

THESE

présentée pour l'obtention du grade de

DOCTEUR DE L'UNIVERSITE LOUIS PASTEUR DE STRASBOURG

par

Elsa WAGNER

NOUVELLES APPROCHES DE LA SPECTROMETRIE DE MASSE EN TANDEM POUR LA CARACTERISATION DE BIOMOLECULES. APPLICATIONS A LA PROTEOMIQUE.

Soutenue le 30 Janvier 2004 devant la commission d'examen :

Prof. Catherine FLORENTZ

Rapporteur interne

Dr. Patrick JOUIN

Rapporteur externe

Prof. Jérôme LEMOINE

Rapporteur externe

Dr. Alain VAN DORSSELAER

Directeur de thèse

Dr. Bruno SCHOENTJES

Examineur

Dr. Emmanuelle LEIZE

Examineur

Merci !

Je tiens à remercier ici chaleureusement toutes les personnes qui ont été présentes à mes côtés durant ces trois années de thèse, partageant mes difficultés mais également les immenses bonheurs de mes petites découvertes...

Merci à la société Bruker qui m'a financée et qui m'a également apporté un soutien scientifique et technique.

Merci à Alain Van Dorsselaer de m'avoir accueillie dans son laboratoire et de m'avoir offert la chance d'apprendre et de travailler dans d'excellentes conditions.

Merci à Emmanuelle Leize pour la qualité de son encadrement scientifique, sa confiance et ses encouragements.

Merci aux personnes avec lesquelles j'ai été amenée à collaborer tout au long de ce travail. En particulier, je remercie Thierry Rabilloud, Luc Moulinier, Catherine Leblanc et Carole Colin pour m'avoir communiqué leur enthousiasme au travers d'échanges constructifs et enrichissants.

Merci à Catherine Florentz, Patrick Jouin, Jérôme Lemoine et Bruno Schoentjes pour avoir accepté d'évaluer mon travail.

Merci à toute l'équipe du LSMBO pour sa bonne humeur au quotidien. Un grand merci à Jean-Marc, Fabrice et Raymond d'avoir partagé avec gentillesse leur savoir et leur expérience. Merci Manuella et Flo, pour votre amitié. Merci Sarah pour ton aide et tes précieux conseils.

Enfin, un immense merci à ma famille de m'avoir épaulée durant toute la durée de mes études. Mes plus tendres remerciements à Xavier pour sa patience, son réconfort et ses encouragements sans lesquels ce travail n'aurait pu aboutir.

A toi, petit frère,

A nos parents.

PLAN GENERAL

<u>PLAN GÉNÉRAL</u>	1
<u>PLAN DÉTAILLÉ</u>	3
<u>LISTE DES PRINCIPALES ABRÉVIATIONS</u>	11
<u>INTRODUCTION GENERALE</u>	13
PARTIE I : INTRODUCTION - BIBLIOGRAPHIE	17
<u>I. La spectrométrie de masse appliquée à l'étude de biomolécules</u>	17
<u>II. L'analyse protéomique</u>	46
<u>III. La caractérisation des modifications post-traductionnelles par spectrométrie de masse</u> ..	60
PARTIE II : <u>MISE EN ŒUVRE EXPERIMENTALE</u>	69
LE COUPLAGE NANOLC-MS/MS AVEC UNE TRAPPE IONIQUE	
<u>I. Introduction</u>	69
<u>II. Les systèmes HPLC à faible débit</u>	71
<u>III. Couplage avec la trappe ionique</u>	78
<u>IV. Traitement des données</u>	81
<u>V. Optimisation des paramètres propres au couplage LC-MS/MS</u>	87
<u>VI. Conclusions et perspectives</u>	97
<u>VII. Note d'application</u>	99
PARTIE III : <u>CARACTÉRISATION PAR APPROCHE PROTÉOMIQUE DE PROTÉINES IMPLIQUÉES DANS LE SYSTÈME DE DÉFENSE DES CELLULES CONTRE LE STRESS OXYDANT</u>	107
<u>I. Introduction : les peroxyrédoxines</u>	107
<u>II. Mise en œuvre d'une nouvelle stratégie protéomique pour la caractérisation du site d'oxydation de peroxyrédoxines lors d'un stress oxydant</u>	109

<u>III. Etude de la régénération des peroxyrédoxines suite à un stress oxydant : marquage <i>in vivo</i> et approche protéomique quantitative</u>	130
<u>BIBLIOGRAPHIE</u>	150
PARTIE IV : <u>CARACTÉRISATION D'ENZYMES PURIFIÉES CHEZ LES ALGUES BRUNES</u> :	152
UNE APPROCHE PROTÉOMIQUE DANS LE CAS OÙ LES PROTÉINES SONT INCONNUES DES BANQUES DE DONNÉES	
<u>I. Introduction</u>	152
<u>II. La stratégie par séquençage <i>de novo</i></u>	155
<u>III. Résultats</u>	162
<u>IV. Conclusion et perspectives</u>	164
<u>V. Présentation de ce travail</u>	165
<u>BIBLIOGRAPHIE</u>	189
PARTIE V : <u>CARACTÉRISATION DE GLYCANNES PAR SPECTROMÉTRIE DE MASSE MS/MS</u>	190
<u>I. Contexte biologique : les gonadotropines</u>	190
<u>II. Caractérisation par MS/MS de N-glycannes</u>	191
<u>III. Développement d'un outil bio-informatique pour assister l'interprétation des spectres de fragmentation des glycannes</u>	211
<u>BIBLIOGRAPHIE</u>	217
<u>CONCLUSION GENERALE</u>	219
<u>PARTIE EXPÉRIMENTALE</u>	223
<u>I. Les spectromètres de masse</u>	223
<u>II. Les systèmes chromatographiques</u>	228
<u>III. L'analyse protéomique</u>	231

PLAN DETAILLE

<u>PLAN GÉNÉRAL</u>	1
<u>PLAN DÉTAILLÉ</u>	3
<u>LISTE DES PRINCIPALES ABRÉVIATIONS</u>	11
<u>INTRODUCTION GENERALE</u>	13
<u>PARTIE I : INTRODUCTION - BIBLIOGRAPHIE</u>	17
<u>I. La spectrométrie de masse appliquée à l'étude de biomolécules</u>	17
<u>1. Spectrométrie de masse à source électrospray (ES)</u>	17
<u>1.1. Introduction</u>	17
<u>1.2. La source électrospray</u>	18
<u>1.2.1. La production de gouttelettes chargées</u>	18
<u>1.2.2. La fission des gouttelettes chargées en gouttelettes filles</u>	19
<u>1.2.3. L'émission des ions en phase gazeuse</u>	19
<u>a. Le modèle de Dole : le modèle de la charge résiduelle</u>	19
<u>b. Le modèle d'Iribarne et Thomson : le modèle de l'évaporation ionique</u>	19
<u>1.2.4. La source nano-électrospray</u>	20
<u>1.3. L'interface</u>	21
<u>1.3.1. Exemples de géométries d'interfaces</u>	21
<u>1.3.2. Les dissociations induites par collisions (CID) dans l'interface</u>	24
<u>1.4. Les analyseurs</u>	24
<u>1.4.1. Introduction</u>	24
<u>1.4.2. La trappe ionique</u>	25
<u>a. Introduction</u>	25
<u>b. Description</u>	25
<u>c. Les étapes de l'analyse MS</u>	26
<u>d. Les étapes de l'analyse MSⁿ</u>	27
<u>1.4.3. Un analyseur hybride : le Q-TOF</u>	29
<u>2. Spectrométrie de masse MALDI-TOF</u>	30
<u>2.1. Introduction</u>	30
<u>2.2. La source MALDI</u>	30
<u>2.2.1. Le rôle de la matrice</u>	30
<u>2.2.2. La préparation de l'échantillon</u>	30
<u>a. Le choix de la matrice</u>	31
<u>b. La préparation du dépôt</u>	31
<u>c. Les cibles utilisées pour le dépôt</u>	32
<u>2.2.3. Le mécanisme d'ionisation MALDI</u>	32
<u>a. Excitation des molécules de matrice</u>	33
<u>b. Emission des ions en phase gazeuse</u>	33

c. Ionisation des molécules en phase gazeuse	33
2.3. L'analyseur à temps de vol	34
2.3.1. Le MALDI-TOF linéaire	34
2.3.2. Le MALDI-TOF avec réflecteur et extraction retardée	35
a. L'extraction retardée	35
b. Le réflecteur électrostatique	35
2.4. La fragmentation en MALDI-TOF	36
2.4.1. Le PSD ou « Post Source Decay »	36
a. Introduction	36
b. Rôle de la matrice en PSD	37
c. Analyse en PSD	37
2.4.2. Le MALDI TOF-TOF	38
a. Introduction	38
b. Principe du LIFT et comparaison avec le PSD	39
II. L'analyse protéomique	46
1. Introduction	46
2. Le gel d'électrophorèse bidimensionnel	47
2.1. Principe du gel 2D-PAGE	47
2.2. La coloration des gels 2D	48
2.3. L'analyse des gels	48
2.4. Récentes améliorations et limites de la technique	49
3. L'apport de la spectrométrie de masse pour l'analyse des peptides en protéomique	49
3.1. La fragmentation des peptides	49
3.2. Nomenclature des fragmentations peptidiques	50
3.3. La détermination de la séquence peptidique	50
4. Place de la spectrométrie de masse dans la stratégie protéomique	51
4.1. La stratégie de l'empreinte peptidique massique	51
4.2. La stratégie par MS/MS	52
4.3. Le couplage avec la chromatographie liquide	52
4.3.1. La stratégie LC-MS/MS à faible débit	52
4.3.2. La LC-LC-MS/MS	53
5. La quantification des protéines	55
5.1. Introduction	55
5.2. La technologie ICAT	55
5.3. Une autre approche : ICNAT	56
6. Conclusion	56
III. La caractérisation des modifications post-traductionnelles par spectrométrie de masse	60
1. Introduction	60
2. Les principales modifications post-traductionnelles	60
3. La glycosylation des protéines	60
3.1. Introduction	60
3.2. Nature et structure des glycosylations	61
3.3. La caractérisation des glycosylations	62
3.3.1. Isolement et purification des glycanes	62
3.3.2. L'analyse par RMN	63
3.3.3. L'analyse par spectrométrie de masse	63
3.4. Les apports de la MS ⁿ pour la caractérisation structurale des oligosaccharides	63
3.4.1. Nomenclature de fragmentation des polysaccharides	64
3.4.2. Les fragmentations des polysaccharides à haute et basse énergie	64
3.4.3. La détermination de la séquence et du motif de branchement des polysaccharides	65
3.4.4. La détermination des isoméries de position	65

LE COUPLAGE NANOLC-MS/MS AVEC UNE TRAPPE IONIQUE

<u>I. Introduction</u>	69
<u>II. Les systèmes HPLC à faible débit</u>	71
<u>1. Le MicroBlotter</u> -----	71
<u>1.1. Le système chromatographique</u> -----	71
<u>1.2. Analyse d'un « mélange peptidique test »</u> -----	71
<u>1.3. Résultats et optimisations</u> -----	72
<u>1.3.1. Le gradient</u> -----	72
<u>1.3.2. Sensibilité et détection UV</u> -----	73
<u>1.4. Difficultés rencontrées</u> -----	73
<u>1.4.1. En chromatographie</u> -----	73
<u>1.4.2. En MS</u> -----	74
<u>1.5. Les limites du MicroBlotter</u> -----	74
<u>2. Le système Agilent</u> -----	75
<u>2.1. Le système chromatographique</u> -----	75
<u>2.1.1. Principe de fonctionnement de la pompe et régulation du débit</u> -----	75
<u>2.1.2. LC capillaire et nanoLC</u> -----	76
<u>2.1.3. Les colonnes chromatographiques</u> -----	76
<u>2.2. Le montage de pré-concentration</u> -----	77
<u>III. Couplage avec la trappe ionique</u>	78
<u>1. Configuration de la source</u> -----	78
<u>2. Acquisition en mode MS/MS automatique</u> -----	79
<u>2.1. Le cycle MS-MS/MS</u> -----	79
<u>2.2. Sélection du précurseur</u> -----	80
<u>2.3. Isolement et fragmentation</u> -----	81
<u>IV. Traitement des données</u>	81
<u>1. Les tracés chromatographiques</u> -----	81
<u>2. Génération de la « peak list »</u> -----	82
<u>2.1. Introduction</u> -----	82
<u>2.2. Détection des composés</u> -----	83
<u>3. Interrogation dans les banques de données</u> -----	84
<u>3.1. Paramètres de la recherche</u> -----	84
<u>3.2. Résultats</u> -----	85
<u>4. Automatisation du traitement des données</u> -----	86
<u>V. Optimisation des paramètres propres au couplage LC-MS/MS</u>	87
<u>1. Le gradient de chromatographie</u> -----	87
<u>2. Tests de sensibilité</u> -----	87
<u>3. Tests de reproductibilité</u> -----	89
<u>4. Les principaux paramètres d'acquisition avec la trappe ionique</u> -----	92
<u>5. Tests des paramètres d'acquisition MS/MS : « Averages et ICC »</u> -----	92
<u>5.1. Introduction</u> -----	92
<u>5.2. Résultats</u> -----	93
<u>5.3. Analyses des résultats et interprétation</u> -----	93

5.3.1. Le nombre d' « averages »	94
5.3.2. L'ICC	94
5.3.3. Conclusion	95
6. Choix du seuil pour la sélection du précurseur	95
6.1. Introduction	95
6.2. Résultats	95
6.3. Analyse et interprétation des résultats	96
VI. Conclusions et perspectives	97
1. Conclusions	97
2. Evaluation du système	97
3. Limites et perspectives	98
VII. Note d'application	99
<u>PARTIE III : CARACTÉRISATION PAR APPROCHE PROTÉOMIQUE DE PROTÉINES IMPLIQUÉES DANS LE SYSTÈME DE DÉFENSE DES CELLULES CONTRE LE STRESS OXYDANT</u>	107
<u>I. Introduction : les peroxyrédoxines</u>	107
<u>II. Mise en œuvre d'une nouvelle stratégie protéomique pour la caractérisation du site d'oxydation de peroxyrédoxines lors d'un stress oxydant</u>	109
1. Observations issues des gels 2D	109
2. Les différents états d'oxydation de la cystéine	110
3. Stratégie développée : choix de la protéase	111
4. Résultats de spectrométrie de masse	112
4.1. Stratégie de caractérisation des peroxyrédoxines par MS	112
4.2. Résultats obtenus en mode d'ionisation positif	112
4.3. Nécessité de réaliser des analyses MALDI-MS en mode d'ionisation négatif	112
4.3.1. Choix de la matrice MALDI pour le mode d'ionisation négatif	113
a. Introduction	113
b. Résultats	113
c. Discussion	113
4.3.2. Compléments d'analyse et nouveaux résultats	115
a. Détection du peptide « complémentaire »	115
b. Détection du peptide « total »	116
5. Résumé des résultats et conclusion	117
5.1. Bilan des résultats	117
5.2. Conclusion	117
6. Présentation de ce travail	118
<u>III. Etude de la régénération des peroxyrédoxines suite à un stress oxydant : marquage <i>in vivo</i> et approche protéomique quantitative</u>	130
1. Résumé de notre approche	130
1.1. Introduction	130
1.2. Expérience n°1	131
1.3. Expériences n°2	133
1.3.1. Sans inhibiteur de la néosynthèse	133
1.3.2. Avec inhibiteur de la néosynthèse	133
2. Difficultés rencontrées lors de cette étude	134
2.1. Le choix du marqueur	134
2.1.1. Introduction	134

2.1.2. La lysine deutérée	135
2.2. La quantification par spectrométrie de masse	135
2.2.1. Introduction	135
2.2.2. Principe du calcul des ratios	136
2.2.3. Calcul des ratios en MALDI-MS	136
2.2.4. Calcul des ratios en nanoLC-MS/MS	138
a. Détection des couples de peptides d'intérêt	138
b. Co-élution des peptides marqués et non marqués	138
c. Calcul des ratios	139
2.2.5. Comparaison des résultats obtenus en MALDI-MS et en nanoLC-MS/MS	139
3. Conclusion	140
4. Présentation de ce travail	141
BIBLIOGRAPHIE	150

PARTIE IV : CARACTÉRISATION D'ENZYMES PURIFIÉES CHEZ LES ALGUES BRUNES : 152

UNE APPROCHE PROTÉOMIQUE DANS LE CAS OÙ LES PROTÉINES SONT INCONNUES DES BANQUES DE DONNÉES

<u>I. Introduction</u>	152
1. <u>Contexte biologique</u>	152
2. <u>Particularités des haloperoxydases de <i>Laminaria digitata</i> pour l'approche protéomique</u>	154
2.1. <u>Quelles sont les informations dont nous disposons?</u>	154
2.2. <u>Objectifs de l'étude et intérêt de la spectrométrie de masse dans ce contexte</u>	154
<u>II. La stratégie par séquençage <i>de novo</i></u>	155
1. <u>Stratégie utilisée pour notre étude</u>	155
2. <u>Mise en œuvre</u>	155
2.1. <u>Séquençage <i>de novo</i> « manuel »</u>	156
2.1.1. <u>Déconvolution</u>	156
2.1.2. <u>Séries b et y</u>	156
2.1.3. <u>Informations complémentaires sur la composition en acides aminés</u>	157
a. <u>Les fragments internes</u>	157
b. <u>Les ions immoniums</u>	157
2.1.4. <u>Ambiguïtés pour la détermination de la séquence peptidique</u>	158
2.2. <u>Séquençage <i>de novo</i> « automatique » pour la LC-MS/MS</u>	158
2.3. <u>Comparaison des résultats MS/MS obtenus avec la trappe ionique et le Q-TOF</u>	159
<u>III. Résultats</u>	162
1. <u>Les bromoperoxydases</u>	162
2. <u>Les iodoperoxydases</u>	164
<u>IV. Conclusion et perspectives</u>	164
<u>V. Présentation de ce travail</u>	165
<u>BIBLIOGRAPHIE</u>	189

PARTIE V : CARACTÉRISATION DE GLYCANNES PAR SPECTROMÉTRIE DE MASSE MS/MS 190

<u>I. Contexte biologique : les gonadotropines</u>	190
<u>II. Caractérisation par MS/MS de N-glycannes</u>	191
1. <u>Mise en œuvre</u>	191
2. <u>La caractérisation du N-glycane NA₂</u>	192
2.1. <u>MS/MS avec la trappe ionique</u>	192
2.1.1. <u>Fragments Y et B</u>	194
2.1.2. <u>Fragments internes</u>	194
2.1.3. <u>Le motif de branchement</u>	194
2.1.4. <u>Clivages intracycliques</u>	195
2.2. <u>MS/MS avec le TOF-TOF</u>	196
2.2.1. <u>Analyse des résultats et comparaison avec la MS/MS sur la trappe ionique</u>	197
2.2.2. <u>Comparaison des modes LID et CID</u>	198
2.2.3. <u>Les isoméries de liaison</u>	199

3. La caractérisation du N-glycane NA ₂ F	200
3.1. MS/MS avec la trappe ionique	200
3.2. MS/MS avec le TOF-TOF	201
4. La caractérisation du N-glycane NA ₃	204
4.1. MS/MS avec la trappe ionique	204
4.2. MS/MS avec le TOF-TOF	206
5. Conclusion	209
III. Développement d'un outil bio-informatique pour assister l'interprétation des spectres de fragmentation des glycanes	211
1. Introduction	211
2. Stratégies possibles	211
2.1. Stratégie 1	211
2.2. Stratégie 2	212
2.3. Stratégie choisie	213
3. Mise en oeuvre	213
4. Présentation du logiciel	214
5. Conclusion	215
BIBLIOGRAPHIE	217
CONCLUSION GENERALE	219
PARTIE EXPÉRIMENTALE	223
I. Les spectromètres de masse	223
1. Spectrométrie de masse MALDI	223
1.1. Appareillage utilisé	223
1.2. Les principaux paramètres de l'analyse MALDI-TOF	223
1.3. Les principaux paramètres de l'analyse MALDI-TOF-TOF	224
1.4. L'étalonnage du spectromètre de masse MALDI-TOF	224
1.5. La préparation de l'échantillon	224
2. Spectrométrie de masse à source électrospray	224
2.1. Le spectromètre de masse Q-TOF	225
2.1.1. Introduction des échantillons pour les analyses nano-ESI et nanoLC-ESI	225
2.1.2. Etalonnage du Q-TOF	225
2.2. La trappe ionique	225
2.2.1. Les performances de l'ESQUIRE 3000+	226
2.2.2. Introduction des échantillons	226
2.2.3. Etalonnage de la trappe ionique	227
2.3. La préparation des échantillons pour l'analyse nanoES	227
II. Les systèmes chromatographiques	228
1. Le MicroBlotter	228
2. Le système Agilent	228
2.1. L'HPLC	228
2.2. Le système de pré-concentration	229
2.3. Les gradients testés	230
3. La CapLC	230

<u>III. L'analyse protéomique</u>	231
<u>1. Réduction / alkylation / digestion</u>	231
<u>1.1. Lavage des spots</u>	231
<u>1.2. La réduction et l'alkylation</u>	232
<u>1.3. La digestion trypsique</u>	232
<u>2. L'extraction des peptides</u>	232

LISTE DES PRINCIPALES ABREVIATIONS

ADN : acide désoxyribonucléique

ADNc : ADN complémentaire

ARN : acide ribonucléique

BPO : bromoperoxydase

CID : dissociation induite par collision

Da : Dalton

ES : électrospray

ESI-MS : spectrométrie de masse à ionisation électrospray

EST : expressed sequence tag (étiquette de séquence transcrite)

eV : électronvolt

HPLC : chromatographie liquide haute performance

IPG : gradient de pH immobilisés

IPO : iodoperoxydase

LC-MS : couplage chromatographie liquide – spectrométrie de masse

m/z : rapport masse sur charge

MALDI : ionisation / désorption laser assistée par matrice

MS : spectrométrie de masse

MS/MS : spectrométrie de masse en tandem

Prx : peroxyrédoxine

PSD : Post Source Decay

Q : quadripôle

RF : radio-fréquence

TOF : temps de vol

INTRODUCTION GENERALE

INTRODUCTION GENERALE

OBJECTIFS DE CE TRAVAIL DE THESE

L'année 2002 a été marquée par l'attribution du prix Nobel de chimie à John B. Fenn et Koichi Tanaka pour le développement, respectivement, de l'ionisation ES (électrospray) et MALDI (ionisation / désorption laser assistée par matrice) qui a révolutionné, en 1988, l'analyse par spectrométrie de masse de macromolécules biologiques. Cette récompense illustre l'importance prise par la spectrométrie de masse ces dernières années pour résoudre des problèmes biologiques, notamment pour l'identification et la caractérisation des protéines.

Au début de mon travail de thèse, la spectrométrie de masse MALDI (MALDI-MS) devenait un outil de choix pour l'approche protéomique permettant l'identification des protéines analysées en quantités picomolaires.

Cependant, devant la diversité et la complexité des problèmes biologiques, une approche protéomique classique, qui consistait en la mesure de la masse des peptides tryptiques par MALDI-MS, n'était pas toujours suffisante.

Ainsi, mon travail de thèse a consisté à évaluer et à optimiser les approches par spectrométrie de masse en tandem (MS/MS), dans le cadre de l'analyse protéomique, afin d'obtenir des informations de séquences nécessaires pour passer du stade d'identification des protéines au stade de leur caractérisation (recouvrement de séquence élevé, détermination des modifications post-traductionnelles ou chimiques).

Pour rendre possible la détermination des séquences protéiques par MS/MS, il nous a fallu concilier sensibilité (quantités subpicomolaires) avec rapidité et reproductibilité d'analyse. Pour cela, nous avons :

- | développé et optimisé le couplage d'un système nanoLC (HP 1100 Series NanoHPLC System, Agilent Technologies) avec un instrument MS/MS de type trappe ionique (ESQUIRE 3000+, Bruker Daltonics) pour l'analyse automatisée de mélanges peptidiques issus de la digestion tryptique des protéines séparées par gel d'électrophorèse.
- | évalué et optimisé les performances d'un instrument MALDI-TOF-TOF (ULTRAFLEX, Bruker Daltonics) pour la caractérisation des glycosylations qui constituent une des modifications post-traductionnelles protéiques majeures.

L'ensemble de ces avancées instrumentales nous ont permis de résoudre différents problèmes biologiques faisant appel à l'analyse protéomique et pour lesquels la seule approche par MALDI-MS aurait été insuffisante.

- | En effet, dans ces études, nous avons été amenés à caractériser, d'une part, des modifications chimiques (oxydation de cystéines) et des modifications post-traductionnelles (glycosylations) nécessitant l'utilisation d'approches MS/MS (trappe ionique, Q-TOF, TOF-TOF).
- | De plus, pour d'autres études (enzymes purifiées chez les algues brunes), aucune information protéique n'était disponible dans les banques de données nécessitant le recours à la recherche par homologie de séquences uniquement possible à partir des données de séquences générées par MS/MS.

Ce manuscrit se compose donc de cinq parties.

Dans une première partie, l'introduction fait, tout d'abord, le point sur l'état des techniques de spectrométrie de masse appliquées à l'étude des biomolécules en analyse protéomique. Nous exposons ensuite les récentes évolutions de la spectrométrie de masse pour l'analyse protéomique et pour la caractérisation des modifications post-traductionnelles en envisageant le cas particulier des glycosylations des protéines.

La deuxième partie de ce travail de thèse est consacrée à **la mise en œuvre expérimentale** du couplage nanoLC-MS/MS avec une trappe ionique. Nous décrivons notre travail d'optimisation des paramètres chromatographiques, de l'interface du système nanoLC avec la trappe ionique et des paramètres d'acquisition de la trappe ionique en MS et en MS/MS automatique (avec notamment, la sélection des ions précurseurs). Nous présentons les résultats obtenus lors de l'évaluation de ce système nanoLC-MS/MS en terme de sensibilité et de reproductibilité pour l'analyse protéomique.

Partie III : Caractérisation par approche protéomique de protéines impliquées dans le système de défense des cellules contre le stress oxydant.

Deux problématiques ont été abordées au cours de cette étude :

- | Dans un premier temps, nous avons développé une nouvelle stratégie protéomique pour la caractérisation du site d'oxydation de peroxyrédoxines (prxs) lors d'un stress oxydant. Notre objectif a consisté à localiser et à caractériser, par spectrométrie de masse, une modification chimique sur ces protéines préalablement séparées par gel d'électrophorèse bi-dimensionnelle.
- | Dans un second temps, nous avons étudié la régénération des prxs suite à un stress oxydant en utilisant un marquage *in vivo* des protéines suivi d'une approche protéomique quantitative.

Partie IV : Caractérisation par séquençage *de novo* d'enzymes purifiées chez les algues brunes.

Pour cette étude, nous avons mis en œuvre une approche protéomique *de novo* dans le cas particulier où les protéines étudiées sont inconnues des banques de données. Nous avons analysé par nanoLC-MS/MS deux types d'haloperoxydases purifiées chez les Laminaires et ayant des activités distinctes. Puis, nous avons séquençé *de novo* les protéines étudiées afin de répondre aux objectifs suivants :

- | Pour la famille des bromoperoxydases, nous disposions d'informations préalables (ADNc et données protéiques pour une activité similaire chez d'autres familles d'algues). Nous avons donc recherché les homologues de séquence entre les protéines analysées et ces données pré-existantes.
- | Pour la famille des iodoperoxydases, nous ne disposions d'aucune information préalable. Aussi nous avons adopté une démarche plus ciblée consistant à obtenir des informations de séquences permettant aux biologistes de construire des sondes (ou « primers dégénérés ») afin de cloner des ADNc correspondant à l'activité iodoperoxydase.

Partie V : Caractérisation de glycosylations par spectrométrie de masse MS/MS.

- | Dans un cadre général, correspondant à l'étude des glycosylations de protéines recombinantes, nous avons entrepris d'évaluer les capacités de deux instruments, une trappe ionique (ESQUIRE 3000+, Bruker Daltonics) et un MALDI TOF-TOF (ULTRAFLEX, Bruker Daltonics) pour l'élucidation par MS/MS de glycosylations connues.
- | Dans ce contexte, face aux difficultés rencontrées pour l'interprétation des spectres de fragmentation obtenus, nous avons souhaité développer un algorithme facilitant l'interprétation des spectres générés par MS/MS de glycanes complexes. Ce travail a été réalisé en collaboration avec un informaticien pour la programmation du logiciel.

Ce travail de thèse a été financé par la société Bruker qui m'a également apporté un support technique et scientifique.

PARTIE I :

INTRODUCTION – BIBLIOGRAPHIE

CHAPITRE I : LA SPECTROMETRIE DE MASSE APPLIQUEE A L'ETUDE DE BIOMOLECULES

CHAPITRE II : L'ANALYSE PROTEOMIQUE

CHAPITRE III : LA CARACTERISATION DES MODIFICATIONS POST-TRADUCTIONNELLES PAR SPECTROMETRIE DE MASSE

INTRODUCTION - BIBLIOGRAPHIE

I. LA SPECTROMETRIE DE MASSE APPLIQUEE A L'ETUDE DE BIOMOLECULES

Ce premier chapitre est consacré aux principales techniques de spectrométrie de masse utilisées dans le cadre de ce travail de thèse pour des approches protéomiques. Nous distinguerons, d'une part, les spectromètres de masse à source électrospray (ES) et d'autre part, les instruments à ionisation / désorption laser assistée par matrice (MALDI).

De façon générale, un spectromètre de masse se compose de quatre régions distinctes où se déroulent successivement les étapes nécessaires à l'analyse des ions :

- | la source où se produisent l'ionisation-désorption des analytes (MALDI) ou la vaporisation (ES)
- | l'interface qui assure la transmission des ions depuis la source vers l'analyseur grâce à un gradient de pression et / ou de potentiel
- | l'analyseur qui a pour rôle de séparer les ions selon leur rapport masse sur charge (m/z)
- | le détecteur.

Dans ce paragraphe, nous décrivons la source, l'interface et les analyseurs d'appareils opérant en mode d'ionisation ES et en mode MALDI. Nous nous intéresserons principalement aux instruments disponibles au laboratoire et qui nous ont été utiles pour développer de nouvelles approches protéomiques par spectrométrie de masse en tandem : une trappe ionique, un Q-TOF et un MALDI-TOF-TOF.

1. Spectrométrie de masse à source électrospray (ES)

1.1. Introduction

Le phénomène d'ionisation par ES, pour générer des ions en phase gazeuse à partir de macromolécules en solution, a été décrit par Dole et coll. en 1968 [Dole et coll. 1968]. Un moment délaissée, cette technique d'ionisation est couplée à un analyseur quadripolaire dans les années 80 [Yamashita et Fenn, 1984]. Rapidement, sont présentés les premiers spectres de masse d'ions multichargés de protéines allant jusqu'à 40 kDa [Meng et coll., 1988]. L'année 1989 est marquée par la commercialisation du premier spectromètre de masse à source ES.

Particulièrement adaptée à l'analyse de composés polaires et thermosensibles, la technique d'ionisation ES a permis d'augmenter considérablement le champ d'application de la spectrométrie de masse, spécialement dans le domaine de l'analyse des biomolécules. En effet, le processus

d'ionisation ES a pour caractéristique de former des ions multichargés ; des composés biologiques de plus en plus lourds (200 kDa) et de plus en plus complexes ont ainsi pu être caractérisés avec une précision de mesure remarquable (0,01 %) en utilisant des appareils qui ne possèdent pas nécessairement une gamme de mesure très étendue.

Par ailleurs, ces espèces multichargées sont produites par électronébulisation d'ions à partir d'une phase liquide à pression atmosphérique. Aussi, la source ES s'est révélée idéale pour le couplage avec la chromatographie liquide (LC/MS) [Hiraoka et coll., 1990].

1.2. La source électrospray

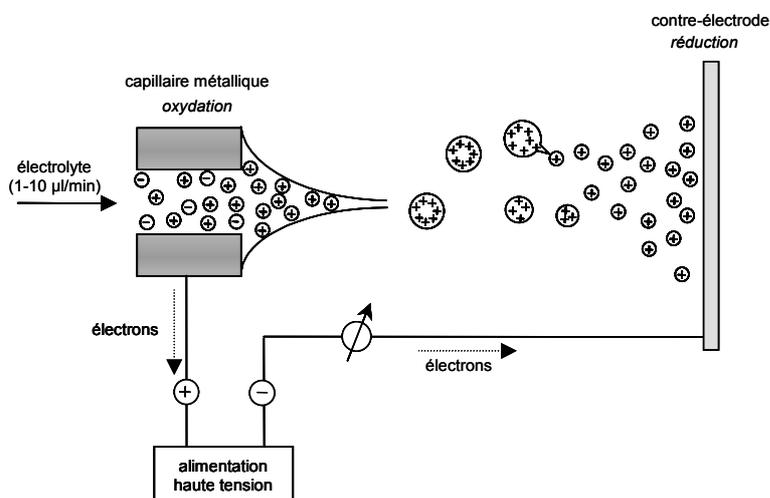
Dans ce type de source, le processus de formation d'ions désolvatés en phase gazeuse à partir des ions initialement présents dans le liquide se produit sous l'effet d'un champ électrique intense à pression atmosphérique. Il peut être décomposé selon trois étapes [Kearle et Ho, 1997 ; Kearle, 2000 ; Kearle et Peschke, 2000] :

- | la production de gouttelettes chargées à partir de l'électrolyte en solution
- | la fission des gouttelettes chargées en gouttelettes de tailles plus petites appelées gouttelettes filles
- | l'émission d'ions désolvatés en phase gazeuse.

1.2.1. La production de gouttelettes chargées

Le processus ES est obtenu par application d'un champ électrique intense (10^6 V.m^{-1}) à la pointe d'un capillaire métallique dans lequel est introduite la solution à analyser. Ce champ électrique provoque la polarisation du liquide et la séparation des charges positives et négatives. A l'extrémité du capillaire se forme un cône de liquide allongé appelé « cône de Taylor » qui éclate en minces gouttelettes chargées (figure 1) [Hayati et coll., 1987 ; Ikonomou et coll., 1991 ; Blades et coll., 1991 ; Cole, 2000]. Ce mécanisme est en général assisté pneumatiquement par un gaz coaxial (par exemple de l'azote).

Figure 1 : représentation schématique de la production des gouttelettes chargées lors du processus ES en mode positif [d'après Kearle et Tang, 1993]. La source ES se comporte comme une cellule électrolytique dans laquelle le capillaire est l'anode et la contre-électrode, la cathode.



1.2.2. La fission des gouttelettes chargées en gouttelettes filles

L'évaporation du solvant contenu dans les gouttelettes mères à la pointe du cône de Taylor induit une diminution progressive de la taille de ces gouttelettes. De ce fait, la densité de charge de la gouttelette augmente pour atteindre un rayon critique appelé rayon limite de Rayleigh (équation 1). Les forces de répulsion coulombiennes deviennent alors supérieures aux forces de cohésion du liquide ce qui entraîne la rupture de la gouttelette mère en gouttelettes filles. Les gouttelettes filles, de tailles hétérogènes, possèdent toutes une densité de charge plus élevée que la gouttelette mère. Ainsi, la fission se poursuit sur plusieurs générations de gouttelettes filles de plus en plus petites [Doyle et coll., 1964 ; Taflin et coll, 1989 ; Wang et Cole, 1995 ; Kebarle et Tang, 1993 ; Gomez et Tang, 1994].

Equation 1 : Relation entre la charge d'une gouttelette et le rayon critique de Rayleigh

$$Q^2 = 64 \pi^2 \varepsilon_0 \gamma R_R^3$$

Avec Q : charge de la gouttelette

ε_0 : permittivité du vide

γ : tension superficielle du liquide

R_R : valeur critique du rayon de Rayleigh

1.2.3. L'émission des ions en phase gazeuse

Le processus selon lequel les ions sont émis en phase gazeuse est encore aujourd'hui assez controversé. Deux modèles sont communément admis : le modèle de Dole [Dole et coll., 1968] et le modèle d'Iribarne et Thomson [Iribarne et Thomson, 1976].

a. Le modèle de Dole : le modèle de la charge résiduelle

Pour Dole, une succession d'explosions coulombiennes conduirait à la formation d'une dernière gouttelette fille ne contenant plus qu'un seul ion. L'ion en phase gazeuse résulterait de l'évaporation du solvant résiduel de cette ultime gouttelette (figure 2).

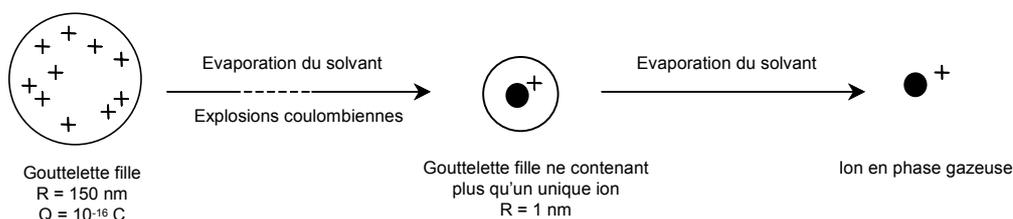
b. Le modèle d'Iribarne et Thomson : le modèle de l'évaporation ionique

Cette théorie suggère que les ions en phase gazeuse sont directement extraits à partir de la surface de la gouttelette lorsque le diamètre de celle-ci est suffisamment petit (inférieur à 10nm) (figure 2).

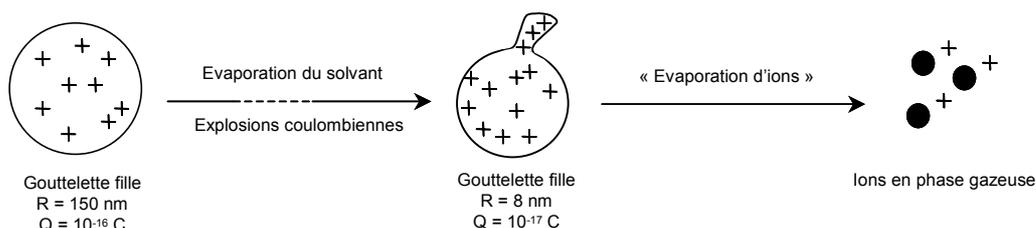
Dans l'état actuel des connaissances, aucune preuve définitive en faveur de l'un ou l'autre de ces deux modèles n'a été fournie. Cependant, des observations expérimentales confirment le modèle de Dole, en particulier pour les molécules de hautes masses [Fernandez de la Mora et coll., 2000]. Mais le débat reste ouvert concernant la désorption des petits ions [Kebarle et Tang, 1993 ; Gamero-Castano et Fernandez de la Mora et coll., 2000].

Figure 2 : émission des ions en phase gazeuse lors du processus d'ionisation/désorption ES. D'après [Leize E., thèse, 1994].

Modèle de Dole



Modèle d'Iribarne et Thomson



1.2.4. La source nano-électrospray

Le principal problème auquel se heurtent les études concernant les biomolécules est la faible quantité d'échantillon disponible pour l'analyse. Cette difficulté a conduit à la miniaturisation des sources ES conventionnelles visant à réduire le débit d'échantillon délivré au niveau de la source.

En 1994, Emmet et Caprioli développent un dispositif micro-électrospray permettant de travailler à des débits inférieurs à $1 \mu\text{L}/\text{min}$ [Emmet et Caprioli, 1994]. Puis Wilm et Mann introduisent la source nano-électrospray [Wilm et Mann, 1996]. Leur traité théorique sur la dispersion électrostatique durant le processus d'ionisation par nanoES montre qu'une réduction de la taille de la zone d'émission au sommet du spray permet de réduire à moins de 200 nm la taille des gouttelettes et d'obtenir un débit de l'ordre de $25 \text{ nL}/\text{min}$ [Wilm et Mann, 1994]. En effet, les gouttelettes de plus petite taille possèdent une plus grande densité de charge ce qui favorise le rendement des phénomènes d'ionisation/désorption des analytes en solution.

Ainsi, la technique nanoES présente le double avantage d'accroître la sensibilité de la mesure et de réduire la consommation d'échantillon. On notera également que, l'ionisation / désorption nanoES, contrairement à l'ES, ne requiert plus d'assistance pneumatique contribuant à la désolvation de l'analyte.

En pratique, on distingue deux types de source nanoES : la source nanoES dite « off-line » et la source nanoES dite « on-line » dont la configuration a été étudiée pour le couplage avec un système chromatographique.

- | La source nanoES off-line est constituée d'un capillaire en quartz métallisé de 1 à $3 \mu\text{m}$ de diamètre interne dans lequel est chargé un faible volume d'échantillon (1 à $5 \mu\text{L}$). Le débit de nébulisation est alors de quelques dizaines de nL/min .

- | La source nanoES on-line comporte un capillaire en silice fondue, métallisé, et connecté à la sortie d'une colonne analytique. Il présente à son extrémité un orifice d'environ 8 μm de diamètre interne. Le débit d'introduction de l'échantillon dans la source est ainsi dépendant de la pompe de chromatographie liquide (de l'ordre de 200 nL / min pour ces travaux de thèse).

Dans les deux cas, le capillaire de nébulisation est ajusté à proximité de l'entrée de l'interface du spectromètre (1 à 2 mm), ce qui nécessite d'adapter les différences de potentiel entre le capillaire métallique et la contre électrode (800 à 1500 V contre 4 kV en ES) afin de prévenir l'apparition de décharges de Corona.

La mise en œuvre expérimentale d'un montage nano-ES sollicite de nombreuses précautions et peut se révéler fastidieuse [Loew, thèse, 1999]. Cependant, la qualité analytique des sources nanoES en terme de sensibilité et de précision de mesure de masse de protéines, sa tolérance aux sels, ont permis une grande avancée ces dernières années dans le domaine de l'analyse protéomique [Körner et coll., 1996 ; Shevchenko et coll., 1996].

1.3. L'interface

L'interface est la région du spectromètre de masse qui assure la transmission et la focalisation des ions depuis la source, à pression atmosphérique, vers l'analyseur où règne un vide poussé de l'ordre de 10^{-6} mbar. L'entraînement des ions a lieu sous les effets combinés :

- | du gradient de pression entre les deux régions
- | des potentiels appliqués sur les lentilles électrostatiques.

