytic effects of triterpenoids isolated from Ginseng and Sea cucumber. *Izv Akad Nauk Ser Biol* 2: 155-164.
11. Kalinin VI, Prokofieava NG, Likhatskaya GN, Schentsova EB, Agafonova IG, et al. (1996) Hemolytic activities of triterpene glycosides from the holothurian order Dendrochirotida: some trends in the evolution of this group of toxins. *Toxicon* 34: 475-483.
12. Guo M, Song F, Liu Z, Liu S (2006) Characterization of triterpenoidic saponin mixture in crude extracts from leaves of *Acanthopanax senticosus* harms by saponin structural correlation and mass spectrometry. *Anal Chim Acta* 557: 198-203.



13. Elyakov GB, Stonik VA, Leniva EV, Slanke VP, Kuznetsova TA, et al. (1973) Glycosides of marine invertebrates-I. A comparative study of the glycosides fraction of Pacific sea cucumbers. *Comp Biochem Physiol B* 44.
14. Kobayashi M, Hori M, Kan K, Yasuzawa T, Matsui M, et al. (1991) Marine Natural Product. XXVII. Distribution of lanostane-type triterpene oligoglycosides in ten kinds of Okinawan sea cucumbers. *Chem Pharm Bull (Tokyo)* 39: 2282-2287.
15. Kalinin VI, Anisimov MM, Prokofieva NG, Avilov SA, Afiyatulloev SH, et al. (1995) Biological activities and biological role of triterpene glycosides from holothuroids. Jangoux, M, Lawrence, LM (Eds), *Echinoderm Studies* 5: 139-181.
16. Fang S, Hao C, Sun W, Liu Z, Liu S (1998) Rapid analysis of steroidal saponin mixture using electrospray ionization mass spectrometry combined with sequential tandem mass spectrometry. *Rapid Commun Mass Spectrom* 12: 589-594.
17. Madl T, Sterk H, Mittelbach M, Rechberger GN (2006) Tandem mass spectrometric analysis of a complex triterpene saponin mixture of *Chenopodium quinoa*. *J Am Soc Mass Spectrom* 17: 795-806.
18. Francese S, Dani FR, Traldi P, Mastrobuoni G, Pieraccini G, et al. (2009) MALDI mass spectrometry imaging, from its origins up to today: the state of the art. *Comb Chem High Throughput Screen* 12: 156-174.
19. Wisztorski M, Croix D, Macagno E, Fournier I, Salzet M (2008) Molecular MALDI imaging: an emerging technology for neuroscience studies. *Dev Neurobiol* 68: 845-858.
20. Franck J, Arafah K, Elayed M, Bonnel D, Vergara D, et al. (2009) MALDI imaging mass spectrometry: state of the art technology in clinical proteomics. *Mol Cell Proteomics* 8: 2023-2033.
21. Hsieh Y, Chen J, Korfmacher WA (2007) Mapping pharmaceuticals in tissues using MALDI imaging mass spectrometry. *J Pharmacol Toxicol Methods* 55: 193-200.
22. Goldstein U, Poretz RD (1986) *The Lectins*; Liener IE, Sharon N, Goldstein IJ, editors. New York: Academic Press.
23. Smith DF (1983) Glycolipid-lectin interactions: detection by direct binding of <sup>125</sup>I-lectins to thin layer chromatograms. *Biochem Biophys Res Commun* 115: 360-367.
24. Vandenspiegel D, Jangoux M (1987) Cuvierian tubules of the holothuroid *Holothuria forskali* (Echinodermata) : A morphofunctional study. *Marine Biology* 96: 263-275.
25. Vandenspiegel D, Jangoux M, Flammang P (2000) Maintaining the line of defense : Regeneration of Cuvierian tubules in the sea cucumber *Holothuria forskali* (Echinodermata, Holothuroidea). *Biol Bull* 198: 34-49.
26. Lemaire R, Wisztorski M, Desmons A, Tabet JC, Day R, et al. (2006) MALDI-MS direct tissue analysis of proteins: Improving signal sensitivity using organic treatments. *Anal Chem* 78: 7145-7153.
27. Seeley EH, Oppenheimer SR, Mi D, Chaurand P, Caprioli RM (2008) Enhancement of protein sensitivity for MALDI imaging mass spectrometry after chemical treatment of tissue sections. *J Am Soc Mass Spectrom* 19: 1069-1077.
28. Van de Plas R, Ojeda F, Dewil M, Van Den Bosch L, De Moor B, et al. (2007) Prospective Exploration of Biochemical Tissue Composition via Imaging

- Mass Spectrometry Guided by Principal Component Analysis. Pacific Symposium on Biocomputing: 458-469.
29. Kalinin VI (2000) System-theoretical (Holistic) approach to the modelling of structural-functional relationships of biomolecules and their evolution: an example of triterpene glycosides from sea cucumbers (Echinodermata, holothurioidea). *J Theor Biol* 206: 151-168.
  30. Hayat MA, editor (1993) *Stains and cytochemical methods*: London Boston.
  31. Lemaire R, Tabet JC, Ducoroy P, Hendra JB, Salzet M, et al. (2006) Solid ionic matrixes for direct tissue analysis and MALDI imaging. *Anal Chem* 78: 809-819.

## Figures

### Figure 1 - Histochemical labeling of the Cuvierian tubules of *Holothuria forskali*

(A, C, D: paraffin sections; B: cryo-section). A: Cross section through a whole tubule stained with the PAS method; B-D: Transverse sections through the tubule wall labeled with the PSA, Con A and LCA, respectively (C and D were counterstained).

ac: adluminal cell layer of the mesothelium; cm: circular muscle; ct: connective tissue; g: granular cell layer of the mesothelium; ict: inner connective tissue layer; ie: inner epithelium; lm: longitudinal muscle; m: mesothelium; ml: muscular layer; ncp: narrow cell processes; oct: outer connective tissue layer; s: spherulocytes.