1.3.1. Exemples de géométries d'interfaces

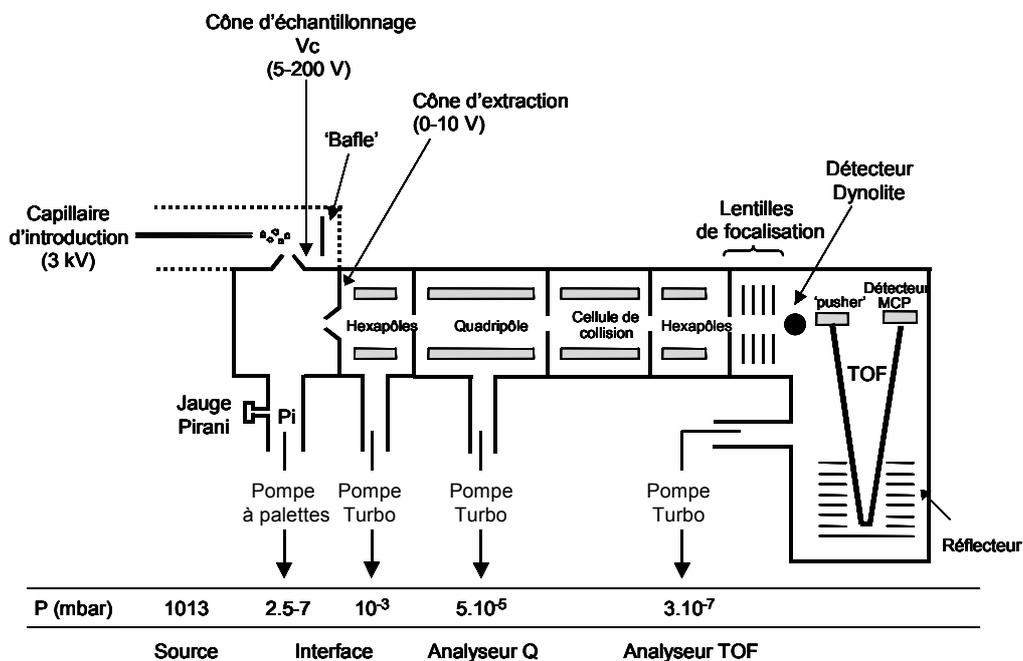
Au cours de ces dernières années, la configuration des interfaces a sans cesse été développée afin d'améliorer la transmission des ions vers l'analyseur.

Des multipôles (quadripôles, hexapôles ou octapôles) à l'entrée de l'analyseur fonctionnent en mode radio-fréquence (« RF-only ») : ils transmettent les ions vers l'analyseur, sans discrimination de masse, en assurant la refocalisation du faisceau ionique.

Par ailleurs, les sources à géométrie orthogonale sont de plus en plus répandues. Dans ce type de configuration, la source et l'interface sont orientées perpendiculairement. Ainsi, seules les espèces chargées, sous l'effet du champ de potentiel, seront déviées en direction de l'analyseur. Les espèces neutres, nuisibles à la sensibilité de la technique, sont donc écartées puisqu'elles poursuivent la trajectoire orientée par la source.

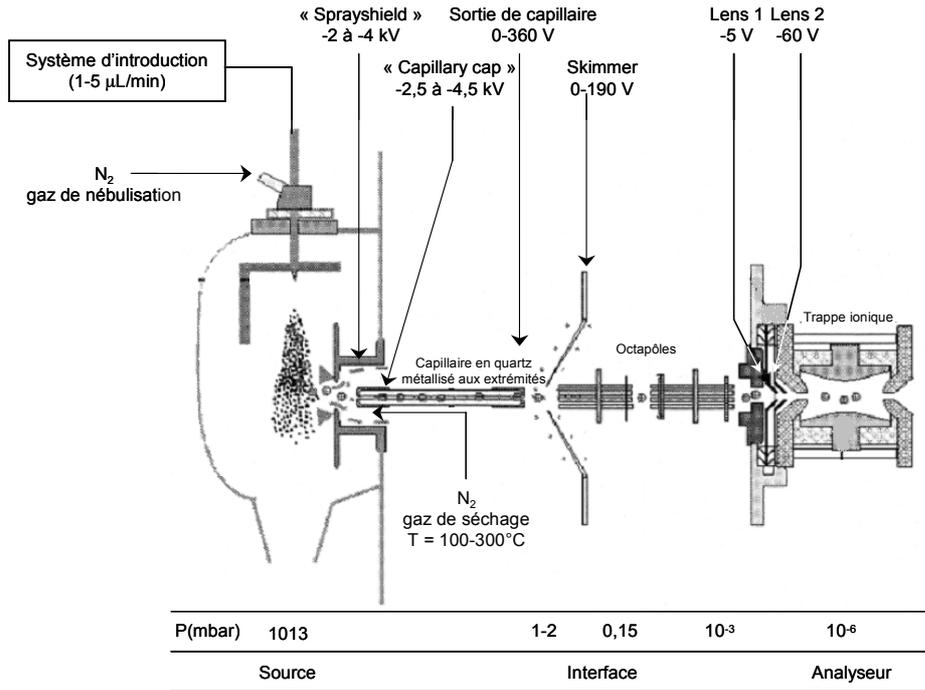
La figure 3 décrit les interfaces des deux instruments à source ES principalement utilisés pour la réalisation de ce travail de thèse : un appareil hybride ES-Q-TOF, le Q-TOFII (Micromass) et une trappe ionique, l'ESQUIRE 3000+ (Bruker Daltonics).

Figure 3 : configuration des interfaces de spectromètres de masse à source ES : A) Q-TOFII (Micromass) d'après [Sanglier, thèse, 2002] ; B) ESQUIRE 3000+ (Bruker Daltonics).



A) Source « Z-Spray » et interface : Q-TOFII (Micromass)

B) Source orthogonale (Hewlett-Packard) et interface : ESQUIRE 3000+ (Bruker Daltonics)



1.3.2. Les dissociations induites par collisions (CID) dans l'interface

Les ions sont accélérés dans la première région de l'interface sous l'effet de la différence de potentiel appliquée entre les premières lentilles électrostatiques du spectromètre (figure 3 : cône d'extraction et cône d'échantillonnage pour le Q-TOF, skimmer et sortie du capillaire métallisé pour la trappe ionique). Dans cette zone intermédiaire de pompage, la pression est encore de quelques millibars. Aussi, les ions peuvent entrer en collision avec des molécules d'azote résiduelles. Ces collisions faiblement énergétiques (quelques eV) contribuent à désolvater les ions et à dissocier les éventuels adduits (sels, métaux...).

L'énergie transmise aux ions lors du choc est d'autant plus importante que leur énergie cinétique est grande. Si la collision est trop énergétique, les molécules analysées peuvent se fragmenter au niveau de leurs liaisons les plus faibles. C'est ce que l'on appelle la dissociation induite par collision (CID). Les tensions d'accélération peuvent être optimisées de manière à favoriser les dissociations dans l'interface afin d'obtenir des informations structurales sur les composés analysés. On peut ainsi s'intéresser aux modifications post-traductionnelles établies sur un chaîne peptidique ou protéique [Light-Wahl et coll., 1993]. Cette approche peut également se révéler utile pour évaluer la stabilité relative de complexes non-covalents en phase gazeuse [Leize, thèse, 1994 ; Potier, thèse, 1996 ; Rogniaux, thèse, 2000 ; Nierengarten, thèse, 2001 ; Sanglier, thèse, 2002].

1.4. Les analyseurs

1.4.1. Introduction

Parmi les analyseurs couplés à une source ES, le triple quadripôle a été l'un des premiers à permettre la spectrométrie de masse en tandem (MS/MS) [Meng et coll., 1988]. Rapidement, les analyseurs à trappe ionique ont aussi été employés [Van Berkel et coll., 1990]. Outre leurs qualités en terme de sensibilité et de résolution, ils offrent la possibilité de réaliser des mesures en spectrométrie de masse multiples de type MS^n au cours desquelles les ions fragments d'un parent préalablement isolé peuvent être re-fragmentés en ions petits-fils et ceci, un nombre n de fois.

Quadripôles et trappes ioniques présentent cependant l'inconvénient d'offrir une gamme de m/z limitée. Aussi, pour l'analyse de composés de haut poids moléculaire faiblement chargés, les sources ES ont été couplées à des analyseurs et TOF (analyseur à temps de vol) [Verentchikov et coll, 1994], qui offrent, en théorie, une gamme de m/z illimitée.

La source ES est également compatible avec les analyseurs à secteurs magnétiques et électrostatiques [Gallagher et coll., 1990] ainsi qu'avec les FT-ICR (analyseurs à résonance cyclotronique et à transformée de Fourier) [Henry et coll., 1989] dont le fonctionnement repose, sur la mesure de la fréquence du courant induit sur une plaque de détection (analyseur offrant une très grande résolution, jusqu'à 1 million).

La tendance actuelle est aux géométries hybrides de type ES-Q-TOF [Morris et coll., 1997] développée pour des applications de MS/MS.

Dans ce paragraphe, nous décrivons les analyseurs des instruments principalement utilisés au cours de ce travail de thèse et exploités pour leur capacité à réaliser de la spectrométrie de masse en

tandem : l'ES-trappe ionique (ESQUIRE 3000+, Bruker Daltonics) et l'ES-Q-TOF (Q-TOFII, Micromass).

1.4.2. La trappe ionique

a. Introduction

Paul et Steinwedel ont décrit pour la première fois, dans les années 50, le filtre de masse quadripolaire appelé trappe ionique [Paul et Steinwedel, 1956]. Il a ensuite été modifié, à des fins commerciales, en analyseur pour la spectrométrie de masse par l'équipe de Stafford de la Finnigan Corporation [Stafford et coll., 1984].

Cet instrument peu encombrant, de coût raisonnable et offrant une grande sensibilité, a connu un rapide succès. Son atout principal est sa capacité à réaliser des analyses de spectrométrie de masse multiple (MS^n).

La trappe ionique a fait l'objet d'une thèse au laboratoire [Carte, thèse, 2001]. Son principe de fonctionnement est largement détaillé dans plusieurs ouvrages et publications [March, 1997 ; Cooks et coll., 1991 ; De Hoffmann et coll., Spectrométrie de masse]. Nous en rappellerons ici les points fondamentaux.

b. Description

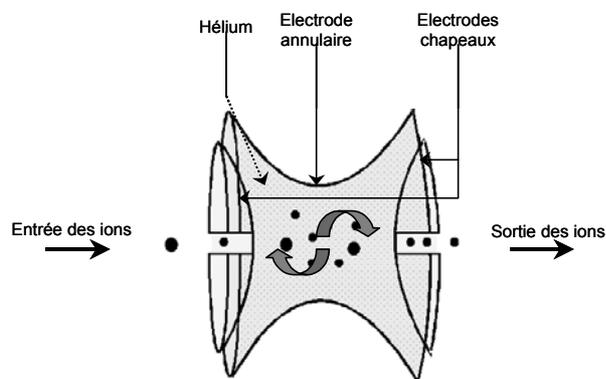


Figure 4 : représentation schématique d'une trappe ionique.

La trappe ionique (figure 4) est constituée d'une électrode annulaire en forme de diabolos couverte par deux électrodes dites « chapeaux » quasi-hyperboliques. Cet assemblage définit un espace d'environ 1 cm^3 . Les ions circulent dans cet espace en empruntant les orifices situés au centre de chaque électrode chapeau. Une pression résiduelle d'hélium, environ $5 \cdot 10^{-3}$ mbar, règne au sein de l'analyseur. La fonction de ce gaz est double : augmenter l'efficacité de piégeage des ions (gaz « tampon ») et assurer le rôle de gaz de collision.

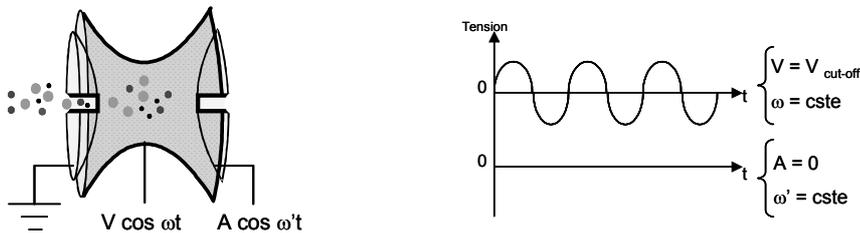
Les différentes étapes nécessaires à l'analyse des ions sont représentées figure 5. Nous y décrivons, d'une part, la trajectoire des ions et, d'autre part, les tensions séquentiellement appliquées sur les différentes électrodes, à savoir :

- | Un courant alternatif principal de type radiofréquence appliqué sur l'électrode annulaire ($V \cos \omega t$), d'amplitude V et de pulsation ω (dit « radiofréquence »).
- | Une tension sinusoïdale auxiliaire appliquée sur l'électrodes chapeau de sortie ($A \cos \omega' t$), d'amplitude A et de pulsation ω' .

c. Les étapes de l'analyse MS

Le piégeage des ions

Figure 5 : étape de piégeage des ions avec la trappe ionique lors de l'analyse MS.

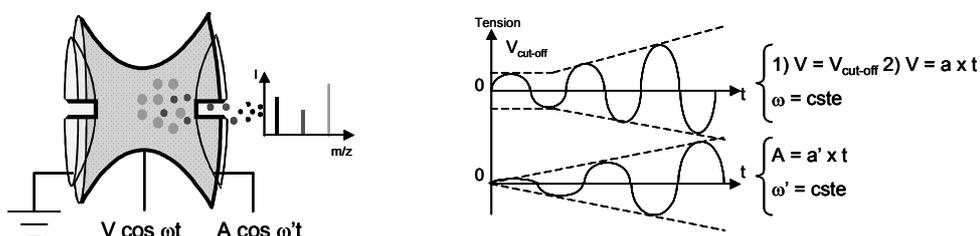


Les ions issus de l'interface sont piégés dans l'analyseur sous l'effet du champ quadripolaire généré par la «radiofréquence» appliquée sur l'électrode annulaire (figure 5). Les ions adoptent alors une trajectoire stable dont la forme est semblable à un huit dans l'espace à trois dimensions. Cette trajectoire peut être décrite par une courbe de Lissajous. Des collisions successives avec les molécules d'hélium ralentissent les ions dont la trajectoire est ainsi refocalisée au sein de l'analyseur.

On peut se représenter les ions comme des billes roulant le long des parois d'un verre à fond rond. Cette cavité représente le puits de potentiel nécessaire au piégeage des ions. C'est l'amplitude de la «radiofréquence» (V) qui détermine la profondeur du puits de potentiel. Plus l'amplitude V augmente, plus le puits de potentiel est profond et meilleur est le piégeage des ions. Par ailleurs, la profondeur du puits de potentiel est inversement proportionnelle au rapport m/z des ions piégés. Aussi, le choix de l'amplitude de la radiofréquence détermine la gamme de m/z de travail. Pour une amplitude V donnée, on définit le « cut-off » comme étant la valeur du plus petit m/z pouvant être piégé dans l'analyseur ; l'amplitude de travail est alors notée $V_{\text{cut-off}}$. Tous les ions de m/z inférieur au cut-off adoptent une trajectoire instable et sont éjectés de la trappe.

L'éjection des ions

Figure 6 : étape d'éjection des ions avec la trappe ionique lors de l'analyse MS.



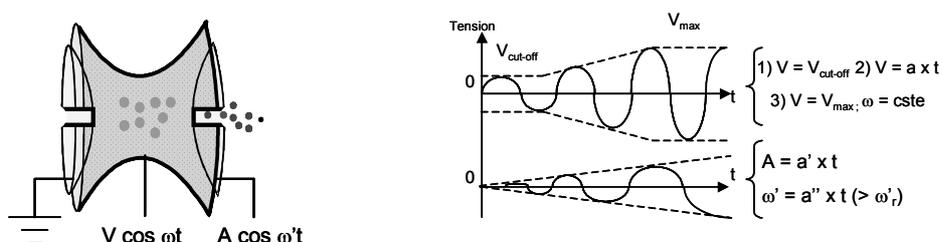
L'éjection des ions (figure 6) se fait de deux manières :

- | L'éjection simple consiste en un balayage croissant de l'amplitude de la «radiofréquence». Au delà de $V_{\text{cut-off}}$, le puits de potentiel devient insuffisant pour piéger les plus petits ions qui sont éjectés de l'analyseur. Le balayage de l'amplitude V permet donc d'éjecter sélectivement les ions selon les m/z croissants. Cependant, cette technique est limitée par la valeur des tensions applicables. Au delà d'une certaine limite (20 kV), on provoque des décharges électriques dans l'analyseur (m/z maximum : 650).
- | L'éjection résonnante consiste à appliquer une tension alternative ($A \cos \omega't$) sur l'électrode chapeau de sortie. Les ions soumis à ce potentiel d'amplitude croissante vont successivement rentrer en résonance et être déstabilisés de leur trajectoire pour être finalement éjectés. Cette approche permet d'augmenter la gamme de m/z d'analyse.

d. Les étapes de l'analyse MSⁿ

L'isolement

Figure 7 : étape d'isolement des ions avec la trappe ionique pour l'analyse MSⁿ.



La maîtrise de la trajectoire des ions dans la trappe ionique permet d'isoler un ion de m/z choisi : c'est l'ion précurseur ou parent (figure 7).

L'isolement résulte de deux étapes successives d'éjection.

- | Les ions de m/z inférieurs à la masse de l'ion parent sont éjectés par balayage croissant de l'amplitude de la «radiofréquence» jusqu'à atteindre V_{max} (amplitude limite pour laquelle le précurseur n'est pas éjecté).
- | Les ions de m/z plus élevés que les ions précurseurs ont des fréquences d'oscillations plus faibles que les ions à isoler. En partant d'une fréquence proche de la fréquence de résonance des ions parents (ω_r), on réalise un balayage décroissant en fréquence de la tension alternative appliquée sur l'électrode chapeau de sortie. Les ions plus lourds que les ions précurseurs seront successivement éjectés de la trappe.

L'excitation et la fragmentation

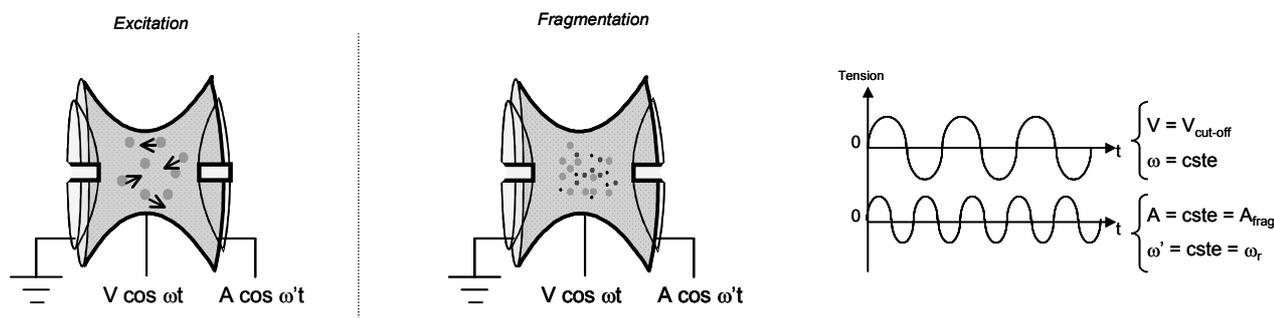


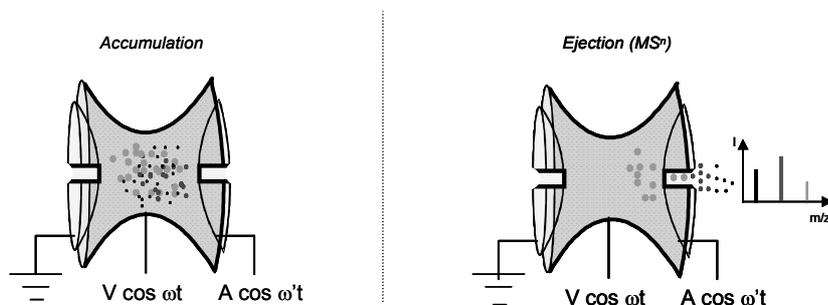
Figure 8 : étapes d'excitation et de fragmentation des ions avec la trappe ionique lors de l'analyse MSⁿ.

La fragmentation (figure 8) est provoquée par l'ajustement de la fréquence de la tension alternative sur l'électrode chapeau de sortie de manière à ce qu'elle coïncide avec la fréquence propre de résonance des ions précurseurs. Ainsi dopés en énergie cinétique, ils rentrent en collision avec les molécules d'hélium. Au fur et à mesure des collisions, leur énergie cinétique est convertie en énergie interne et l'ion parent finit par se fragmenter. L'amplitude du signal de résonance est maintenue constante et égale à une valeur inférieure à l'amplitude d'éjection des ions, ceci permet de conserver dans l'analyseur les ions fils qui ont été générés.

L'étape de fragmentation est suffisamment longue (< 50 ms) pour permettre aux ions d'emprunter les voies de fragmentation les plus stables énergétiquement, par opposition à d'autres modes de fragmentation comme le PSD MALDI-TOF ($\approx 50 \mu\text{s}$) ou la CID avec un triple quadripôle (de 1 à 10 ms). L'avantage de la MSⁿ avec une trappe ionique est une transmission exclusive de l'énergie de fragmentation aux ions précurseurs. Théoriquement, les ions fils issus d'une fragmentation ne peuvent pas se fragmenter simultanément en ions petits fils, l'énergie qui est transmise aux ions précurseurs n'étant délivrée qu'à leur seule fréquence de résonance.

Les autres étapes de l'analyse MSⁿ

Figure 9 : étapes d'accumulation et d'éjection avec la trappe ionique pour l'analyse MSⁿ.

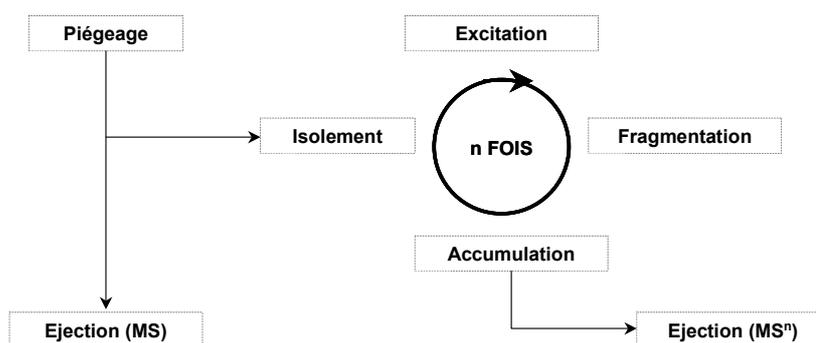


Suite à l'accumulation des ions fragments dans la trappe, une étape de spectrométrie de masse multiple de degré supérieur peut être réalisée en isolant à nouveau un ion précurseur parmi les ions fils obtenus et en reproduisant les étapes décrites précédemment (figure 9).

Ces étapes d'isolement, d'excitation et de fragmentation peuvent être réalisées n fois (figure 10), dans la limite de sensibilité de l'instrument (femtomolaire pour l'ESQUIRE 3000+, Bruker), pour obtenir des informations structurales très complètes et caractéristiques des molécules analysées.

Toutefois, le nombre n de cycles de fragmentation est limité par la quantité maximale d'ions qui peuvent être emmagasinés dans la trappe avant la fragmentation.

Figure 10 : résumé des différentes étapes de l'analyse des ions dans une trappe ionique.



1.4.3. Un analyseur hybride : le Q-TOF

Au laboratoire, nous disposons d'un appareil de type ES-Q-TOF, le Q-TOFII (Micromass) présenté figure 3. Cet instrument hybride est composé d'un premier analyseur quadripolaire couplé à un analyseur à temps de vol orienté perpendiculairement. Entre les deux analyseurs, se trouve une cellule de collision.

- | En mode MS, le quadripôle fonctionne en mode « RF only » (il laisse passer tous les ions) ; les collisions avec les molécules de gaz dans la cellule de collision ne sont pas suffisamment énergétiques pour fragmenter les ions.
- | En mode MS/MS, le quadripôle assure la sélection de l'ion parent qui est ensuite fragmenté en présence de gaz (de l'Argon) dans la cellule de collision.

Pour les deux modes de fonctionnement, les ions sont ensuite focalisés à l'entrée du TOF grâce à un système de lentilles électrostatiques. Le « pusher » donne l'impulsion nécessaire pour envoyer les ions, par paquets, dans l'analyseur. Cette étape permet d'adapter une source ES (qui produit des ions de façon continue) à un analyseur TOF plus approprié à une source d'ions pulsés comme la source MALDI.

Les principaux avantages du TOF dans cette configuration sont le gain en sensibilité et en résolution ainsi que le fait de pouvoir travailler sur une gamme de m/z en théorie illimitée. Nous reviendrons plus en détails sur ce type d'analyseur dans le paragraphe suivant consacré à la spectrométrie de masse MALDI-TOF.

2. Spectrométrie de masse MALDI-TOF

2.1. Introduction

La spectrométrie de masse MALDI-TOF, aujourd'hui incontournable, peut être appliquée à l'étude d'une large gamme de molécules d'intérêts biologiques. Elle permet de travailler sur une grande gamme de m/z (en théorie illimitée, en pratique jusqu'à environ 500 kDa) avec une bonne précision de mesure, une mise en œuvre assez simple et une relative tolérance vis à vis des sels et des tampons peu volatils.

Le mécanisme d'ionisation / désorption laser assistée par matrice (MALDI) a été introduit à la fin des années 80 par Karas et Hillenkamp [Karas et coll., 1987 ; Karas et coll., 1989] et, au même moment, par une équipe japonaise [Tanaka et coll., 1988]. La technique MALDI consiste à irradier avec un faisceau laser pulsé de longueur d'onde donnée (généralement dans l'UV) un dépôt cristallin contenant une matrice organique et l'échantillon à analyser.

2.2. La source MALDI

2.2.1. Le rôle de la matrice

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

La matrice doit avoir pour propriété d'absorber de façon résonnante à la longueur d'onde du laser. Co-cristallisée avec les analytes et en large excès, la matrice a un rôle de protection de l'échantillon par rapport aux effets destructeurs du faisceau laser incident. Elle permet également de transférer une partie de l'énergie du laser aux analytes pour former les ions en phase gazeuse.

2.2.2. La préparation de l'échantillon

Il est communément admis que la qualité des spectres de masse dépend de certains facteurs liés à la préparation de l'échantillon :

- | la sélection et la nature de la matrice
- | la méthode de préparation des cristaux
- | la présence de tensioactifs et / ou de tampons
- | le pH des solutions
- | les caractéristiques intrinsèques de l'échantillon
- | les solvants utilisés.

a. Le choix de la matrice

Il n'existe pas de matrice « universelle ». La matrice doit former des cristaux incluant l'analyte. Elle doit être soluble dans les solvants utilisés pour dissoudre l'échantillon et doit absorber à la longueur d'onde du laser.

Parmi les matrices les plus communément utilisées on distingue [Kusmann et coll., 1997] :

- | l'acide α -cyano-4-hydroxycinnamique (α -CHCA) pour l'analyse des peptides et des protéines
- | l'acide 2,5 dihydroxybenzoïque (DHB) pour les oligosaccharides, les glycoprotéines, les lipides, les peptides et les protéines
- | l'acide sinapinique (SA) pour les peptides, les protéines et les glycoprotéines.

Certaines techniques de préparation font appel à l'utilisation d'une co-matrice qui peut influencer la qualité de l'analyse en terme de résolution, de sensibilité, et de reproductibilité. Par exemple, la présence des sels d'ammonium dans les dépôts en DHB ou α -CHCA améliore le signal pour l'analyse des peptides et des oligonucléotides [Tang et coll, 1994].

b. La préparation du dépôt

Parmi les modes de dépôts les plus usuels, on distingue :

| La méthode de la goutte séchée

Cette technique consiste à mélanger préalablement à la solution contenant l'échantillon une solution saturée de matrice. Le mélange est déposé sur la cible et évaporé sous vide ou à l'air afin qu'il cristallise. Généralement, les cristaux produits en employant cette méthode sont relativement grossiers et hétérogènes ce qui peut avoir une incidence négative sur la résolution et la reproductibilité des résultats.

| La méthode en couche mince

Dans un premier temps, une goutte de solution saturée de matrice est déposée sur la cible. Généralement, le solvant utilisé pour dissoudre la matrice est très volatil (acétone). Aussi, l'évaporation quasi instantanée de ce solvant conduit à la formation d'une couche de cristaux de matrice très homogène que l'on appelle « couche mince ». La solution contenant l'échantillon est déposée sur cette couche mince et évaporée.

| La méthode en sandwich

Dérivée de la précédente méthode, elle présente une étape supplémentaire dans laquelle on recouvre la solution d'analytes par une solution saturée de matrice en prenant soin à ne pas utiliser des solvants susceptibles de re-dissoudre la première sous-couche de matrice.

Les cristaux résultants de ces deux dernières techniques de préparation sont petits et homogènes ce qui permet une amélioration notable de la reproductibilité et de la résolution des spectres de masse.

La technique MALDI est particulièrement tolérante aux sels. Cette tolérance est liée à la nature solide du dépôt. Dans le cas où la matrice est peu soluble dans l'eau (α -CHCA par exemple), le dessalage peut être réalisé par simple ajout d'une goutte d'eau légèrement acidifiée sur le dépôt, chassée ensuite par un souffle d'air ou par aspiration.

c. Les cibles utilisées pour le dépôt

Les supports utilisés pour réaliser le dépôt sont généralement en métal inoxydable (souvent à base d'aluminium), conducteur et traité en surface de manière à être chimiquement inerte par rapport à l'échantillon. Il existe également des cibles en verre recouvertes d'une couche de métal conducteur.

La surface des dépôts (en goutte séchée) sur ces cibles varie de 5-15 mm². Seule, une minuscule fraction de cette surface (0,002 à 0,03 mm²) est irradiée par le laser : une part importante de l'échantillon est inutilisée. Souvent, les dépôts sont hétérogènes ce qui conduit l'utilisateur à accumuler un grand nombre de tirs laser en balayant la surface du dépôt pour obtenir un spectre de masse représentatif de l'échantillon. Par conséquent, avec les cibles classiques, ce défaut de reproductibilité d'un tir à l'autre est difficilement compatible avec l'automatisation de la technique.

Il a été montré que ces problèmes pouvaient être résolus en réduisant nettement la taille du dépôt (à 0,05 mm²) avec seulement quelques nL d'un échantillon hautement concentré. Par ce biais, une large portion de l'échantillon peut être irradiée simultanément, la sensibilité est accrue et la reproductibilité des tirs est meilleure [Little et coll., 1997 ; Gobom et coll., 1999].

Une nouvelle technologie dite « AnchorChip » (développée par Bruker Daltonics) a fait récemment son apparition et permet d'augmenter la sensibilité des analyses d'un facteur 10 à 100 [Schuerenberg et coll., 2000]. Les cibles « AnchorChip » sont couvertes d'un revêtement hydrophobe à base de téflon sur lequel se répartissent des sites hydrophiles appelés « ancras » de faible diamètre (200 à 600 μ m). En séchant, les dépôts rétrécissent et se positionnent spontanément au cœur de la région hydrophile. Ce phénomène, indépendant du volume d'échantillon déposé (0,1 à 4 μ L), permet de concentrer l'échantillon sur une très petite surface et, ainsi, d'augmenter la quantité d'ions désorbés au cours du tir laser. La taille et l'espacement des ancras hydrophiles ont été optimisés pour permettre la préparation robotisée de 384 dépôts sur la cible, ce qui est particulièrement utile dans le contexte de l'analyse à haut débit.

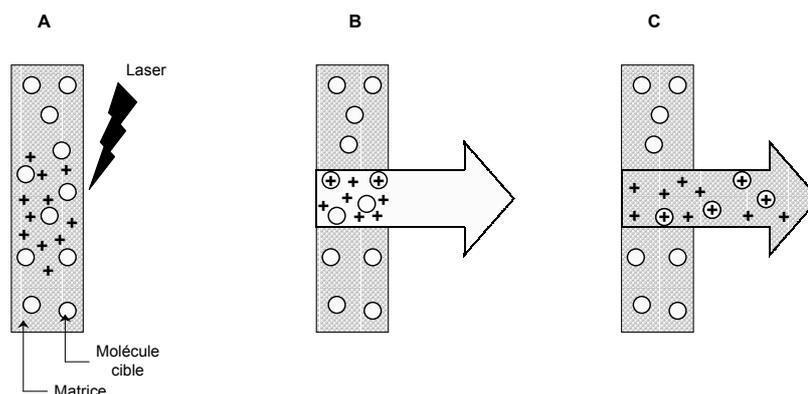
2.2.3. Le mécanisme d'ionisation MALDI

Le processus d'ionisation MALDI s'effectue sous vide ($\approx 10^{-7}$ mbar). Les ions sont produits en phase gazeuse sous l'effet d'une impulsion laser qui irradie la surface cristalline contenant la matrice et l'échantillon à analyser. Ils sont ensuite accélérés par application d'un champ électrique intense entre la cible et la première lentille électrostatique du spectromètre appelée aussi lentille d'extraction.

Les phénomènes d'ionisation MALDI ne sont toujours pas, à l'heure actuelle, très bien compris. Plusieurs modèles ont été proposés mais de nombreuses ambiguïtés concernant certains aspects de ce processus subsistent. Selon les hypothèses les plus communément admises, il serait constitué de trois événements majeurs (figure 11) :

- | l'excitation des molécules de matrice par les photons du laser
- | l'émission des molécules cibles dans la phase gazeuse
- | l'ionisation des molécules en phase gazeuse.

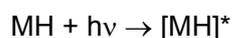
Figure 11 : schéma montrant la formation des ions en phase gazeuse en MALDI mode d'ionisation positif (A : absorption de la radiation laser par la matrice et ionisation de la matrice ; B : dissociation de la matrice et transfert de charges aux molécules cibles ; C : formation d'un jet supersonique et transfert de charges aux molécules



cibles). D'après [Fenselau, 1997].

a. Excitation des molécules de matrice

Après irradiation laser, les molécules de matrice notées MH (notation indiquant la présence d'un atome d'hydrogène labile) sont excitées par absorption résonnante de photons ($h\nu$).



L'excès d'énergie interne emmagasinée par les molécules de matrices excitées $[MH]^*$ peut être redistribué immédiatement à la molécule cible mais peut également provoquer la fragmentation de cristaux de matrice.

b. Emission des ions en phase gazeuse

Cette redistribution d'énergie abaisse les forces de cohésion des cristaux et provoque une co-éjection de molécules de matrice et de molécules cibles dans la phase gazeuse. Le gaz très dense qui résulte de cette éjection de matière subit une expansion adiabatique dans le vide sous forme d'un jet supersonique formant ainsi une plume [Beavis et Chait, 1991].

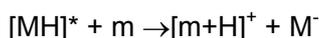
c. Ionisation des molécules en phase gazeuse

Il semblerait que ce soit dans ce gaz en expansion qu'intervienne l'ionisation des analytes. L'état des connaissances sur cet aspect du processus MALDI, encore mal défini, a été résumé par Zenobi et Knochenmuss [Zenobi et Knochenmuss, 2003].

Deux hypothèses ont été retenues :

- | l'ionisation pourrait se produire par transfert d'un proton lorsque les ions de matrice entrent en collision avec la molécule cible en phase gazeuse

| l'ionisation pourrait être le résultat d'une réaction acide/base entre la molécule cible (m) et des molécules de matrice excitées [MH]*



Par conséquent, contrairement à l'ionisation ES, la plupart des ions observés en MALDI sont monochargés.

2.3. L'analyseur à temps de vol

La source MALDI est en général couplée à un analyseur à temps de vol qui est bien adapté à l'ionisation pulsée par désorption laser. Le mode d'ionisation MALDI génère essentiellement des ions monochargés pour lesquels l'analyseur TOF offre une gamme de m/z théoriquement illimitée.

Le MALDI peut également être couplé, entre autres, à un analyseur à résonance ionique cyclotronique à transformée de Fourier [Sheng et coll., 1994], à une trappe ionique [Schwartz et Bier, 1993], à un analyseur à secteur magnétique [Hill et coll., 1991].

2.3.1. Le MALDI-TOF linéaire

Les premiers analyseurs à temps de vol ont été décrits par Wiley et McLaren en 1955 [Wiley et McLaren, 1955]. Le principe de base d'un analyseur TOF consiste à mesurer le temps de vol d'un ion (préalablement accéléré dans un champ électrostatique E) au travers d'une région libre de champ : le tube de vol (figure 12).

Dans le cas idéal où les ions ont tous la même énergie cinétique initiale, le temps de vol des ions est proportionnel à la racine carrée de leur rapport masse sur charge (m/z) (équation 4).

Soit l'énergie cinétique E_c d'un ion de masse m et de charge z accéléré dans un champ E avec e : charge d'un électron, $1,602 \cdot 10^{-19}\text{C}$ - les unités sont celles du système international :

$$E_c = zeE \quad \text{(équation 2)}$$

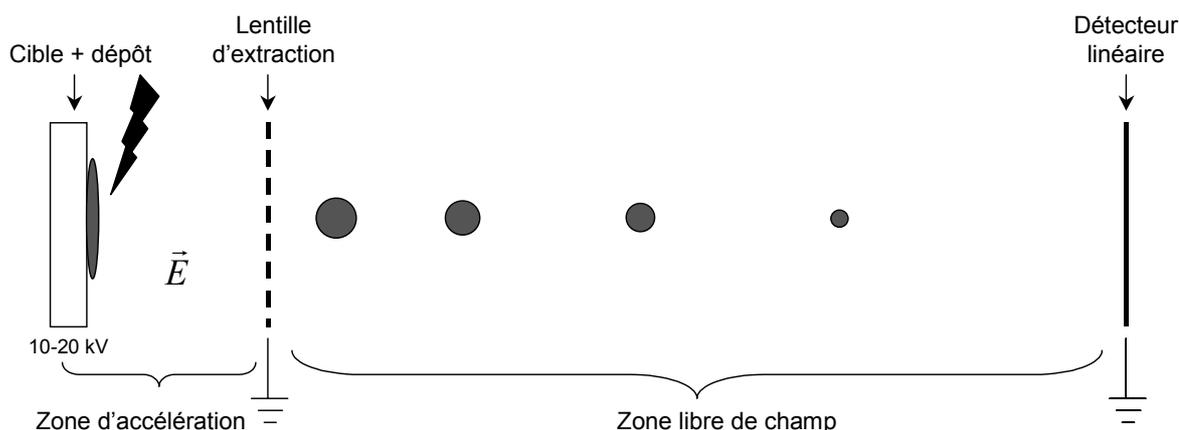
La vitesse v de cet ion est donnée par l'équation suivante :

$$v = \sqrt{\frac{2zeE}{m}} \quad \text{(équation 3)}$$

Le temps nécessaire à un ion pour traverser le tube de vol de longueur L , sachant que la vitesse $v = L/t$, peut donc être déduit des équations précédentes :

$$t = L\sqrt{\frac{m}{2zeE}} \quad \text{(équation 4)}$$

Pour un même état de charge, un ion de masse plus élevée passe donc plus de temps dans le tube de vol qu'un ion de masse plus faible. La mesure du temps de vol d'un ion permet ainsi la détermination de son rapport m/z.

Figure 12 : schéma représentant le principe de fonctionnement du MALDI-TOF en mode linéaire.

2.3.2. Le MALDI-TOF avec réflecteur et extraction retardée

Les ions issus de la source MALDI ne sont pas toujours générés ni au même moment ni au même endroit. Deux ions de même rapport m/z n'acquièrent donc pas forcément la même énergie cinétique initiale. Cette dispersion en énergie cinétique initiale affecte la capacité de l'analyseur à distinguer deux ions de m/z voisins : m/z et $m/z + \delta(m/z)$. La résolution R d'un analyseur de masse est donnée par :

$$R = m/z / \delta(m/z)$$

Deux dispositifs peuvent contribuer à améliorer la résolution de l'analyseur (figure 13) :

- | l'extraction retardée (« Delay Extraction ») [Vestal et coll., 1995 ; Brown et Lennon, 1995]
- | le réflecteur électrostatique [Mamyrin et coll, 1973].

a. L'extraction retardée

Lorsque le nuage, composé de molécules de matrice et d'analytes, est éjecté de la surface du cristal, les ions trouvent une résistance à leur accélération dans le champ électrique appliqué à cause des collisions avec les molécules de matrice. La durée pendant laquelle les ions sont piégés dans le nuage en expansion est proportionnelle à la taille de ce nuage et dépend de l'énergie fournie par le laser. Ces observations ont conduit au développement de la source à extraction retardée.

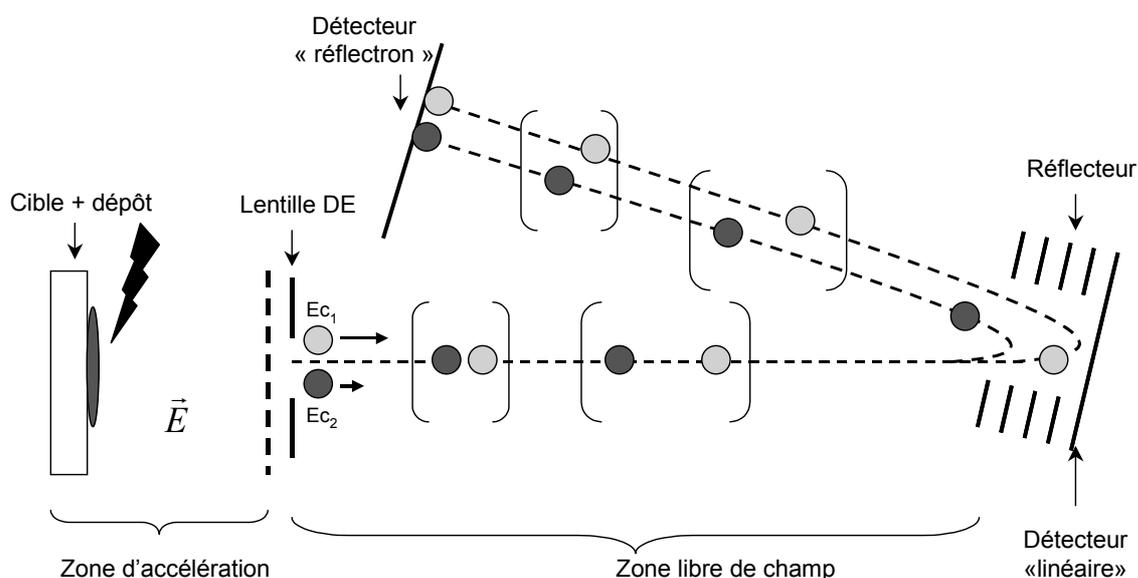
Cette méthode consiste à introduire, après la source, une lentille supplémentaire appelée lentille d'extraction retardée (DE). Pendant un bref délai suivant le tir laser (quelques centaines de nanosecondes), la lentille DE est portée à un potentiel légèrement supérieur à celui de la cible s'opposant ainsi à la progression des ions dans l'analyseur. Ceci permet la dissipation du nuage gazeux. Les ions sont alors focalisés dans l'espace : ils se placent sur « une ligne de départ ». Après le délai qui leur est imposé, ils pénètrent donc simultanément dans le tube de vol. Ce dispositif va aussi contribuer à minimiser la dispersion en énergie cinétique initiale.

b. Le réflecteur électrostatique

Ce réflecteur, situé à l'extrémité du tube de vol, est composé d'une série d'anneaux ou de grilles sur lesquels sont appliqués des potentiels croissants. Il agit comme un miroir électrostatique imposant

un champ électrique s'opposant à la progression des ions. Les ions font demi-tour et sortent du réflecteur avec une vitesse longitudinale de sens opposée à leur vitesse initiale. Pour une famille d'ions de m/z donné, les ions les plus énergétiques pénètrent plus profondément dans le réflecteur et mettent plus de temps à être réfléchis, leur temps de vol est allongé. Finalement, tous les ions de même m/z arrivent simultanément au détecteur : ils sont focalisés temporellement. Par ce biais, on corrige les effets de la dispersion en énergie cinétique et on améliore la résolution de l'analyseur.