### Figure 2 - Detection of Saponins by MALDI direct tissue analysis

Comparison of an average mass spectrum of the Cuvierian tubules of *Holothuria forskali* obtained by direct MALDI-MSI (A) and a mass spectrum obtained by classical MALDI-MS analysis after extraction and purification of saponins from the tubules (B). Saponin ions are marked by an asterisk.

### Figure 3 – Direct tissue MS/MS analysis of the saponin ions at $m/z$ 1287.

Characteristics' signatures peaks of saponins ( $m/z$  503 and  $m/z$  523) are marked by an asterisk.

### Figure 4 – Comparison of saponins composition between stress or relaxed holothuroids.

(A) Pseudogel view of 300 mass spectra from the Cuvierian tubules of *Holothuria forskali*: spectra from stressed holothuroids (upper part) and from relaxed holothuroids (inner part). (B and C): scores plot of first, second and third principal component considering all the ions (B) or only the saponin ions (C) in the MALDI spectra. Each data point represents one mass spectrum (red, relaxed; green, stressed). (D) Scores plot of first, second and third loading considering only the saponin ions. Each data point represents one  $m/z$  ratio.

**Figure 5 – Repartition of saponin groups depending on the state of stress.**

(A) Optical image of sections through bundles of Cuvierian tubules. The red dots depict the positions of each mass spectrum. (B, C and D) Localization of group of saponins  $m/z$  1125 and 1141 (B); group of saponins  $m/z$  1287 and 1303 (C); and group of saponins  $m/z$  1433, 1449, 1463 and 1479 (D) for tubule frozen sections from stressed (left part) or relaxed (right part) individuals.

**Figure 6 – MALDI-Imaging results shown in false colour representation.**

Colors reflect intensities of selected mass signals: green, the ions at  $m/z$  843 used as a counterstain; red, the different saponin ions for which the  $m/z$  ratios are indicated below the images.