Figure 13 : schéma représentant le principe de fonctionnement du MALDI-TOF avec un réflecteur électrostatique et un système d'extraction retardée. Deux ions de même m/z et d'énergie cinétique différente ($E_{c1} > E_{c2}$) pénètrent plus ou moins profondément dans le réflecteur, sont réfléchis, puis arrivent simultanément au détecteur. Le dispositif permet de porter la résolution d'un ion de m/z 6000 à une valeur supérieure à 10000.



2.4. La fragmentation en MALDI-TOF

2.4.1. Le PSD ou « Post Source Decay »

a. Introduction

Le processus de dégradation des ions dans le tube de vol, zone libre de champ, est désigné par le terme PSD pour Post Source Decay [Spengler et coll., 1992]. Les principaux mécanismes d'activation à l'origine de ce phénomène sont décrits dans le tableau 1 et concernent essentiellement :

- | la dégradation métastable d'ions ayant accumulé beaucoup d'énergie interne sur la cible ou par collisions multiples dans la source
- | la dissociation induite par collisions dans l'analyseur.

Tableau 1 : principaux mécanismes d'activation rencontrés en PSD [Spengler, 1997].

Zone d'activation	Cible - Surface de l'échantillon	Source - Nuage en expansion	Tube de vol
-------------------	----------------------------------	-----------------------------	-------------

Processus d'activation	Interaction photon-molécule	Collisions multiples à basse et haute énergie	Collisions à haute énergie
	Chaleur Energie d'ionisation		
Echelle de temps	ps - ns	ns- μ s	μ s-ms
Distance de la cible	0	< 200 μ m	> zone accélération

b. Rôle de la matrice en PSD

La nature de la matrice semble jouer un rôle important dans le processus PSD. A ce titre, on distingue les matrices « chaudes », c'est à dire les plus énergétiques, des matrices « froides » [Karas et coll., 1995]. Cette classification a été corrélée à l'affinité protonique de ces molécules [Jorgensen et coll., 1994 ; Burton et coll., 1997] (tableau 2). Pour qu'un transfert de proton puisse se produire, il faut que l'affinité protonique de l'accepteur (échantillon) soit supérieure à celle du donneur (matrice). Pour les protéines et les peptides, la présence d'une liaison amide, d'affinité protonique élevée, est très favorable à ce transfert et permet d'obtenir un ion pseudomoléculaire $[M+H]^+$. Une matrice de faible affinité protonique (comme α -CHCA) est une matrice « chaude » qui peut transférer beaucoup d'énergie, puisque la différence d'affinité protonique entre matrice et analyte est maximale. Ce type de matrice favorise la dégradation de l'échantillon après ionisation et est donc appropriée à l'analyse en PSD.

Tableau 2 : affinités protoniques de quelques matrices [Jorgensen et coll., 1994 ; Burton et coll., 1997].

Matrices	Affinités protoniques mesurées (kcal / mol)
Acide α -cyano-4-hydroxycinnamique (CHCA)	183
Acide 2,5-dihydrobenzoïque (DHB)	204
Acide sinapinique	204
2',4',6'-trihydroxyacétophénone (THAP)	210,8

c. Analyse en PSD

Lorsqu'un ion se dissocie pendant son vol dans la zone libre de champ, il génère des fragments de même vitesse mais d'énergie cinétique plus faible. En effet, l'énergie cinétique de l'ion qui se dissocie est redistribuée aux fragments en fractions proportionnelles à leur masse. Ainsi, en mode linéaire, fragments et précurseur arrivent simultanément au détecteur. L'analyseur TOF en mode linéaire n'est pas capable de différencier ces fragments.

Pour les analyses en PSD, le réflecteur est utilisé en tant que filtre en énergie cinétique. Les fragments sont réfléchis sur les surfaces équipotentielles du réflecteur plus ou moins profondément

selon leur énergie cinétique. Les fragments les plus légers sont réfléchis plus rapidement et ont un temps de vol plus court que les fragments plus lourds.

Cependant, le réflecteur n'est capable d'analyser les ions fragments avec une résolution correcte que sur une certaine gamme d'énergie. Aussi, un spectre complet des ions fragments sur toute la gamme de m/z doit être acquis en plusieurs étapes durant lesquelles les potentiels appliqués dans le réflecteur sont progressivement diminués. Finalement, toutes les acquisitions sont juxtaposées pour reconstituer le spectre d'ions fragments complet : le spectre PSD est un spectre recomposé.

Par ailleurs, l'étalonnage des spectres PSD n'est pas aussi facile que dans le cas des spectres MALDI-TOF classiques. Il n'y a pas de relation simple entre le temps de vol et la masse des ions fragments : de nombreux paramètres instrumentaux sont à prendre en considération.

Par conséquent, malgré de récents progrès dans le domaine [Franzen et coll., 2000], l'analyse PSD présente l'inconvénient majeur d'être longue et fastidieuse à mettre en œuvre. Peu sensible (au moins 10 picomoles d'échantillon sont nécessaires), elle offre, de plus, une résolution très limitée (100 à 500 sur une gamme de 1000 m/z). L'apparition d'une nouvelle génération d'instruments appelés TOF-TOF, décrits dans le paragraphe suivant, a permis de simplifier les expériences MS/MS avec une source MALDI tout en conservant une bonne qualité d'analyse.

2.4.2. Le MALDI TOF-TOF

a. Introduction

La technologie TOF-TOF a été décrite pour la première fois en 2000 [Medzihradszky et coll., 2000] et a été développée par plusieurs constructeurs de spectromètres de masse, parmi lesquels : Applied Biosystem, Bruker Daltonics, PE Biosystem.

Cette nouvelle technique, alternative au PSD, est aujourd'hui principalement utilisée dans le domaine de l'analyse protéomique pour le séquençage de peptides et de protéines ainsi que la caractérisation de modifications post-traductionnelles [Yergey et coll., 2002 ; Schnaible et coll, 2002 ; Suckau et coll, 2003].

Comme le PSD, la technique TOF-TOF, s'intéresse aux informations obtenues par dégradation d'ions métastables dans l'analyseur à temps de vol. De géométrie variable, ces instruments se composent généralement de la façon suivante :

- | la première partie du tube de vol assure la sélection de l'ion précurseur formé dans la source MALDI
- | la région suivante fait office de cellule de collision, la présence de gaz permet de doper les ions en énergie interne et d'accroître le nombre des fragmentations métastables
- | les fragments sont ré-accélérés dans la seconde partie du tube de vol qui, avec le réflecteur, permet l'analyse des fragments métastables.

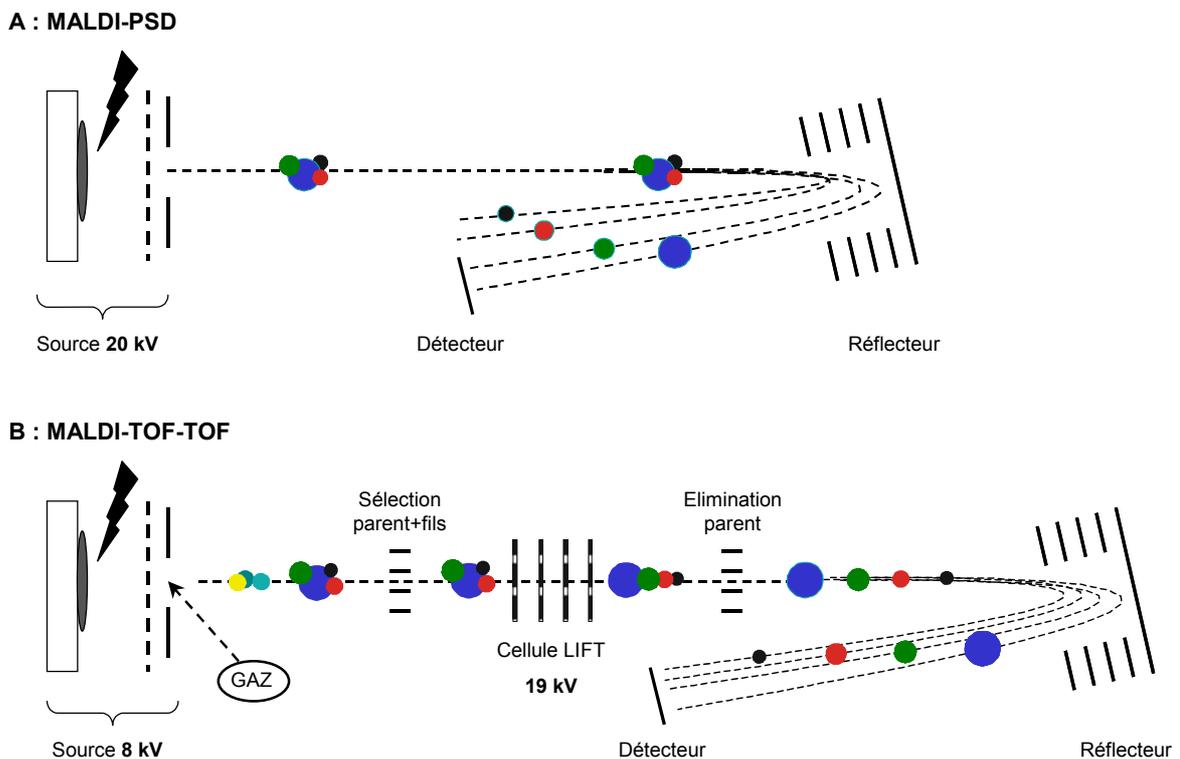
Le principal avantage de ces instruments est d'assurer la focalisation des fragments en énergie cinétique ce qui permet l'acquisition d'un spectre MS/MS à un potentiel fixe du réflecteur. Ainsi, le

temps d'analyse est aussi court qu'une analyse MS (quelques secondes), la résolution et la précision sur la mesure de masse sont nettement améliorées, et la sensibilité est de loin supérieure à celle obtenue en PSD (le gain en sensibilité est d'environ un facteur 100, d'après les spécifications données par Bruker pour leurs instruments).

b. Principe du LIFT et comparaison avec le PSD

Dans cette partie, nous décrivons le principe de fonctionnement de l'instrument Ultraflex TOF-TOF (Bruker) principalement utilisé lors de ce travail de thèse. Il repose sur la technologie appelée « LIFT » (pour « ascenseur à potentiel ») [Frey et coll., 2000]. Les avantages de cet instrument par rapport au PSD sont schématisés figure 14.

Figure 14 : différences entre l'analyse MS/MS par MALDI-PSD et MALDI-TOF-TOF. La cellule LIFT permet la focalisation des ions fragments en énergie cinétique : l'acquisition d'un spectre MS/MS avec le TOF-TOF se fait en une unique étape contrairement au PSD.



Principe de fonctionnement du TOF-TOF

Les ions formés dans la source MALDI sont accélérés à un potentiel d'environ 8kV et sont fragmentés dans la première région du tube de vol. A ce niveau, précurseur et fragments volent à la même vitesse. Leur sélection est assurée par une grille déflectrice sur la base de leur temps de vol. Au niveau de la cellule LIFT, les ions sont soumis à un potentiel d'accélération supplémentaire au moins deux fois égal au potentiel initial (≈ 19 kV). Les disparités en énergie cinétique entre les différents fragments deviennent par conséquent très petites comparées à l'énergie cinétique totale qui leur est conférée. Dans la seconde partie du tube de vol, les ions fragments évoluent selon leur masse (les plus légers prennent de l'avance sur les plus lourds) et le réflecteur, qui agit comme un filtre en énergie cinétique, peut focaliser les fragments en une seule étape.

La résolution sur ce type d'instrument, en mode MS/MS peut atteindre des valeurs de 2000 à 3000 pour un parent de m/z allant jusqu'à 2500. La sensibilité, qui dépend du parent, est généralement picomolaire.

Utilisation d'un gaz de collision

Un gaz de collision (de l'argon) peut être introduit à la sortie de la source. La présence ou non de gaz permet de distinguer deux modes de fonctionnement :

- | LID pour « Laser Induced Dissociation », la fragmentation des ions métastables est principalement liée à l'excès d'énergie interne conférée dans la source, notamment sous l'effet de l'irradiance du laser
- | CID pour « Collision Induced Dissociation », la présence de molécules d'argon dans la première partie du tube de vol renforce les collisions et favorise les fragmentations à haute énergie.

L'étalonnage

Les spectres MS/MS obtenus avec le MALDI-TOF-TOF doivent nécessairement être calibrés après l'acquisition. Cette étalonnage fait appel à des spectres MS/MS de composés de référence (de m/z représentatifs de la gamme de m/z de travail) réalisés à l'installation de l'instrument. Chacun de ces spectres permet le calcul des constantes d'une équation polynomiale associée à une valeur de m/z d'un précurseur donné. Lors de l'étalonnage, qui peut être réalisé en mode automatique, les constantes de cette équation sont recalculées sur la base du ratio entre la masse du précurseur expérimental et celle du composé de référence (choisi parmi les plus proches).

Cette étape d'étalonnage est très rapide et permet d'obtenir une précision d'environ 200 ppm sur les fragments. L'avantage par rapport au PSD est qu'il n'est pas nécessaire de calibrer la gamme de m/z travail par segments successifs.

BIBLIOGRAPHIE

Beavis R., Chait B.

Velocity distributions of intact high mass polypeptide molecule ions produced by matrix assisted laser desorption, *Chem. Phys. Lett.*, **1991**, 181, 479.

Blades A.T., Ikononou M.G., Kebarle P.

Mechanism of electrospray mass spectrometry. Electrospray as an electrolysis cell, *Anal. Chem.*, **1991**, 63, 2109-2114.

Brown R.S., Lennon J.J.

Mass resolution by incorporation of pulsed ion extraction in a matrix-assisted laser desorption/ionization linear time-of-flight mass spectrometer, *Anal. Chem.*, **1995**, 67, 1998.

Burton R.D., Watson C.H., Eyer J.R., Lang G.L., Powell D.H., Avery M.Y.

Proton affinities of eight matrices used for matrix-assisted laser desorption/ionization, *Rapid Commun. Mass Spectrom.*, **1997**, 11, 443-446.

Carte N.

La trappe ionique et l'ionisation électrospray : un nouveau potentiel pour la caractérisation de biomolécules, *Thèse de l'université Louis Pasteur de Strasbourg*, **2001**.

Cole R.B.

Some tenets pertaining to electrospray ionization mass spectrometry, *J. Mass Spectrom.*, **2000**, 35, 763-772.

Cooks R.G., Glish G.L., McLuckey S.A., Kaiser R.E.

Ion trap mass spectrometry, *Chemical & Engineering News*, **1991**, Mars 25, 26-41.

De Hoffmann E., Charette J., Stroobant V.

Spectrométrie de masse, cours et exercices corrigés, **2^e édition**.

Dole M., Mack L. L., Himes R. L., Mobley R. C., Ferguson L. D., Alice M. B.

Molecular beams of macroions, *J. Chem. Phys.*, **1968**, 49, 2240-2249.

Doyle A., Moffett D.R., Vonnegut B.

Behavior of evaporating electrically charged droplets, *J. Colloid Sci.*, **1964**, 19, 136-143.

Emmet M. R., Caprioli R. M.

Micro-electrospray mass spectrometry : ultra-high-sensitivity analysis of peptides and proteins., *J. Am. Soc. Mass Spectrom.*, **1994**, 5, 605-613.

Fenselau C.

MALDI MS and strategies for protein analysis, *Anal. Chem. News and Features*, **1997**, November 1, 661-665.

Fernandez de la Mora J., Van Berkel G.J., Enke C.G., Cole R.B., Martinez-Sanchez M., Fenn J.B.

Special feature : discussion. Electrochemical processes in electrospray ionization mass spectrometry, *J. Mass Spectrom.*, **2000**, 35, 939-952.

Franzen J., Frey R., Holle A., Kräuter K.O.

Recent progress in matrix-assisted laser desorption ionization postsorce decay, *Int. J. Mass Spectrom.*, **2000**, 206, 275-286.

Frey R., LaRotta A., Holle A., Köster C., Franzen J.

A reflector TOF system with MS/MS capability, **2000**, Proceedings of the 48th ASMS conference, Long Beach, Ca, US.

Gallagher R.T., Chapman J.R., Mann M.

Design and performance of an electrospray ionization source for a doubly-focusing magnetic sector mass spectrometer, *Rapid Commun. Mass Spectrom.*, **1990**, 4, 369-372.

Gamero-Castano M., Fernandez de la Mora J.

Mechanisms of electrospray ionization of singly and multiply charged salt clusters, *Anal. Chim. Acta.*, **2000**, 406, 67-91.

Gobom J., Nordhoff E., Mirgorodskaya E., Ekman R., Roepstorff P.

Sample purification and preparation technique based on nano-scale reversed-phase columns for the sensitive analysis of complex peptide mixtures by matrix-assisted laser desorption/ionization mass spectrometry, *J. Mass Spectrom.*, **1999**, 34, 105-116.

Gomez A., Tang K.

Charge and fission of droplets in electrostatic sprays, *Phys. Fluids*, **1994**, 6, 404-414.

Hayati I., Bailey A.I., Tadros T.F.

Investigations into the mechanisms of electrodynamic spraying of liquids. Effects of electric field and the environment on pendant drops and factors affecting the formation of stable jets and atomization, *J. Colloid Interface Sci.*, **1987**, 117, 205-221.

Henry K.D., Williams E.R., Wang B.H., McLafferty F.W., Shabanowitz J., Hunt D.F.

Fourier-transform mass spectrometry of large molecules by electrospray ionization, *PNAS*, **1989**, 86, 9075-9078.

Hill J.A., Annan R.S., Biemann K.

Matrix-assisted laser desorption ionization with a magnetic mass spectrometer, *Rapid Commun. Mass Spectrom.*, **1991**, 5, 395-399.

Hiraoka K., Kudaka I.

Electrospray interface for liquid chromatography/mass spectrometry, *Rapid Commun. Mass Spectrom.*, **1990**, 4, 519-526.

Ikonomou M.G., Blades A.T., Kebarle P.

Electrospray-ion spray : a comparison of mechanisms and performances, *Anal. Chem.*, **1991**, 63, 1989-1998.

Iribarne J. V., Thomson B. A.

On the evaporation of small ions from charged droplets, *J. Chem. Phys.*, **1976**, 64, 2287-2294.

Jorgensen T.J.D., Vulpius T., Bojesen G.

The order of proton affinities of five common MALDI matrices, *In proceedings of the 13th international MS conference, Budapest, Hungary*, **1994**, Wiley, John&Sons.

Karas M., Bachmann D., Bach U., Hillenkamp F.

Matrix-assisted ultraviolet laser desorption of non-volatile compounds, *Int. J. Mass Spectrom. Ion Proc.*, **1987**, 78, 53-68.

Karas M., Bach U., Hillenkamp F.

UV laser matrix desorption/ionization mass spectrometry of proteins in the 100 000 dalton range, *Int. J. Mass Spectrom. Ion Proc.*, **1989**, 92, 231-242.

Karas M., Bach U., Strupat K., Hillenkamp F., Tsarbopoulos A., Pramanik B.N.

Matrix dependence of metastable fragmentation of glycoproteins in MALDI TOF mass spectrometry, *Anal. Chem.*, **1995**, 67, 675-679.

Kebarle P., Tang L.

From ions in solution to ions in the gas phase, The mechanism of electrospray mass spectrometry, *Anal. Chem.*, **1993**, 65, 972-86.

Kebarle P., Ho Y.

On the mechanism of electrospray mass spectrometry. Electrospray Ionization Mass Spectrometry : Fundamentals, Instrument and Applications, *John Wiley and Sons, R.B. Cole Ed.*, **1997**, 3-63.

Kebarle P.

A brief overview of the present status of the mechanism involved in electrospray mass spectrometry, *J. Mass Spectrom.*, **2000**, 35, 804-817.

Kebarle P., Peschke M.

On the mechanisms by which the charged droplets produced by electrospray lead to gas phase ions, *Anal. Chim. Acta.*, **2000**, 406, 11-35.

Knochenmuss R., Zenobi R.

MALDI ionization : the role of in-plume processes, *Chem. Rev.*, **2003**, 103 (2), 441-452.

Körner R., Wilm M., Morand K., Schubert M., Mann M.

Nano electrospray combined with a quadrupole ion trap for the analysis of peptides and protein digests, *J. Am. Soc. Mass Spectrom.*, **1996**, 7, 150-156.

Kussmann M., Nordhoff E., Rabek-Nielsen H., Haebel S., Rossel-Larsen M., Jakobsen L., Gbom J., Mirgorodskaya E., Kroll-Kristensen A., Palm L., Roepstorff P.

Matrix-assisted laser desorption-ionization mass spectrometry sample preparation techniques designed for various peptide and protein analytes, *J. Mass Spectrom.*, **1997**, 32, 593-601.

Leize E.

Caractérisation d'édifices supramoléculaires par spectrométrie de masse avec ionisation électrospray, *Thèse de l'université Louis Pasteur de Strasbourg*, **1994**.

Light-Wahl K.J., Loo J.A., Edmonds C.G., Smith R.D., Witkowska H.E., Shackleton C.H., Wu C.S.

Collisionally activated dissociation and tandem mass spectrometry of intact hemoglobin beta-chain variant proteins with electrospray ionization, *Biol. Mass Spectrom.*, **1993**, 22, 112-120.

Little D.P., Cornish T.J., O'Donnell J., Braun A., Cotter R.J., Köster H.

MALDI on a chip : analysis of arrays of low-femtomole to subfemtomole quantities of synthetic oligonucleotides and DNA diagnostic products dispensed by a piezoelectric pipet, *Anal. Chem.*, **1997**, 69, 4540-4546.

Loew D.

Gain de sensibilité en spectrométrie de masse des biomolécules, *Thèse de l'université Louis Pasteur de Strasbourg*, **1999**.

Mamyrin B.A., Karataev V.I., Shmikk D.V., Zagulin V.A.

Mass reflectron. New nonmagnetic time-of-flight high-resolution mass spectrometer, *Zh. Eksp. Teor. Fiz.*, **1973**, 64, 82.

March R.E.

An introduction to quadrupole ion trap mass spectrometry, *J. Mass Spectrom.*, **1997**, 32, 351-369.

Medzihradszky K.F., Campbell J.M., Baldwin M.A., Falick A.M., Juhasz P., Vestal M.L., Burlingame A.L.

The characteristics of peptide collision-induced dissociation using a high-performance MALDI-TOF-TOF tandem mass spectrometer, *Anal. Chem.*, **2000**, 72, 552-558.

Meng C. K., Mann M., Fenn J. B.

Of protons or proteins, *Z. Phys. D.*, **1988**, 10, 361-368.

Morris H.R., Paxton T., Panico M., McDowell R., Dell A.

A novel geometry mass spectrometer, the Q-TOF, for low-femtomoles/attomoles-range biopolymer sequencing, *J. Protein Chem.*, **1997**, 16, 469-479.

Nierengarten H.

Nouvelles approches pour l'étude de supramolécules synthétiques par spectrométrie de masse, *Thèse de l'université Louis Pasteur de Strasbourg*, **2001**.

Paul W., Steinwedel H.

German Patent, **1956**, 944900.

Potier N.

Caractérisation de complexes biologiques non covalents par spectrométrie de masse avec ionisation électrospray, *Thèse de l'université Louis Pasteur de Strasbourg*, **1996**.

Rogniaux H.

Etude d'interactions faibles intra- et inter-moléculaires de biomolécules par spectrométrie de masse, *Thèse de l'université Louis Pasteur de Strasbourg*, **2000**.

Sanglier S.

La spectrométrie de masse : un nouvel outil pour l'étude des interactions faibles en biologie, *Thèse de l'université Louis Pasteur de Strasbourg*, **2002**.

Schnaible V., Wefing S., Resemann A., Suckau D., Bücker A., Wolf-Kümmeth S., Hoffmann D.

Screening for disulphide bonds in proteins by MALDI in-source decay and LIFT-TOF/TOF-MS, *Anal. Chem.*, **2002**, 74, 4980-4988.

Schuerenberg M., Luebbert C., Eickhoff H., Kalkum M., Lehrach H., Nordhoff E.

Prestructured MALDI-MS sample supports, *Anal. Chem.*, **2000**, 72, 3436-3442.

Schwartz J.C., Bier M.E.

Matrix-assisted laser desorption of peptides and proteins using a quadrupole ion trap mass spectrometer, *Rapid Commun. Mass Spectrom.*, **1993**, 7, 27-32.

Sheng L.S., Covey J.E., Shew S.L., Winger B.E., Campana J.E.

Matrix-assisted laser desorption ionization Fourier-transform mass spectrometry, *Rapid Commun. Mass Spectrom.*, **1994**, 8, 498-500.

Shevchenko A., Wilm M., Vorm O., Mann M.

Mass spectrometric sequencing of proteins from silver-stained polyacrylamide gels, *Anal. Chem.*, **1996**, 68, 850-858.

Spengler B., Kirsch D., Kaufmann R.

Peptide sequencing by matrix-assisted laser desorption mass spectrometry, *Rapid Commun. Mass Spectrom.*, **1992**, 6, 105-108.

Spengler B.

Post-source decay analysis in matrix-assisted laser desorption/ionization mass spectrometry of biomolecules, *J. Mass Spectrom.*, **1997**, 32, 1019-1036.

Stafford G.C., Kelley P.E., Syka J.E.P., Reynolds W.E., Todd J.F.J.

Recent improvements in and analytical applications of advanced ion trap technology, *Int. J. Mass Spectrom. Ion Proc.*, **1984**, 60, 85-98.

Suckau D., Resemann A., Schuerenberg M., Hufnagel P., Franzen J., Holle A.

A novel MALDI-TOF/TOF mass spectrometer for proteomics, *Anal. Bioanal. Chem.*, **2003**, Sous presse.

Taflin D.C., Ward T.L., Davis E.J.

Electrified droplet fission and the Rayleigh limit, *Langmuir*, **1989**, 5, 376-384.

Tanaka K., Waki H., Ido Y., Akita S., Yoshida Y., Yoshida T.

Protein and polymer analysis up to m/z 100.000 by ionization time-of-flight mass spectrometry, *Rapid Commun. Mass Spectrom.*, **1988**, 2, 151-152.

Tang K., S. L. Allman, Chen C.H., Chang L.Y., Schell M.

Matrix-assisted laser desorption/ionization of restriction enzyme-digested DNA, *Rapid Commun. Mass Spectrom.*, **1994**, 8, 183-186.

Van Berkel G.J., Glish G.L., McLuckey S.A.

Electrospray ionization combined with ion trap mass spectrometry, *Anal. Chem.*, **1990**, 62, 1284-1295.

Verentchikov A.N., Ens W., Standing K.G.

Reflecting time-of-flight mass spectrometer with an electrospray ion source and orthogonal extraction, *Anal. Chem.*, **1994**, 66, 126-133.

Vestal M.L., Juhasz P., Martin S.A.

Delayed extraction matrix-assisted laser desorption time-of-flight mass spectrometry, *Rapid Commun. Mass Spectrom.*, **1995**, 9, 1044.

Wang G., Cole R.D.

Mechanistic interpretation of the dependence of charge states distributions on analyte concentrations in electrospray ionization mass spectrometry, *Anal. Chem.*, **1995**, 67, 2892-2900.

Wiley W.C., McLaren I.H.

Time-of-flight mass spectrometry with improved resolution, *Rev. Sci. Instrum.*, **1955**, 26, 1150-1157.

Wilm M., Mann M.

Electrospray and Taylor-cone theory, Dole's beam of macromolecules at last?, *Int. J. Mass Spectrom. Ion Proc.*, **1994**, 136, 167-180.

Wilm M., Mann M.

Analytical properties of the nanoelectrospray ion source, *Anal. Chem.*, **1996**, 68, 1-8.

Yamashita M., Fenn J.

Electrospray ion source. Another variation on the free-jet theme, *J. Phys. Chem.*, **1984**, 88, 4451-4460.

Yergey A.L., Coorssen J.R., Backlund P.S., Blank P.S., Humphrey G.A., Zimmerberg J., Campbell J.M., Vestal M.L.

De novo sequencing of peptides using MALDI/TOF/TOF, *J. Am. Soc. Mass Spectrom.*, **2002**, 13, 784-791.

II. L'ANALYSE PROTEOMIQUE

1. Introduction

Le terme protéome a été proposé pour la première fois en 1995 pour désigner « l'ensemble des produits fonctionnels exprimés par un génome » [Kahn, 1995]. Rapidement, avec l'avancée des techniques dans ce domaine, cette définition a évolué : N.L. Anderson et N.G. Anderson introduisent le terme « proteomics » et définissent l'analyse protéomique par « l'utilisation de la quantification au niveau de la protéine comme mesure objective de l'expression génique caractérisant un processus biologique donné et comme un moyen de décodage des mécanismes contrôlant cette expression » [N.L. Anderson et N.G. Anderson, 1998].

Ces dernières années, le décryptage de nombreux génomes, en particulier du génome humain, a grandement facilité l'émergence du protéome. Les raisons qui expliquent l'importance prise par l'analyse protéomique sont les suivantes :

- | Les niveaux d'expression protéiques ne sont pas le simple reflet des niveaux d'expression des ARN messagers. En effet, des études ont montré, que pour une population de gènes considérée comme relativement homogène en terme d'expression des ARN messagers et de demi-vie des protéines codées, des variations significatives du niveau d'expression des protéines peuvent être observées [Anderson et Seilhamer, 1997 ; Gygi et coll., 1999a].
- | Les protéines peuvent subir une série de modifications qui ne sont pas forcément identifiables à partir de la seule séquence de leurs gènes. Beaucoup de ces molécules ne parviennent à leur forme biologiquement active qu'à la suite d'une étape de maturation post-traductionnelle comme, par exemple, la glycosylation ou la phosphorylation.
- | L'analyse protéomique est dynamique. En effet, un même génome peut conduire à différents protéomes selon les étapes du cycle cellulaire ou de la différenciation, de la réponse à des signaux biologiques ou physiques, de l'état physiopathologique. Lottspeich illustre parfaitement ces phénomènes en rapprochant les images d'une chenille d'un papillon [Lottspeich, 1999].

Ainsi, l'analyse protéomique se propose de donner une image globale de ces systèmes biologiques et de leur complexité. Mais ce développement n'aurait pas été possible sans les avancées technologiques qui permettent l'étude du protéome telles que :

- | le séquençage de génomes complets qui alimentent les banques de données
- | l'évolution des techniques de séparation des protéines et de préparation de l'échantillon avec les gels d'électrophorèse (1D et 2D), l'apparition de systèmes de micro et nano-chromatographie liquide et leur couplage avec la spectrométrie de masse
- | le développement des spectromètres de masse en terme de sensibilité, de précision et de résolution ainsi que l'apparition de nouvelles générations d'instrument (MALDI-TOF-TOF, cf chapitre I, 2.4.2.)

des outils bio-informatiques de plus en plus performants et puissants.

Aujourd'hui, la protéomique a démontré qu'elle était indispensable à une meilleure compréhension des processus normaux ou pathologiques, qu'elle pouvait conduire au développement de nouveaux biomarqueurs pour le dépistage de diverses maladies et accélérer la recherche pharmaceutique [Sam Hanash, 2003]. Dans ce contexte, des approches analytiques à haut débit, basées sur la spectrométrie de masse, sont développées permettant, entre autres, l'étude des interactions protéine-protéine ainsi que la quantification précise des différentes espèces protéiques [Aebersold et Mann, 2003]. Le succès de la protéomique dans les prochaines années passera, sans aucun doute, par des campagnes d'analyse organisées à l'échelle mondiale sur des thèmes d'intérêts vitaux comme la malaria ou le cancer [Tyers et Mann, 2003].

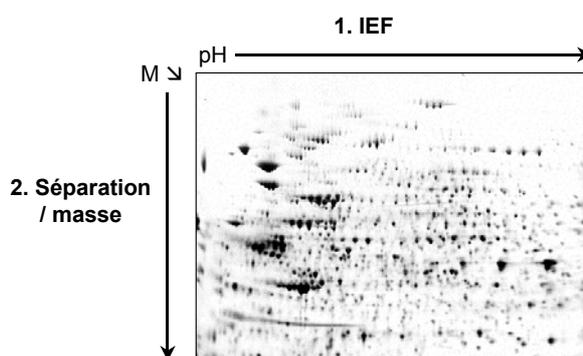
Cette partie fait le point sur les récents développements dans le domaine de l'analyse protéomique.

2. Le gel d'électrophorèse bidimensionnel

2.1. Principe du gel 2D-PAGE

Les travaux de O'Farrel en 1975 [O'Farrel, 1975] ont conduit à la mise au point d'une technique aujourd'hui au cœur de l'analyse protéomique : l'électrophorèse bidimensionnelle sur gel de polyacrylamide ou 2D-PAGE. Cette technique permet de séparer successivement les protéines d'un mélange complexe selon deux propriétés distinctes : leur charge et leur masse moléculaire (figure 15).

Figure 15 : schéma de principe et photo d'un gel 2D-PAGE d'un extrait de cellules Hela (d'après nos travaux en collaboration avec T. Rabilloud concernant l'étude des peroxyrédoxines, cf partie Résultats).



Dans un premier temps, les protéines sont soumises à une électrophorèse dans un gel présentant un gradient de pH continu. Au cours de cette étape, nommée isoélectrofocalisation (IEF), les protéines soumises à un champ électrique constant migrent dans le gel jusqu'à atteindre une valeur de pH égale à leur point isoélectrique pour laquelle la charge globale de la protéine est nulle. Ensuite, une deuxième séparation orientée perpendiculairement à l'IEF est réalisée dans un gel réticulé constitué de polyacrylamide en présence d'un agent chimique dénaturant, le sodium dodécyl sulfate (SDS). La réticulation du gel permet de séparer les molécules suivant leur taille. Le SDS interagit avec les protéines, les dénature et élimine l'influence de leur charge pour la séparation selon la seconde

dimension (en les « tapissant » de charges négatives). Dès lors, les protéines sont séparées en fonction de leur masse.

2.2. La coloration des gels 2D

Le choix du colorant permettant de révéler les protéines sur le gel conditionne la sensibilité de la technique. Parmi les colorants les plus fréquemment utilisés on distingue :

- | le bleu de Coomassie pour lequel l'intensité de la coloration est proportionnelle à la quantité d'échantillon sur une gamme de 10 à 200 ng / mm² [Radola et coll., 1987]
- | la coloration à l'argent, plus sensible et linéaire pour des concentrations de 0,04 à 2 ng / mm², dont l'inconvénient est de dépendre des propriétés intrinsèques de la protéine [Giometti et coll., 1991] et par conséquent de ne pas refléter correctement la quantité de protéines de natures différentes en présence.

De nouvelles techniques de coloration ont vu le jour récemment. Il s'agit de colorations fluorescentes comme le Sypro Ruby, le ruthenium [Rabilloud et coll., 2001]. Elles présentent les avantages d'être à la fois sensibles, rapides à mettre en œuvre et, également, de ne pas fixer les protéines au gel, ce qui facilite leur extraction et permet, éventuellement, leur transfert sur membrane.

Par ailleurs, le choix de la coloration conditionne les résultats obtenus lors de l'analyse des tâches (spots) protéiques par spectrométrie de masse. Les mécanismes d'interférence de la coloration à l'argent avec la spectrométrie de masse ont été étudiés et ont montré que l'utilisation de formaldéhyde dans le protocole de coloration était néfaste à l'identification des protéines par MS. A l'issue de ce travail, de nouveaux protocoles « formaldéhyde free » ont été proposés permettant de conserver la sensibilité de la coloration à l'argent et d'offrir une compatibilité avec la MS comparable avec les colorations fluorescentes ou les colorations au bleu [Richert et coll., 2003].

2.3. L'analyse des gels

La méthode d'analyse des gels a évolué grâce aux progrès combinés de l'informatique et de l'analyse d'image.

Aujourd'hui, la première étape de l'analyse est la digitalisation des gels, c'est-à-dire la transformation de l'image expérimentale en une information numérique utilisable par ordinateur. Plusieurs types d'appareils permettent de réaliser cette acquisition : scanners, caméras, densitomètres. Ils reproduisent l'intensité, la taille et la position des spots protéiques. Divers logiciels (ex. : PDQuest, Biorad ; MELANIE, Proteomics Solutions) permettent ensuite de traiter les images obtenues.

Dans le cadre de l'analyse protéomique, les informations pertinentes sont alors issues d'analyses différentielles obtenues par comparaison des gels entre eux.

2.4. Récentes améliorations et limites de la technique

La technique des gels 2D a été considérablement améliorée ces dernières années. Parmi les récents progrès apparus dans ce domaine, on peut noter :

- | l'apparition des gradients immobilisés de pH (IPG) [Bjellqvist et coll., 1982] qui permettent, d'une part, de stabiliser le gradient de pH et donc d'augmenter la reproductibilité des gels et, d'autre part, d'augmenter la capacité de chargement des gels (quelques mg de protéines) [Bjellqvist et coll., 1993]
- | l'apparition de nouvelles techniques de préparation de l'échantillon qui permettent d'améliorer la solubilisation des protéines afin de faciliter leur pénétration dans le gel avec l'incorporation d'agents de solubilisation, de surfactants, de détergents tels que la thiourée, le CHAPS [Rabilloud, 1996 ; Rabilloud et coll., 1997]
- | l'apparition des « zoom gels » [Westbrook et coll, 2001], c'est à dire des gels avec un gradient de pH restreint, offrant une plus grande résolution et permettant de s'intéresser à des protéines présentes en faible quantité dans une zone de pH réduite
- | l'amélioration de la séparation des protéines basiques par blocage des cystéines avec différents agents alkylants [Luche et coll., 2003].

Malgré ces améliorations, la technique des gels 2D présente une capacité à séparer les protéines encore limitée : le plus souvent les spots colorés ne correspondent pas à une seule protéine. De plus, cette technique reste trop sélective [Rabilloud, 2001] :

- | l'IEF est difficile pour les protéines très basiques ou très acides
- | les protéines membranaires (hydrophobes), d'intérêt thérapeutique majeur, sont difficiles à solubiliser [Santoni et coll., 2000].

Aussi, dans ce contexte, des stratégies alternatives de séparation des protéines sont envisagées pour palier à ces diverses limitations. Leur principe sera développé dans les paragraphes suivants.

3. L'apport de la spectrométrie de masse pour l'analyse des peptides en protéomique

3.1. La fragmentation des peptides

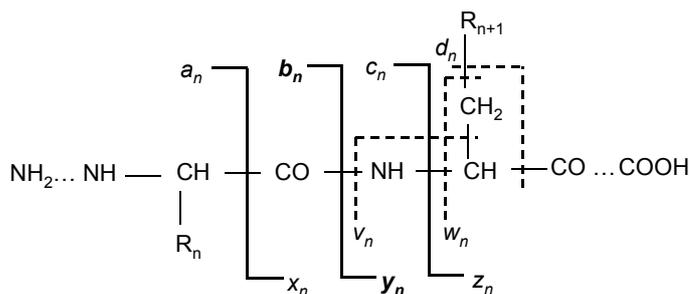
Les fragmentations peptidiques se produisent essentiellement au niveau des liaisons amides du squelette du peptide. On peut également observer, dans certaines conditions, des clivages de la chaîne latérale des acides aminés : il s'agit soit de fragments de deuxième génération (ions « petit-fils »), soit de fragments obtenus par MS/MS à haute énergie [Biemann, 1990a]. Ces observations ont permis à Roepstoff et Fohlman d'établir une nomenclature des fragmentations peptidiques [Roepstoff et Fohlman, 1984]. Cette nomenclature complétée ensuite par Biemann sert aujourd'hui de référence [Biemann, 1990b].

3.2. Nomenclature des fragmentations peptidiques

On différencie deux types d'ions issus de la fragmentation du squelette peptidique (figure 16) :

- | Les ions pour lesquels la charge positive est portée par la partie contenant l'acide aminé N-terminal (ions de série a, b, c).
- | Les ions pour lesquels la charge positive est portée par la partie contenant l'acide aminé C-terminal (ions de série x, y, z).

Figure 16 : nomenclature des ions issus de la fragmentation du squelette peptidique d'après Biemann (en mode d'ionisation positive).



Les ions b et y sont rencontrés lors des fragmentations à basse ou à haute énergie et sont généralement parmi les plus abondants. Les ions obtenus par fragmentation des chaînes latérales (ions de série d, v, w) apparaissent essentiellement lors des fragmentations à haute énergie. Ils résultent de clivages homolytiques des liaisons carbone-carbone des chaînes latérales ou de mécanismes radicalaires avec réarrangement sur un site éloigné de la charge : ce mécanisme est appelé « charge remote fragmentation ».

3.3. La détermination de la séquence peptidique

Les spectres de fragmentations peptidiques à basse énergie donnent majoritairement des ions de série y et b. La séquence partielle ou totale du peptide est déterminée par la succession de clivages consécutifs appartenant à une même série (y ou b). Ainsi, deux clivages consécutifs induisent la formation d'ions fragments dont la différence de masse équivaut à la masse moléculaire caractéristique d'un acide aminé.

Il existe cependant des ambiguïtés lors de l'interprétation des spectres de fragmentation. Il est, par exemple, difficile de distinguer :

- | les acides aminés isobares, leucine et isoleucine
- | les acides aminés de masses très proches comme la lysine et la glutamine ($\Delta M = 0,036$ Da)
- | certaines combinaisons de résidus de masse égale à celle d'un autre résidu (par exemple deux glycines ont la même masse qu'une asparagine).

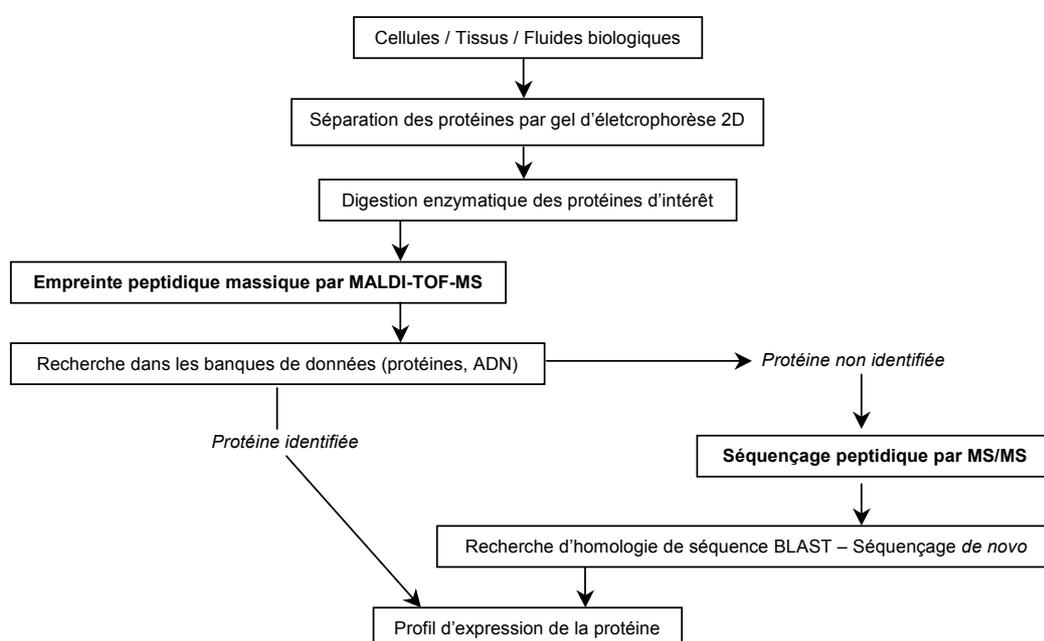
Les acides aminés isobares ou de masses proches peuvent être discriminés en générant des fragmentations de leurs chaînes latérales (ions de série d) par spectrométrie de masse en tandem à haute énergie [Biemann, 1990a].

Dans le cas particulier des peptides issus de la digestion trypsique, les fragments sont porteurs d'une charge permanente apportée par la lysine ou l'arginine en position C-terminale. La présence de cette charge a pour effet de favoriser la détection des fragments de type y lors de l'analyse par spectrométrie de masse en tandem. Par conséquent, l'interprétation des spectres de fragmentation est plus simple et permet d'obtenir un meilleur recouvrement des séquences primaires.

4. Place de la spectrométrie de masse dans la stratégie protéomique

La stratégie générale d'analyse protéomique décrite figure 17 comprend deux alternatives complémentaires : l'empreinte peptidique massique par MALDI-TOF-MS et le séquençage peptidique par MS/MS.

Figure 17 : stratégie générale d'analyse protéomique



4.1. La stratégie de l'empreinte peptidique massique

Cette stratégie a été décrite par plusieurs équipes comme alternative au séquençage des peptides par la dégradation chimique d'Edman [Henzel et coll., 1993 ; James et coll., 1993 ; Pappin et coll. 1993 ; Yates et coll., 1993].

Les protéines séparées par gel 2D sont découpées et digérées « in gel » par une enzyme, souvent la trypsine. Les peptides protéolytiques sont extraits du gel puis analysés par MALDI-TOF-MS afin de générer une liste des masses expérimentales des peptides présents dans l'échantillon. Grâce à un outil informatique appelé moteur de recherche, cette liste est comparée aux listes de masses théoriques obtenues par digestion *in silico* (théorique) des protéines présentes dans les banques de données. Les critères pouvant permettre d'aboutir à l'identification de la protéine sont les suivants :

- | l'écart de m/z entre la valeur théorique et la valeur expérimentale
- | le nombre de peptides reconnus

- | le pourcentage de recouvrement de la séquence protéique.

Néanmoins, l'identification par cette stratégie peut échouer pour les raisons suivantes :

- | la présence de modifications post-traductionnelles
- | la protéine n'est pas répertoriée dans les banques de données.

Dans ce cas, le recours à la spectrométrie de masse en tandem s'impose afin d'obtenir des informations sur la séquence en acides aminés des peptides.

4.2. La stratégie par MS/MS

Les techniques de spectrométrie de masse décrites dans les paragraphes précédents permettent d'obtenir des informations de séquence (MALDI-TOF-TOF, ES-Ion Trap et ES-Q-TOF). Il s'agit de techniques de choix par le biais desquelles il est possible de sélectionner spécifiquement un peptide dans un mélange et de le fragmenter tout en conservant une bonne sensibilité.

Il existe, de la même manière que pour la stratégie par empreinte peptidique, des moteurs de recherche capables de comparer la liste de masses des fragments peptidiques obtenus avec une liste de masses théoriques [Yates et coll., 1995]. Certains logiciels permettent également l'interrogation dans les banques de données avec une information de séquence partielle, appelée « peptide sequence tag » [Wilm et Mann, 1996].

Dans les cas où la protéine n'est pas répertoriée dans les banques de données, l'approche par séquençage *de novo* s'avère indispensable. Cette approche consiste, par définition, à reconstituer la séquence en acides aminés d'un peptide à partir de son spectre de fragmentation issu de l'analyse par spectrométrie de masse en tandem. Très souvent réalisé « à la main » sur la base des règles de fragmentation des peptides (cf § II.3.2.), le séquençage *de novo* peut également être assisté par un logiciel adapté (Biotools®, Bruker ; PEAKS®, Bioinformatics Solutions Inc).

Les techniques citées précédemment sont, certes très efficaces dans le cadre de l'approche protéomique par MS/MS, mais elles présentent cependant des inconvénients :

- | la technique nanoES-MS/MS off-line est longue à mettre en œuvre et difficilement automatisable
- | la complexité des échantillons peut conduire à des phénomènes de suppression du signal de certains peptides présents dans le mélange.

Dans ce contexte, le développement du couplage nano-LC-MS/MS s'est avéré extrêmement utile pour la stratégie protéomique comme nous nous proposons de l'expliquer dans le paragraphe suivant. Nous avons développé cette approche au laboratoire pendant notre travail de thèse (cf. partie II, mise en œuvre expérimentale).

4.3. Le couplage avec la chromatographie liquide

4.3.1. La stratégie LC-MS/MS à faible débit

Face à la complexité des mélanges peptidiques analysés et dans l'objectif d'obtenir un maximum d'informations de séquences en acides aminés, une étape préalable de fractionnement chromatographique a été introduite. Ainsi, la LC-MS/MS est devenue une approche complémentaire ou alternative à la stratégie associant gel 2D et MALDI-MS.

Le développement de colonnes de chromatographie capillaires (entre 100 μm et 1 mm de diamètre interne) et de nano-colonnes (diamètre interne inférieur à 100 μm), ainsi que l'apparition de systèmes HPLC pouvant travailler à des débits faibles (de 200 nL / min à 4 μL / min) ont permis le couplage à la spectrométrie de masse tout en conservant une bonne sensibilité (femtomolaire).

L'introduction du couplage LC-MS/MS a permis l'automatisation de l'analyse protéomique à différents niveaux :

- | injection des échantillons avec un passeur automatique
- | automatisation de la séparation chromatographique sur micro ou nano-colonnes
- | sélection et fragmentation automatiques des ions précurseurs par le spectromètre de masse
- | traitement des données.

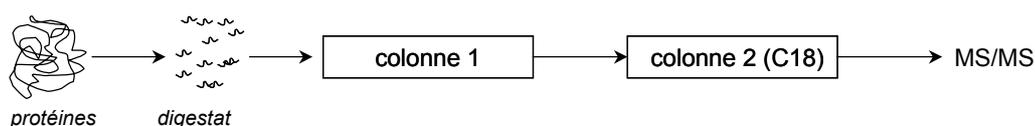
La LC-MS/MS à faible débit a ainsi révolutionné le domaine de l'analyse protéomique permettant d'accéder à un très grand nombre d'informations de séquences peptidiques par digestat de protéine analysé : jusqu'à 2000 peptides peuvent être isolés et fragmentés au cours d'une analyse d'une heure environ.

La mise au point d'un tel système ayant fait l'objet de ce travail de thèse, les différents aspects du couplage LC-MS/MS à faible débit seront détaillés dans la partie II de ce manuscrit (mise en œuvre expérimentale).

4.3.2. La LC-LC-MS/MS

Dans le cadre de l'analyse protéomique, la chromatographie multi-dimensionnelle a été développée comme une stratégie alternative aux gels 2D qui, nous l'avons décrit au paragraphe II.2.4., présente certaines limitations. La combinaison de plusieurs étapes chromatographiques permet de séparer avec une résolution satisfaisante des mélanges protéiques complexes, d'isoler dans ces mélanges des espèces en faible quantité ainsi que des protéines hydrophobes et de pI « extrêmes » (<4 ou >9). La 2D-LC-MS/MS (figure 18) présente, de plus, l'avantage d'être complètement automatisable.

Figure 18 : schéma de principe général de la 2D-LC-MS/MS. Les peptides sont élués par paliers successifs de la colonne 1. A chacun de ces paliers correspond une étape de fractionnement sur la phase inverse (colonne 2).



La combinaison de plusieurs méthodes chromatographiques permet de séparer les peptides protéolytiques selon différentes propriétés :

- | Les colonnes échangeuses de cations ou d'anions permettent d'analyser des peptides respectivement basiques ou acides.
- | Les colonnes de phase inverse (C18) permettent une séparation selon l'hydrophobicité.
- | Certaines colonnes d'affinité sont destinées à isoler une classe de peptides, contenant, par exemple, une histidine (cf technique ICAT décrite au paragraphe suivant) ou porteurs d'une phosphorylation.

Parmi les différentes combinaisons possibles, le couplage associant une colonne d'échange d'ions et une colonne de phase inverse est le plus fréquemment rencontré.

La technique décrite par Link [Link et coll., 1999], dénommée DALPC (« direct analysis of protein complexes using mass spectrometry ») repose sur l'utilisation d'une colonne contenant successivement ces deux types de phases. Les peptides sont élués de la phase échangeuse d'ions par paliers correspondant à des injections de solvant aqueux contenant des quantités croissantes de sels volatils (par exemple l'acétate d'ammonium). A chacun de ces paliers est associée une étape de fractionnement sur la phase inverse à l'aide d'un gradient de solvant organique. Les peptides élués sont directement analysés par le spectromètre de masse.

Cette technique reprise par Washburn et coll. sous le nom de MudPIT (Multidimensional Protein Identification Technology) s'est avérée particulièrement puissante pour l'analyse du protéome de *Saccharomyces cerevisiae* : en 15 étapes de force ionique croissante, près de 1484 protéines ont été détectées et identifiées, parmi lesquelles 131 étaient des protéines membranaires [Washburn et coll., 2001].

Parallèlement à ce type de montage associant deux colonnes en série, il existe également des systèmes intercalant entre les deux types de phase une colonne supplémentaire dite « de pré-concentration » [Davis et coll., 2001].

Ces systèmes requièrent l'utilisation de deux systèmes de pompage et d'une vanne permettant, dans un premier temps, le chargement des peptides de la colonne d'échange d'ions sur la pré-colonne de concentration, leur rinçage, et, dans un second temps, l'éluion sur la colonne de phase inverse. L'avantage principal d'un tel montage est d'éviter la présence, dans la source du spectromètre de masse, de sels pouvant être nuisibles à la détection.

La chromatographie multi-dimensionnelle, ou LCⁿ, a certes étendu le domaine de l'analyse par gel 2D, cependant la résolution complète de tous les composés d'un protéome complexe, comme c'est le cas pour les mammifères, reste un défi. De telles problématiques requièrent certainement un nombre n d'étapes de fractionnement supérieur à 2.

5. La quantification des protéines

5.1. Introduction

Les méthodes décrites précédemment peuvent être appliquées à l'analyse protéomique différentielle qui consiste, par exemple, à comparer le protéome d'un individu sain et celui d'un malade en observant l'apparition ou la disparition de protéines. Toutefois, ces techniques ne permettent pas la quantification précise des protéines identifiées. Il existe différentes stratégies basées sur la MS permettant de répondre à ce type de question.

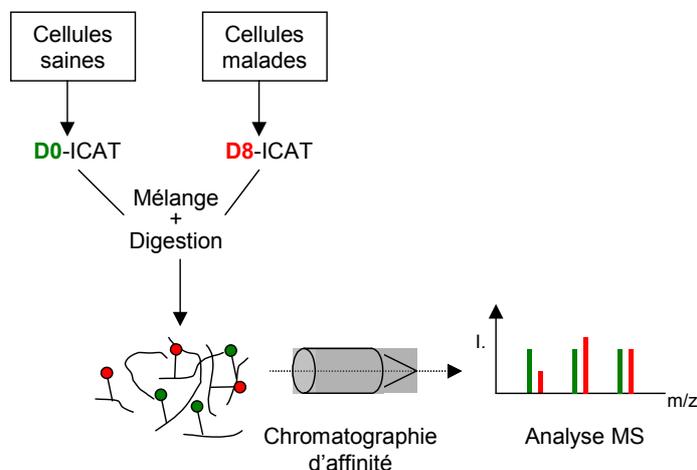
Le développement d'une technique combinant un marquage *in vivo* des lysines avec un isotope stable et la quantification par MS fait l'objet d'un chapitre de cette thèse et sera décrit par la suite.

Récemment, l'équipe de Rudi Aebersold a décrit un marquage isotopique différentiel, appelé ICAT pour « Isotope Coded Affinity Tag ».

5.2. La technologie ICAT

La stratégie ICAT, également appelée stratégie de la signature en acides aminés est résumée figure 19 [Gygi et coll., 1999b ; Gygi et coll., 2000 ; Griffin et coll., 2001]. Le principe de cette approche consiste à marquer sélectivement les cystéines d'un mélange de protéines.

Figure 19 : Stratégie ICAT, un exemple d'analyse différentielle entre deux extraits cellulaires, l'un correspondant à des cellules saines l'autre à des cellules malades.



Dans le cadre d'une approche protéomique différentielle, les deux extraits cellulaires à comparer sont d'abord marqués avec un réactif : pour l'un le réactif est léger (d0), pour l'autre le réactif contient du deuterium (d8). Puis, les deux extraits sont mélangés et soumis à une hydrolyse enzymatique. Le marqueur utilisé possède à son extrémité un groupement biotine permettant ensuite la sélection spécifique des peptides contenant au moins une cystéine après passage du mélange sur une colonne d'affinité (colonne avidine). La dernière étape consiste à analyser le mélange peptidique par MALDI-MS ou par LC-MS/MS. Deux peptides « homologues » issus chacun d'un des deux extraits cellulaires analysés se distinguent par un écart de 8 Da. Ainsi, le calcul du ratio de l'intensité entre les deux formes permet la quantification relative des protéines présentes dans le mélange.

L'avantage principal de cette technique est de s'affranchir de la complexité des extraits cellulaires puisque l'on ne conserve que les peptides comportant au moins une cystéine. Il est ainsi possible de quantifier des protéines présentes en faible quantité.

5.3. Une autre approche : ICNAT

Un autre type d'approche, de type ICNAT (« Isotope Coded Non Affinity Tag ») a été développée au laboratoire. Elle consiste également à marquer sélectivement les cystéines mais permet de s'affranchir de certaines contraintes liées à la stratégie ICAT décrite précédemment.

La chromatographie d'affinité est remplacée par une étape de séparation des protéines sur gel 1D suivie d'une digestion trypsique. Ainsi, on évite de perdre certaines protéines (basiques, hydrophobes). Le mélange peptidique est analysé par MALDI-MS ou par LC-MS/MS. Les peptides sans cystéine sont conservés et contribuent à l'identification protéique.

L'agent alkylant (N-Ethylodoacétamide) utilisé est relativement simple à synthétiser et peu coûteux. Contrairement au réactif utilisé dans l'approche ICAT, il n'entraîne pas de modification trop importante de la masse de la protéine (85,1 Da par cystéine pour le réactif léger ou 90,1 Da pour la forme deutérée). D'autre part, les couples de peptides marqués coéluent lors de la séparation chromatographique : il n'y a pas de retard d'élution de la forme légère par rapport à la forme deutérée comme observé avec la stratégie ICAT.

6. Conclusion

Nous avons ici décrit l'état actuel des connaissances en protéomique. Les études qualitatives et quantitatives menées dans ce cadre permettent d'établir un premier lien entre le produit des gènes et les processus biologiques. Mais le travail est loin d'être achevé. En effet, plus de 60% des protéines sont glycosylées. Il a été démontré que ces glycosylations jouaient un rôle important dans la structure et la fonction des protéines [Rud et Dwek, 1997 ; Van Den Steen et coll., 1998]. Par ailleurs, d'autres modifications post-traductionnelles, telles que les phosphorylations, sont également au centre de l'activité protéique [Liao et coll., 1994].

Par conséquent, la caractérisation des modifications post-traductionnelles représente une étape clé de l'analyse protéomique. Nous nous proposons maintenant de détailler les méthodes utilisées afin de localiser et caractériser ces modifications. Nous verrons que la spectrométrie de masse en tandem est un outil puissant particulièrement adapté à cette étude.

BIBLIOGRAPHIE

Aebersold R., Mann M.

Mass spectrometry-based proteomics, *Nature*, **2003**, 422, 198-206.

Anderson L., Seilhamer J.

A comparison of selected mRNA and protein abundances in human liver, *Electrophoresis*, **1997**, 18, 533-537.

Anderson N.L., Anderson N.G.

Proteome and proteomics : new technologies, new concepts, and new words, *Electrophoresis*, **1998**, 19, 1853-1861.

Biemann K.

Sequencing of peptides by tandem mass spectrometry and high energy collision induced dissociation, *Meth. Enzymol.*, **1990a**, 193, 455-479.

Biemann K.

Nomenclature for peptide fragment ions (positive ions), *Meth. Enzymol.*, **1990b**, 193, 886-887.

Bjellqvist B., Ek K., Righetti P.G., Gianazza E., Gorg A., Westermeier R., Postel W.

Isoelectric focusing in immobilized pH gradients : principle, methodology and some applications., *J. Biochem. Biophys. Methods*, **1982**, 6 (4), 317-339.

Bjellqvist B., Sanchez J.C., Pasquali C., Ravier F., Paquet N., Frutiger S., Hughes G.J., Hochstrasser D.

Micropreparative two-dimensional electrophoresis allowing the separation of samples containing milligram amounts of proteins, *Electrophoresis*, **1993**, 14 (12), 1375-8.

Davis M.T., Beierle J., Bures E.T., McGinley M.D., Mort J., Robinson J.H., Sparh C.S., Yu W., Luethy R., Patterson S.D.

Automated LC-LC-MS-MS platform using binary ion exchange and gradient reversed-phase chromatography for improved proteomic analyses, *J. Chromatogr. B*, **2001**, 752, 281-291.

Giometti C.S., Gemmell M.A., Tollaksen S.L., Taylor J.

Quantitation of human leucocyte proteins after silver staining : a study with two-dimensional electrophoresis, *Electrophoresis*, **1991**, 12, 536-543.

Griffin T.J., Gygi S.P., Rist B., Aebersold R.

Quantitative proteomic analysis using a MALDI quadrupole time of flight mass spectrometer, *Anal. Biochem.*, **2001**, 73, 978-986.

Gygi S.P., Rochon Y., Franz B.R., Aebersold R.

Correlation between protein and mRNA abundance in yeast, *Mol. Cell. Biol.*, **1999a**, 19, 1720-1730.

Gygi S.P., Rist B., Gerber S.A., Turecek F., Gelb M.H., Aebersold R.

Quantitative analysis of complex protein mixtures using isotope-coded affinity tags, *Nat. Biotechnol.*, **1999b**, 17, 994-999.

Gygi S.P., Rist B., Aebersold R.

Measuring gene expression by quantitative proteome analysis, *Curr. Opin. in Biotech.*, **2000**, 11, 396-401.

Hanash Sam

Disease proteomics, *Nature*, **2003**, 422, 226-232.

Henzel W.J., Billeci T.M., Stults J.T., Wong S.C., Grimley C., Watanabe C.

Identifying proteins from two-dimensional gels by molecular mass searching of peptide fragments in protein sequence databases, *Proc. Natl. Acad. Sci. USA*, **1993**, 90, 5011-5015.

James P., Quadroni M., Carafoli E., Gonnet G.

Protein identification by mass profile fingerprinting, *Biochem. Biophys. Res. Commun.*, **1993**, 195, 58-64.

Kahn P.

From genome to proteome : looking at a cell's proteins, *Science*, **1995**, 270, 369-370.

Liao P.C., Leykam J., Andrews P.C., Gage D.A., Allison J.

An approach to locate phosphorylation sites in a phosphoprotein : mass mapping by combining specific enzymatic degradation with matrix-assisted laser desorption/ionization mass spectrometry, *Anal. Biochem.*, **1994**, 219, 9-20.

Link A.J., Eng J., Schieltz D.M., Carmack E., Mize G.J., Morris D.R., Garvik B.M., Yates J.R.

Direct analysis of protein complexes using mass spectrometry, *Nature Biotechnology*, **1999**, 17, 676-682.

Lottspeich F.

Proteome analysis : a pathway to the functional analysis of proteins, *Angew. Chem. Int. Ed.*, **1999**, 38, 2476-2492.

Luche S., Diemer H., Tastet C., Chevallet M., Van Dorsselaer A., Leize E., Rabilloud T.

About thiol derivatization and resolution of basic proteins in two-dimensional electrophoresis, *Proteomics*, **2003**, Sous presse.

O'Farrel

High resolution two-dimensional electrophoresis of proteins, *J. Biol. Chem.*, **1975**, 250, 4007-4021.

Pappin D.J.C., Hojrup P., P and Bleasby A.J.

Rapid identification of proteins by peptide-mass fingerprinting, *Curr. Biol.*, **1993**, 3, 327-332.

Rabilloud T.

Solubilization of proteins for electrophoretic analyses, *Electrophoresis*, **1996**, 17, 813-829.

Rabilloud T., Adessi C., Giraudel A., Lunardi J.

Improvement of the solubilization of proteins in two-dimensional electrophoresis with immobilized pH gradients, *Electrophoresis*, **1997**, 18, 307-316.

Rabilloud T., Strub J.M., Luche S., Van Dorsselaer A., Lunardi J.

A comparison between Sypro Ruby and ruthenium II tris (bathophenanthroline disulfonate) as fluorescent stains for protein detection in gels, *Proteomics*, **2001**, 1, 699-704.

Rabilloud T.

Protéomes et analyse protéomique : réalités et perspectives, *Spectra analyse*, **2001**, 219, 9-15.

Radola B.J., Dunn M.J., Chrambach A.

Advances in electrophoresis, **1997**, 1-109.

Richert S., Luche S., Van Dorselaer A., Leize E., Rabilloud T.

About the mechanism of interference of silver staining with peptide mass spectrometry and how to cure it, *Proteomics*, **2003**, sous presse.

Roepstorff P., Fohlman J.

Proposal for a common nomenclature for sequence ions in mass spectra of peptides, *Biomed. Mass Spectrom.*, **1984**, 11, 601.

Rudd P.M., Dwek R.A.

Glycosylation : heterogeneity and the 3D structure of proteins, *Crit. Rev. Biochem. Mol. Biol.*, **1997**, 32, 1-100.

Santoni V., Molloy M., Rabilloud T.

Membrane proteins and proteomics : un amour impossible?, *Electrophoresis*, **2000**, 21, 1054-1070.

Tyers M., Mann M.

From genomics to proteomics, *Nature*, **2003**, 422, 193-197.

Van Den Steen P., Rudd P.M., Dwek R.A., Opdenakker G.

Concepts and principles of O-linked glycosylation, *Crit. Rev. Biochem. Mol. Biol.*, **1998**, 33, 151-208.

Washburn M.P., Wolters D., Yates J.R.

Large-scale analysis of the yeast proteome by multidimensional protein identification technology, *Nature Biotechnology*, **2001**, 19, 242-247.

Westbrook J.A., Yan J.X., Wait R., Welson S.Y., Dunn M.J.

Zooming-in on the proteome : very narrow-range immobilised pH gradients reveal more protein species and isoforms, *Electrophoresis*, **2001**, 22, 2865-2871.

Wilm M., Mann M.

Analytical properties of the nanoelectrospray ion source, *Anal. Chem.*, **1996**, 68, 1-8.

Yates J.R., Eng J.K., McCormack A., Schieltz D.

Method to correlate tandem mass spectra of modified peptides to amino acid sequences in the protein database, *Anal. Chem.*, **1995**, 67, 1426-1436.

Yates J.R. 3rd, Speicher S., Griffin P.R., Hunkapiller T.

Peptide mass maps : a highly informative approach to protein identification, *Anal. Biochem.*, **1993**, 214, 397-408.

III. LA CARACTERISATION DES MODIFICATIONS POST-TRADUCTIONNELLES PAR SPECTROMETRIE DE MASSE

1. Introduction

Les modifications chimiques covalentes ainsi que la perte d'acides aminés par l'action d'exopeptidases sur les protéines [Biemann, 1990] sont appelées modifications post-traductionnelles. Elles ne sont pas codées par l'ADN cellulaire mais se produisent lors de la biosynthèse des protéines et peuvent intervenir de manière spécifique tout au long de la vie de la cellule. Bien que leurs rôles ne soient pas encore toujours établis, les modifications naturelles des protéines sont nécessaires à l'activité protéique. Les modifications post-traductionnelles peuvent notamment influencer le repliement des protéines ainsi que leur régulation.

2. Les principales modifications post-traductionnelles

Les principales modifications post-traductionnelles ainsi que la différence de masse qui les caractérise (par rapport à la protéine non modifiée) sont résumées dans le tableau 3.

Tableau 3 : nature des principales modifications post-traductionnelles. D'après [De Hoffmann et coll., 1994].

Nature des modifications post-traductionnelles	Variation de masse (Da)	Nature des modifications post-traductionnelles	Variation de masse (Da)
Pyroglutamique à partir de Gln	- 17,03	Carboxyméthylation	+ 58,04
Réduction d'un pont disulfure	- 2,02	Phosphorylation	+ 79,98
Déamidation de Asn ou Gln	+ 0,98	Sulfonation	+ 80,06
Méthylation	+ 14,03	Ethylpyridylation	+ 105,12
Hydroxylation	+ 16,00	Cystéinylation	+ 119,14
Oxydation de Met	+ 16,00	Acide lipoïque	+ 188,30
Hydrolyse d'une liaison peptidique	+ 18,02	Farnésylation	+ 204,36
Formylation	+ 28,01	Myristoylation	+ 210,36
Acétylation	+ 42,04	Biotinylation	+ 226,29
Propylation	+ 42,08	Pyridoxal phosphate	+ 231,14
Carboxylation de Asp ou Glu	+ 44,01	Palmitoylation	+ 238,41
Carbamidométhylation	+ 57,03	Stéaroylation	+ 266,47

Les prochains paragraphes seront consacrés aux glycosylations qui sont des modifications très fréquemment rencontrées et qui ont, par ailleurs, fait l'objet d'une partie de ce travail de thèse.

3. La glycosylation des protéines

3.1. Introduction

La glycosylation est l'une des modifications post-traductionnelles la plus répandue. Elle modifie les propriétés physico-chimiques des protéines (solubilité, stabilité et conformation) [Lis et Sharon, 1993]. Elle joue également un rôle de médiateur lors de la reconnaissance des protéines intra et intercellulaires [Stryer, 1988 ; Hegy, 1994]. Les glycosylations, généralement spécifiques d'une espèce,

d'un tissu ou d'une cellule, sont cruciales pour le développement, la croissance et la survie d'un organisme [Varki, 1993].

C'est pourquoi, caractériser les formes glycosidiques des protéines est un enjeu considérable pour comprendre leurs fonctions biologiques. Ces études sont capitales pour l'industrie pharmaceutique qui, par exemple, commercialise des glycoprotéines recombinantes [Gervais et coll., 2002].

3.2. Nature et structure des glycosylations

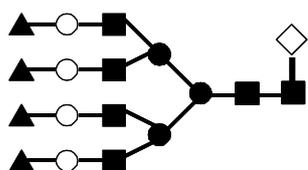
Les polysaccharides sont fixés par liaison covalente sur la protéine et se distinguent en deux catégories suivant la nature de cette liaison : les N- ou les O-glycannes.

- | Les N-glycosylations se fixent sur le squelette protéique par l'intermédiaire de l'azote d'un résidu asparagine.
- | Les O-glycosylations sont liées à la protéine par le groupement hydroxyle des résidus sérine, thréonine, hydroxyproline ou hydroxylysine.

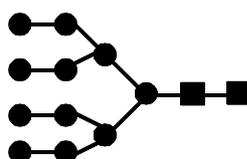
Les N-glycannes possèdent un squelette commun, appelé aussi « core », constitué de trois résidus mannose et de deux résidus N-acétylglucosamine. Le core peut être substitué, soit de mannoses, soit de séquences galactose/N-acétylglucosamine en combinaison avec des acides sialiques et des fucoses. Ces différentes combinaisons forment respectivement des N-glycannes de type oligomannosidiques ou des N-glycannes complexes [Kornfeld et Kornfeld, 1985] (figure 20). Il existe également des structures N-glycannes hybrides possédant les deux types de motifs précédemment cités.

Figure 20 : modèles de structures de N-glycannes et tableau des masses des principaux oses rencontrés.

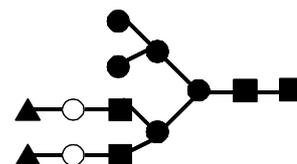
N-glycane de type complexe



N-glycane de type oligomannosidique



N-glycane de type hybride



	Glycanes	Masse (Da)
●	Mannose	162,14
○	Galactose	162,14
◇	Fucose	146,14
■	N-acétylglucosamine	203,19
▲	Acide N-acétylneuraminique	291,16

Certaines séquences peptidiques particulières sur la protéine semblent favoriser les glycosylations, elles sont appelées séquences consensus :

- | Pour les N-glycannes, la séquence tripeptidique, Asn–X–Ser/Thr avec X ≠ Pro, Trp, Glu, et Asp est connue comme site préférentiel de glycosylation [Mellquist et coll., 1998].
- | Pour les O-glycannes, les sites consensus font intervenir des Prolines positionnées de façon bien particulière par rapport au site de glycosylation, à savoir : Pro-Ser/Thr-X-X-Pro [Wilson et coll., 1991] et Pro-Ser/Thr-Pro [Elliott et coll., 1994].

D'autre part, il semblerait également que la structure tridimensionnelle de la protéine influence la O-glycosylation [Elliott et coll., 1994].

3.3. La caractérisation des glycosylations

La diversité structurale des glycosylations est considérable du fait :

- | des combinaisons possibles de l'enchaînement des unités monosaccharidiques
- | de l'isomérisation de position de chaque liaison glycosidique
- | des ramifications
- | de la configuration anomérique de chacun des sucres
- | des modifications possibles sur les monomères (sulfatation ou phosphorylation par exemple).

De plus, les glycoprotéines sont souvent hétérogènes : une même protéine peut présenter plusieurs glycoformes pour un même site de glycosylation.

C'est pourquoi, la mise en œuvre de la caractérisation structurale des glycoprotéines peut s'avérer très complexe et nécessite, pour être complète, de recourir à différentes techniques analytiques.

3.3.1. Isolement et purification des glycannes

L'analyse des glycosylations nécessite le plus souvent d'isoler et de purifier préalablement les glycoconjugués ou les glycannes eux-mêmes [Scherz et Bonn, 1998].

- | Les glycannes peuvent être séparés de leur partie glycoconjuguée suivant des procédures chimiques comme l'hydrolyse par l'acide fluorhydrique, l'hydrazinolyse ou la β -élimination avec NaIO_4 .
- | Il existe également des protocoles enzymatiques avec, par exemple, la PNGase F, les N-acétylglucosaminidase et N-acétylgalactosaminidase ou la O-glycosydase.

Les glycannes ainsi libérés de leur glycoconjugués sont ensuite isolés par chromatographie, électrophorèse capillaire ou sur gel. Suivant la technique choisie, ils peuvent être soumis à une dérivatisation permettant leur détection : acétylation, perméthylation ou introduction d'un chromophore par exemple.

3.3.2. L'analyse par RMN

La technique analytique de choix pour faire face à la complexité structurale des oligosaccharides reste la RMN. Elle permet l'élucidation de la séquence primaire des polysaccharides, la détermination de la nature des liaisons glycosydiques (isoméries de position) et des anoméries [Pope et coll., 1992 ; Nimitz et coll., 1995]. Cependant cette technique présente l'inconvénient d'être longue et fastidieuse à mettre en oeuvre. Mais surtout, elle est peu sensible : une analyse requiert quelques micromoles de glycanne de grande pureté.

3.3.3. L'analyse par spectrométrie de masse

L'analyse structurale par spectrométrie de masse, plus tolérante quant à la pureté des glycannes, se révèle être une approche alternative et complémentaire.

Initialement, les oligosaccharides ont été analysés par FAB-MS et FAB-MS/MS avec une sensibilité de l'ordre de 10 à 100 picomoles [Gillece-Castro et Burlingame, 1990]. Ces techniques permettent l'interprétation de la séquence oligosaccharidique et donnent également des informations quant à l'isométrie de position des liaisons glycosidiques [Garozzo et coll., 1990 ; Reinhold et Santikarn, 1988]. Toutefois, la faible basicité / acidité des oligosaccharides engendre un rendement d'ionisation médiocre en FAB. Le développement de techniques de dérivatisation de l'extrémité réductrice des sucres a permis de pallier ce problème en conférant une charge permanente aux glycannes [Poulter et Burlingame, 1990]. Outre l'amélioration de l'efficacité d'ionisation, la dérivatisation des oligosaccharides facilite l'obtention d'ions fragments plus facilement détectables [Gage et coll., 1992].

Aujourd'hui, les techniques d'ionisation MALDI et ES couplées à la spectrométrie de masse multiple (PSD, MSⁿ) sont utilisées pour la caractérisation des oligosaccharides. Elles permettent l'identification des glycoconjugués (glycopeptides et glycoprotéines) et offrent la possibilité d'obtenir la signature de l'hétérogénéité des diverses glycoformes [Jaquinod et coll., 1999]. Par ailleurs, l'émergence de la MSⁿ par trappe ionique couplée au nano-électrospray a permis d'analyser les glycannes avec une grande sensibilité (de l'ordre d'une centaine d'atomes) [Bahr et coll., 1997].

Toutefois, la MSⁿ ne peut, à elle seule, différencier toutes les différentes formes isomères constitutives d'un glycanne. Seule une stratégie conjointe d'hydrolyses sélectives des liaisons interglycosydiques par des exoglycosidases et d'analyse par spectrométrie de masse offre une solution à ce problème [Mechref et Novotny, 1998 ; Zhao et coll., 1997]. Cette technique s'avère performante mais présente l'inconvénient d'être longue à réaliser.

Nous verrons par la suite que l'apparition de la nouvelle génération d'instruments MALDI-TOF-TOF ouvre une voie nouvelle pour l'étude des oligosaccharides par spectrométrie de masse en tandem. L'analyse de composés de type N-glycannes par cette technique ayant fait l'objet d'une partie de ce travail de thèse, les résultats seront détaillés dans un prochain chapitre.

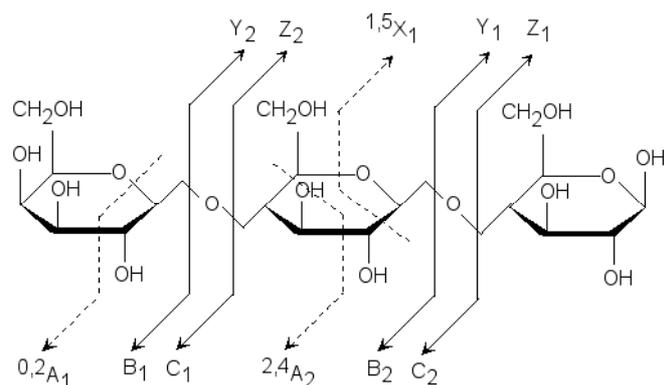
3.4. Les apports de la MSⁿ pour la caractérisation structurale des oligosaccharides

3.4.1. Nomenclature de fragmentation des polysaccharides

Les fragmentations des polysaccharides par spectrométrie de masse se produisent au niveau des liaisons interglycosidiques (Bi, Ci, Yi, Zi) ou intracycliques (Ai et Xi). La nomenclature proposée par Domon et Costello [Domon et Costello, 1988] définit ces fragments (figure 21). On distingue :

- | les fragments qui contiennent l'extrémité réductrice (Xi, Yi, Zi)
- | ceux qui contiennent l'extrémité non-réductrice (Ai, Bi, Ci).

Figure 21 : nomenclature de fragmentation des polysaccharides selon Domon et Costello.



3.4.2. Les fragmentations des polysaccharides à haute et basse énergie

Les processus de fragmentation des polysaccharides sont distincts selon qu'ils se produisent à haute ou à basse énergie :

- | A haute énergie, l'énergie qui est délivrée aux ions précurseurs est suffisamment intense pour que tous les chemins de fragmentation soient empruntés sans aucune sélectivité. C'est le cas des spectres de fragmentations obtenus avec un instrument de type secteur magnétique mais aussi, comme nous le verrons par la suite, avec un instrument TOF-TOF.
- | A basse énergie (instruments triple quadripôle, trappe ionique), l'énergie d'activation est plus faible mais plus longuement communiquée aux ions précurseurs. Les fragments thermodynamiquement les plus stables sont favorisés.

Par conséquent, les ions fragments issus de la rupture des liaisons interglycosidiques (Bi, Ci, Yi, Zi) sont simultanément présents lors des MS/MS à haute ou à basse énergie. Par contre, les ions issus des ruptures intracycliques (Ai, Xi) ne sont favorisés que lors des fragmentations à haute énergie et restent minoritaires, voire quasi inexistantes, à basse énergie [Lemoine et coll., 1993]. Or, l'information structurale que renferment les ions fragments de type Ai et Xi est d'une grande importance puisqu'elle peut permettre de caractériser l'isomérisation de position des liaisons interglycosidiques.

Ainsi, la nature des informations obtenues est variable suivant le choix du spectromètre de masse. La caractérisation structurale des sucres peut se faire à des niveaux plus ou moins avancés : la détermination de la séquence du polysaccharide, le motif de branchement et l'isomérisation de position des liaisons glycosidiques.

3.4.3. La détermination de la séquence et du motif de branchement des polysaccharides

Les fragmentations interglycosydiques (Bi, Ci, Yi, Zi) renseignent sur la séquence en oses et les branchements du polysaccharide. Toutefois, il peut être délicat de déterminer sans ambiguïté si les ions fragments contiennent ou non la partie réductrice du sucre. C'est notamment le cas pour les O-glycannes non dérivés qui présentent une symétrie. Pour les N-glycannes, la partie réductrice correspond à l'extrémité du core (N-acétylgalactosamine ou N-acétylglucosamine) ce qui facilite l'interprétation des spectres de fragmentation.

La définition du sens de lecture de la séquence du sucre peut être favorisée par la dérivatisation chimique de la partie réductrice du polysaccharide. L'incrément de masse apporté par l'agent de dérivatisation permet de localiser l'extrémité réductrice. Les dérivatisations chimiques utilisées sont généralement des chromophores détectables par UV ou par fluorescence (également nécessaires lors de l'étape de séparation et de purification des glycannes). Des travaux récents [Harvey, 2000] ont conduit à l'optimisation des propriétés physico-chimiques de tels agents chimiques. Ces composés sont sélectionnés pour leur grande affinité protonique. Ils favorisent la formation et l'ionisation des ions fragments contenant la partie réductrice du sucre (Yi, Zi).

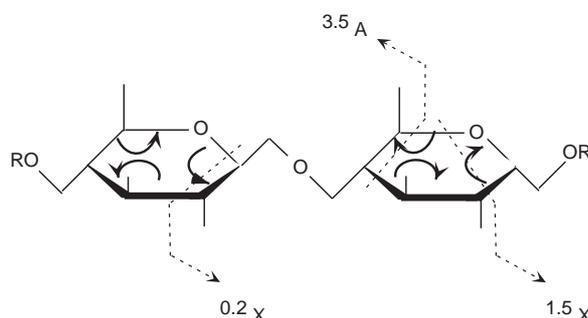
Par ailleurs, les polysaccharides peuvent être marqués sur leur partie réductrice par de l'oxygène 18 (dilution du polysaccharide dans de l' $H_2^{18}O$ pure) [Viseux et coll., 1997]. Les fragments contenant la partie réductrice marquée sont ainsi identifiables par leur profil isotopique.

3.4.4. La détermination des isoméries de position

L'obtention des ions fragments de type Ai et Xi résulte de réarrangements intracycliques (figure 22) et peut dans certains cas être déterminante pour définir la nature des isoméries de position. De nombreuses stratégies ont été développées pour exacerber ce type de clivage :

- | addition de métaux (lithium, cobalt) [König et Leary, 1998]
- | addition d'alcalino-terreux [Lemoine et coll., 1993].

Figure 22 : réarrangements intracycliques à l'origine des ions Ai et Xi.