## Tables

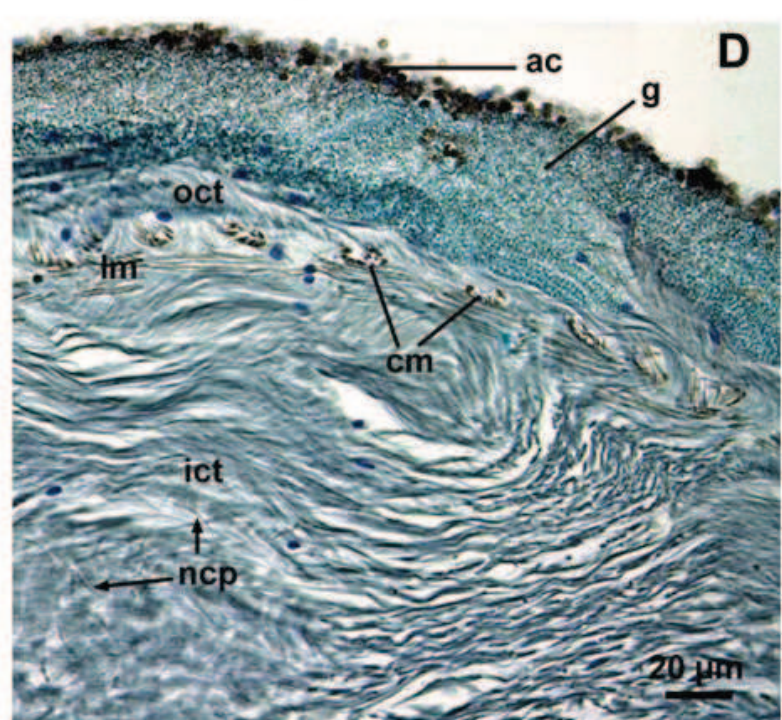
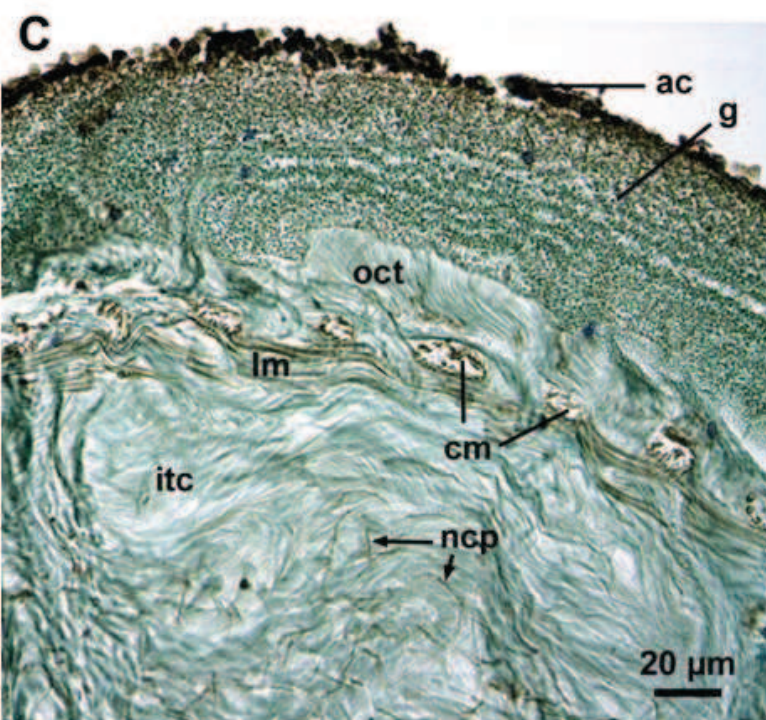
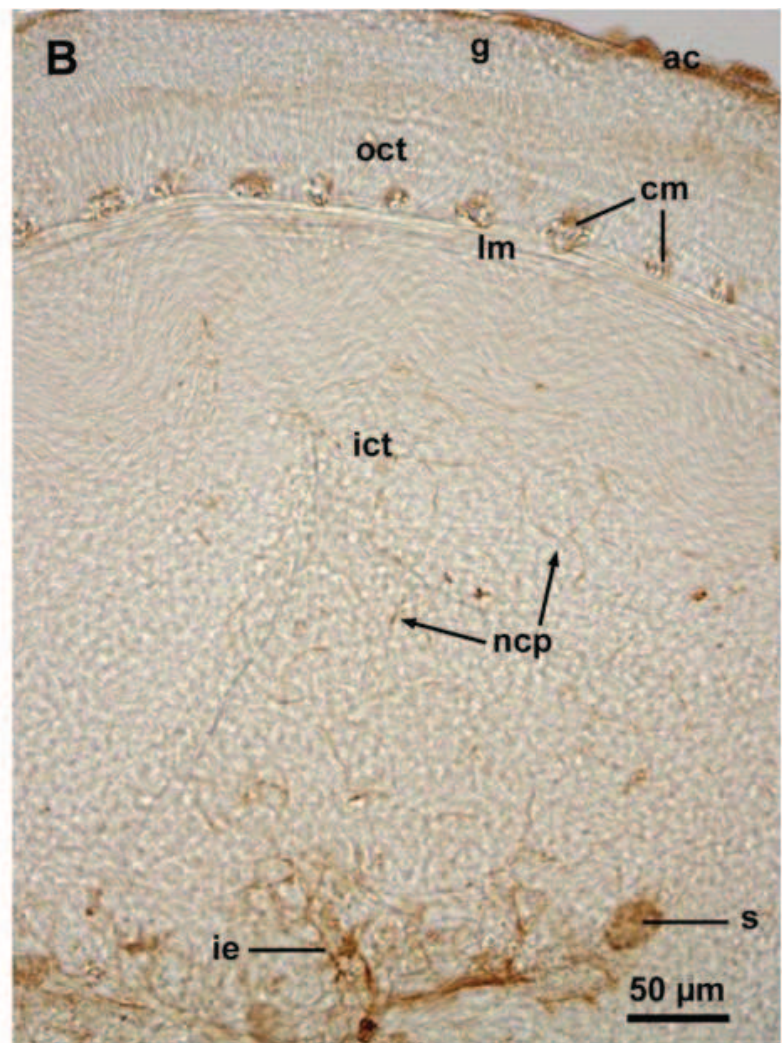
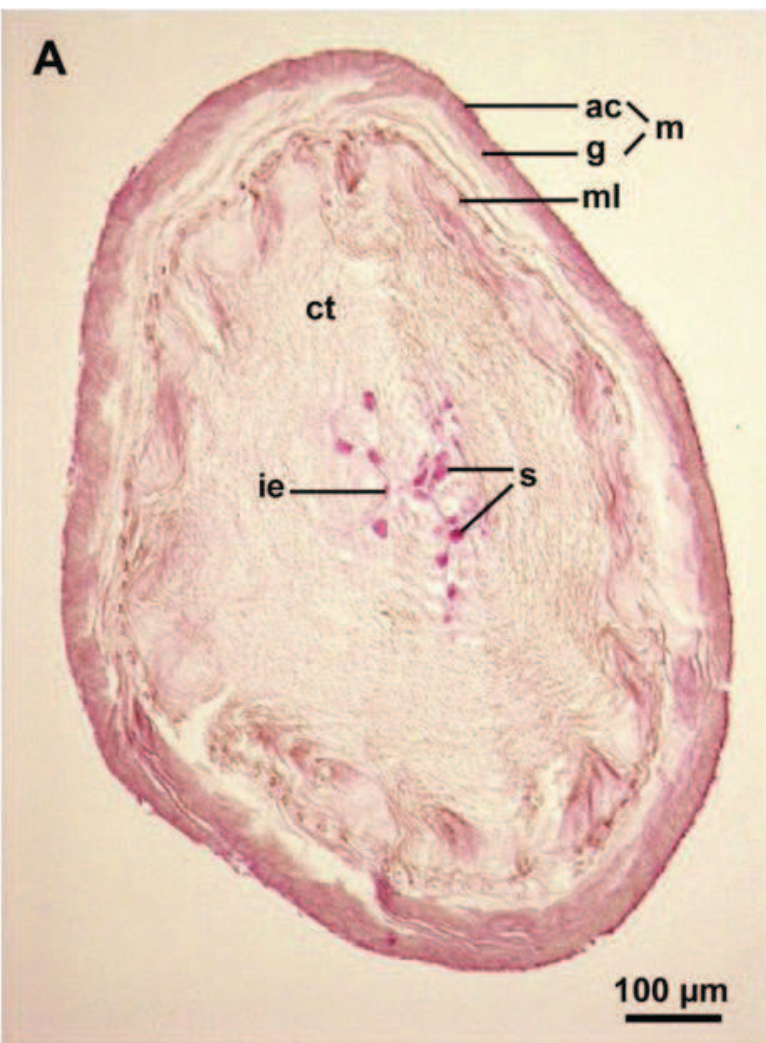
**Table 1- Complete list of saponins detected by MALDI Direct Tissue Analysis of a Cuvierian tubules section of *Holothuria forskali*.**

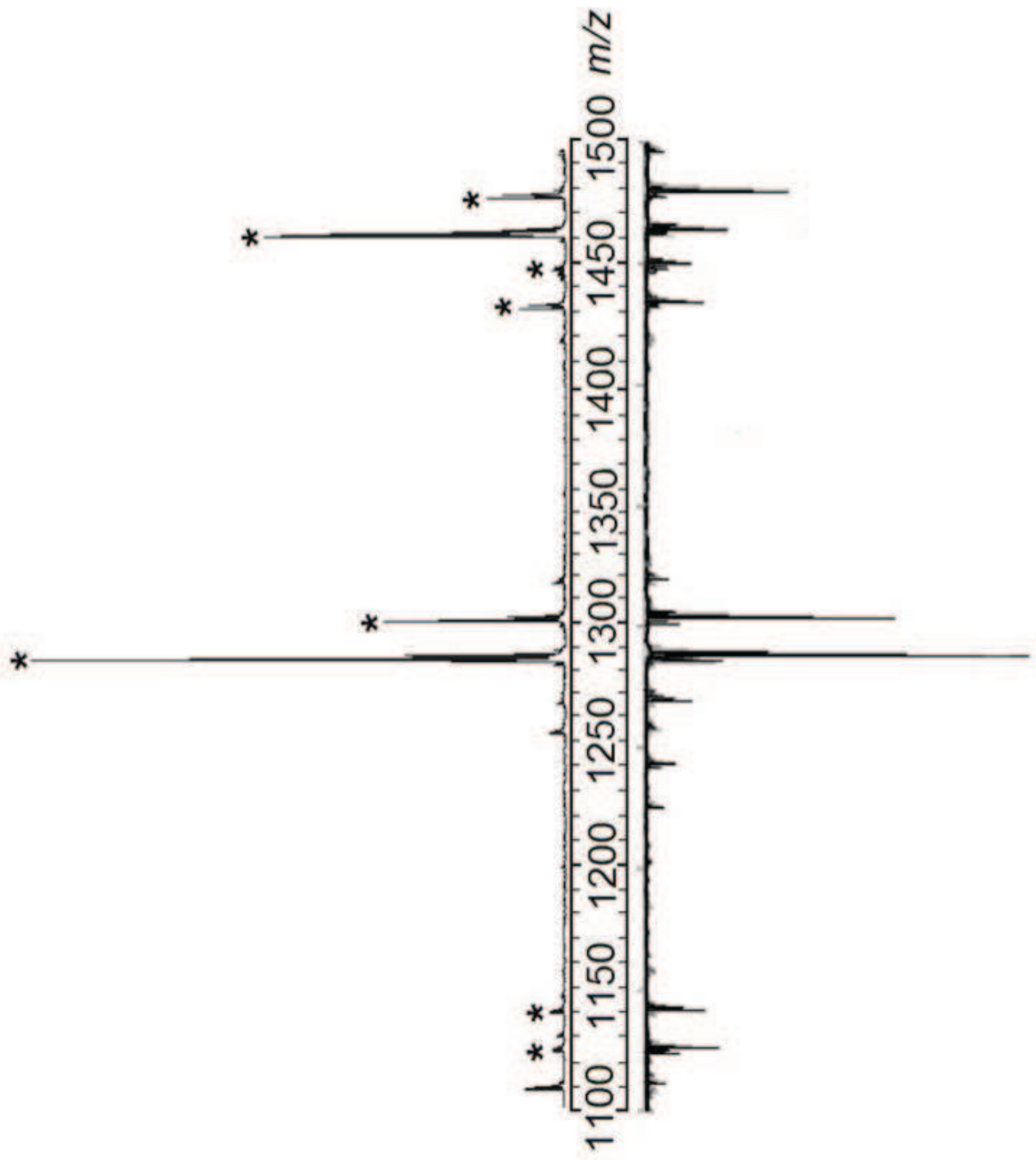
$[M+Na^+]$	Name	Molecular Formula
1125	Holothurinoside C Holothurinoside C <sub>1</sub>	C <sub>54</sub> H <sub>86</sub> O <sub>23</sub>
1141	Desholothurin A Desholothurin A <sub>1</sub>	C <sub>54</sub> H <sub>86</sub> O <sub>24</sub>
1287	Holothurinoside E Holothurinoside E <sub>1</sub>	C <sub>60</sub> H <sub>96</sub> O <sub>28</sub>
1303	Holothurinoside A Holothurinoside A <sub>1</sub>	C <sub>60</sub> H <sub>96</sub> O <sub>29</sub>
1433	Holothurinoside F Holothurinoside F <sub>1</sub>	C <sub>66</sub> H <sub>106</sub> O <sub>32</sub>
1449	Holothurinoside G Holothurinoside G <sub>1</sub>	C <sub>66</sub> H <sub>106</sub> O <sub>33</sub>
1463	Holothurinoside H Holothurinoside H <sub>1</sub>	C <sub>67</sub> H <sub>108</sub> O <sub>33</sub>
1479	Holothurinoside I Holothurinoside I <sub>1</sub>	C <sub>67</sub> H <sub>108</sub> O <sub>34</sub>

**Table 2 - Localization of saponins in the Cuvierian tubules of relaxed and stressed individuals of *Holothuria forskali*.**