D'autre part, la perméthylation des polysaccharides renseigne sur les isoméries de positions des liaisons interglycosydiques. Les fragments obtenus à partir de polysaccharides perméthylés sont différenciables suivant qu'ils sont issus d'un simple ou d'un double clivage ce qui évite toute confusion lors de l'interprétation des fragments [Viseux et coll., 1998].

BIBLIOGRAPHIE

Bahr U., Pfenninger A., Karas M., Stahl B.

High-sensitivity analysis of neutral underivatized oligosaccharides by nanoelectrospray mass spectrometry, *Anal. Chem.*, **1997**, 69, 4530-4535.

Biemann K.

Peptides and proteins : overview and strategy, *Meth. Enzymol.*, **1990**, 193, 351-360.

De Hoffmann E., Charette J., Stroobant V.

Spectrométrie de masse, **1994**, 7, 206-215.

Domon B., Costello C.E.

A systematic nomenclature for carbohydrate fragmentations in FAB-MS/MS spectra of glycoconjugates, *Glycoconj. J.*, **1988**, 5, 397-409.

Elliott S., Bartley T., Delorme E., Derby P., Hunt R., Lorenzini T., Parker V., Rohde M.F., Stoney K.

Structural requirements for addition of O-linked carbohydrate to recombinant erythropoietin, *Biochem.*, **1994**, 33, 37, 11237-45.

Gage D.A., Rathke E., Costello C.E., Jones M.Z.

Determination of sequence linkage of tissue oligosaccharides in caprine beta-mannosidosis by fast atom bombardment, collisionally activated dissociation tandem mass spectrometry, *J. Glycoconj.*, **1992**, 9,3, 126-131.

Garozzo D., Giuffrida M., Impallomeni G., Ballistreri A., Montaudo G.

Determination of linkage position and identification of the reducing end in linear oligosaccharides by negative ion fast atom bombardment mass spectrometry, *Anal. Chem.*, **1990**, 62, 279-286.

Gervais A., Hammel Y.A., Pelloux S., Lepage P., Baer G., Carte N., Sorokine O., Strub J.M., Koerner R., Leize E., Van Dorselaer A.

Glycosylation of human recombinant gonadotrophins : characterization and batch to batch consistency, *Glycobiology*, **2002**, 13 (3), 179-89.

Gillece-Castro B., Burlingame A.L.

Oligosaccharide characterization with high energy collision induced dissociation mass spectrometry, *Meth. Enzymol.*, **1990**, 193, 689-712.

Harvey D.J.

Electrospray mass spectrometry and fragmentation of N-linked carbohydrates derivatized at the reducing terminus, *J. Am. Soc. Mass Spectrom.*, **2000**, 11, 900-915.

Hegy G.

Etude des glycoprotéines par spectrométrie de masse avec ionisation électrospray, *Thèse de l'université Louis Pasteur de Strasbourg*, **1994**.

Jaquinod M., Holtet T.L., Etzerodt M., Clemmensen I., Thogersen H.C., Roepstorff P.

Mass spectrometric characterization of post-translational modification and genetic variation in human tetranectin, *Biol. Chem.*, **1999**, 11, 1307-1314.

Kornfeld R., Kornfeld S.

Assembly of asparagine-linked oligosaccharides, *Ann. Rev. Biochem.*, **1985**, 54, 631-664.

König S., Leary J.A.

Evidence for linkage position determination in cobalt coordinated pentasaccharides using ion trap mass spectrometry, *J. Am. Soc. Mass Spectrom.*, **1998**, 11, 1125-1134.

Lemoine J., Fournet B., Despeyroux D., Jennings K.R., Rosenberg R., De Hoffmann E.

Collision induced dissociation of alkali metal cationized and permethylated oligosaccharides : influence of the collision energy and of the collision gas for the assignment of linkage position, *J. Am. Soc. Mass Spectrom.*, **1993**, 4, 197-203.

Lis H., Sharon N.

Protein glycosylation. Structural and functional aspects, *Eur. J. Biochem.*, **1993**, 218, 1-27.

Mechref Y., Novotny M.V.

Mass spectrometric mapping and sequencing of N-linked oligosaccharides derived from submicrogram amounts of glycoproteins, *Anal. Chem.*, **1998**, 3, 455-463.

Mellquist J.L., Kasturi L., Spitalnik S.L., Shakin-Eshleman S.H.

The amino acid following an Asn-X-Ser/Thr sequon is an important determinant of N-linked core glycosylation efficiency, *Biochem.*, **1998**, 37, 6833-6837.

Nimtz M., Wray V., Rüdiger A., Conradt H.S.

Identification and structural characterization of a mannose-6-phosphate containing oligomannosidic N-glycan from human erythropoietin secreted by recombinant BHK-21 cells, *FEBS Lett.*, **1995**, 365, 2-3, 203-208.

Poppe L., Stuïke-Prill R., Meyer B., Van Halbeek S.A.

The solution conformation of sialyl-alpha (2-6)-lactose studied by modern NMR techniques and Monte Carlo simulations, *J. Biomol. NMR*, **1992**, 2, 2, 109-136.

Poulter L., Burlingame A.L.

Desorption mass spectrometry of oligosaccharides coupled with hydrophobic chromophores, *Meth. Enzymol.*, **1990**, 193, 661-689.

Reinhold V.N., Santikarn S.

Structural characterization of complex carbohydrates, *Adv. Exp. Med. Biol.*, **1988**, 228, 727-758.

Scherz H., Bonn G.

Analytical chemistry of carbohydrates, *George Thieme Verlag Stuttgart ed. New York*, **1998**, **Stryer L.**

Biochemistry, ed. *W.H. Freeman, New York*, **1988**, 331-348.

Varki A.

Biological roles of oligosaccharides : all of theories are correct, *Glycobiol.*, **1993**, 3, 97-130.

Viseux N., De Hoffmann E., Domon B.

Structural analysis of permethylated oligosaccharides by electrospray tandem mass spectrometry, *Anal. Chem.*, **1997**, 16, 3193-3198.

Viseux N., De Hoffmann E., Domon B.

Structural assignment of permethylated oligosaccharide subunits using sequential tandem mass spectrometry, *Anal. Chem.*, **1998**, 23, 4951-4959.

Wilson I.B., Gavel Y., Von Heijne G.

Amino acid distributions around O-linked glycosylation sites, *J. Biochem.*, **1991**, 275, 529-534.

Zhao Y., Kent S.B., Chait B.T.

Rapid, sensitive structure analysis of oligosaccharides, *Proc. Natl. Acad. Sci. USA*, **1997**, 5, 1629-1633.

PARTIE II :

MISE EN ŒUVRE EXPERIMENTALE

LE COUPLAGE NANOLC-MS/MS AVEC UNE TRAPPE IONIQUE

MISE EN ŒUVRE EXPERIMENTALE

LE COUPLAGE NANOLC-MS/MS AVEC UNE TRAPPE IONIQUE

I. INTRODUCTION

Avec l'apparition de la chromatographie liquide capillaire, la diminution du débit (1 à 10 $\mu\text{L}/\text{min}$) liée à l'utilisation de colonnes de faible diamètre interne (150 à 500 μm) a largement contribué à augmenter la sensibilité analytique en LC. En effet, plus le débit chromatographique est faible, plus le volume d'élution des analytes diminue ce qui permet d'augmenter leur concentration à la sortie de la colonne et donc au niveau du détecteur.

Par conséquent, ces avantages, liés à la miniaturisation des colonnes chromatographiques, ont permis d'accroître la sensibilité de l'analyse protéomique par LC capillaire-MS/MS [Griffin et coll., 1991 ; Yates et coll., 1996].

Par ailleurs, la miniaturisation de la source ES en source nano-ES a contribué à augmenter les performances de l'analyse par spectrométrie de masse en termes de tolérance aux sels, de sensibilité et de précision de mesure [Wilm et Mann, 1996]. Les qualités d'un tel dispositif (nanoES off-line, cf. Introduction - Bibliographie) ont été exploitées, notamment pour développer l'approche protéomique [Shevchenko et coll., 1996].

Le développement récent des colonnes chromatographiques avec l'apparition de nano-colonnes (diamètre interne inférieur à 150 μm) permet de travailler à des débits très faibles (10 à 1000 nL/min) compatibles avec la spectrométrie de masse nanoES. Ainsi, le couplage nanoLC-MS/MS on-line offre une sensibilité inégalée (femtolaire).

Dans ce cadre, notre travail a consisté à mettre au point les couplages LC capillaire-MS/MS (4 $\mu\text{L}/\text{min}$) et nanoLC-MS/MS (200 nL/min) avec une trappe ionique. Cette étude comprend l'optimisation des paramètres chromatographiques, de l'interface avec le spectromètre de masse ainsi que des paramètres d'acquisition de la trappe ionique dans le but d'automatiser toutes les étapes de l'analyse.

Ce travail a été réalisé en collaboration avec les sociétés Bruker et Agilent pour le développement, en protéomique, des conditions d'analyse inhérentes au couplage entre la trappe ionique Esquire

3000+ et les systèmes chromatographiques HP 1100 Series Capillary HPLC System et HP 1100 Series Nano HPLC System.

II. LES SYSTEMES HPLC A FAIBLE DEBIT

1. Le MicroBlotter

1.1. Le système chromatographique

Lorsque j'ai débuté mon travail de thèse au laboratoire, nous disposions d'un système HPLC appelé MicroBlotter commercialisé par la société Applied Biosystem. Cet appareil était conçu à l'origine pour séparer les peptides issus de la digestion sur une colonne de phase inverse et pour les collecter sur une membrane de PVDF qui pouvait être placée directement dans un séquenceur de protéines [Hsi, 1997]. En 1997, le MicroBlotter apportait donc une miniaturisation dans le domaine de la préparation de l'échantillon pour le micro-séquençage protéique.

Ce système, conçu pour délivrer des micro-débits (quelques $\mu\text{L}/\text{min}$) m'a été utile, dans un premier temps, pour appréhender les contraintes analytiques liées aux faibles débits. Cependant, l'arrivée de la nanoHPLC pour les stratégies de séquençage par spectrométrie de masse ont rendu l'instrument obsolète. En effet, bien qu'il soit envisageable d'installer un diviseur de débit sur le MicroBlotter, ce montage n'offre pas la stabilité de certains systèmes nanoHPLC aujourd'hui commercialisés.

La configuration du système HPLC utilisé pour nos travaux est décrite dans la partie expérimentale de cette thèse. Nous l'avons équipé d'une colonne de phase inverse adaptée à l'analyse des peptides (0.3×150 mm) et travaillé à un débit de $4 \mu\text{L}/\text{min}$ compatible avec la source microspray de la trappe ionique (cf. partie expérimentale).

Pour le couplage, le système HPLC est connecté à la source du spectromètre de masse avec un capillaire en silice de $50 \mu\text{m}$ de diamètre interne ajusté au plus court afin de limiter les volumes morts affectant la qualité de l'analyse MS (re-mélange des espèces chromatographiées conduisant à l'élargissement des pics chromatographiques en détection MS).

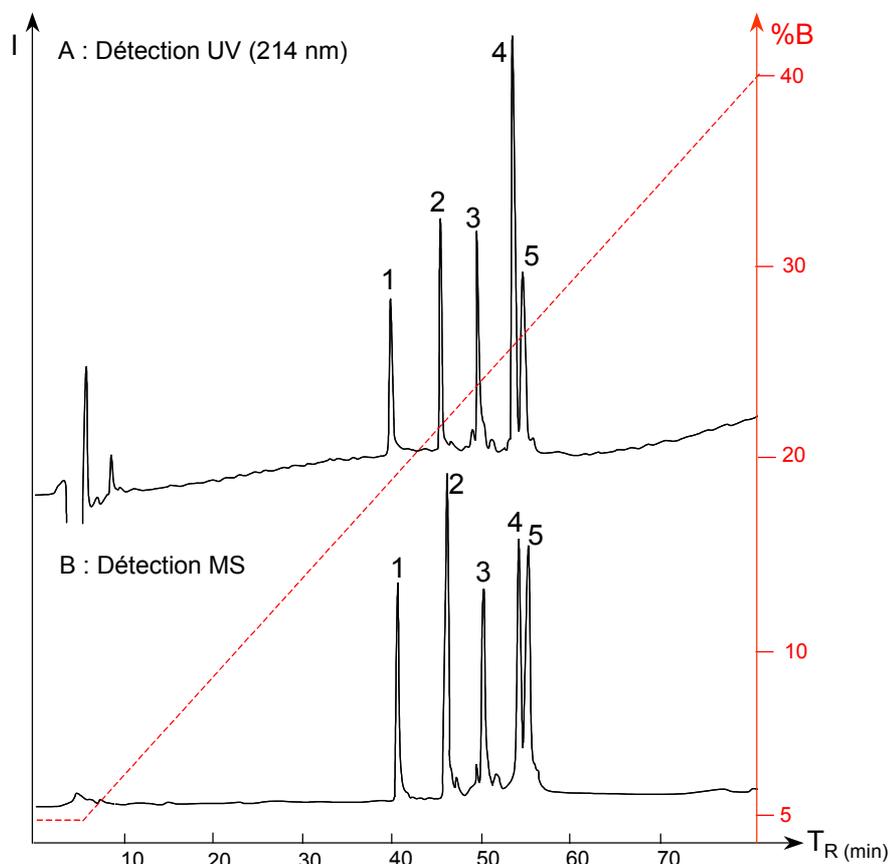
Au fur et à mesure de leur élution de la colonne, les peptides traversent d'abord la cellule UV (détection à 214 nm) puis sont analysés successivement par le spectromètre de masse en mode MS/MS automatique.

1.2. Analyse d'un « mélange peptidique test »

Pour les premiers essais de couplage nous avons injecté un mélange contenant les peptides utilisés habituellement pour étalonner la trappe ionique (cf. partie expérimentale).

Un exemple des profils chromatographiques obtenus en UV et en MS est présenté figure 1. On peut remarquer que les facteurs de réponse obtenus pour chacun des peptides varient selon le mode de détection. Par ailleurs, les temps de rétention en MS sont décalés d'environ 1,5 minutes ce qui correspond au volume mort entre la cellule de détection UV et la source du spectromètre de masse.

Figure 1 : profils chromatographiques obtenus après injection du mélange peptidique composé de 20 picomoles de leu-enképhaline-arginine (1), 12,5 picomoles d'angiotensine (2), 25 picomoles de substance P (3), 20 picomoles de bombésine (4), 17,5 picomoles d'ACTH (5).



1.3. Résultats et optimisations

Les résultats obtenus avec le MicroBlotter nous ont conduit à réaliser quelques optimisations et observations préliminaires utiles pour la suite de notre travail avec d'autres systèmes capillaires et nanoHPLC.

1.3.1. Le gradient

Nous avons, dans un premier temps, étudié le comportement chromatographique des peptides analysés. Avec la colonne utilisée, les peptides du mélange test sont élués avec 20 à 30% d'ACN. Nous verrons par la suite que le comportement des peptides tryptiques est relativement voisin de ceux-ci : ils sont rarement élués au delà de 60% de solvant organique (pour une phase stationnaire équivalente). Les peptides très hydrophobes élués au delà de cette limite sont d'ailleurs difficilement détectés en MS. Nous avons donc adapté le gradient en conséquence (tableau 1). Ce gradient sera utilisé pour la suite de ce travail de thèse.

Le temps nécessaire à l'équilibration du système afin de démarrer une nouvelle analyse a été évalué, après une série de tests, à 20. Ce temps est certes considérable, mais nécessaire, lorsque l'on travaille à faible débit, quelque soit le système HPLC utilisé.

Tableau 1 : gradient utilisé pour la séparation chromatographique de peptides tryptiques en LC capillaire-MS/MS (colonne C18, 0.3 ×150 mm).

Temps (min)	%B	Débit ($\mu\text{L} / \text{min}$)
5.0	5.0	4
65.0	60.0	4
70.0	90.0	4
80.0	5.0	4

1.3.2. Sensibilité et détection UV

Des injections de quantités croissantes de mélange peptidique nous ont permis de tester la sensibilité du système :

- | Dans nos conditions (cellule UV de 1 μL), la détection UV est adaptée à la détection de quantités picomolaires de composés, en dessous elle n'est pas suffisamment sensible.
- | La MS nous a permis de détecter quelques centaines de femtomoles de peptides, selon leur nature.

Par conséquent, dans le cadre de l'analyse protéomique, le détecteur UV utilisé n'offre pas la sensibilité nécessaire à la détection des peptides généralement présents dans le mélange en quantités femtomolaires. La cellule UV et les capillaires de connexion présentent, de plus, l'inconvénient d'introduire des volumes morts supplémentaires affectant la résolution chromatographique.

Cependant, un détecteur UV dans un tel couplage reste un atout non négligeable : en cas de disfonctionnements, il permet de tester indépendamment le système chromatographique en s'affranchissant de la complexité de la détection MS. Ainsi, l'injection régulière de mélange test (en concentration suffisante pour être détecté en UV) permet de vérifier la qualité de la séparation chromatographique. Par ailleurs, dans le cadre de notre travail d'optimisation des paramètres MS et MS/MS inhérents à ce couplage, la détection UV a servi de référence et nous a permis de vérifier systématiquement que seuls les paramètres testés étaient à l'origine des modifications observées sur le chromatogramme en mode de détection MS et MS/MS.

Finalement, l'idéal est de pouvoir disposer d'un détecteur UV que l'on puisse facilement « court-circuiter » du système analytique selon les besoins. Dans cet objectif, les constructeurs proposent des cellules UV plus petites (quelques nL) adaptées à la détection de très faibles concentrations d'analytes. Le système HPLC capillaire d'Agilent que nous avons utilisé par la suite est équipé d'un détecteur UV comprenant une cellule de 500 nL.

1.4. Difficultés rencontrées

1.4.1. En chromatographie

L'utilisation du MicroBlotter nous a permis d'appréhender les principales difficultés rencontrées plus généralement lors de la mise au point d'un couplage capillaire ou nanoLC-MS/MS.

Ces difficultés sont principalement liées à la miniaturisation du système chromatographique (colonne, capillaires et connexions) et imposent de travailler avec beaucoup de rigueur :

- | Il faut éviter la moindre contamination (des solvants par exemple) qui, à ce niveau de sensibilité, nuit à l'analyse et peut être à l'origine de bouchages du système.
- | La mise en route, la purge et l'équilibration sont des étapes longues et délicates. Pour un fonctionnement optimal, il est impératif d'alimenter en permanence le système avec des solvants.
- | Il faut veiller à limiter au maximum les volumes morts et les pertes en charges, particulièrement au niveau des connexions.
- | A de tels débits, les fuites éventuelles ne sont pas détectables à l'œil nu. Ceci nécessite, par conséquent, de porter une attention particulière à l'évolution de la pression qui est le seul reflet du fonctionnement du système.

1.4.2. En MS

Par ailleurs, si les critères permettant d'évaluer la qualité de la séparation sont ceux classiquement utilisés en chromatographie (résolution, reproductibilité, allure des pics), il faut néanmoins admettre que les chromatogrammes obtenus en mode de détection MS et MS/MS ne sont pas le reflet exact de la qualité d'analyse attendue.

En effet, les pics chromatographiques issus de la détection MS et MS/MS de faibles quantités de produit (femtomolaires) sont parfois peu distincts. Comme nous l'expliquerons par la suite (§IV.1.), la digitalisation d'un pic chromatographique correspond à des étapes successives incluant des séries de MS et de MS/MS. Et, nous verrons que la représentation du signal choisie pour visualiser le chromatogramme (courant d'ions MS + MS/MS ou courant d'ion MS) influe considérablement sur son aspect.

Aussi, après s'être assuré du bon fonctionnement du système chromatographique, les critères qui nous ont permis d'évaluer la qualité de notre couplage LC-MS/MS pour la protéomique sont les suivants :

- | nombre de peptides détectés
- | pourcentage de recouvrement de la protéine identifiée
- | sensibilité et reproductibilité
- | précision de la mesure MS et MS/MS.

Nous proposons d'illustrer ceci, par la suite, au travers des résultats obtenus lors de l'analyse nanoLC-MS/MS de mélanges issus de la digestion trypsique de protéines.

1.5. Les limites du MicroBlotter

Face à notre objectif d'automatiser toutes les étapes de l'analyse LC-MS/MS à faible débit en protéomique, les capacités du MicroBlotter se sont avérées limitées.

Faute d'être équipé d'un injecteur automatique, et à cause, également, des fluctuations du débit liées au système, les résultats obtenus souffrent, en effet, d'un manque de reproductibilité.

Par la suite, nous avons travaillé en collaboration avec la société Agilent, sur un système LC capillaire puis nanoLC. Le principe de fonctionnement de ces instruments associé au développement de méthodes analytiques adaptées à l'analyse protéomique font l'objet des prochains paragraphes.

2. Le système Agilent

2.1. Le système chromatographique

Les principaux éléments constitutifs du système Agilent sont décrits dans la partie expérimentale.

2.1.1. Principe de fonctionnement de la pompe et régulation du débit

Le débit de travail (en LC capillaire ou en nanoLC) est régulé par un diviseur de débit dynamique constitué d'une vanne électromagnétique et d'une cellule de mesure. Son principe de fonctionnement est schématisé figure 2.

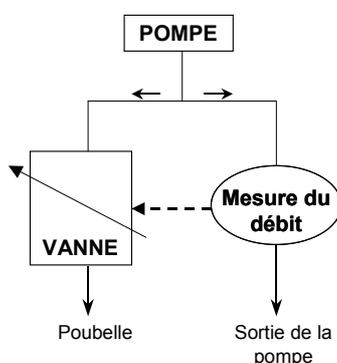
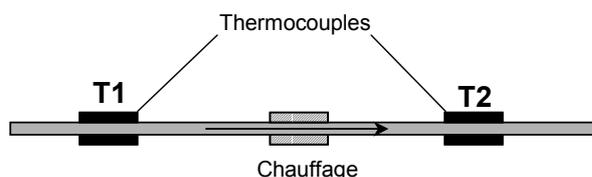


Figure 2 : principe de fonctionnement du diviseur de débit qui équipe les pompes capillaire et nanoLC Agilent.

La cellule de mesure comprend un système de chauffage encadré par deux thermocouples montés en série qui permettent de mesurer la différence de température entre deux points de la ligne (figure 3). Cette mesure dépend de la viscosité des solvants et est directement proportionnelle au débit de phase liquide. Ainsi, le débit à la sortie du diviseur est mesuré en permanence et, cette information est transmise à la vanne électromagnétique dont l'ouverture est ajustée en conséquence. Ce système de contrôle électronique est conçu pour réguler finement le débit : les fluctuations observées dépassent rarement 0.1% du débit final dans nos conditions de travail (4 μ L/min en LC capillaire et 200 nL/min en nanoLC).

Figure 3 : principe de fonctionnement de la cellule de mesure. La différence de température mesurée est proportionnelle au débit et dépend de la nature du solvant.



Il s'agit de choisir un compromis entre la vitesse de formation du gradient et la consommation en solvants (non négligeable dans nos conditions). Cependant, le débit primaire n'altère en rien la qualité de la séparation chromatographique puisque le débit de travail, régulé par le diviseur de débit, reste constant.

Le débit primaire de la pompe est défini par l'utilisateur. Pour nos analyses, nous avons opté pour les gammes de débit suivantes :

- | en LC capillaire, 500 à 800 $\mu\text{L}/\text{min}$
- | en nanoLC, 200 à 500 $\mu\text{L}/\text{min}$.

2.1.2. LC capillaire et nanoLC

Les deux systèmes sur lesquels nous avons travaillé (LC capillaire et nanoLC) fonctionnent sur le même principe. Ils diffèrent par la gamme de débit qu'ils sont capables de délivrer liée à la technologie de la cellule de mesure qui a été améliorée afin d'atteindre les nano-débits.

- | Pour la LC capillaire, la gamme de travail optimale est de 1 à 20 $\mu\text{L}/\text{min}$.
- | Le système nanoLC permet de travailler entre 0.1 à 1 $\mu\text{L}/\text{min}$.

Quelques modifications ont également été apportées au système nanoLC, essentiellement afin de limiter les volumes morts critiques à de si faibles débits :

- | La chambre de mélange à la sortie de la pompe binaire a été retirée.
- | Les capillaires de 50 μm de diamètre interne utilisés en LC capillaire ont été remplacés par des capillaires de 25 μm .

2.1.3. Les colonnes chromatographiques

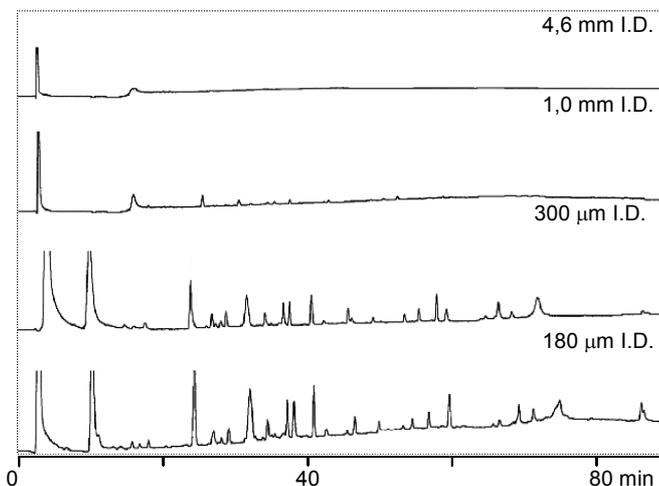
Les pompes décrites précédemment représentent certes une technologie nouvelle à la clé de la qualité analytique des couplages capillaire et nanoLC-MS/MS. Cependant, il n'en serait rien sans les récentes évolutions qui ont concerné les colonnes de chromatographie. Le développement de nouvelles colonnes de diamètre interne de plus en plus petit a permis de travailler à des débits de plus en plus faibles (tableau 2) et d'accéder à une sensibilité femtomolaire dans le domaine du couplage LC/MS. Le gain en sensibilité obtenu avec des colonnes de faible diamètre interne est illustré figure 4.

Tableau 2 : classification des colonnes de chromatographie liquide selon leur diamètre [Vissers et coll., 1997].

Nom générique	Diamètre interne	Débit
Préparative	> 4,6 mm	> 3 mL / min
Conventionnelle	3,2 à 4,6 mm	0,5 à 2 mL / min
Analytique	1,5 à 3,2 mm	100 à 500 μL / min
Micro-LC	0,5 à 1,5 mm	10 à 100 μL / min
LC capillaire	150 à 500 μm	1 à 10 μL / min

Nano-LC | 10 à 150 μm | 10 à 1000 nL / min

Figure 4 : injection de 2 picomoles de myoglobine digérée à la trypsine et détection UV à 206 nm. Le gain de sensibilité entre deux colonnes est défini par le ratio entre les carrés des diamètres interne ($= \text{I.D.1}^2 / \text{I.D.2}^2$).
D'après LC Packings.



Pour notre étude, nous avons utilisé :

- | en LC capillaire (4 $\mu\text{L}/\text{min}$) une colonne de 300 μm de diamètre interne (C18, 150 mm)
- | en nanoLC (200 nL/min) une colonne de 75 μm de diamètre interne (C18, 150 mm).

2.2. Le montage de pré-concentration

La présence de sels dans les mélanges tryptiques en protéomique est fréquente et nuisible à la sensibilité de la détection en MS. Par ailleurs, les quantités analysées sont généralement si faibles qu'elles requièrent, avant la séparation des analytes sur la colonne, d'être concentrées afin d'éviter leur dispersion dans le système chromatographique. Pour ces raisons, la mise au point d'un système permettant d'enrichir et de dessaler les échantillons s'avère indispensable dans le montage LC-MS/MS à faible débit.

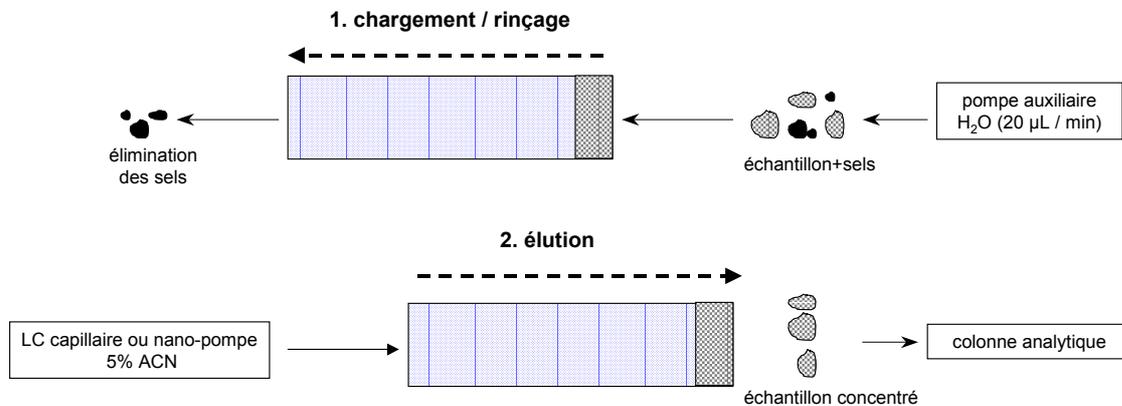
Le principe de fonctionnement de ce système est schématisé figure 5 et repose sur l'utilisation d'une pré-colonne (C18, 0.3 \times 5 mm) et d'une pompe délivrant un solvant aqueux appelée pompe auxiliaire (la configuration du montage est détaillée dans la partie expérimentale).

Dans un premier temps, l'échantillon est entraîné sur la pré-colonne grâce au débit délivré par la pompe auxiliaire (20 $\mu\text{L}/\text{min}$). Les peptides sont retenus et concentrés sur la phase en tête de pré-colonne et les sels sont élués vers la poubelle.

La seconde étape se déroule en sens inverse. La pré-colonne rentre dans le circuit de la LC capillaire (ou de la nano-pompe) qui délivre un solvant constitué de 5% de phase organique suffisant à décrocher les peptides piégés sur la pré-colonne. Ceux-ci sont entraînés vers la colonne analytique qui permettra ensuite leur séparation au fur et à mesure du gradient d'élution.

L'avantage d'un tel système, outre de permettre le dessalage de l'échantillon, est de pouvoir injecter de relativement larges volumes d'échantillon (jusqu'à 16 μL dans notre configuration). Aussi, nous ne sommes pas limités par le volume d'extraction des peptides lorsque la digestion trypsique est effectuée « *in gel* ».

Figure 5 : principe de fonctionnement du montage de pré-concentration.



La mise au point du système de pré-concentration avec une LC capillaire et une trappe ionique a été réalisée en collaboration avec l'équipe du laboratoire d'application de la société Agilent (Waldbronn, A.) et a fait l'objet d'une note d'application présentée à la fin de ce chapitre.

III. COUPLAGE AVEC LA TRAPPE IONIQUE

1. Configuration de la source

La configuration de la source de la trappe ionique dépend du débit de travail :

- En **mode LC capillaire** (colonne de 300 μm diamètre interne), le débit de 4 $\mu\text{L}/\text{min}$ est compatible avec l'utilisation de la source microspray (cf. description de la source microspray, partie expérimentale). La colonne est connectée à la source au moyen d'un capillaire de 50 μm de diamètre interne. Ce montage est relativement simple et robuste.
- En **mode nanoLC** (colonne de 75 μm de diamètre interne), le débit (200 nL/min) impose de travailler avec la source nanospray on-line. Pour éviter les volumes morts, l'aiguille de nébulisation est montée directement à la sortie de la nano-colonne grâce à une connexion assortie d'une férule conductrice qui permet l'écoulement des charges électrostatiques. Nous avons utilisé pour ce montage des aiguilles New Objective en silice étirée couverte d'un revêtement métallisé. Leur diamètre interne est de 50 μm sur la longueur et 8 μm à la pointe. Ce montage est loin d'être trivial et nécessite beaucoup de minutie. L'aiguille est extrêmement fragile, elle souffre des effets du champ électrique (900 eV) et sa durée de vie est très variable. De ce fait, l'approche nanoLC n'est pas très robuste.

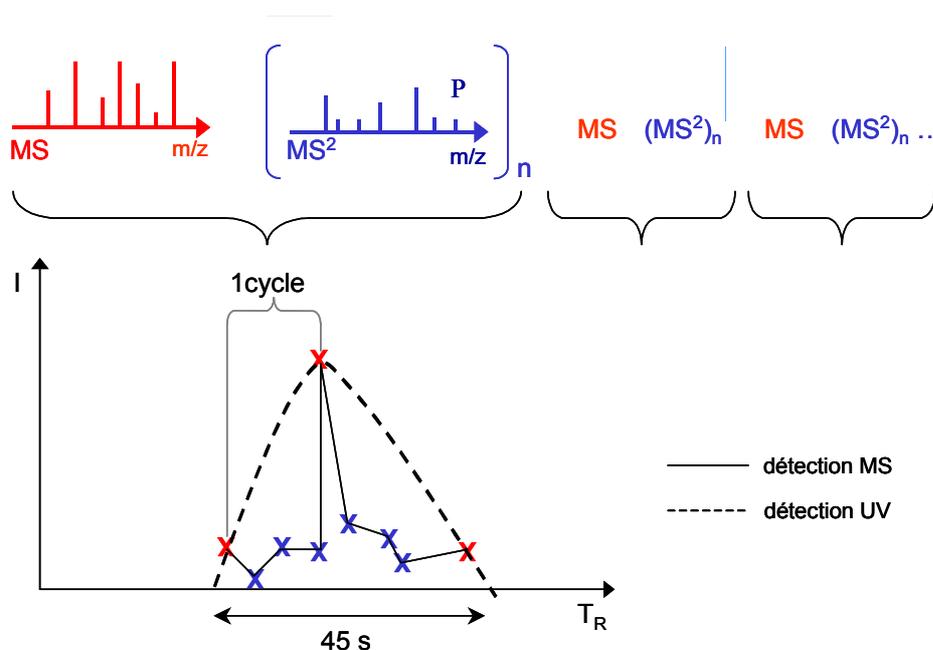
2. Acquisition en mode MS/MS automatique

2.1. Le cycle MS-MS/MS

En mode d'acquisition MS/MS automatique, la sélection, l'isolement et la fragmentation des ions est réalisée de manière automatique selon des critères définis par l'utilisateur qui seront expliqués au prochain paragraphe et qui ont faits l'objet d'une série d'optimisation comme nous le verrons ensuite.

La digitalisation d'un pic chromatographique en mode de détection MS/MS automatique correspond à une série de points qui comprennent successivement une étape de MS, utile à la sélection de n ions précurseurs, et n étapes de MS/MS correspondant à la fragmentation de chacun d'entre eux. Un cycle équivaut à une MS suivie de n MS/MS. Ces cycles se répètent tout au long de la séparation chromatographique (figure 6).

Figure 6 : représentation schématique des cycles d'acquisition en mode MS/MS automatique et allure du chromatogramme correspondant. Le tracé UV (théorique) a été ajouté pour mieux comprendre les effets de la digitalisation en MS/MS automatique.



La durée consacrée à chaque cycle doit, par conséquent, être adaptée à la durée d'éluion des peptides séparés sur la colonne, à savoir environ 45 secondes dans nos conditions d'analyse (largeur à la base d'un pic chromatographique). Idéalement, un cycle toutes les 45 secondes est suffisant pour obtenir des informations de séquence sur la majorité des peptides d'intérêt présents dans le mélange. Cependant, en pratique, il est intéressant d'écourter sa durée (un cycle toutes les 20 à 25 secondes soit deux cycles pour un pic chromatographique) pour les raisons suivantes :

Un cycle de 45 secondes peut se chevaucher sur deux pics chromatographiques. Si, en effet, un spectre MS est acquis alors que les ions précurseurs correspondants sont déjà partiellement élués, ces derniers risquent de ne plus être détectés en quantité suffisante lors des étapes de

MS/MS suivantes et les spectres de fragmentation seront moins informatifs. Par contre, si la durée du cycle est divisée par deux, il est plus probable qu'aucune information ne soit laissée de côté.

Les mélanges peptidiques analysés sont parfois très complexes et il est rare qu'un pic chromatographique corresponde à l'élution d'une seule espèce. Aussi, il n'est pas inutile de multiplier le nombre de cycles afin d'obtenir des informations de séquence sur un plus grand nombre de peptides.

2.2. Sélection du précurseur

Les ions détectés en mode MS doivent remplir plusieurs critères, définis préalablement par l'utilisateur, pour être sélectionnés et fragmentés en mode MS/MS automatique :

Le seuil d'intensité ou « threshold » : l'intensité (absolue ou relative) des ions détectés doit être supérieure à une valeur seuil afin que ceux-ci puissent être sélectionnés comme ions parents. Plus la valeur du threshold est haute, meilleure sera la qualité des spectres MS/MS correspondants. Cependant, au delà d'une certaine limite de threshold, on devient trop sélectif et l'on risque de perdre des informations. A l'inverse, si l'on définit un threshold trop bas, l'intensité des précurseurs sélectionnés est voisine de celle du bruit de fond ce qui ne permet pas d'obtenir d'informations de séquence significatives.

Le nombre n de précurseurs sélectionnés : à chaque cycle, n ions sont choisis parmi les plus intenses dépassant le seuil d'intensité défini précédemment. Plus le nombre d'ions n est grand, plus le cycle sera long. Dans nos conditions d'analyse, nous avons vérifié que la sélection de 3 ions précurseurs par spectre MS permettait d'obtenir le maximum d'informations de séquence dans un temps raisonnable.

La charge de l'ion : les peptides issus de la digestion trypsique sont généralement doublement ou triplement chargés lorsqu'ils sont analysés par spectrométrie de masse en mode d'ionisation ES positif (ils présentent au moins deux sites de protonation, un sur l'amine N-terminale et l'autre sur la lysine ou l'arginine en position C-terminale qui sont des acides aminés basiques). Par ailleurs, les ions monochargés sont assez difficiles à fragmenter, il faut leur conférer beaucoup d'énergie pour les casser. Aussi, ces derniers sont systématiquement écartés et la préférence est donnée aux ions di et tri-chargés pour le choix du précurseur.

L'exclusion temporaire d'un ion de masse donnée : il arrive que le signal correspondant à un peptide soit suffisamment intense pour que celui-ci soit sélectionné à plusieurs reprises au cours de cycles successifs. L'information MS/MS générée est alors redondante. Pour éviter ce problème, il est possible d'exclure temporairement les ions qui ont été sélectionnés comme précurseur. Pendant un délai fixé à une minute suivant la MS/MS d'un ion de m/z donné, le mode de sélection automatique exclu la masse correspondante.

Inclusion et exclusion : il est possible d'inclure ou d'exclure préférentiellement une liste de masse de façon à orienter le choix de l'ion précurseur. De façon générale, la gamme de masse utilisée pour cette sélection s'étend de 250 à 2000 Da. Cependant, il peut être intéressant d'exclure spécifiquement les peptides ne présentant pas d'intérêt pour l'analyse comme, par exemple, ceux correspondant à l'auto-digestion de la trypsine. D'autre part, l'inclusion peut s'avérer utile pour obtenir des informations de séquence sur un peptide ayant un faible facteur de réponse en MS ou présent en faible concentration dans le mélange. Dans ce cas, les masses définies dans la liste d'inclusion sont prioritairement sélectionnées et fragmentées sans considération d'intensité autre que d'être supérieure à la valeur du threshold.

2.3. Isolement et fragmentation

Les précurseurs sélectionnés en mode MS/MS automatique sont tour à tour isolés et fragmentés. La fenêtre d'isolement est fixée à 4 Da, soit +/- 2 Da autour de la masse du précurseur. Cette fenêtre est suffisamment sélective et permet de conserver l'information liée au massif isotopique.

L'énergie nécessaire à la fragmentation d'un peptide dépend de plusieurs critères : sa masse, sa charge et sa composition en acides aminés. Afin de favoriser la fragmentation du plus grand nombre de peptides possibles, nous avons recours, en mode automatique, à une rampe de fragmentation. Cette rampe consiste, pour chaque ion parent, à balayer pendant un court délai (40 ms) toutes les valeurs d'énergie de fragmentation correspondant à une fenêtre définie autour d'une valeur moyenne.

Sur la trappe ionique, l'utilisateur peut moduler l'énergie de fragmentation d'un ion en ajustant l'amplitude du signal de résonance sur les électrodes chapeaux (cf. Introduction - Bibliographie §II.1.4.2.). La valeur moyenne choisie pour la MS/MS automatique a été fixée à 1,2 V et la fenêtre définissant la rampe consiste à balayer les amplitudes comprises entre 30% et 200% de cette valeur.

IV. TRAITEMENT DES DONNEES

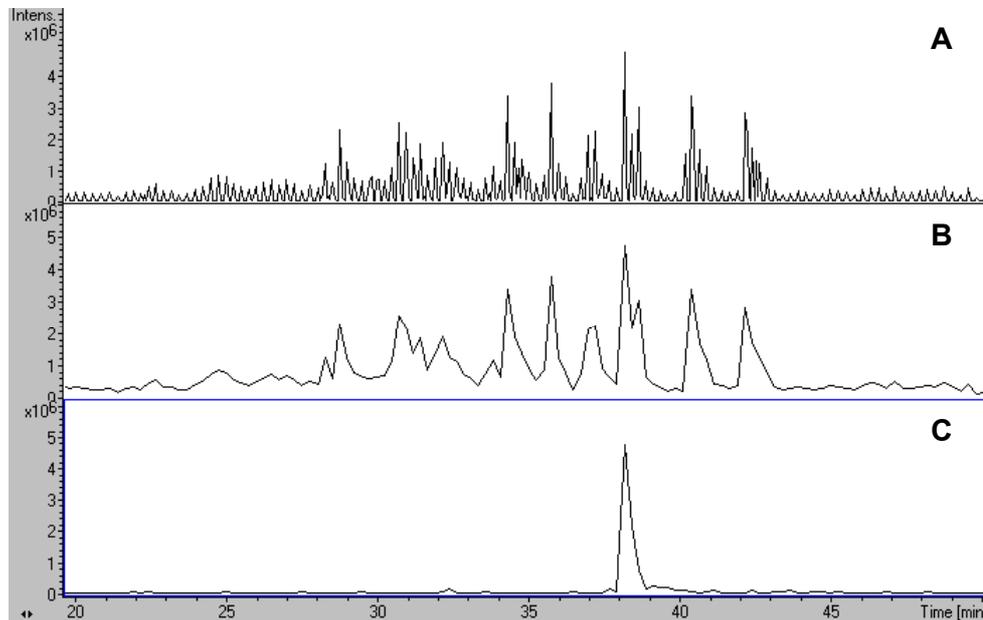
1. Les tracés chromatographiques

Comme nous venons de l'expliquer la digitalisation d'un chromatogramme en mode MS/MS automatique correspond à un enchaînement de cycles [MS + n MS/MS].