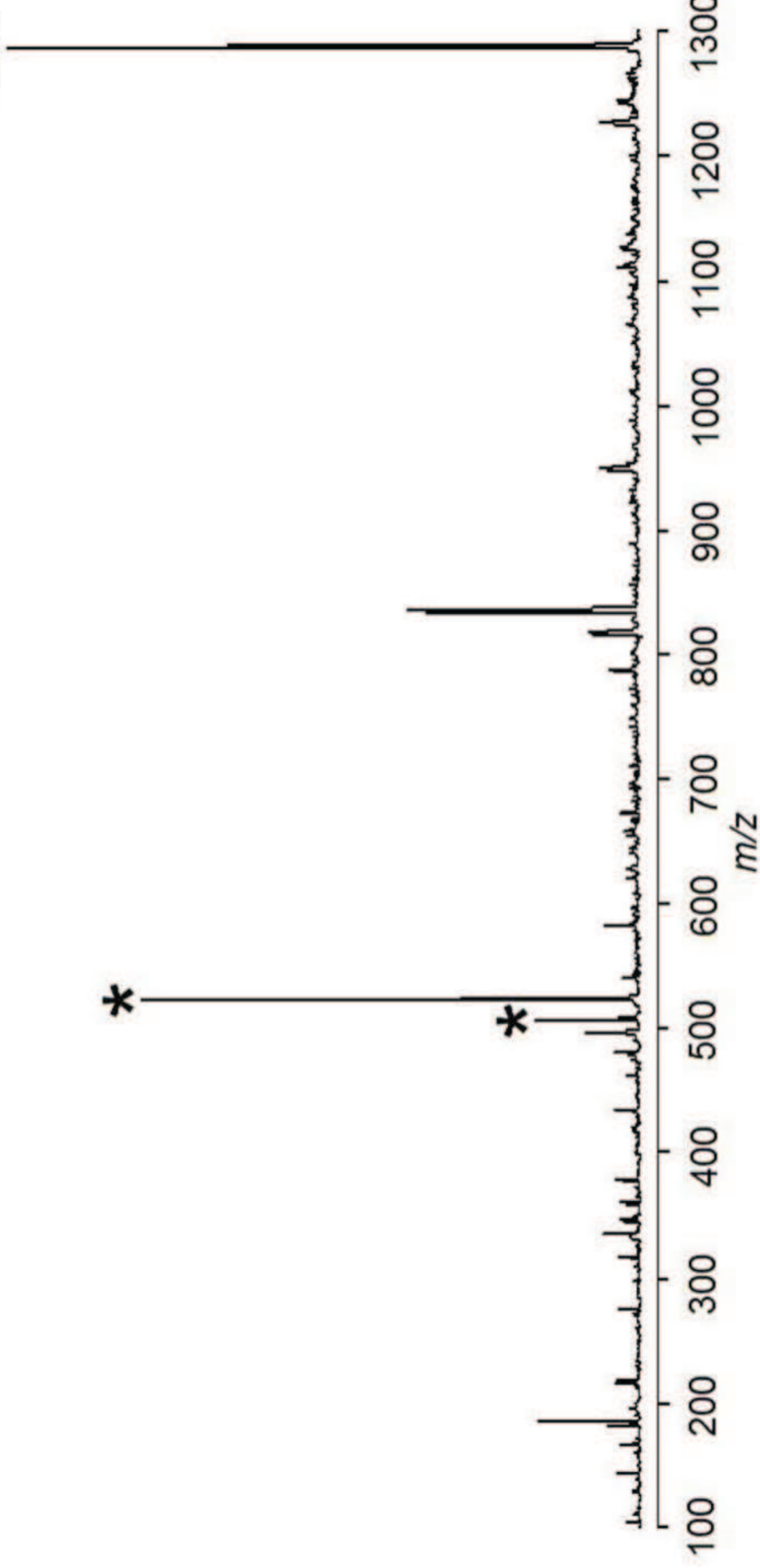
The saponins ions are present (+) or absent (-).

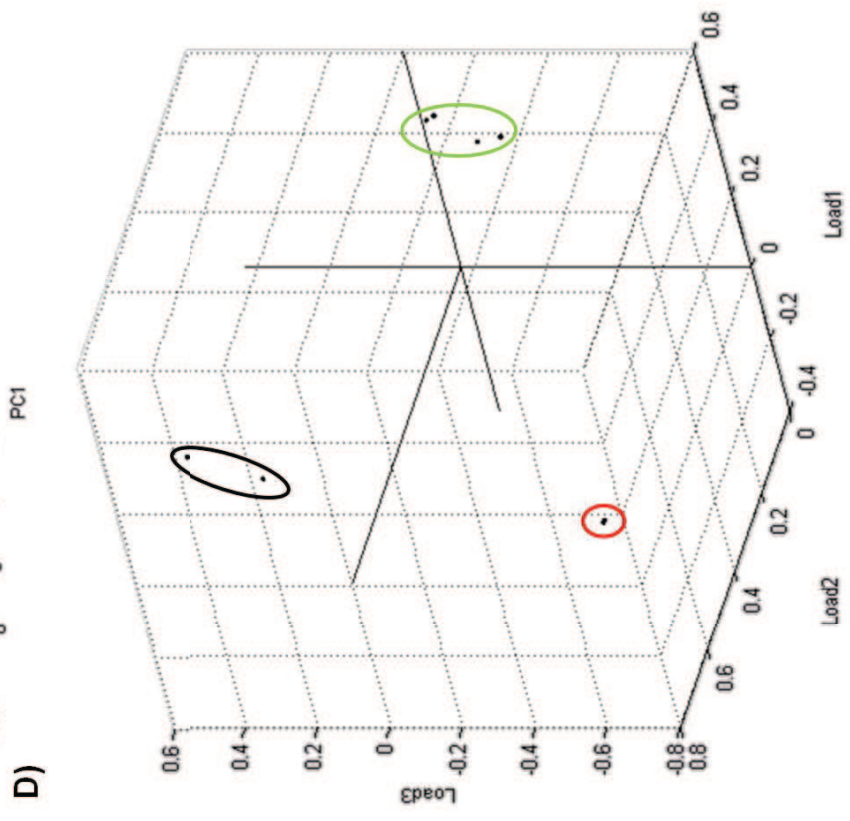
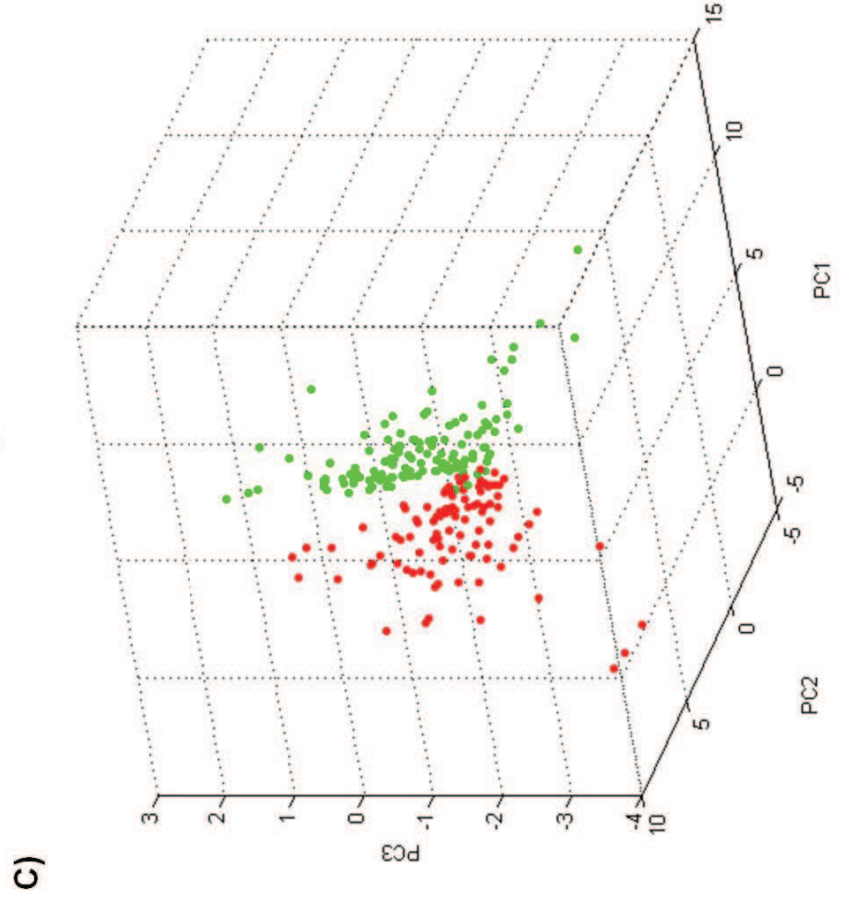
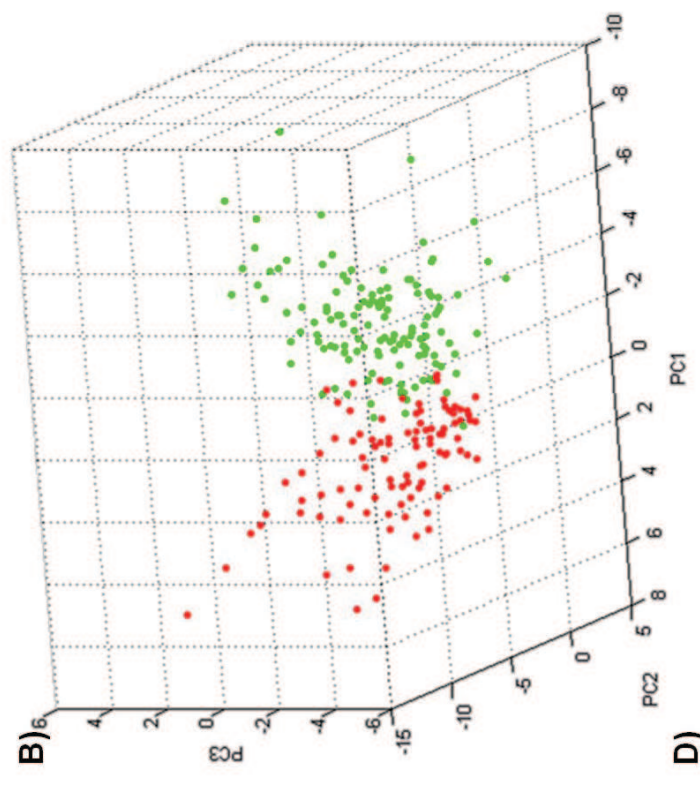
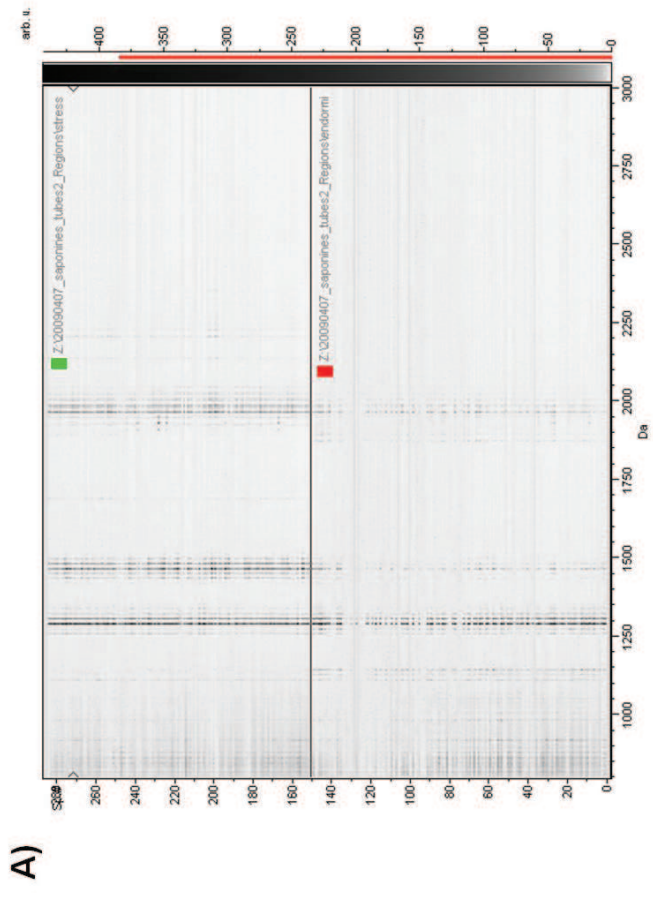
Saponins [M + Na] <sup>+</sup>	Relaxed	Stressed
<b>1125</b>	+	-
<b>1141</b>	+	-
<b>1287</b>	+	+
<b>1303</b>	+	+
<b>1433</b>	-	+
<b>1449</b>	-	+
<b>1463</b>	-	+
<b>1479</b>	-	+