Les spectres MS/MS qui résultent d'étapes d'isolement et de fragmentation d'un ion sont beaucoup moins intenses que les spectres MS qui, eux, sont le reflet de tous les ions piégés dans la trappe à un instant donné. Aussi, le chromatogramme obtenu en mode de détection MS/MS automatique est caractéristique par son allure crénelée : aux sommets se trouvent les points correspondant à un spectre MS, dans les creux ceux représentatifs des étapes de spectrométrie de masse en tandem.

Il est cependant possible de visualiser ces chromatogrammes de manières différentes suivant que l'on choisisse de s'intéresser au signal MS seul ou à un courant d'ions extrait (figure 7).

Figure 7 : allure du signal chromatographique obtenu en mode d'acquisition MS/MS automatique après injection de 500 femtomoles d'un mélange tryptique de BSA et de GAPDH. A) Courant d'ions total (MS + MS/MS). B) Courant d'ions (MS seule). C) Courant d'ions extrait (MS) : m/z 582.3 (2+) correspondant au peptide 66-75 de la BSA.



2. Génération de la « peak list »

2.1. Introduction

Une « peak list » (pour « liste de pics ») est un fichier dans lequel sont répertoriées toutes les informations issues de l'analyse protéomique par spectrométrie de masse en tandem qui sont nécessaires à l'interrogation dans les banques de données. La peak list générée à partir d'une analyse LC-MS/MS est une liste de masses (ions parents et fragments) comprenant des informations relatives à leur intensité et leur état de charge (tableau 3).

Le format informatique de la peak-list dépend du logiciel utilisé pour le traitement des données et du moteur de recherche servant à interroger les banques.

Pour notre étude, nous avons utilisé le logiciel Data Analysis® (Bruker) pour interpréter les données de LC-MS/MS et générer les peak-list. Le moteur de recherche MASCOT® a été utilisé pour la recherche dans les banques.

Nous allons voir dans le paragraphe suivants quels sont les paramètres nécessaires à l'obtention d'une peak list reflétant correctement la qualité de l'analyse.

Tableau 3 : exemple de « peak list » générée à partir d'une analyse LC-MS/MS. La taille de cette liste dépend du nombre d'ions précurseurs détectés.

m/z	Intensités relatives	Etats de charge
Parent 1188.5643	1056.562	2
Liste des fragments		2
400.093	0.467	
423.954	1.287	
424.443	0.676	
426.902	38.560	2
609.438	4.712	1
947.021	0.324	
...
Parent 517.239	945.762	3
Liste des fragments		
206.563	6.943	3
298.076	2.453	
298.398	1.323	
600.456	0.076	2
...

2.2. Détection des composés

Pour le traitement des données, le logiciel Data Analysis permet la détection automatique de l'ensemble des « composés » représentatifs de l'analyse LC-MS/MS. Il s'agit des principaux ions précurseurs auxquels sont associées les informations suivantes (figure 8) :

- | ratio m/z
- | temps de rétention
- | spectre MS dont est issu le parent
- | spectre MS/MS associé
- | éventuellement, informations sur l'état de charge des ions en MS et en MS/MS.

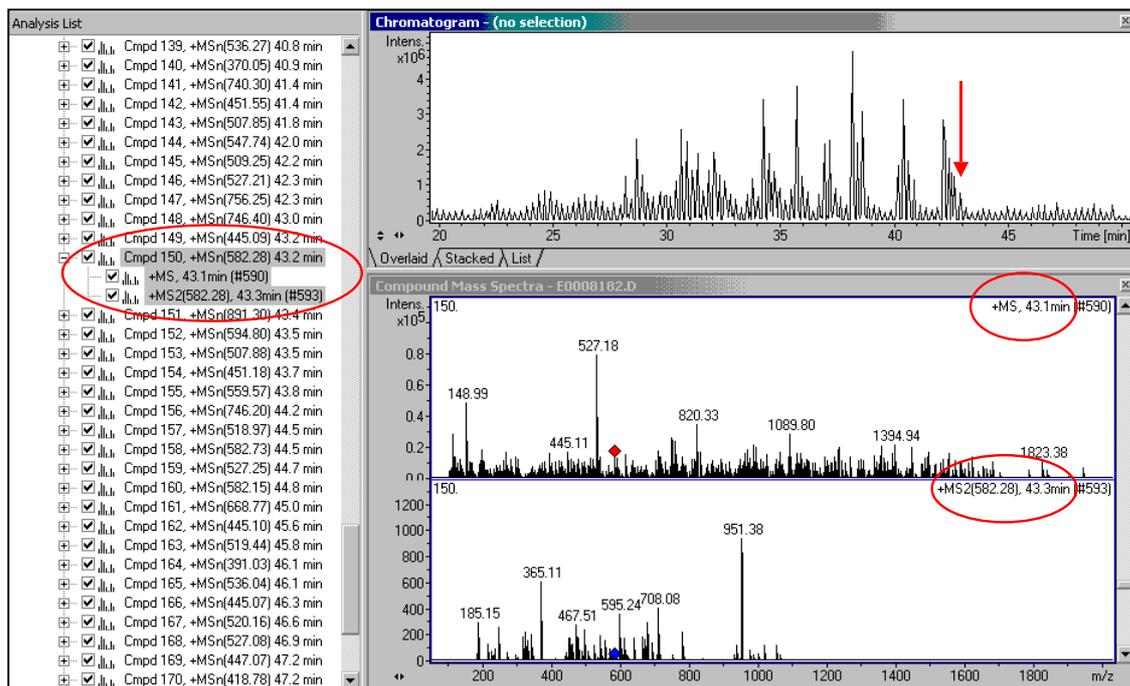
Pour une détection optimale de ces composés il faut ajuster certains paramètres :

- | La détection est limitée par un seuil relatif à l'intensité du signal chromatographique de chacun des composés. Ce seuil est défini par l'utilisateur et sert à « filtrer » la quantité d'informations exportées ensuite dans la peak list (l'export est limité en nombre de composés).
- | Les pics issus de la détection MS et MS/MS doivent être correctement annotés afin de conserver une bonne précision sur la mesure de masse. Nous annotons les pics en mode centroïde à 80% de hauteur.
- | Il est possible d'obtenir des informations sur les états de charge des ions détectés. Ces états de charge sont calculés grâce à un logiciel de déconvolution qui se base, soit sur la mesure de l'écart de masse entre deux pics consécutifs d'un massif isotopique, soit sur la mesure de

masse de plusieurs états de charge d'un même ion. Ce logiciel a préalablement été testé afin d'en optimiser les paramètres pour l'analyse des peptides.

Le délai nécessaire pour traiter les données issues de l'analyse LC-MS/MS dépend de la puissance du système informatique utilisé. Une fois générée, la liste de composés est exportée, sous forme de peak list, vers le moteur de recherche MASCOT.

Figure 8 : exemple de résultats issus de la détection des composés avec le logiciel Data Analysis (Bruker) après une analyse LC-MS/MS. A gauche, la liste des composés classés dans l'ordre croissant de leur temps de rétention. A droite, le chromatogramme de courant d'ions total (MS + MS/MS) avec, en dessous, le spectre MS



ayant permis la sélection du précurseur de m/z (582.3) à 43.1 min et le spectre MS/MS associé à cet ion.

3. Interrogation dans les banques de données

3.1. Paramètres de la recherche

Le moteur de recherche (figure 9) est paramétré selon plusieurs critères parmi lesquels :

- | La base de données à laquelle sont confrontés les résultats. Il peut s'agir de banques de séquences protéiques (ex : SwissProt), de banques de séquences nucléotidiques (ex : EMBL, GenBank) ou bien de banques composites contenant ces deux types d'information (ex : NCBI nr).
- | La taxonomie permettant d'affiner la recherche en fonction de l'organisme étudié.
- | Le type d'enzyme utilisée pour la digestion protéique.

- | L'état de charge des peptides.
- | Les modifications chimiques pouvant affecter la masse des peptides. Par exemple, pour nos analyses, les protéines ont été réduites et alkylées afin de rompre les ponts disulfures entre les cystéines. Au cours de cette étape, les cystéines sont carbamidométhylées ce qui entraîne un incrément de masse de 57.03 Da par cystéine.
- | La tolérance de masse accordée pour la recherche, en MS et en MS/MS. Il s'agit de l'écart de masse que l'on autorise entre la mesure expérimentale de la masse d'un peptide et la valeur théorique issue de la digestion *in silico* d'une protéine (MS) ou bien, entre la mesure expérimentale de la masse d'un fragment et le calcul théorique correspondant (MS/MS). Pour nos résultats, cette tolérance est fixée à 0.2 Da en MS et 0.5 Da en MS/MS.

Figure 9 : paramètres utilisés pour l'interrogation dans les banques de données avec le moteur de recherche

Mascot: MS/MS Ions Search

Your name: Elsa Email: WagnerE@ecpm.u-strasl

Search title: (BSA+GAPDH) 500fmol

Database: Sprout

Taxonomy: All entries

Enzyme: Trypsin Allow up to: 1 missed cleavages

Fixed modifications: AB_ol_d_ICATd0 (C), AB_ol_d_ICATd8 (C), Acetyl (K), Acetyl (N-term), Amide (C-term)

Variable modifications: Biotin (N-term), Carbamidomethyl (C), Carbamyl (K), Carbamyl (N-term), Carboxymethyl (C)

Protein mass: kDa ICAT:

Peptide toL ±: 0.2 Da MS/MS toL ±: 0.5 Da

Peptide charge: 1+, 2+ and 3+ Monoisotopic: Average:

Data file: C:\Documents and Settings\dd\Bureau Parcourir...

Data format: Mascot generic Precursor: m/z

Instrument: ESI-TRAP

Overview: Report top: 20 hits

Start Search ... Reset Form

MASCOT.

3.2. Résultats

Les résultats de la recherche (figure 10) sont présentés sous forme d'une liste de candidats (« hits ») auxquels est attribué un score. Ces scores permettent de classer les protéines proposées pour l'identification selon un indice de probabilité qui tient compte des éléments suivants :

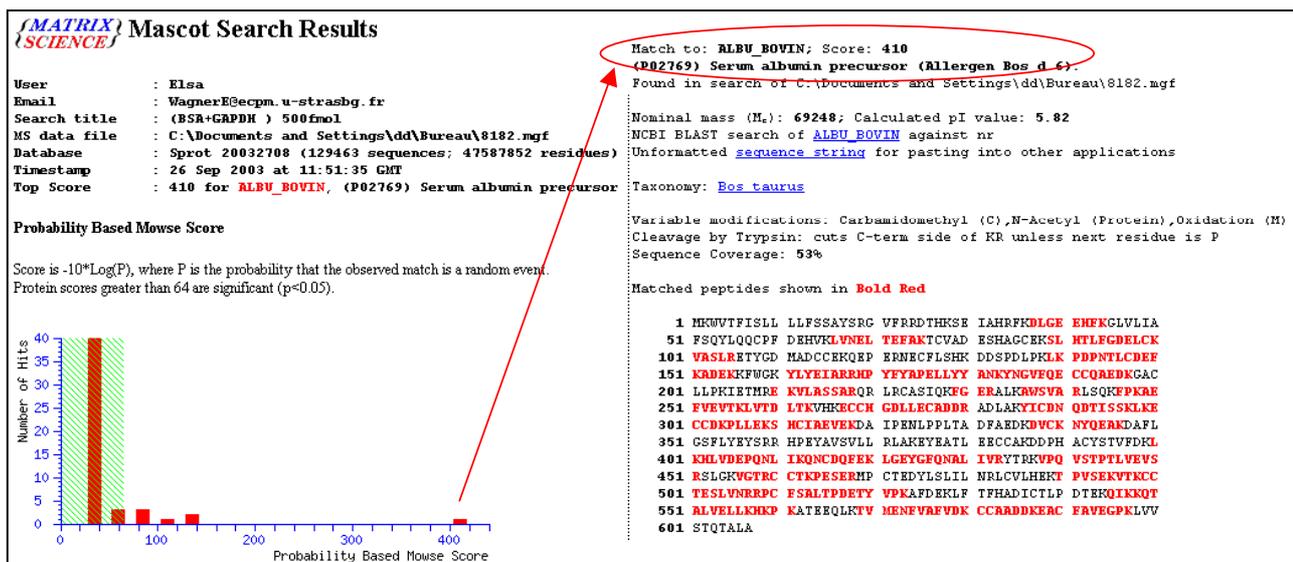
- | le nombre de peptides ayant « matché »
- | le pourcentage de recouvrement de la protéine identifiée
- | la précision de la mesure de masse.

Il est important de vérifier la qualité et la cohérence des résultats obtenus :

- | Les peptides identifiés doivent être distribués largement sur toute la longueur de la protéine et non seulement à une extrémité.
- | Une tolérance trop grande au niveau des écarts de masse enlève toute spécificité à l'identification.
- | Il faut vérifier la cohérence du résultat avec le contexte biologique dont provient l'échantillon.
- | Lorsque la protéine est issue d'un gel 2D, il est utile de vérifier la concordance entre le pI et la masse théorique de la protéine identifiée avec ceux mesurés sur le gel.

Il ne faut cependant pas oublier que les valeurs expérimentales du pI et de la masse moléculaire peuvent être affectées par les conditions d'analyse par gel 2D ainsi que par la présence de modifications post-traductionnelles sur la protéine. De plus, la protéine proposée par les banques peut inclure une séquence précurseur, parfois absente de la protéine étudiée, modifiant ainsi la masse théorique.

Figure 10 : exemple de résultats issus de l'interrogation de la banque SwissProt avec le moteur de recherche MASCOT après analyse par nanoLC-MS/MS de 500 femtomoles d'un mélange tryptique de BSA et de GAPDH.



4. Automatisation du traitement des données

Toutes les étapes décrites précédemment comprenant la détection des composés, leur déconvolution, la génération d'une peak list et l'interrogation dans les banques de données ont été automatisées pour une série d'analyses. Cette automatisation représente un gain de temps considérable dans le cadre de l'analyse protéomique à haut débit.

Elle repose sur l'utilisation d'un programme informatique en langage Visual Basic appelé script. Ce programme permet de regrouper et d'exécuter une série de commandes et peut être modifié selon les besoins de l'analyse.

V. OPTIMISATION DES PARAMETRES PROPRES AU COUPLAGE LC-MS/MS

Cette partie est consacrée au travail réalisé afin d'optimiser les différents paramètres inhérents aux couplages nanoLC avec une trappe ionique.

Les résultats présentés concernent les aspects suivants : optimisation du gradient, tests de sensibilité et de reproductibilité, qualité des spectres MS et MS/MS et sélection des précurseurs. Ces séries d'essais ont été réalisées sur le système nanoHPLC Agilent (HP 1100 Series Nano HPLC System) couplé avec la trappe ionique Bruker (ESQUIRE 3000+).

1. Le gradient de chromatographie

Comme nous l'avons expliqué au paragraphe II.1.3.1., le gradient (GRADIENT LONG, cf. partie expérimentale) proposé a été optimisé sur la base des résultats obtenus en mode de détection UV et MS, avec un mélange peptidique standard, pour permettre la séparation efficace des peptides issus de protéolyse trypsique.

Cependant, malgré la qualité des résultats obtenus, le temps d'analyse dans ces conditions est relativement long : 100 minutes en tenant compte de la période d'équilibration de la colonne. Dans l'idée d'accroître le débit de nos analyses en protéomique, nous avons augmenté la pente du gradient et ainsi testé un gradient plus court, d'une durée totale de 63 minutes (GRADIENT COURT, cf. partie expérimentale).

Les tests présentés ensuite concernant la sensibilité et la reproductibilité de l'analyse ont été réalisés avec ces deux types de gradient afin d'en évaluer la qualité.

2. Tests de sensibilité

Ces tests ont été réalisés à partir d'un mélange équimolaire en solution de BSA (Bovine Serum Albumin, Sigma) et de GAPDH (Glyceraldehyde-3-Phosphate Deshydrogenase, Sigma) digérés à la trypsine. Des injections en quantités croissantes de ce mélange ont permis de tester la sensibilité du couplage nanoLC-MS/MS sur une large gamme : de 5 femtomoles à 1 picomole.

Les résultats obtenus à l'issue de l'interrogation dans les banques de données avec le moteur de recherche MASCOT sont présentés dans les tableaux 4 et 5, ils concernent :

- | le pourcentage de recouvrement de la protéine (% rec.)
- | le nombre de peptides différents fragmentés en MS/MS et contribuant à identifier la protéine (Nb peptides)
- | le score
- | l'erreur moyenne sur la mesure.

N/ID indique que la protéine n'a pas été identifiée dans MASCOT parmi les 20 meilleurs scores.

Tableau 4: résultats des tests de sensibilité en nanoLC-MS/MS après analyse de mélanges de quantités croissantes de mélanges tryptiques de BSA et de GAPDH en solution équimolaire. GRADIENT LONG.

<i>Quantité (fmol)</i>	<i>% rec.</i>	<i>Nb peptides</i>	<i>Score</i>	<i>Erreur (ppm)</i>
BSA				
5	0	0	n/ID	n/ID
10	32	6	217	143
20	29	6	171	181
25	44	11	282	133
50	59	17	525	165
75	52	19	701	196
100	64	23	818	137
200	61	24	858	85
500	65	26	894	88
1000	80	23	915	158
GAPDH				
5	0	0	n/ID	n/ID
10	30	2	54	134
20	6	2	43	85
25	0	0	n/ID	n/ID
50	39	4	93	123
75	0	0	n/ID	n/ID
100	32	5	111	55
200	29	5	186	71
500	35	6	156	125
1000	48	7	155	193

Tableau 5: résultats des tests de sensibilité en nanoLC-MS/MS après analyse de mélanges de quantités croissantes de mélanges tryptiques de BSA et de GAPDH en solution équimolaire. GRADIENT COURT.

<i>Quantité (fmol)</i>	<i>% rec.</i>	<i>Nb peptides</i>	<i>Score</i>	<i>Erreur (ppm)</i>
BSA				
5	0	0	n/ID	n/ID
10	26	3	90	106
20	42	9	250	139
50	56	15	415	160
100	59	17	514	163
200	65	28	1010	128
500	65	21	795	95
1000	76	22	707	92
GAPDH				
5	0	0	n/ID	n/ID
10	0	0	n/ID	n/ID
20	0	0	n/ID	n/ID
50	0	0	n/ID	n/ID
100	9	3	74	225
200	13	5	145	163
500	11	4	101	40
1000	8	2	85	55

Ces expériences montrent une relative linéarité des résultats étudiés (pourcentage de recouvrement, nombre de peptides identifiés et score) en fonction de la quantité de peptides injectée.

Les résultats concernant la GAPDH sont globalement moins bons que ceux obtenus avec la BSA, quelle que soit la quantité injectée. En effet, la qualité des informations obtenues par LC-MS/MS varie selon les caractéristiques des peptides analysés (hydrophobicité, capacité à se protonner, facteur de réponse en MS).

Dans notre cas, la quantité minimale de protéine analysée permettant d'aboutir à son identification est :

- | 10 femtomoles pour les deux mélanges peptidiques avec le gradient long
- | 10 femtomoles pour la BSA et 100 femtomoles pour la GAPDH avec le gradient court.

Les chromatogrammes correspondant à l'injection d'une picomole de mélange tryptique présentent des pics légèrement élargis laissant penser à un phénomène de surcharge, cependant les « blancs » (gradient avec injection de solvant) réalisés à la suite de ces analyses ne présentent pas de traces des composés étudiés.

Si les résultats obtenus pour la BSA sont similaires avec les deux types de gradient ce n'est pas le cas pour la GAPDH dont l'analyse est meilleure avec le gradient long. Par conséquent, il est préférable d'utiliser le gradient court pour l'analyse de mélanges peptidiques peu complexes et en quantité suffisante (supérieure à 100 femtomoles).

3. Tests de reproductibilité

Pour ces essais, nous avons analysé 500 femtomoles d'un mélange tryptique de BSA et de GAPDH en concentration équimolaire selon la séquence suivante : trois injections consécutives suivies d'un « blanc ». Les deux types de gradients, court et long, ont été testés.

Les résultats obtenus sont présentés dans les tableaux 6 et 7.

Tableau 6 : résultats des tests de reproductibilité en nanoLC-MS/MS après analyse de 500 femtomoles de mélange tryptique de BSA et de GAPDH en solution équimolaire. GRADIENT LONG.

<i>Gradient</i>	<i>% rec.</i>	<i>Nb peptides</i>	<i>Score</i>	<i>Erreur (ppm)</i>
BSA				
1	73	25	919	78
2	70	23	930	110
3	70	24	933	112
GAPDH				
1	48	7	156	62
2	45	5	177	100
3	41	6	177	83

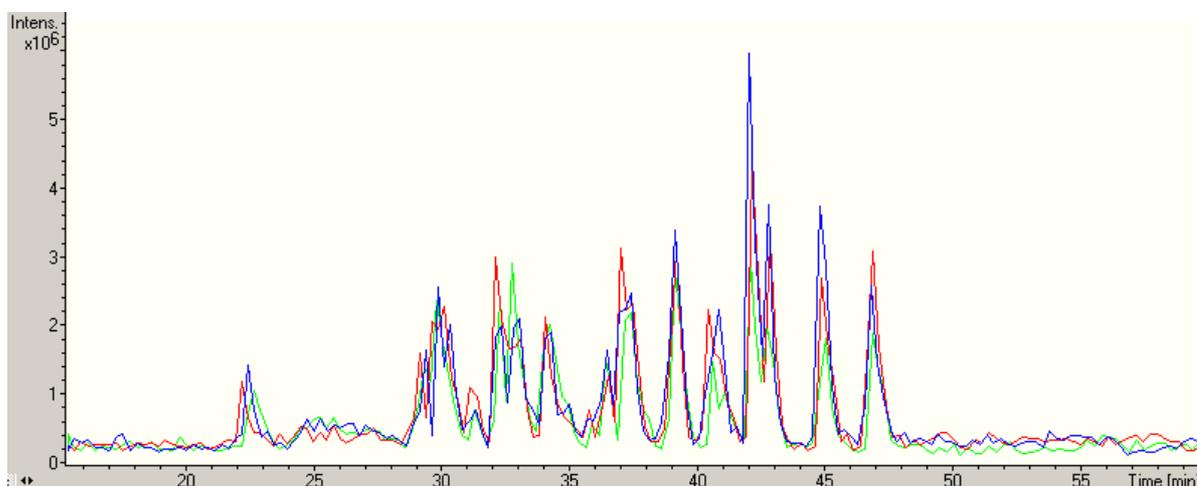
Tableau 7 : résultats des tests de reproductibilité en nanoLC-MS/MS après analyse de 500 femtomoles de mélange tryptique de BSA et de GAPDH en solution équimolaire. GRADIENT COURT.

Gradient	% rec.	Nb peptides	Score	Erreur (ppm)
BSA				
1	72	23	786	118
2	76	26	1050	105
3	65	22	716	96
GAPDH				
1	42	3	83	60
2	36	3	77	90
3	38	3	66	23

- | Aucun phénomène de surcharge (traces de peptides résiduels) n'a été observé dans les blancs ce qui prouve la validité des résultats obtenus.
- | Les mêmes profils d'élution ont été observés pour les 3 analyses (figure 11).
- | Le temps d'élution de chacun des peptides identifiés fluctue de +/- 40 secondes d'une analyse à l'autre pour le gradient long et de +/- 25 secondes pour le gradient court.
- | La largeur des pics chromatographiques (à la base) est de 55 à 80 secondes pour le gradient long et de 25 à 50 secondes pour le gradient court.
- | Le pourcentage de recouvrement est approximativement le même d'une analyse à l'autre.

En conclusion, ces résultats montrent la très bonne reproductibilité du système chromatographique, quelque soit le gradient utilisé.

Figure 11 : superposition des profils chromatographiques correspondant à trois analyses nanoLC-MS/MS de 500 femtomoles de mélange tryptique de BSA et de GAPDH. Courant d'ions MS. Gradient long.



4. Les principaux paramètres d'acquisition avec la trappe ionique

Les paramètres d'acquisition avec l'ESQUIRE 3000+ ont été optimisés pour l'analyse protéomique de peptides tryptiques en mode nanoLC-MS/MS. Nous résumerons ici les principaux paramètres utilisés ; ceux qui concernent spécifiquement l'acquisition MS/MS font l'objet de tests décrits au paragraphe suivant.

Les **tensions appliquées au skimmer et à la sortie du capillaire** (« cap exit ») sont respectivement fixées à 40V et 100V (cf. schéma de la trappe ionique, Introduction – Bibliographie). Parmi les valeurs que nous avons testées, ces paramètres sont ceux qui permettent d'obtenir une bonne transmission des ions tout en limitant les fragmentations, fréquentes dans cette région de l'interface.

La **valeur du cut-off** (également appelé « trap drive », cf. §I.1.4.2. Introduction – Bibliographie) la plus adaptée à l'analyse des peptides tryptiques, sur une gamme de 200 à 2000 m/z, a été fixée à 110V.

Nous avons travaillé en **mode de balayage** « enhanced » (cf. partie expérimentale, les performances de l'ESQUIRE 3000+), pour lequel la résolution à mi-hauteur sur la gamme de masse de travail est de 0,45 m/z. La qualité des spectres ainsi générés permet la reconnaissance de l'état de charge de la majorité des peptides analysés (avec le logiciel de déconvolution). La vitesse de balayage en mode enhanced (5500 m/z / s) est certes plus lente qu'en mode normal (13000 m/z / s), cependant nous avons remarqué que la résolution obtenue en mode normal (0,6 m/z à mi-hauteur) était souvent insuffisante pour la déconvolution des spectres.

La valeur limite correspondant au **temps d'accumulation** des ions dans la trappe a été fixée à 200 ms. Il s'agit du délai maximum pour atteindre la valeur de l'ICC (cf. paragraphe suivant).

5. Tests des paramètres d'acquisition MS/MS : « Averages et ICC »

5.1. Introduction

Deux paramètres, dont la valeur peut être fixée par l'utilisateur, influent considérablement la qualité des spectres MS/MS :

Le nombre d' « averages » représente le nombre de scans nécessaire à l'acquisition d'un spectre MS/MS. Par exemple, si ce nombre est fixé à 8, à chaque étape de fragmentation, l'instrument réalise 8 scans en mode MS/MS sur le même parent (soient 8 étapes successives d'isolement, de fragmentation et d'éjection des ions vers le détecteur). Le spectre affiché à l'écran correspond au calcul du spectre moyen correspondant à ces 8 scans. Ainsi, plus le nombre d' « averages » est grand, plus le spectre résultant est représentatif du signal étudié (le rapport signal sur bruit de chacun des pics est accru). Cependant, la multiplication du nombre de scans rallonge d'autant le temps d'acquisition. Par conséquent, il faut trouver un compromis satisfaisant à la fois nos conditions d'analyse en chromatographie et la qualité de la représentation du signal.

La valeur de l'ICC (Ion Current Control) correspond à la quantité d'ions piégés dans la trappe à chaque scan. Les ions provenant de la source pénètrent dans l'analyseur jusqu'à atteindre un certain quota (valeur de l'ICC). Alors, l'application d'un potentiel à l'entrée du piège empêche le passage des ions jusqu'au scan suivant permettant ainsi les étapes nécessaires à l'analyse (isolement, fragmentation, éjection). Le temps mis pour atteindre la valeur de l'ICC choisie dépend du flux d'ions : si peu d'ions sont produits au niveau de la source, le délai est plus long. Si l'ICC est trop faible, il en résulte un signal appauvri et difficilement exploitable. Si, au contraire, il est trop élevé, on observe des phénomènes d'espace de charges (décrits au paragraphe suivant) qui affectent, entre autres, la précision de la mesure.

Un série d'injections d'un mélange tryptique de BSA en nanoLC-MS/MS nous a permis de faire varier successivement ces deux paramètres afin de déterminer quelles sont les valeurs d' « averages » et d'ICC permettant d'obtenir une qualité d'analyse optimale en MS/MS.

Nous avons analysé 200 femtomoles de mélange ce qui correspond à la quantité statistique de peptides tryptiques extraits d'un spot coloré au bleu de Coomassie en protéomique.

5.2. Résultats

Nous avons testé ces paramètres en faisant varier les valeurs d'ICC et le nombre d' « averages ». Les résultats obtenus sont regroupés dans le tableau 8.

Tableau 8 : analyse nanoLC-MS/MS de 200 femtomoles de mélange tryptique de BSA.

ICC	Nb averages	% rec.	Nb peptides	Score	Erreur (ppm)
2×10^4	8	45	23	776	112
4×10^4	8	47	20	680	169
5×10^4	8	45	17	554	171
1×10^5	8	34	2	115	80
2×10^4	10	51	22	723	106
4×10^4	10	35	14	492	129
5×10^4	10	31	11	451	148
1×10^5	10	0	0	n/ID	n/ID
2×10^4	15	49	22	697	114
4×10^4	15	49	18	749	170
5×10^4	15	44	17	606	59
1×10^5	15	16	2	76	388
2×10^4	20	46	18	613	95
4×10^4	20	40	18	484	93
5×10^4	20	34	13	443	90
1×10^5	20	0	0	n/ID	n/ID

5.3. Analyses des résultats et interprétation

5.3.1. Le nombre d' « averages »

En dessous de 8, le nombre d' « averages » est généralement insuffisant : les spectres MS/MS sont de mauvaise qualité et difficiles à interpréter. Le rapport signal sur bruit des fragments générés est faible et leur massif isotopique peu discernable. Par ailleurs, la réduction du nombre d' « averages » diminue le temps nécessaire au cycle (MS + nMS/MS) ce qui a pour effet d'accroître le nombre de ces cycles par analyse et d'ainsi générer des fichiers très volumineux. Le traitement de ces fichiers, lié à la puissance du système informatique, peut s'en trouver long et difficile.

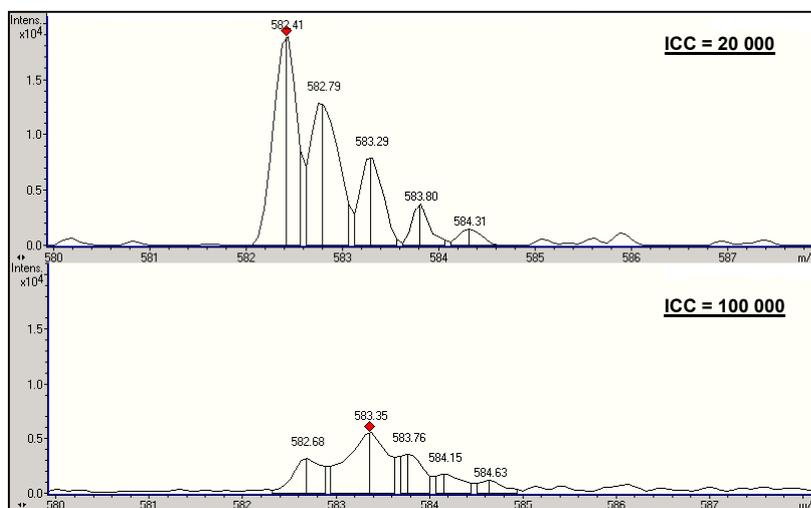
Nos données montrent que l'augmentation du nombre d' « averages » tend à améliorer les résultats obtenus avec MASCOT. Cependant, au delà de 15, une interprétation plus fine des résultats révèle que les cycles deviennent si longs qu'ils dépassent la largeur des pics chromatographiques. Ceci conduit à perdre une partie de l'information incluse sous ces pics (plus particulièrement dans le cas d'espèces présentes en faible abondance dans un mélange peptidique complexe). Les résultats montrent que l'augmentation du nombre d' « averages » de 15 à 20 (pour une valeur d'ICC de 20000) entraîne la perte de 4 peptides d'intérêt parmi les peptides de BSA sélectionnés et fragmentés en MS/MS.

5.3.2. L'ICC

En dessous de 20000, les ions sont piégés en quantité insuffisante dans la trappe pour donner un résultat satisfaisant. Les signaux enregistrés sont de faible intensité et les spectres MS/MS sont assez difficiles à interpréter.

Au dessus de 40000, la sur-accumulation des ions dans la trappe conduit à des phénomènes d'espace de charges affectant la résolution, la sensibilité et la précision de la mesure (cf. figure 12). A une valeur de 100000, ces phénomènes sont si intenses que l'identification via MASCOT est presque impossible, la mesure de masse est parfois entachée d'une erreur de plus de 2 Da.

Figure 12 : observation des phénomènes d'espace de charge sur le peptide 66-75 (2+, m/z théorique 582.3) issu des analyses nanoLC-MS/MS de 200 femtomoles d'un mélange tryptique de BSA.



5.3.3. Conclusion

Finalement, le nombre optimal d' « averages » est compris entre 10 et 15 et la valeur d'ICC qui donne de meilleurs résultats est comprise entre 2×10^4 et 4×10^4 .

Comme nous l'avons vu, ces paramètres influencent beaucoup la qualité des spectres MS/MS et il est nécessaire de les ajuster au mieux afin d'obtenir le maximum d'informations de séquences, surtout pour des mélanges complexes dans lesquels certaines espèces sont très faiblement représentées.

Dans ce contexte, on peut noter que la dernière version de la trappe Bruker (appelée HCT pour « high capacity trap ») présente une capacité de stockage des ions accrue par rapport à l'ESQUIRE 3000+ utilisé pour nos expériences (10 à 20 fois supérieure, d'après le constructeur). Elle est plus sensible et permet également de réaliser des cycles [MS + nMS/MS] plus courts grâce à une vitesse de balayage de la gamme de m/z allant jusqu'à 26 000 m/z /s (mode « ultrascan ») avec une résolution inférieure à 0,6 m/z (largeur des pics à mi-hauteur). Les spectres MS et MS/MS ainsi générés sont donc de meilleure qualité et l'analyse par couplage nanoLC-MS/MS est plus rapide.

6. Choix du seuil pour la sélection du précurseur

6.1. Introduction

La sélection des ions précurseurs destinés à être fragmentés est fonction d'une valeur seuil (ou « threshold », cf. §II.2.2.) définie par l'utilisateur. Le choix de cette valeur conditionne le nombre de MS/MS qui sont réalisées au cours d'une analyse et, par conséquent, la quantité d'informations de séquences qui sont générées.

Nous avons cherché à estimer l'impact de cette valeur sur la qualité de nos résultats pour l'identification dans les banques de données : 3 valeurs de threshold ont été testées (2×10^3 , 5×10^4 et 1×10^5) lors de l'analyse par nanoLC-MS/MS de quantités croissantes de mélanges tryptiques de BSA (50, 100, 200 et 500 femtomoles).

6.2. Résultats

Le tableau 9 présente un récapitulatif des tests consistant à faire varier l'intensité du seuil permettant la sélection des ions précurseurs.

Tableau 9 : analyses nanoLC-MS/MS de quantités croissantes de mélange tryptique de BSA.

<i>Quantités (fmol)</i>	<i>Threshold</i>	<i>% rec.</i>	<i>Nb peptides</i>	<i>Score</i>	<i>Erreur (ppm)</i>
50	2×10^3	0	0	n/ID	n/ID
50	5×10^4	0	0	n/ID	n/ID
50	1×10^5	0	0	n/ID	n/ID
100	2×10^3	18	4	220	81
100	5×10^4	12	2	86	108
100	1×10^5	0	0	n/ID	n/ID
200	2×10^3	25	11	339	119
200	5×10^4	18	8	291	63
200	1×10^5	8	5	205	67
500	2×10^3	40	19	632	149
500	5×10^4	44	23	883	105
500	1×10^5	26	11	458	103

6.3. Analyse et interprétation des résultats

Nos résultats montrent que les valeurs de threshold utilisées ne sont pas adaptées à l'analyse de 50 femtomoles de mélange tryptique de BSA car elles ne permettent la sélection d'aucun des peptides contribuant à l'identification de la protéine.

Pour les quantités allant de 100 à 200 femtomoles, l'augmentation de cette valeur seuil rend l'analyse plus sélective. Plus le seuil est haut, plus le nombre de peptides sélectionnés pour la MS/MS diminue. La qualité des spectres de fragmentation correspondant est cependant bien meilleure.

Cette tendance est un peu différente pour l'analyse de 500 femtomoles de BSA, puisque c'est la valeur de threshold intermédiaire (5×10^4) qui donne les meilleurs résultats. Dans ce cas, la sélectivité générée par la hausse du threshold rend l'analyse plus spécifique : il n'y a pas de pertes de temps engendrées par la fragmentation d'espèces minoritaires non caractéristiques de la BSA.

Finalement, pour l'analyse protéomique, nous utilisons deux types de méthodes :

- | l'une avec un threshold bas privilégiant la quantité de résultats (en terme de nombre de peptides séquencés) et contribuant à renforcer la fiabilité de l'identification protéique
- | l'autre avec un threshold plus haut privilégiant la qualité des MS/MS générées, intéressante pour l'approche par séquençage *de novo* par exemple.

VI. CONCLUSIONS ET PERSPECTIVES

1. Conclusions

Ce travail de mise au point et d'optimisation du couplage LC-MS/MS à faibles débits nous a permis de développer un système performant aujourd'hui incontournable pour l'analyse protéomique. Nos efforts ont porté sur la qualité des données générées sans perdre de vue nos objectifs initiaux : automatiser l'analyse et atteindre des sensibilités inégalées avec les techniques usuelles de séquençage des peptides. Nous avons, en effet, optimisé les paramètres instrumentaux (cf. tableau 10) pour l'analyse de 10 femtomoles à 1 picomole de peptides standards ce qui correspond aux quantités de peptides statistiquement extraits des gels colorés au bleu ou à l'argent (coloration intense).

Aujourd'hui, l'analyse protéomique par spectrométrie de masse ne se contente plus d'identifier les protéines mais également de les caractériser, notamment en terme de modifications post-traductionnelles (souvent liées à l'activité biologique). Il s'agit également de caractériser des protéines dont le génome est inconnu (approche par séquençage *de novo*). Aussi, nous avons mis l'accent sur la qualité des spectres de fragmentation générés par nanoLC-MS/MS et optimisé nos méthodes d'analyse à ces fins.

Tableau 10 : récapitulatif des principaux paramètres optimisés pour l'analyse nanoLC-MS/MS de peptides tryptiques en protéomique (gamme de m/z de 200 à 2000).

skimmer	cap Exit	trap drive	ICC	averages	temps maximum d'accumulation	nombre de précurseurs
40 V	100 V	110 V	2×10^4	10	200 ms	3

2. Evaluation du système

Dans le cadre d'une étude ultérieure, nous avons cherché à évaluer notre système par rapport à un montage équivalent disponible au laboratoire couplant une nanoHPLC (CapLC, Waters) et un Q-TOF (Q-TOFII, Micromass) [Richert, thèse, 2003].

Des tests ont été réalisés en parallèle sur les deux types d'instrument avec un large panel de protéines modèles et d'échantillons biologiques dans les mêmes conditions d'analyse chromatographiques. Les résultats obtenus sur les deux systèmes sont très similaires, en terme de recouvrement de séquence et permettent, dans la large majorité des cas, l'identification des protéines étudiées. Les deux spectromètres de masse, pourtant différents dans leur principe de fonctionnement, génèrent des fragments peptidiques de même nature (y et b).

Lors de cette étude, les potentialités du MALDI-TOF-TOF (ULTRAFLEX, Bruker) pour la fragmentation peptidique ont également été exploitées. Cette approche, automatisable, s'avère intéressante et complémentaire puisqu'elle permet de générer, en plus des fragments y et b, des

fragments de type w (produits à plus haute énergie) permettant, par exemple, de distinguer des acides aminés isobares (comme la Leucine et l'Isoleucine). La technique est cependant moins sensible (environ 1 picomole d'ion parent est nécessaire) et moins résolutive ($R = 2000$ sur les fragments) que l'approche nanoLC-MS/MS décrite précédemment.

3. Limites et perspectives

Les limites du couplage nanoLC-MS/MS, tel que nous l'avons envisagé, résident dans sa capacité à analyser des mélanges peptidiques très complexes (issus par exemple de gels 1D). Ces limites peuvent être dépassées par l'ajout d'étapes de fractionnement supplémentaires grâce à la **chromatographie multi-dimensionnelle**.

Nous avons ajouté une étape de chromatographie supplémentaire et ainsi réalisé quelques essais de 2D-nanoLC-MS/MS en associant une colonne d'échange d'ions avec une colonne de phase inverse (figure 13). Dans cette approche, les peptides sont d'abord élués de la phase échangeuse d'ions par paliers correspondant à des injections de solvant aqueux contenant des quantités croissantes de sels volatils (20 à 500 mM KCl). A chacun de ces paliers est associée une étape de fractionnement sur la phase inverse à l'aide d'un gradient de solvant organique.

Cette approche est certes plus résolutive, cependant elle génère des fichiers très lourds équivalents à n fichiers de 1DnanoLC-MS/MS, où n représente le nombre de paliers d'injection de sels. Par conséquent, au delà de sa difficulté technique, la mise en œuvre de cette approche nécessite d'être équipé de systèmes informatiques très puissants pour le traitement des données.

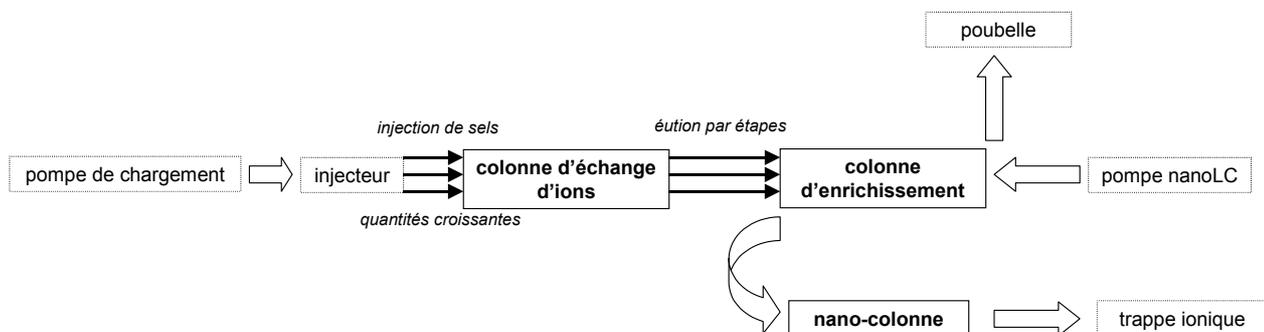


Figure 13 : schéma de l'approche 2D-LC.