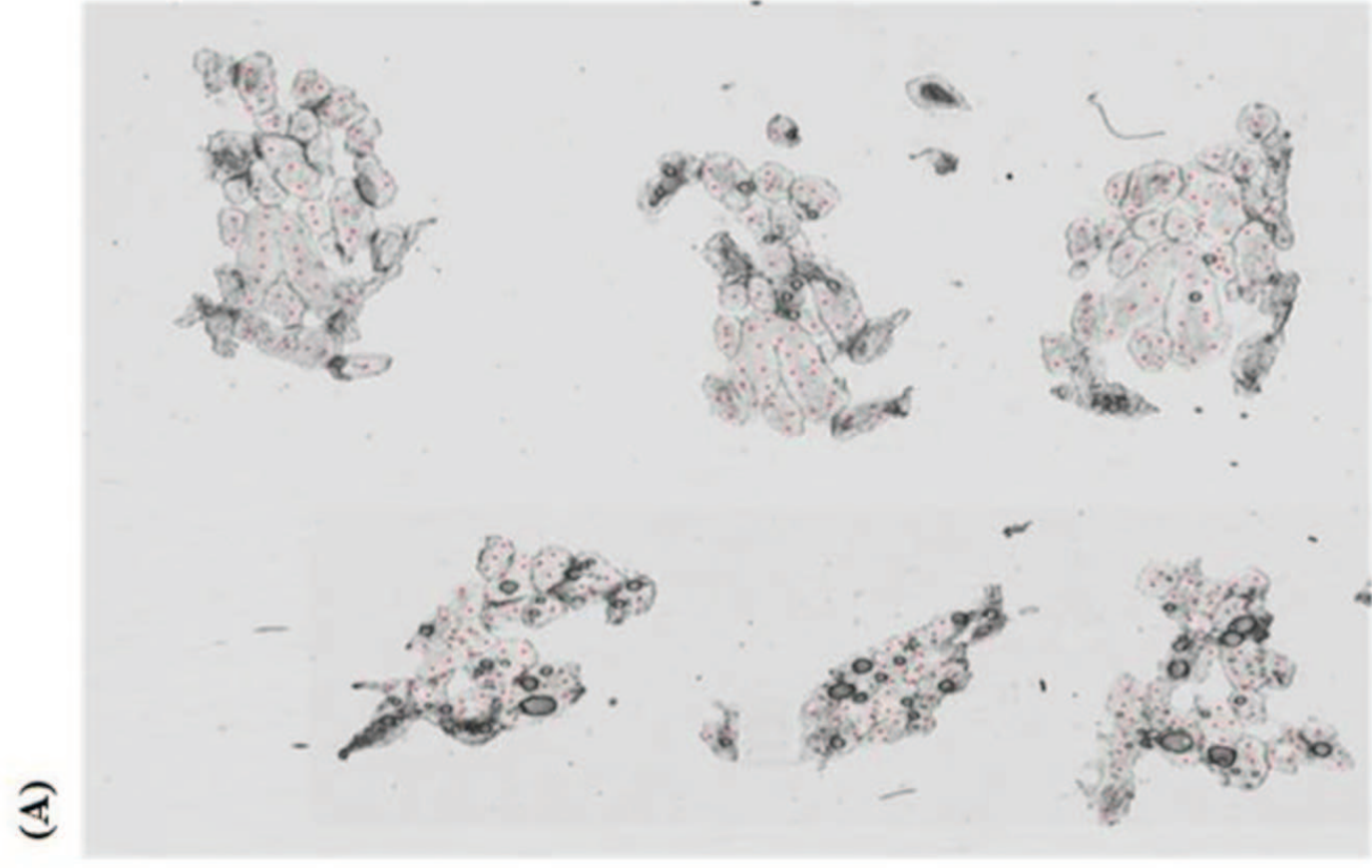


1287





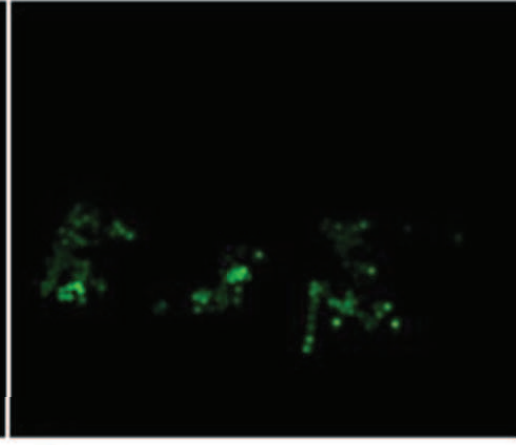




Relax



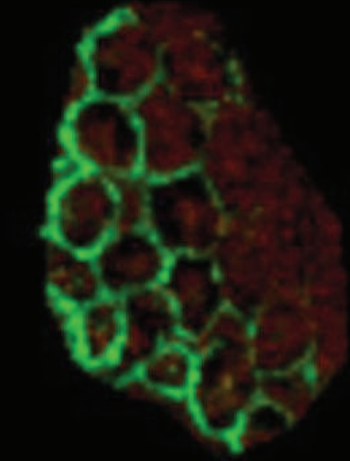
Stress



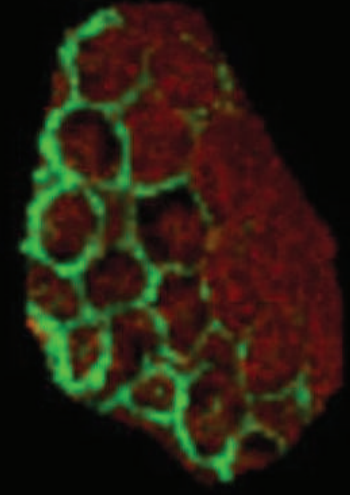
(B)

(C)

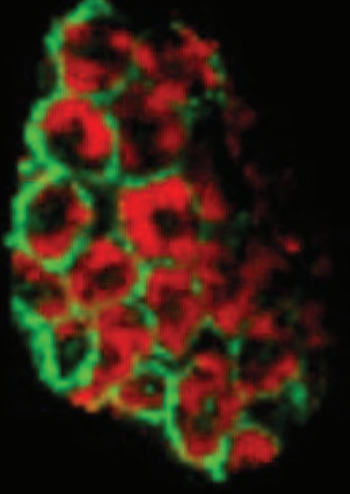
(D)



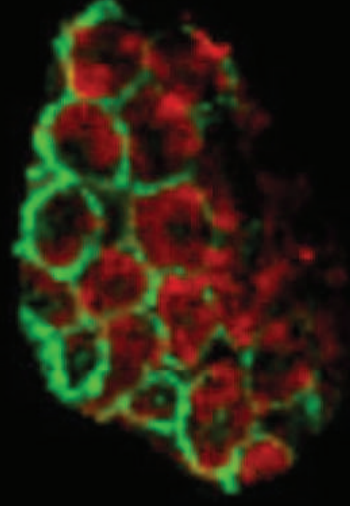
$m/z$  1125



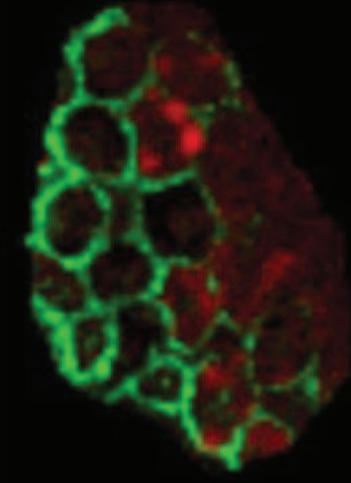
$m/z$  1141



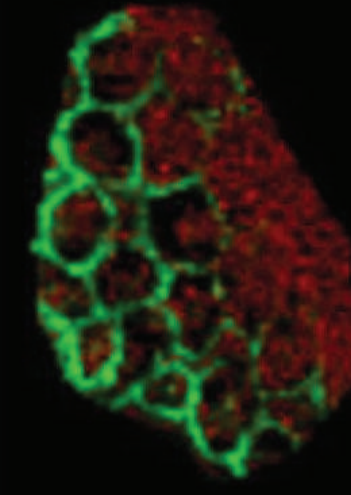
$m/z$  1287



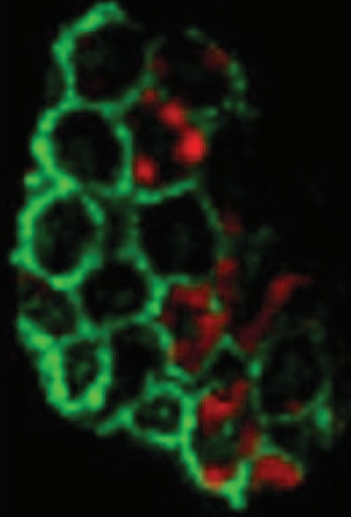
$m/z$  1303



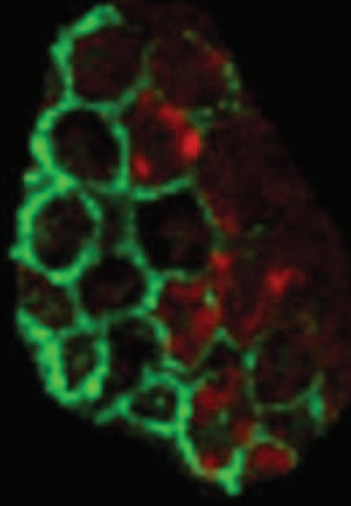
$m/z$  1433



$m/z$  1449



$m/z$  1463



$m/z$  1479

## **Résumé :**

Mon travail de thèse a été consacré à l'optimisation et l'utilisation de techniques en protéomique clinique, et plus particulièrement de l'imagerie MALDI, pour la recherche et l'identification de marqueurs pathologiques. C'est pourquoi nous avons, dans un premier temps, développé et appliqué des outils d'analyse statistique basés sur la PCA et la classification hiérarchique, appelés PCA-SDA. Ceux-ci offraient une combinaison intéressante avec l'imagerie MALDI et permettaient une simplification des données, une recherche fine des variations moléculaires au sein du tissu et une classification sur la base des profils moléculaires obtenus localement sur les tissus. Appliquée ensuite sur des études de biopsies de cancer de l'ovaire, cette approche nous a permis de détecter et d'identifier plusieurs marqueurs potentiels jouant un rôle dans la réponse du système immunitaire, l'adhésion et l'invasion tumorale. Or, ces mécanismes sont connus pour impliquer des protéases, les proprotéines convertases, dans la maturation des différentes protéines impliquées dans le développement tumoral. Dans ce contexte, nous avons étudié leur expression dans le cancer de la prostate. Il s'est avéré que seule PACE4 était surexprimée dans cette pathologie et nous avons pu établir son rôle primordial à la fois dans la prolifération cellulaire à l'aide d'études *in vitro*, et dans l'évasion apoptotique par approche protéomique, suite à l'identification de TRPS1, un facteur de transcription impliqué dans l'apoptose. Le rôle prépondérant de PACE4 fait de cette enzyme une cible thérapeutique potentielle du cancer.

---

## **Abstract:**

My PhD's work has been completely devoted to the optimization and the use of technologies in clinical proteomics, especially MALDI imaging, for research and identification of disease markers. Therefore we have initially developed and applied the tools of statistical analysis based on PCA and hierarchical clustering, called PCA-SDA, which offered an interesting combination with MALDI imaging and allow simplification of data, fine search of molecular changes within the tissue and a classification based on molecular profiles obtained locally on the tissues. Then applied on ovarian cancer biopsies study, this approach allowed us to detect and identify several potential markers playing a role in immune response, adhesion and tumor invasion. However, these mechanisms are known to involve proteases, like proprotein-convertases, in the maturation of various proteins implicated in tumor development. In this context, we studied their expression in prostate cancer. It pointed that only PACE4 was over-expressed in this disease and we were able to establish its role in cell proliferation using *in vitro* analysis and in apoptotic evasion with the identification of TRPS1, a transcription factor involved in apoptosis, by proteomics approach. So, PACE4 is a potential therapeutic target for cancer due to its leading role in tumor cell capacities.

---

**MOTS-CLEFS :** Protéomique, Spectrométrie de masse, Imagerie MALDI, Proprotéine-convertase, PACE4, PCA, Classification hiérarchique, Protéine, Peptide, Identification, Cancer, Ovaire, Prostate.

---

**Laboratoire de Neuroimmunologie et  
Neurochimie Evolutives  
Equipe d'Imagerie MALDI  
FRE-CNRS 3249, IFR 147, Bât SN3  
Université de Lille 1  
59655 Villeneuve d'Ascq Cedex  
France**

**Institut de Pharmacologie de Sherbrooke  
Université de Sherbrooke  
Faculté de Médecine et des Sciences de la  
Santé  
3001, 12ème Avenue Nord.  
Sherbrooke, Québec, J1H 5N4  
Canada**