L'amélioration des performances du couplage nanoLC-MS/MS passe également par **le choix de la colonne analytique** :

- | Une façon, alternative, d'améliorer le pouvoir résolutif du couplage nanoLC-MS/MS consiste à utiliser des **colonnes de chromatographie plus longues**. En augmentant le nombre de plateaux théoriques (lié à la longueur de la colonne), on améliore la résolution chromatographique ce qui permet d'accroître le nombre de cycles de MS/MS par gradient d'éluion et d'obtenir plus d'informations de séquences.

- | Il serait également intéressant de travailler avec des **colonnes monolithiques** qui possèdent une phase stationnaire continue et homogène (une seule structure contrairement aux phases classiques qui contiennent des particules). Leurs atouts sont d'offrir un plus grand nombre de plateaux théoriques (meilleure séparation chromatographique) et de supporter des pressions plus élevées que les colonnes classiques permettant ainsi une analyse à des débits plus élevés, donc plus rapide.
- | Par ailleurs, nous pourrions tester des **colonnes de plus faible diamètre interne**. Aujourd'hui, sont commercialisées des colonnes de 50 µm pour lesquelles il est préconisé de travailler à un débit de 50 à 200 nL/min (compatible avec notre système nanoLC). Il s'agirait d'évaluer le gain en sensibilité apporté par l'utilisation de telles colonnes et de vérifier leur utilité pour la détection et l'analyse MS/MS d'espèces très faiblement représentées dans les mélanges issus de la digestion trypsique de protéines.

VII. NOTE D'APPLICATION

Une partie de ce travail a fait l'objet d'une note d'application, jointe ci-après, réalisée en collaboration avec la société Agilent et intitulée « Sensitivity Improvement by On-line Sample Enrichment with Capillary LC/Ion Trap Mass Spectrometry ».

SENSITIVITY IMPROVEMENT BY ON-LINE SAMPLE ENRICHMENT WITH CAPILLARY LC/ION TRAP MASS SPECTROMETRY

Application Note

Alain van Dorsselaer, Christine Schaeffer, and Elsa Wagner
Centre National de la Recherche Scientifique, Strasbourg
Christian Sauber and Friedrich Mandel
Agilent Technologies

Introduction

Low flow HPLC techniques, in combination with powerful data-dependent ion trap mass spectrometry, have proven an ideal tool for the identification of proteins. Samples from cell lysates typically contain a large number of different proteins, and one-dimensional or two-dimensional gel electrophoresis is commonly used as a first step in their separation. After staining, the gel spots of interest are removed and enzymatically digested before LC/MSn analysis. The process from staining to digestion can be fully automated by liquid handling systems, resulting in a final sample volume of up to 50 μl . For highest concentration sensitivity, it is important to inject as much volume of each sample as possible, without sacrificing the chromatographic separation. We describe a methodology, including capillary LC/MS/MS with on-line sample enrichment, for the routine analysis of protein digests at the low femtomole per microliter range.

Experimental

All experiments were performed using an Agilent 1100 Series LC/MSD Trap SL ion trap mass spectrometer coupled to an Agilent 1100 Series capillary LC system. The column-switching system used the six-port valve from the 1100 column compartment and an additional 1100 isocratic pump. The complete system was fully controlled by Agilent ChemStation software. Mascot software from Matrix Science Ltd. was used for protein database searching.

Two different samples were prepared. Bovine serum albumin (BSA) was diluted with water to a concentration of 7.5 pmol/ μl and digested with trypsin at 37°C overnight. A protein/trypsin ratio of 50 to 1 was used for the digestion. For the sample measurements, the resulting tryptic digest was subsequently diluted with water to a concentration of 50 fmol/ μl .

Proteins from human mitochondria were separated by 2D PAGE and then stained with ruthenium. This sample was prepared by T. Rabilloud (CEA Grenoble). The spot of interest was excised and in-gel digestion was performed as described by Rabilloud.¹ The sample volume was reduced under vacuum to 3 μl and then 27 μl of 5% formic acid was added for LC/MS/MS analyses.

Samples were loaded onto a 0.5 x 35 mm ZORBAX SB-C18 column at a flow rate of 50 $\mu\text{l}/\text{min}$ with 0.01% aqueous trifluoroacetic acid (TFA) for 5 minutes. The six-port valve of the column compartment was then switched and the trapped peptide fragments were backflushed onto the analytical column with a gradient consisting of 0.01% aqueous TFA and acetonitrile. The analytical column was a ZORBAX 300 SB-C18 0.3 x 150 mm and the flow rate was 4 $\mu\text{l}/\text{min}$.

The ion trap mass spectrometer was operated with the electrospray (ESI) source in positive ion mode. Automated data-dependent MS/MS with active exclusion and maximum resolution scan (MaxResScan) was used to maximize the MS and MS/MS information acquired from a single chromatographic run.

ANALYSIS METHOD

LC/MS/MS

LC Conditions

Enrichment column : Agilent ZORBAX SB-C18,
0.5 x 35 mm, 5 µm
(part number 5064-8254)

Flow rate : 50 µl/min
Injection volume : 20 µl
Mobile phase : 0.01% TFA in water
Column : Agilent ZORBAX 300SB-C18,
0.3 x 150 mm, 5 µm
(part number 5064-8263)

Flow rate : 4 µl/min
Mobile phase : A = 0.01% TFA in water
Mobile phase : B = acetonitrile
Gradient : 5% B for 7 minutes
60%B at 67 minutes

MS Conditions

Ionization mode : positive ESI
Drying gas flow : 6.0 l/min
Nebulizer : 15 psi
Drying gas temperature : 300°C
Skimmer : 40.0 V
Capillary exit : 150 V
Trap drive : 80
ICC : On
Maximum accumulation time : 50 ms
Target : 10000
Scan range : m/z 300 – m/z 2200
Auto-MS/MS : On
Isolation width : 4.0 m/z
Fragmentation amplitude : 1.5 V
No. of parents : 1
Active exclusion : 2 spectra, 0.5 minutes
MaxResScan : On

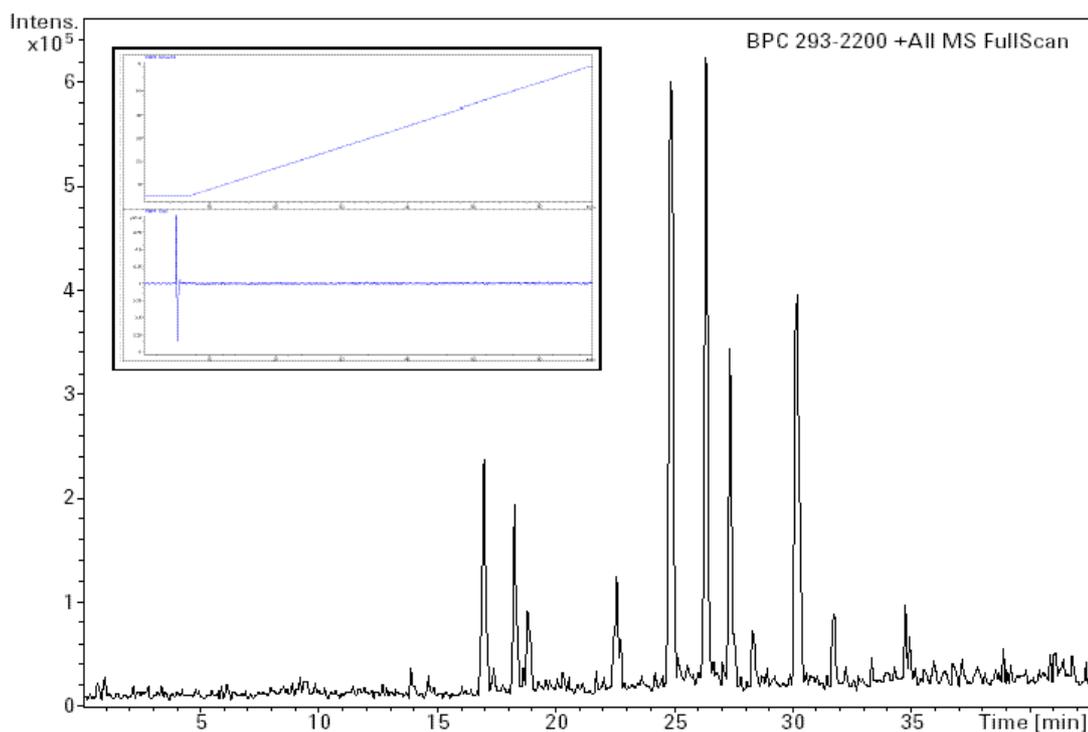


Figure 1. Base peak chromatogram (All MS 300-2200) of 50 fmol/µl BSA digest. Inset shows the gradient profile and the measured flow rate (4 µl/min).

Results and Discussion

The digestion of protein samples typically leads to final volumes of up to 50 μ l, especially when liquid handling robot systems are used. For best LC/MS results, it is necessary to have a methodology that enables injecting a reasonable portion of this sample without compromising the capillary chromatography. A column switching method was developed with the standard 1100 capillary LC system that allows up to 40 μ l of sample to be injected. Moreover, salts and other unwanted polar contaminants are directed to waste during the loading phase and thus do not contaminate the ESI source of the ion trap mass spectrometer.

Figure 1 shows the base peak chromatogram (BPC) of the BSA digest at a concentration of 50 fmol/ μ l with an injection volume of 20 μ l. In the inset, the gradient profile and the measured flow rate are shown. The brief distortion of the flow rate at 5 minutes is a result of switching the six-port valve to bring the enrichment column into the analytical flow path.

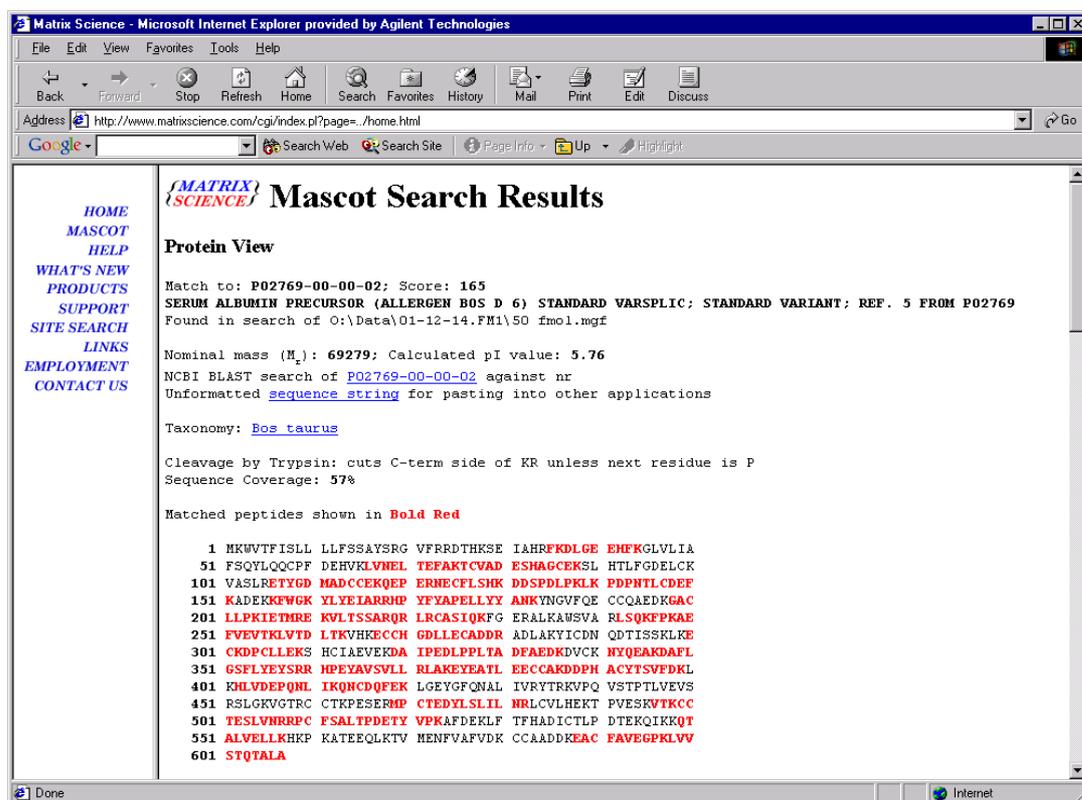


Figure 2. Database search result of 50 fmol/ μ l BSA with Mascot. BSA was found with a MOWSE probability score of 236 and sequence coverage of 94%.

The acquisition and processing of the peptide map was fully automated, including export of the data to Mascot for protein database searching². BSA was found with a MOWSE probability score of 236 and sequence coverage of 94% (Figure 2).

Figure 3 demonstrates the effects of automated, data-dependent acquisition in determining the precursor ions selected for collision-induced dissociation (CID). A preferred charge state of +2 caused the doubly charged ion at m/z 462 to be selected for CID over the more abundant singly charged ion at m/z 922.4. Active exclusion was also applied, increasing the amount of unique data acquired by preventing acquisition of MS/MS data from the same ion more than a user-specified number of times during a specified time interval.

A MaxResScan spectrum is also shown in Figure 3 for this significant tryptic fragment of BSA (sequence: AEFVEVTK). MaxResScan scans a narrow window around the selected m/z value at the highest resolution (0.3 amu FWHM). MaxResScan data is stored for post-acquisition analysis and enables easy determination of charge states up to +4.

The sample containing proteins from human mitochondria was first analyzed by MALDI-TOF mass spectrometry, but the number of peptides present in the spectrum was insufficient for an unambiguous identification by database searching. Reanalysis using the described capillary LC/MS/MS method yielded much better results. The protein, ATP synthase, was correctly identified with a probability score of 117 and sequence coverage of 69% (Figure 4).

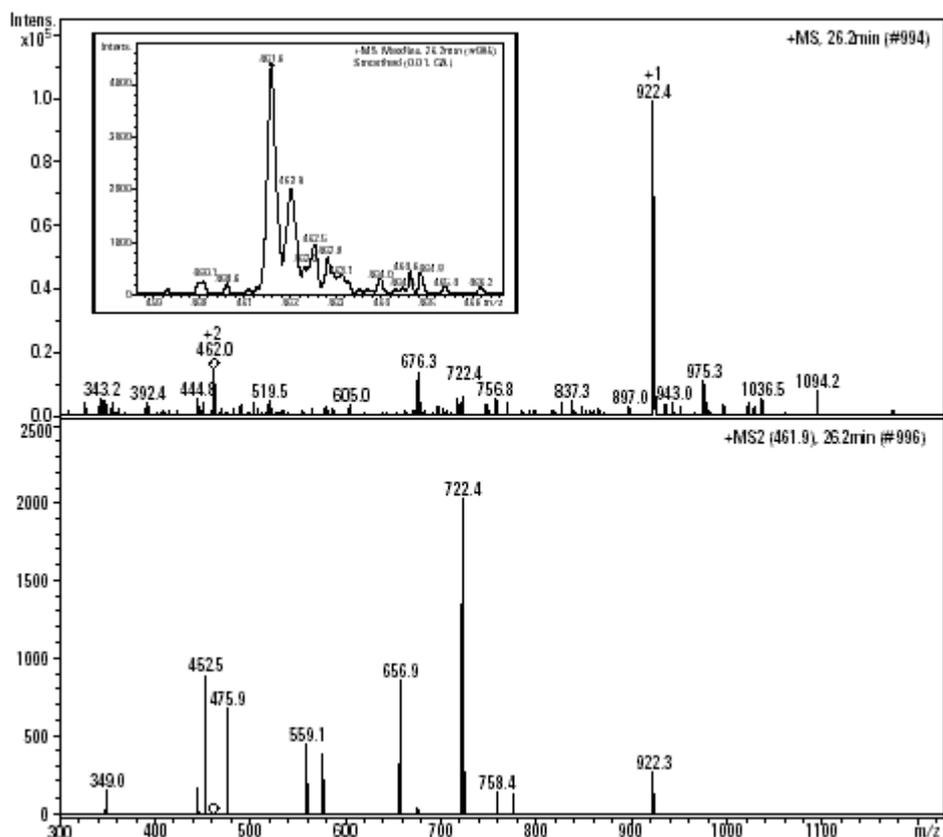


Figure 3. Representative set of MS, MaxResScan, and MS/MS spectra of a significant tryptic peptide fragment of BSA. The sequence of the peptide was AEFVEVTK, matched with an ion score of 17.

Conclusions

Capillary liquid chromatography and ion trap mass spectrometry with protein database searching has proven to be a powerful set of tools for robust and easy-to-use protein identification.

However, for spots from 1D or 2D gels, the concentration of the sample could be too low to produce a positive database result by direct injection. With the fully automated column switching approach described here, the injection volume can be increased without compromising chromatographic performance. This produces significantly higher concentration sensitivity and enables samples in the low- to mid-femtomole per microliter range to be routinely analyzed with excellent results.

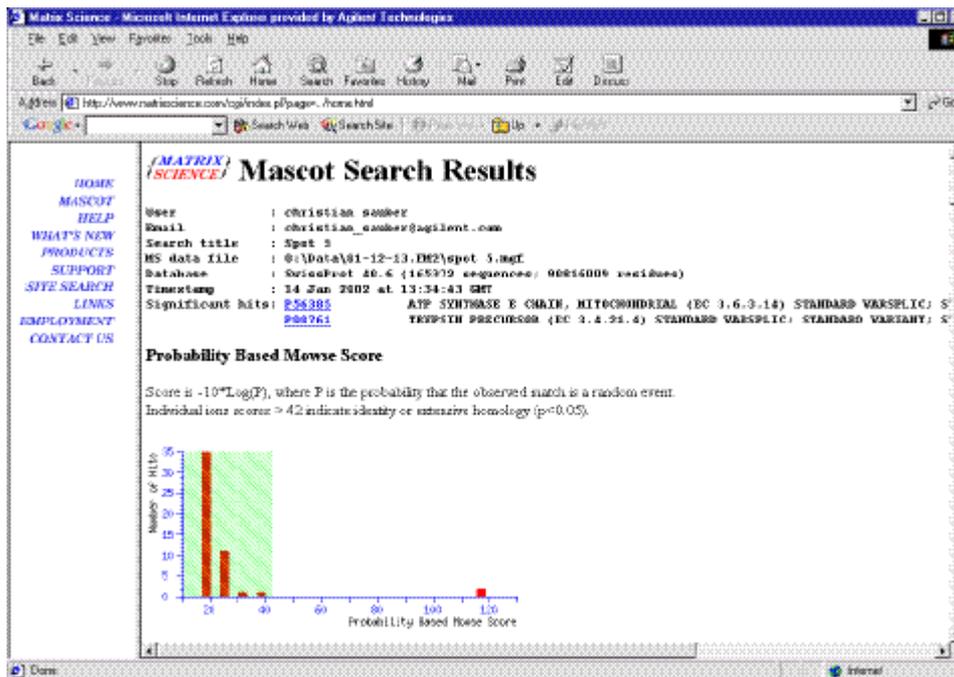


Figure 4. Mascot database search result for 2D gel sample. ATP synthase was found with a probability-based MOWSE score of 117 and sequence coverage of 69%.

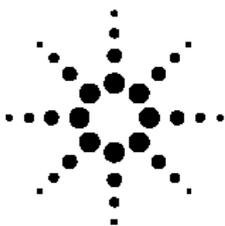
References

1. Rabilloud, T., Strub, J.M., Luche, S., Girardet, J-L., van Dorsselaer, A., and Lunardi, J. "A comparison between Sypro Ruby and Ruthenium II tris (bathophenanthroline disulfonate) as fluorescent stains for protein detection in gels." *Proteomics*, 2001, **1**, 699–704.
2. Perkins, D.N., Pappin, D.J., Creasy, D.M. and Cottrell, J.S. "Probability-based protein identification by searching sequence databases using mass spectrometry data," *Electrophoresis*, 1999, **20**, 3551–3567.

Authors

Dr. Christian Sauber and **Dr. Friedrich Mandel** are applications chemists at Agilent Technologies in Waldbronn, Germany.

Prof. Dr. Alain van Dorsselaer is the scientific director, **Dr. Christine Schaeffer** is a research chemist, and **Elsa Wagner** is a doctoral candidate at the Centre National de la Recherche Scientifique, Strasbourg, France.



Agilent Technologies

www.agilent.com/chem

Copyright © 2002
Agilent Technologies

Information, descriptions and specifications in this publication are subject to change without notice. Agilent Technologies shall not be liable for errors contained herein or for incidental or consequential damages in connection with the furnishing, performance or use of this material.

All rights reserved. Reproduction, adaptation or translation without prior written permission is prohibited, except as allowed under the copyright laws.

Printed in the U.S.A. April 11, 2002
5988-6178EN

BIBLIOGRAPHIE

Griffin P.R., Coffman J.A., Hood L.E., Yates J.R.

Structural analysis of proteins by capillary HPLC electrospray tandem mass spectrometry, *J. Mass Spectrom. Ion Processes*, **1991**, 111, 131-149.

K.L. Hsi, M.L. Kochersperger, W.E. Werner, C.H. Grimley, L.R. Zieske, P.M. Yuan

Development of a capillary LC blotting systems for low and sub-picomole sequencing sample preparation, *Protein and Peptide Letters*, **1997**, 4, 1.

Richert S.

De l'identification à la caractérisation des protéines : développement de la spectrométrie de masse dans le cadre de l'approche protéomique, *Thèse de l'université Louis Pasteur de Strasbourg*, **2003**.

Shevchenko A., Wilm M., Vorm O., Mann M.

Mass spectrometric sequencing of proteins from silver-stained polyacrylamide gels, *Anal. Chem.*, **1996**, 68, 850-858.

Vissers J.P.C., Claessens H.A., Cramers C.A.

Microcolumn liquid chromatography : instrumentation, detection and applications, *J. Chromatography A*, **1997**, 779, 1-28.

Wilm M., Mann M.

Analytical properties of the nanoelectrospray ion source, *Anal. Chem.*, **1996**, 68, 1-8.

Yates J.R., McCormack A.L., Link A.J., Schieltz D., Eng J., Hays L.

Future prospects for the analysis of complex biological systems using micro-column liquid chromatography-electrospray tandem mass spectrometry, *Analyst*, **1996**, 121, 65R-76R.

PARTIE III :

CARACTERISATION PAR APPROCHE PROTEOMIQUE DE PROTEINES IMPLIQUEES DANS LE SYSTEME DE DEFENSE DES CELLULES CONTRE LE STRESS OXYDANT

CHAPITRE I : MISE EN ŒUVRE D'UNE NOUVELLE STRATEGIE PROTEOMIQUE
POUR LA CARACTERISATION DU SITE D'OXYDATION DE
PEROXYREDOXINES LORS D'UN STRESS OXYDANT

CHAPITRE II : ÉTUDE DE LA REGENERATION DES PEROXYREDOXINES SUITE A UN
STRESS OXYDANT : MARQUAGE *IN VIVO* ET APPROCHE
PROTEOMIQUE QUANTITATIVE

CARACTERISATION PAR APPROCHE PROTEOMIQUE DE PROTEINES IMPLIQUEES DANS LE SYSTEME DE DEFENSE DES CELLULES CONTRE LE STRESS OXYDANT

Ce travail a été réalisé en collaboration avec l'équipe du Dr. T. Rabilloud (Laboratoire de Bioénergétique Cellulaire et Pathologique, CEA Grenoble).

I. INTRODUCTION : LES PEROXYREDOXINES

Les cellules vivant en conditions aérobies sont confrontées à la formation d'espèces réactives de l'oxygène qui sont issues de la réduction incomplète de l'oxygène pendant la respiration cellulaire ou bien produites lors d'une exposition à des radiations ionisantes, à la lumière, aux métaux ou à certaines substances [Sies, 1993 ; Fridovich et Freeman, 1986]. Ces espèces, également appelées ROI (pour redox-active oxygen intermediates), sont soit un radical superoxyde ($O_2^{\bullet-}$), soit du peroxyde d'hydrogène (H_2O_2) ou bien un radical hydroxyle (OH^{\bullet}). Les ROI réagissent avec divers constituants de la cellule, comme l'ADN, les protéines ou les lipides, causant des dommages irréversibles qui conduisent à la mort cellulaire. Parmi les moyens dont elle dispose, la cellule peut se protéger de ces agressions en détruisant les ROI grâce à des peroxydases un peu particulières appelées peroxyrédoxines (prxs).

Les prxs contiennent une cystéine activée au niveau de leur site actif [Choi et coll., 1998] et certaines possèdent une autre cystéine conservée [Hirotsu et coll., 1999 ; Declercq et coll., 2001].

Lors du cycle catalytique normal, les prxs réduisent les peroxydes toxiques. La cystéine contenue au niveau du site actif est alors oxydée en acide cystéine sulfénique [Choi et coll., 1998] ou en cystine [Chae et coll., 1994]. Les prxs ainsi inactivées sont ensuite re-réduites en forme thiol active grâce, notamment, à un système appelé thiorédoxine réductase [Chae et coll., 1994].

Il semblerait que les prxs soient également impliquées dans des situations de stress oxydant liées à des processus physiopathologiques comme, par exemple, la différenciation cellulaire, la transformation cancéreuse ou l'artériosclérose [Phelan et coll., 1998 ; Prospero et coll., 1993 ; Rabilloud et coll., 1995]. Ces enzymes apparaissent alors sous deux formes protéiques distinctes (acide et

basique) dont la proportion varie en fonction du degré d'oxydo-réduction de la cellule, et notamment, en faveur de la forme la plus acide dans des conditions oxydantes.

Notre travail a débuté à partir de ces observations et s'est articulé en deux étapes :

- | Premièrement, nous nous sommes intéressés aux différentes prxs d'un modèle cellulaire et avons entrepris de démontrer que les deux formes (basique et acide), dont la proportion varie lors d'un stress oxydant, correspondaient respectivement à la forme normale et à la forme oxydée de la protéine étudiée. Pour vérifier cette hypothèse, nous avons développé une nouvelle stratégie basée sur l'approche protéomique consistant à caractériser le site d'oxydation des prxs.
- | Nous avons ensuite étudié les mécanismes de régénération des prxs suite au stress oxydant. Pour cela, nous avons utilisé un marquage *in vivo* des protéines suivi d'une approche protéomique quantitative.

II. MISE EN ŒUVRE D'UNE NOUVELLE STRATEGIE PROTEOMIQUE POUR LA CARACTERISATION DU SITE D'OXYDATION DE PEROXYREDOXINES LORS D'UN STRESS OXYDANT

Ce travail a fait l'objet d'une publication présentée à la fin de ce chapitre. Dans les prochains paragraphes, nous rappellerons brièvement les principales étapes de cette étude et nous détaillerons quelques résultats complémentaires.

1. Observations issues des gels 2D

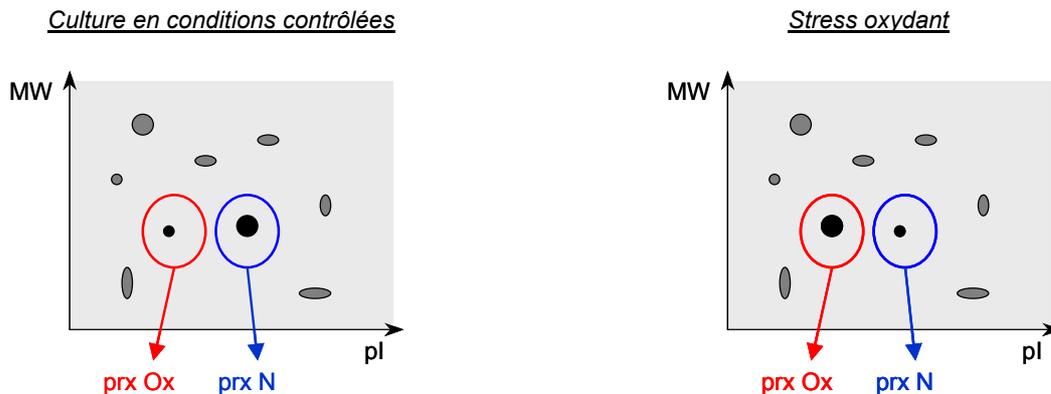
Nous avons travaillé à partir d'un modèle de cellules appelées HeLa issues d'une lignée cellulaire épithélioïde (cancer utérin humain). Ce modèle permet l'étude d'un panel de prxs de localisations subcellulaires variées : prx I, prx II, 1cys prx (cytosoliques) ; prx III (mitochondriale) et prx IV (compartiments sécrétoires).

Le point de départ de ce travail concerne l'observation des gels 2D correspondant aux extraits de cellules HeLa cultivées dans des conditions contrôlées ou bien soumises à un stress oxydant (75 – 150 μ M t-butyl hydroperoxyde, 1-6 h). Les prxs apparaissent chacune sous forme d'un doublet de «spots» sensibles au stress oxydant (figure 1) :

- | le spot le plus basique est toujours le plus intense dans des conditions de culture contrôlées
- | le spot le plus acide est majoritaire dans des conditions de stress oxydant.

Les analyses MALDI-MS confirment que les deux spots du doublet correspondent à la même protéine ce qui prouve que la différence entre ces deux spots est due à une modification post-traductionnelle. Le calcul théorique des pI des différentes prxs analysées et leur corrélation avec ceux observés sur le gel 2D suggèrent que le spot le plus basique du doublet correspond à la protéine native (prx N) et le spot le plus acide à une forme modifiée, vraisemblablement oxydée (prx Ox).

Figure 1 : représentation schématique de l'analyse par gel 2D d'extraits de cellules Hela cultivées dans des conditions contrôlées ou ayant été soumises à un stress oxydant. Les deux spots entourés correspondent à un doublet de peroxyrédoxines (selon notre hypothèse : prx Ox = forme oxydée ; prx N = forme native).

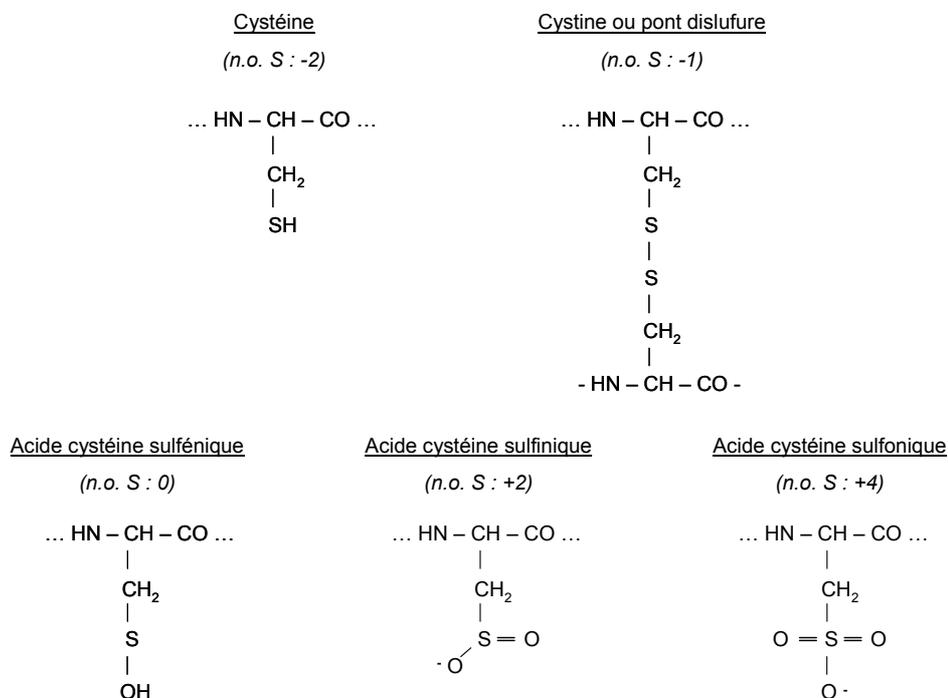


A ce stade de l'étude, les hypothèses émises par l'équipe de biologistes avec laquelle nous avons collaboré, est la suivante : la forme modifiée correspondrait à l'oxydation irréversible de la cystéine du site actif en acide cystéine sulfinique ou sulfonique. C'est ce que nous proposons de vérifier par spectrométrie de masse.

2. Les différents états d'oxydation de la cystéine

Les différents états d'oxydation de la cystéine sont représentés figure 2.

Figure 2 : la cystéine et ses différentes formes oxydées. Le degré d'oxydation du soufre est précisé (n.o. S).



Dans ce manuscrit, les formes acide cystéines sulfiniques et sulfoniques, pour lesquelles le degré d'oxydation du soufre est supérieure à 0, sont qualifiées de « sur-oxydées ».

La présence de ponts disulfures sur une protéine est gênante pour l'analyse protéomique car elle induit un repliement de la molécule. Les sites de clivage deviennent moins accessibles à l'enzyme de digestion. Le rendement d'extraction du gel des peptides liés par pont disulfure est faible et leur identification est plus compliquée.

Aussi, nous avons procédé, avant la digestion *in gel*, à un traitement au DTT (dithiothreitol) qui permet de réduire ces ponts puis nous avons ensuite alkylé les cystéines avec de l'iodoacétamide pour empêcher les ponts disulfures de se reformer (cf. partie expérimentale, l'analyse protéomique).

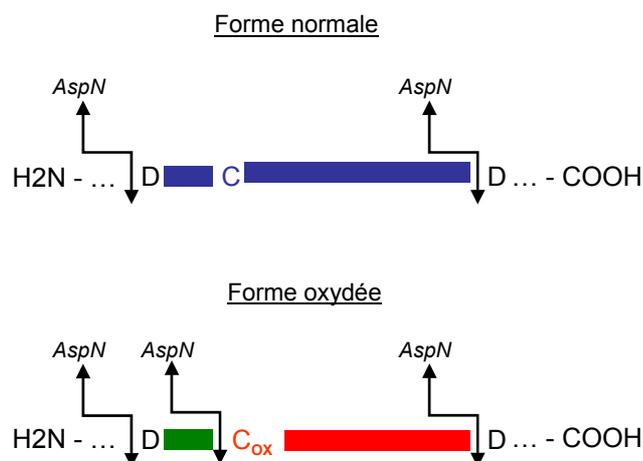
Le DTT est également capable de réduire la forme acide cystéine sulfénique. Cependant, il est important de souligner pour la suite de ce travail que cet agent réducteur est sans effet sur les formes acides sulfonique et sulfonique dont les degrés d'oxydation sont plus élevés.

3. Stratégie développée : choix de la protéase

Il s'agit de montrer que la forme modifiée observée sur les gels 2D (spot acide du doublet de prxs) correspond à l'oxydation de la cystéine du site actif.

Afin de vérifier cette hypothèse, nous avons digéré les deux isoformes de chaque prx avec de l'AspN dont la spécificité est de cliver du côté N-terminal d'un acide aminé à courte chaîne latérale acide, à savoir l'acide aspartique mais aussi les acides cystéines sulfonique et sulfonique [Watabe et coll., 1994 ; Drapeau, 1980]. Ainsi, si notre hypothèse est correcte, les profils de digestion obtenus avec les deux spots doivent être différents (figure 3).

Figure 3 : schéma illustrant la stratégie de digestion à l'AspN. En bleu, le peptide généré par la protéase à partir de la forme normale. Il correspond à la séquence partielle de la prx contenant la cystéine du site actif (alkylée par de l'iodoacétamide). Chez la forme oxydée, la présence de la forme oxydée génère un nouveau site de clivage pour l'AspN avec un court peptide du côté N-terminal (en vert) et le peptide complémentaire du côté C-terminal (en rouge) contenant la cystéine oxydée.



4. Résultats de spectrométrie de masse

4.1. Stratégie de caractérisation des peroxyrédoxines par MS

Les peptides générés par la digestion à l'AspN des deux isoformes protéiques de chaque prx ont été analysés par spectrométrie de masse. Les peptides attendus, caractéristiques de chacune des deux formes, sont décrits dans le tableau 1.

Tableau 1 : peptides attendus après analyse par spectrométrie de masse des peptides issus de la digestion à l'AspN des deux formes de prx. Les couleurs attribuées aux différents peptides font référence à la figure 3.

	Forme Normale	Forme Oxydée
Peptide total	oui	non
Peptide N-terminal	non	oui
Peptide complémentaire	non	oui

4.2. Résultats obtenus en mode d'ionisation positif

Après analyse du mélange peptidique par MALDI-MS en mode d'ionisation positif, nous avons pu vérifier la présence du peptide total dans la forme normale et non dans la forme oxydée.

La détection du court peptide N-terminal (M = 578 ou 628 Da selon le type de prx) en MALDI-MS est rendue difficile par la présence des signaux liés à la matrice dans cette région de la gamme de m/z. Nous avons, par conséquent, réalisé des analyses complémentaires en nanoES-MS/MS sur la trappe ionique et le Q-TOF. Ce peptide a été détecté en MS seulement dans la forme modifiée. Les fragments obtenus en MS/MS à partir de ce peptide nous ont permis de vérifier sa séquence.

Ces résultats, détaillés dans la publication ci-après, nous ont permis de confirmer notre hypothèse : la modification observée sur les gels est bien due à l'oxydation de la cystéine du site actif en acide cystéine sulfinique ou sulfonique.

4.3. Nécessité de réaliser des analyses MALDI-MS en mode d'ionisation négatif

La détection du peptide complémentaire (contenant la cystéine oxydée) n'a pas été possible lors des expériences décrites précédemment. En effet, l'oxydation de la cystéine induit la présence d'une forte charge négative sur le peptide modifié ce qui diminue considérablement son efficacité d'ionisation en mode positif (pK_a des acides cystéines sulfiniques et sulfoniques < 2).

C'est pourquoi nous avons analysé les digests peptidiques en MALDI-MS mode d'ionisation négatif. Les résultats obtenus nous ont permis, à nouveau, de confirmer l'oxydation de la cystéine contenue au niveau du site actif mais également, de déterminer son état d'oxydation.

4.3.1. Choix de la matrice MALDI pour le mode d'ionisation négatif

a. Introduction

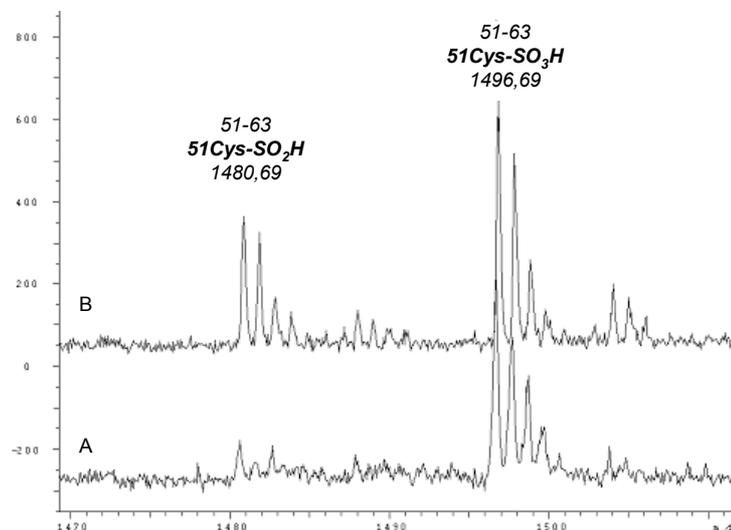
Pour ces analyses, nous avons utilisé une matrice constituée d'un mélange de 2,6-dihydroxyacétophénone (DHAP) et de citrate de 2-hydrogéoammonium (DAHC) dilué dans une phase organique [Gorman J.J. et coll., 1996]. Nous avons voulu tester l'efficacité de cette préparation en mode d'ionisation négatif par rapport à une matrice plus acide utilisée classiquement, l'acide α -cyano-4-hydroxycinnamique ou CHCA.

b. Résultats

Les résultats obtenus pour prx II en mode négatif avec les matrices DHAP+DAHC et CHCA sont présentés figure 4.

- | Les spectres de masse obtenus avec les deux types de matrice nous permettent d'observer le peptide complémentaire (non détecté en mode positif) contenant la forme acide cystéine sulfonique.
- | La matrice DHAP+DAHC nous permet de détecter une seconde forme, absente du spectre de la matrice CHCA, correspondant au peptide complémentaire contenant la forme acide cystéine sulfonique.

Figure 4 : spectres MALDI-MS en mode de détection négatif de la forme modifiée de prx II digérée à l'AspN. A :

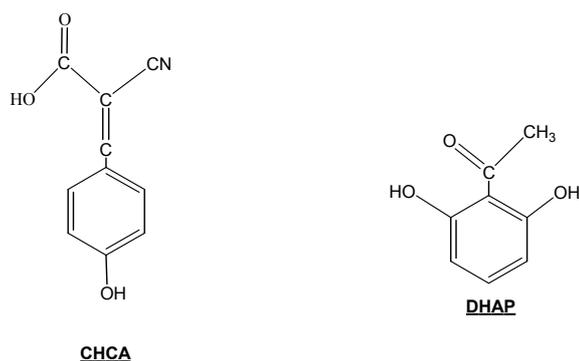


matrice CHCA; B : matrice DHAP + DAHC.

c. Discussion

Afin d'interpréter les différences observées sur ces deux spectres, nous nous sommes intéressés à l'acidité des matrices utilisées (figure 5) ainsi qu'à celle des acides cystéine sulfonique et sulfonique.

Figure 5 : structures chimiques des matrices CHCA (acide α -cyano-4-hydroxycinnamique) et DHAP (dihydroxyacétophénone).



- | L'acidité de la matrice DHAP est liée aux hydrogènes des fonctions phénols. Le pK_a en solution de cette structure chimique est compris entre 8 et 10.
- | La matrice CHCA a un caractère acide plus fort, lié à la fonction acide carboxylique, et renforcé par la présence du groupement cyano (effet inductif attracteur) et par l'effet mésomère de la double liaison conjuguée et du cycle aromatique.
- | Les acides cystéines sulfiniques et sulfoniques ont respectivement les valeurs de pK_a 1,5 et -1,5.

NB : Les valeurs de pK_a citées sont issues du site <http://sakura.cpe.fr/pka.html>.

Par conséquent, nous proposons un classement des structures issues de l'oxydation de la cystéine par ordre d'acidité relative et en fonction des matrices utilisées (figure 6).

Figure 6 : classement des acides cystéine sulfinique et sulfonique ainsi que des matrices utilisées sur une échelle



d'acidité relative.

Ce classement pourrait expliquer les différences observées sur nos spectres :

- | L'acide cystéine sulfonique étant un acide plus fort que CHCA et DHAP, il est majoritairement présent dans ces préparations sous forme sulfite ($Cys-SO_3^-$) détectable, dans les deux cas, en mode d'ionisation négatif.
- | Si l'acidité de l'acide cystéine sulfinique est intermédiaire entre celle de CHCA et DHAP, il est sous forme protonnée ($Cys-SO_2H$) non détectable avec CHCA et sous forme déprotonnée ($Cys-SO_2^-$) détectable avec DHAP.

Cependant, ces hypothèses qui cherchent à rationaliser les résultats obtenus doivent être considérées avec prudence. En effet, les pK_a reflètent l'acidité d'un produit en solution. Pour être plus précis, il faudrait considérer les constantes d'acidité de nos différents produits en phase gazeuse (pas toutes disponibles dans la littérature). Il faudrait également tenir compte de la présence de la co-matrice

DAHC pour la préparation du DHAP. Par ailleurs, lors du processus d'ablation laser, les phénomènes d'agrégation propre à chacune des matrices sont susceptibles d'affecter l'ionisation dans le plasma MALDI [Fournier et coll., 2003].

4.3.2. Compléments d'analyse et nouveaux résultats

Les résultats de la publication présentée à la fin de ce chapitre correspondent aux analyses MALDI-MS réalisées sur un instrument de type BIFLEX III (Bruker). Nous avons par la suite fait l'acquisition au laboratoire d'un nouveau spectromètre de masse MALDI, l'ULTRAFLEX (Bruker), qui est plus sensible que le précédent. Aussi, afin de compléter nos résultats, nous avons refait l'analyse en mode négatif de certaines prxs : prx I, prx II, prx III et 1cys prx. Les spectres obtenus font l'objet des figures 7 et 8.

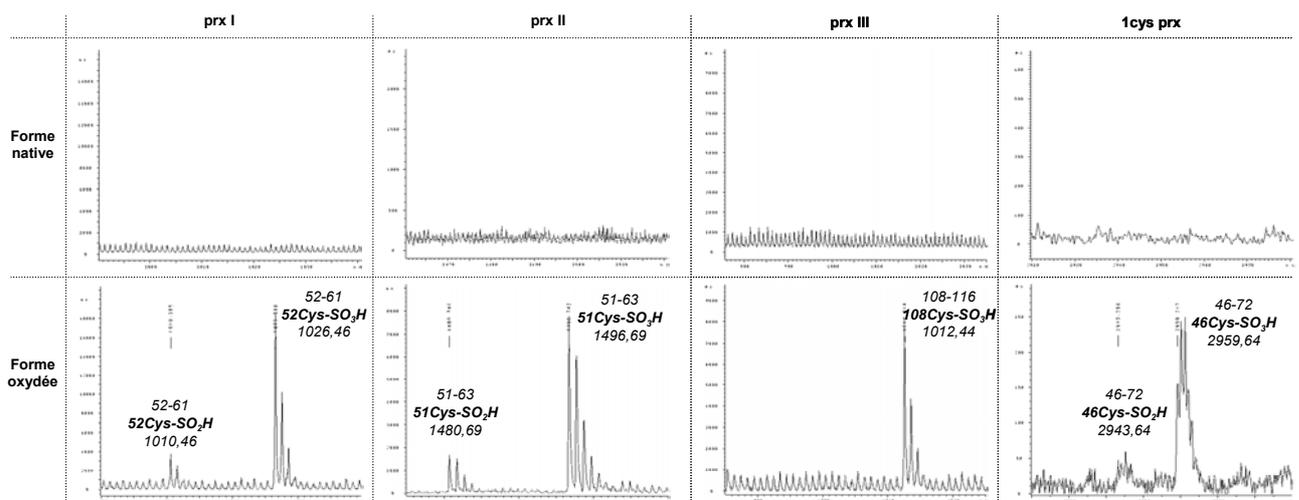
a. Détection du peptide « complémentaire »

Comme précédemment, nous avons recherché dans la forme modifiée des prx la présence du peptide contenant la cystéine oxydée et nous avons ainsi pu déterminer l'état d'oxydation de cette cystéine. Nous avons également vérifié que ce peptide était absent des spectres correspondant à la forme native (figure 7).

Nous pouvons ainsi compléter nos résultats par rapport à ceux rapportés dans la publication :

- | Nous détectons, pour la forme modifiée de prx I, deux états d'oxydation de la cystéine : acide sulfonique et sulfonique.
- | De même pour 1 cys prx, pour laquelle les premières expériences n'avaient pas permis la détection du « peptide complémentaire ».

Figure 7 : spectres MALDI-MS en mode négatif obtenus après analyse des formes modifiée et native de prx I, prx II, prx III et 1cys prx digérées à l'AspN. Recherche du peptide complémentaire contenant la cystéine oxydée.



b. Détection du peptide « total »

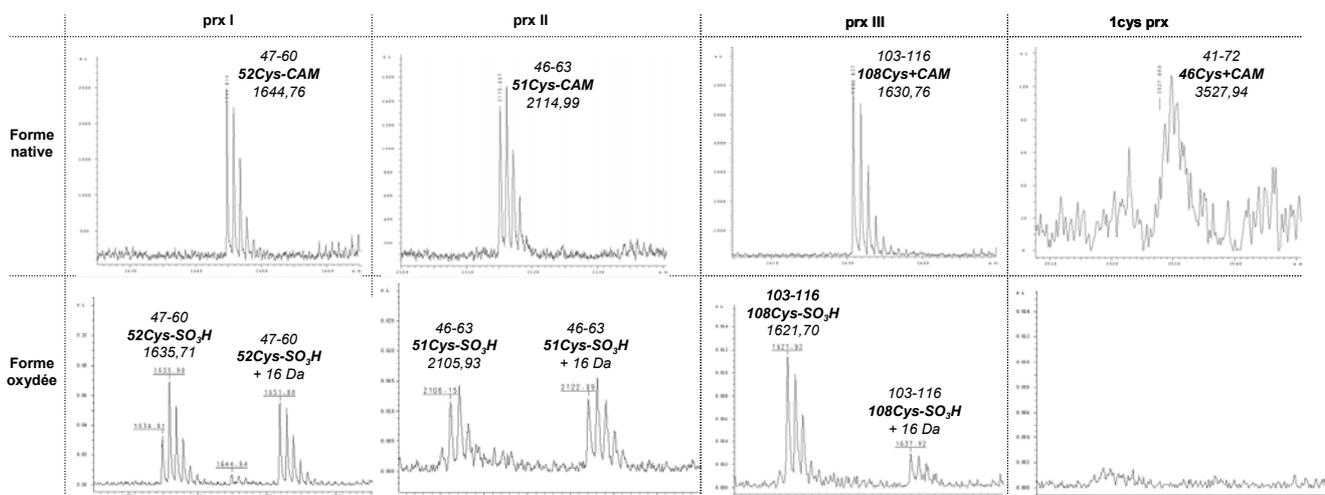
Nous nous sommes également intéressés à la détection en mode négatif du peptide « total » correspondant, pour la forme native de la protéine, aux clivages par l'AspN des deux acides aspartiques encadrant la cystéine du site actif. Celle-ci est alors non oxydée et alkylée (figure 8).

Ce peptide total s'avère plus facilement détectable en mode d'ionisation négatif dans nos nouvelles conditions d'analyse que lors de nos précédentes expériences en mode positif. Nous le détectons, avec un bon rapport signal sur bruit, pour trois des prxs analysées (le pK_a du peptide total pour prx I, II et III avec la cystéine du site actif sous forme carbamidométhylée est compris entre 3 et 5) et, comme attendu, seulement chez la forme native. Ce peptide est également détecté, de façon moins intense, chez 1cys prx native ($pK_a \sim 9$).

Les spectres correspondant à la forme oxydée obtenus pour prx I, II et III révèlent la présence de deux espèces inattendues dans cette zone de la gamme de m/z :

- | L'une dont la masse correspond au peptide total contenant la cystéine du site actif sous forme acide sulfonique. Dans ce cas, l'AspN n'aurait pas clivé du côté N-terminal de la cystéine oxydée. Ces cas de clivage incomplet sont fréquemment observés quelle que soit l'enzyme de digestion utilisée.
- | L'autre espèce possède une masse de 16 Da supérieure à la précédente. Nous émettons l'hypothèse qu'il s'agit de la forme de digestion incomplète décrite ci-dessus comprenant une autre modification post-traductionnelle, par exemple une hydroxylation, sur la proline adjacente à la cystéine du site actif.

Figure 8 : spectres MALDI-MS en mode négatif obtenus après analyse des formes modifiée et native de prx I, prx II, prx III et 1cys prx digérées à l'AspN. Recherche du peptide total contenant la cystéine du site actif non oxydée et alkylée (carbamidométhylation : CAM).



5. Résumé des résultats et conclusion

5.1. Bilan des résultats

L'ensemble de ces résultats permet de compléter la table 1 de la publication présentée à la fin de ce chapitre (tableau 2).

Tableau 2 : résumé des résultats obtenus lors de l'analyse par spectrométrie de masse des prxs. Les dénominations « peptide total », « peptide N-terminal » et « peptide complémentaire » font référence à la figure 3 et au tableau 1 de ce chapitre. Les résultats présentés en gras correspondent à des informations complétant la table 1 de la publication présentée ci-après.

<i>prx</i>	<i>spot</i>	<i>peptide total</i>	<i>peptide N-terminal</i>	<i>peptide complémentaire</i>	<i>état d'oxydation de la cystéine du site actif</i>
prx I	normal	oui	non	non	
prx I	modifié	non	oui	oui	acide sulfinique + acide sulfonique
prx II	normal	oui	non	non	
prx II	modifié	non	oui	oui	acide sulfinique + acide sulfonique
prx III	normal	oui	non	non	
prx III	modifié	non	oui	oui	acide sulfonique
prx IV	normal	non	non	non	
prx IV	modifié	non	oui	oui	
1 cys prx	normal	oui	non	non	
1 cys prx	modifié	non	oui	oui	acide sulfinique + acide sulfonique

5.2. Conclusion

Dans le cadre de l'étude de peroxyrédoxines, nous avons mis au point une stratégie permettant, de façon plus générale, de détecter, de localiser et de caractériser la « sur-oxydation » d'une cystéine contenue dans une séquence protéique. Cette approche peut être reproduite dans d'autres contextes : elle permet de vérifier sans ambiguïté l'état d'oxydation de cystéines appartenant à des isoformes protéiques se distinguant par leur point isoélectrique lors de l'analyse par gel 2D.

Le clivage à l'AspN s'est avéré être très spécifique des acides sulfoniques et sulfiniques. Pour cette étude, nous avons combiné plusieurs techniques qui ont chacune, à différents niveaux, contribué à vérifier notre hypothèse de départ : MALDI-MS en mode positif et négatif, ES-MS et MS/MS. Par ailleurs, les difficultés d'ionisation rencontrées en MALDI-MS pour l'analyse des peptides contenant la cystéine oxydée nous ont amené à tester une matrice originale (DHAP + DAHC) qui semble particulièrement intéressante pour l'analyse en mode d'ionisation négatif.

6. Présentation de ce travail

Ce travail a fait l'objet d'une publication dans *Biochemical Journal* présentée à la page suivante :
« A method for detection of overoxidation of cysteines : peroxiredoxins are oxidized *in vivo* at the active-site cysteine during oxidative stress ».

Nous avons également présenté ces résultats, sous forme de communication par affiche, au 50^{ème} congrès de la Société Américaine de Spectrométrie de Masse (50th ASMS) qui s'est déroulé à Orlando (Floride – USA) en Juin 2002.

III. ETUDE DE LA REGENERATION DES PEROXYREDOXINES SUITE A UN STRESS OXYDANT : MARQUAGE *IN VIVO* ET APPROCHE PROTEOMIQUE QUANTITATIVE

Pour cette seconde partie de notre étude, nous nous sommes intéressés à la régénération des prxs après un stress oxydant en utilisant un marquage *in vivo* des protéines suivi d'une approche protéomique quantitative.

Ce travail a fait l'objet d'un article dans The Journal of Biological Chemistry joint à la fin de ce chapitre. Nous résumerons ici schématiquement la stratégie employée et nous expliquerons les principales difficultés rencontrées lors de cette étude quantitative par spectrométrie de masse.

1. Résumé de notre approche

1.1. Introduction

Comme précédemment, le point de départ de cette étude repose sur le fait que les prxs apparaissent sur les gels 2D sous forme d'un doublet de spots qui sont sensibles au stress oxydant. Le spot basique du doublet correspond aux formes impliquées dans le cycle catalytique normal (thiol, cystine ou acide sulfénique ; cf. § I.). Le spot acide correspond aux formes sur-oxydées de la protéine pour lesquelles la cystéine du site actif est convertie en acide cystéine sulfinique ou sulfonique comme nous l'avons démontré dans la première partie de cette étude.

Nous avons donc utilisé le pouvoir de résolution qu'offrent les gels 2D pour suivre l'évolution des formes normales et sur-oxydées après un stress oxydant :

- | Tout d'abord, la mesure de l'intensité de chacun des deux spots du doublet (colorés à l'argent ou au bleu de Coomassie colloïdal) nous a permis d'estimer la vitesse de régénération des prxs à différents stades de la phase de récupération des cellules suivant le stress. Cette première étape montre déjà des disparités dans les vitesses de régénération des différentes prxs étudiées.
- | Ensuite, dans le but d'appréhender ces processus de régénération plus en détails, nous avons réalisé une série d'expériences consistant à marquer *in vivo* nos prxs avec un isotope radioactif : la lysine deutérée. Après digestion trypsique des protéines, la spectrométrie de masse nous a permis de mesurer les ratios entre les peptides marqués et non marqués et de suivre leur évolution durant la période de régénération.

Les conclusions de ce travail nous ont permis de montrer que prx I, III et 1cys prx se régénèrent grâce à un mécanisme de néosynthèse. Par ailleurs, de façon surprenante, nous avons mis en évidence que le processus de régénération concernant prx II implique, en plus de la néosynthèse, la rétro-réduction chimique de la forme sur-oxydée.

NB : néosynthèse et synthèse protéique désignent le même processus cellulaire comprenant deux principales étapes : transcription et traduction. La transcription est le processus par lequel un morceau particulier d'ARN, appelé ARNm pour ARN messenger, est construit à l'aide d'une séquence génétique

(ADN) comme matrice. La traduction est le processus de fabrication d'une protéine en fonction de l'information contenue dans l'ARNm, à partir des acides aminés disponibles dans la cellule.

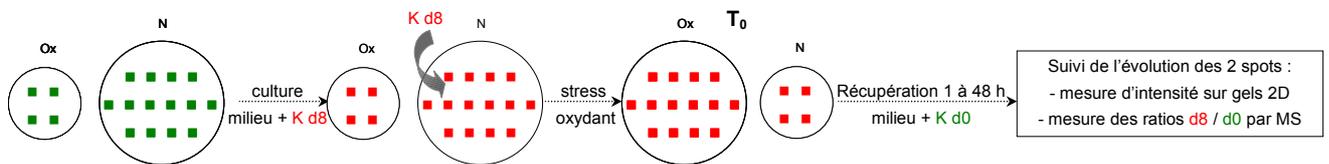
Deux types d'expériences ont été nécessaires pour aboutir à ces conclusions. Nous proposons, dans les prochains paragraphes, d'en expliquer le cheminement de façon schématique.

Dans les figures 9 et 10, des cercles représentent les doublets de spots de prxs visualisés sur les gels 2D (N pour la forme normale, Ox pour la forme sur-oxydée). La taille de ces cercles reflète la proportion relative des deux formes de prxs aux différents stades de l'expérience. Les carrés de couleurs schématisent la population de protéines présentes dans chacun des spots : la forme non marquée (d0) en vert, les protéines marquées avec de la lysine deutérée (d8 ou d4) en rouge. Les flèches épaisses signalent l'incorporation de lysine (d0, d4 ou d8) ainsi que les phénomènes de néosynthèse et de rétro-réduction.

1.2. Expérience n°1

Dans la première partie de nos expériences, les cellules ont poussé dans un milieu enrichi en lysine deutérée (d8). Elles ont ensuite été soumises à un stress oxydant, rincées et finalement placées dans un milieu contenant seulement de la lysine normale (d0) (figure 9A).

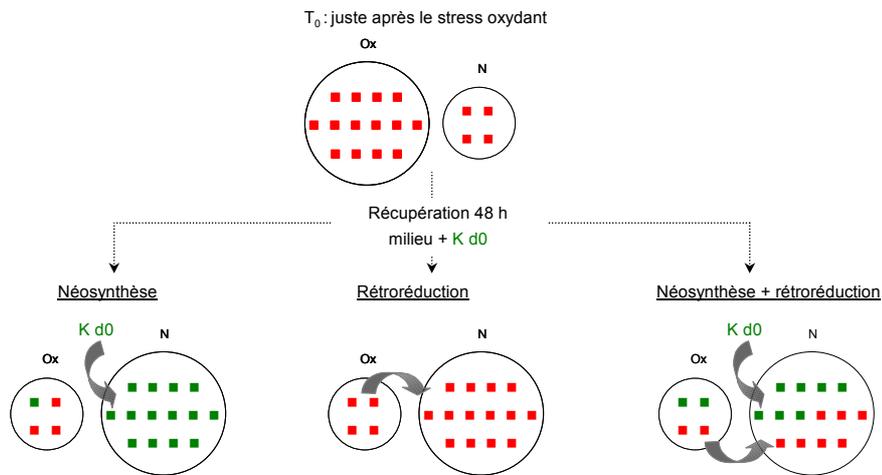
Figure 9A : culture en milieu d8 suivie d'un stress oxydant puis d'une phase de récupération en milieu d0.



La grande majorité des prxs sont marquées à l'issue de leur culture dans le milieu enrichi en lysine deutérée. Le stress oxydant induit la suroxydation des formes natives : le spot acide devient plus intense que le spot basique. Les deux formes de prx sont entièrement marquées au temps T_0 qui correspond au début de la phase de récupération.

Le devenir des prxs au terme de la période de récupération ($T = 24$ à 48 h) est envisagé selon différents modes de régénération (figure 9B).

Figure 9B : devenir des prxs au terme de la phase de récupération selon différents modes de régénération (dans les conditions préalables décrites figure 9A).



Néosynthèse

Les protéines de la forme native ont été entièrement renouvelées par néosynthèse : ces protéines nouvellement synthétisées ont incorporé la lysine $d0$ du milieu de culture. En l'absence de conditions oxydantes, le processus de sur-oxydation est très lent. Aussi, les formes oxydées non marquées correspondantes sont en très faible proportion par rapport aux espèces oxydées marquées pré-existantes.

Rétroréduction

Dans le cas (extrême) où l'on envisage seulement un mécanisme de rétroréduction, les espèces sur-oxydées marquées sont réduites en formes normales marquées. La lysine $d0$ contenue dans le milieu de culture pendant la phase de récupération est inutilisée. Les deux spots correspondent à des prxs complètement marquées. Il faut souligner que ce cas est envisagé pour faciliter la modélisation de notre système biologique et aider à sa compréhension mais il n'a pas été observé expérimentalement.

Néosynthèse + rétroréduction

D'un côté, le mécanisme de néosynthèse contribue à augmenter la proportion de forme normale non marquée. D'autre part, certaines protéines sur-oxydées marquées sont rétroréduites en forme normales marquées. Finalement, la proportion d'espèces marquées dans la forme native est assez importante et supérieure à celle observée chez une protéine témoin qui aurait la même capacité de régénération par néosynthèse.

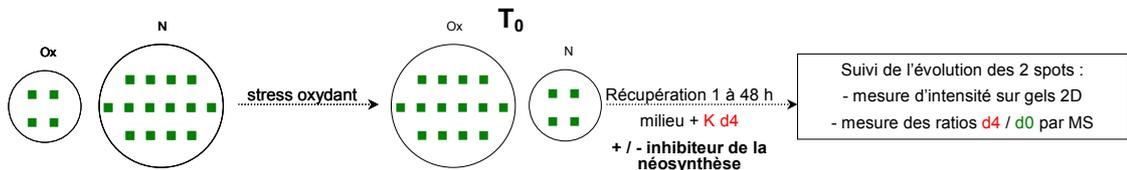
Les résultats obtenus lors de l'expérience n°1 laissent penser que prx I, III et 1cys prx se renouvèlent par néosynthèse plus ou moins rapidement selon le type de prx. Il semblerait que prx II, elle, soit régénérée beaucoup plus rapidement grâce à un mécanisme combinant à la fois néosynthèse et rétroréduction.

Afin de pouvoir vérifier ces observations et discriminer les mécanismes de néosynthèse et de rétroréduction, nous avons réalisé un second type d'expérience avec un schéma de marquage inverse.

1.3. Expériences n°2

Dans cette expérience, les cellules sont d'abord cultivées dans un milieu contenant de la lysine normale (milieu d0). Elles sont ensuite soumises à un stress oxydant, rincées et finalement placées dans un milieu contenant de la lysine deutérée (d4) (figure 10A). Cette expérience a été réalisée de deux manières : avec ou sans inhibiteur de la néosynthèse (figures 10B et 10C).

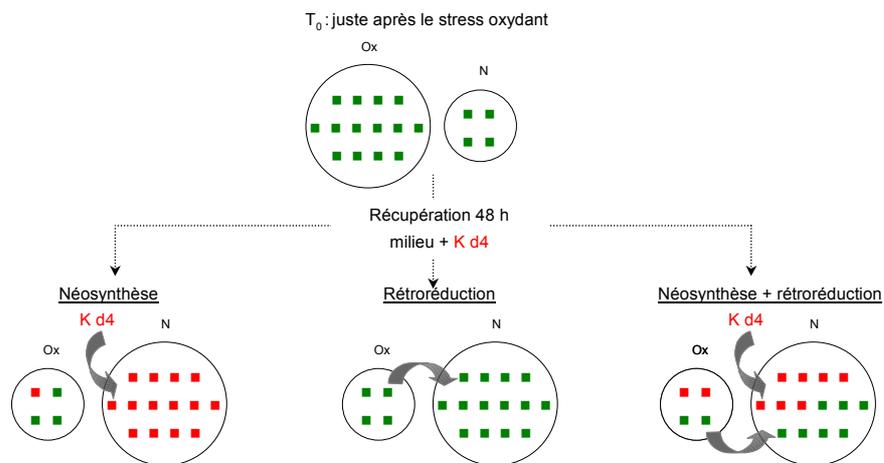
Figure 10A : culture en milieu d0 suivi d'un stress oxydant puis d'une phase de récupération en milieu d4.



De la même manière que précédemment, nous pouvons envisager le comportement des prxs au terme de la phase de récupération selon différents mode de régénération.

1.3.1. Sans inhibiteur de la néosynthèse

Figure 10B : devenir des prxs au terme de la phase de récupération selon différents modes de régénération (dans les conditions préalables décrites figure 10A). Sans inhibiteur de la néosynthèse.

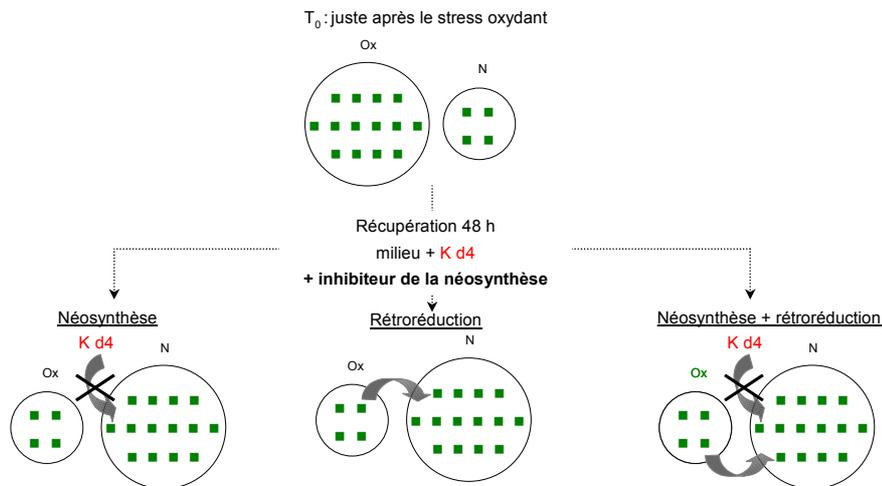


Dans le cas où nous n'utilisons pas d'inhibiteur de néosynthèse (figure 2B), les résultats sont simplement inversés par rapport à l'expérience n°1 décrite précédemment :

- | Pour un mécanisme de **néosynthèse**, la totalité des protéines du spot normal est marquée à la lysine deutérée. Les protéines suroxydées sont majoritairement sous forme d0.
- | Si l'on tient compte du seul mécanisme de **rétro réduction**, aucune des prxs des deux spots n'est marquée.
- | Dans le cas où l'on envisage un mécanisme combinant **néosynthèse et rétro réduction**, les protéines de deux spots coexistent sous formes marquées et non marquées.

1.3.2. Avec inhibiteur de la néosynthèse

Figure 10C : devenir des prxs au terme de la phase de récupération selon différents modes de régénération (dans les conditions préalables décrites figure 10A). Avec inhibiteur de la néosynthèse.



La spectrométrie de masse nous a permis de vérifier que l'ajout d'un inhibiteur de néosynthèse bloque complètement l'incorporation de lysine marquée (d4) (figure 10C) : aucun des peptides analysés à l'issue de cette expérience ne se présente sous la forme marquée (d4).

Si l'on corrèle cette information à l'observation des gels 2D, seul un processus de rétroréduction peut expliquer la régénération des protéines telle qu'elle est visualisée pour prx II.

2. Difficultés rencontrées lors de cette étude

2.1. Le choix du marqueur

2.1.1. Introduction

La méthode de marquage *in vivo* a été décrite pour la première fois par Oda et son équipe [Oda et coll., 1999]. Dans cette expérience, des levures sont cultivées dans deux milieux de culture distincts dont l'un contient des isotopes lourds (¹⁵N). Les deux milieux sont ensuite mélangés. Les protéines du mélange sont extraites, fractionnées sur gel 2D et digérées à la trypsine avant l'analyse MS. Les quantités relatives des protéines sont alors déterminées à partir du calcul des ratios correspondant à la distribution des isotopes dans le mélange.

Aujourd'hui, la littérature présente de nombreux exemples de quantification employant des techniques similaires. Parmi ceux-ci, les travaux de Gu reposent sur l'incorporation de lysines lourdes (marquées au deutérium) dans les protéines durant la culture cellulaire [Gu et coll., 2003]. Outre l'avantage de permettre la quantification, cette technique facilite le séquençage *de novo* car elle permet de distinguer sans ambiguïté l'extrémité C-terminale des peptides tryptiques [Gu et coll., 2002].

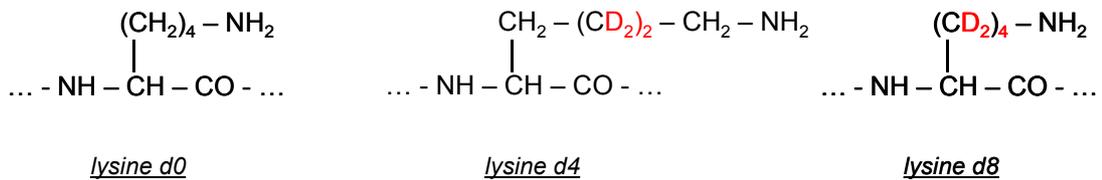
Notre approche s'inspire de celle de Gu dans un contexte très particulier. En effet, il ne s'agit pas de quantifier une famille de protéines dans un mélange mais de suivre la cinétique du renouvellement cellulaire des prxs suite à un stress oxydant afin de caractériser leur mode de régénération.

2.1.2. La lysine deutérée

Le choix de la lysine deutérée comme marqueur pour notre étude s'est avéré intéressant pour les raisons suivantes :

- | Les résidus lysines dans les prxs sont assez abondants.
- | La lysine marquée au deutérium (figure 11) a les mêmes propriétés physiques qu'un résidu de lysine naturel. Aussi, son incorporation cellulaire est tout aussi efficace et spécifique.
- | La lysine est un acide aminé essentiel. Par conséquent, la seule source de lysine pour la cellule est le milieu de culture utilisé sans quoi nos résultats de quantification seraient faussés.
- | La trypsine digère les protéines du côté C-terminal des lysines et des arginines. Ainsi, on génère majoritairement des peptides contenant une lysine unique ce qui évite les marquages multiples qui compliqueraient la détection des peptides correspondants.

Figure 11 : la lysine et ses homologues marqués utilisés pour nos expériences.



N.B. : Le milieu de culture contenant la lysine d8 utilisé pour la première partie de nos expériences contient des traces de résidus lysine d0. Bien que plus onéreux, le milieu enrichi en lysine d4 est, lui, complètement exempt de lysine d0 et s'est avéré indispensable pour la seconde partie de notre étude quantitative qui requiert plus de sensibilité (§ 1.4.).

2.2. La quantification par spectrométrie de masse

2.2.1. Introduction

La spectrométrie de masse nous permet de mesurer la proportion relative de protéines marquées (d4 ou d8) et non marquées (d0) dans chacun des spots du doublet de prxs à différentes étapes de la phase de récupération suivant le stress oxydant.

Pour cette étude quantitative, le mélange issu de la digestion trypsique des prxs a été analysé parallèlement selon deux méthodes :

- | par MALDI-MS (conditions expérimentales décrites dans la publication ci-joint)
- | par nanoLC-MS/MS (couplage entre le système nanoLC Agilent et la trappe ionique Bruker, cf. Mise en œuvre expérimentale).

Dans les deux cas, les informations issues de l'analyse nous ont généralement permis de confirmer l'identité protéique des spots étudiés et de calculer le ratio d4 (ou d8) /d0 correspondant au mélange des formes marquées et non marquées.

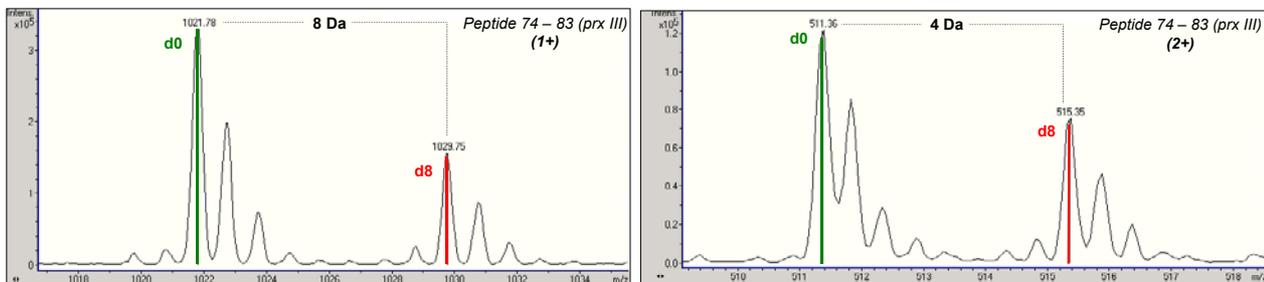
2.2.2. Principe du calcul des ratios

Le calcul des ratios entre les formes marquées et non marquées consiste à mesurer les intensités absolues des pics mono-isotopiques de chacune des deux formes et à en faire le rapport.

Selon la technique utilisée, les peptides sont détectés, soit à l'état simplement chargé (en MALDI-MS et parfois en nanoLC-MS/MS), soit doublement chargé (essentiellement en nanoLC-MS/MS) (figure 12).

- | Lorsqu'ils sont mono-chargés, les paires de peptides (marqué – non marqué) sont séparés de 4 ou 8 Da, selon le marqueur utilisé, lysine d4 ou d8 respectivement.
- | Lorsqu'ils sont doublement chargés, ces peptides sont espacés de 2 unités m/z pour le marquage d4 ou 4 m/z pour le marquage d8.

Figure 12 : formes d0 et d8 du peptide 74 – 83 de prx III (GTAVVNGEFK) à l'état simplement ou doublement chargé. Spectres issus de l'analyse nanoLC-MS/MS.



2.2.3. Calcul des ratios en MALDI-MS

Nous avons utilisé un programme informatique (figure 13), appelé HMTag (pour « Home Made Tag »), développé au laboratoire afin de faciliter l'interprétation des données générées par MALDI-MS lors d'approches quantitatives nécessitant le calcul des ratios entre deux formes se distinguant par la masse d'un marqueur prédéfini.

Dans un premier temps, l'assignation des pics détectés par MALDI-MS permet de générer une liste des m/z caractéristiques du spectre obtenu. A chaque valeur de m/z est associée l'intensité absolue du pic correspondant. Cette liste est ensuite soumise au logiciel HMTag qui permet la détection automatique des paires de peptides séparés par une valeur de m/z multiple de la masse du marqueur.

Figure 13 : le logiciel « Hmtag ».

Ind.	M	M'	Int Abs	Ratio
6	765.37		1.5401e+03	
7		773.42	2.6431e+03	Ratio: 1.7161872605675
11	862.48		3.2283e+03	
12		870.53	5.2774e+03	Ratio: 1.63473035343679
23	1023.53		1.1137e+03	
24		1031.58	1.6830e+03	Ratio: 1.51117895303942
28	1179.64		4.5806e+03	
29		1187.69	7.0176e+03	Ratio: 1.53202637209099
34	1233.66		8.1736e+02	
36		1249.62	7.0940e+02	Ratio: 0.867916218067926
58	1863.04		6.6504e+03	
59		1871.09	9.8942e+03	Ratio: 1.48776013472874
60	1928.95		3.0241e+02	
61		1936.99	4.2340e+02	Ratio: 1.40008597599286
65	2072.25		1.8814e+02	
66		2080.27	2.6312e+02	Ratio: 1.39853300733496
73	2699.39		1.3925e+03	
74		2715.41	6.4345e+02	Ratio: 0.462082585278276
77	2827.53		3.5145e+02	
78		2835.59	5.6068e+02	Ratio: 1.59533361786883

Moyenne des Ratios : 1.20011067002064
Ecart Type P : 0.460179716581907

Dans le cadre de notre étude, nous avons paramétré le logiciel selon les critères suivants :

- | l'incrément de masse lié au marqueur, (4 ou 8 Da) pour la lysine deutérée selon l'expérience
- | la tolérance de masse accordée pour la recherche des paires de peptides, soit 50 ppm (équivalente à la précision de mesure)
- | « le degré d'itération » fixé à 2, c'est-à-dire le nombre maximum de marquages possibles par peptide.

NB : nous tenons compte des cas éventuels où la digestion trypsique est incomplète et génère des peptides contenant deux lysines.

Finalement, HMTAG calcule les ratios correspondant aux paires de peptides de la liste répondant à nos critères. Le logiciel calcule également la moyenne et l'écart type relatifs à l'ensemble des ratios calculés.

Nous vérifions manuellement la cohérence des résultats obtenus. Nous écartons :

- | les couples de masses ne correspondant pas à des peptides issus de la digestion *in silico* des prxs étudiées contenant 1 ou 2 lysines
- | les paires de peptides dont le ratio est trop éloigné de la valeur moyenne (après vérification sur le spectre, il s'agit généralement de la superposition de profils isotopiques de peptides de masses voisines).

Après ces vérifications, nous considérons que le ratio moyen calculé est le reflet de la proportion entre la quantité de protéines marquées et non marquées présentes dans l'échantillon.

2.2.4. Calcul des ratios en nanoLC-MS/MS

Lorsque nous avons effectué ces analyses, nous ne disposions pas d'outils informatiques permettant la détection automatique, à partir des résultats obtenus en nanoLC-MS/MS, des couples de peptides marqués et non marqués ainsi que le calcul des ratios correspondant.

C'est pourquoi, le traitement des données générées par ces analyses a principalement été réalisé de façon manuelle.

a. Détection des couples de peptides d'intérêt

Nous avons paramétré le moteur de recherche MASCOT afin de pouvoir détecter la présence des peptides modifiés par les isotopes lourds de la lysine. En effet, nous avons défini dans MASCOT, au même titre qu'une modification post-traductionnelle (cf. partie II, §III.3.1.), l'incrément de masse correspondant à la présence de lysine deutérée dans les peptides analysés (+ 4 ou + 8 Da).

Ainsi, les résultats obtenus, après interrogation dans les banques de données protéiques avec le moteur de recherche MASCOT, nous ont indiqué quels couples de peptides (forme marquée + non marquée) utiliser pour le calcul des ratios.

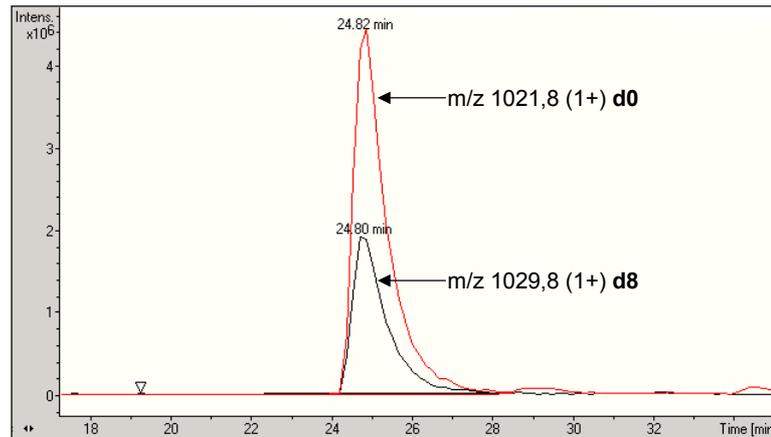
b. Co-élution des peptides marqués et non marqués

Pour une quantification correcte, il est important que les couples de peptides (marqués et non marqués) soit coélus lors de la séparation chromatographique.

Or, des études ont montré que les isotopes deutérés de certains agents utilisés pour le marquage en protéomique quantitative pouvaient avoir des comportements chromatographiques très différents de l'isotope léger correspondant. La réactivité des isotopes deutérés par rapport à la phase stationnaire dépend, entre autres, de la structure chimique du marqueur, du nombre d'atomes de deutérium et de leur emplacement. Ces phénomènes semblent être relativement limités lorsque le nombre de deutérium est faible et qu'ils sont positionnés à proximité des fonctions à caractère hydrophile de l'agent de marquage [Zhang et coll., 2002].

Par conséquent, dans le cadre de notre étude, nous avons vérifié la coélution des peptides contenant de la lysine d0 avec leurs homologues marqués sous forme d4 ou d8. Pour l'ensemble de nos expériences, nous n'avons pas mis en évidence de différences significatives dans les temps de rétention de ces espèces. Un exemple est présenté figure 14.

Figure 14 : coélution des peptides marqué et non marqué en nanoLC-MS/MS. Courant d'ions extrait du peptide 74 – 83 de prx III.



c. Calcul des ratios

Nous avons recherché les pics chromatographiques correspondant à l'élution de chacun des couples de peptides d'intérêt préalablement identifiés lors de l'interrogation dans les banques de données protéiques avec MASCOT. Le spectre de masse pris au sommet du pic chromatographique nous a permis d'obtenir les intensités des peptides étudiés et de calculer le ratio correspondant.

La moyenne des ratios ainsi calculés pour une analyse reflète la proportion entre la quantité de protéines marquées et non marquées présentes dans l'échantillon.

2.2.5. Comparaison des résultats obtenus en MALDI-MS et en nanoLC-MS/MS

Les résultats obtenus nous permettent de tracer les courbes caractéristiques de la cinétique de régénération des prxs après un stress oxydant. Nous avons superposé les courbes obtenues avec les techniques MALDI-MS et nanoLC-MS/MS afin d'en comparer les résultats. Un exemple est présenté figure 15.

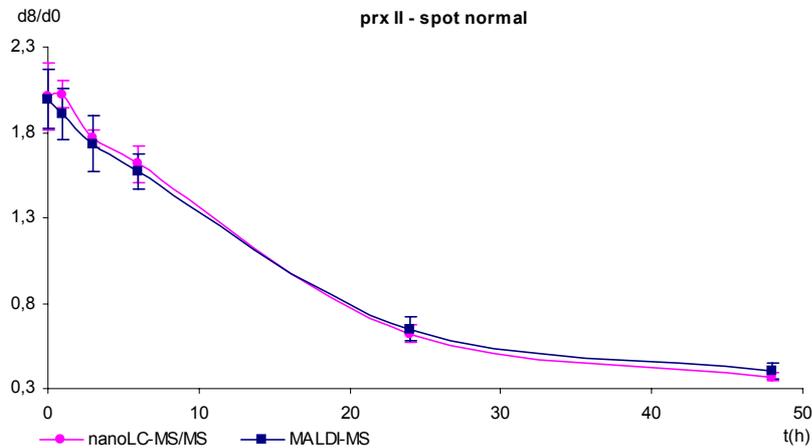
De façon générale, la comparaison de ces courbes montrent que les deux techniques utilisées donnent des résultats similaires pour le calcul des ratios et permettent indifféremment d'étudier le comportement des prxs.

Cependant, un examen plus approfondi des résultats révèle que le nombre de peptides détectés en nanoLC-MS/MS ayant contribué au calcul des ratios isotopiques est souvent supérieur à celui obtenu en MALDI-MS.

En effet, dans ce type d'approche, l'avantage de la technique nanoLC-MS/MS par rapport à l'analyse MALDI-MS est de permettre la séparation préalable du mélange peptidique, rendu d'autant plus complexe par le marquage isotopique. Cette étape de fractionnement de l'échantillon, permet d'analyser successivement les peptides en limitant les phénomènes d'interférence nuisibles à la qualité de leur détection. Par ailleurs, les informations de séquences obtenues par MS/MS permettent de renforcer la confiance portée aux résultats en terme d'identification protéique (cf. Introduction – Bibliographie §II.1.2.2).

Aussi, pour notre étude, la technique nanoLC-MS/MS a permis de compléter les résultats obtenus par MALDI-MS notamment dans le cas où les prxs étaient très faiblement représentées dans le spot analysé (pour certaines formes sur-oxydées par exemple).

Figure 15 : courbe illustrant la régénération de la forme normale de prx II après un stress oxydant. Suivi de



l'évolution du ratio d8/d0 dans le temps (dans les conditions de l'expérience décrite figure 9A).

3. Conclusion

Sur le plan biologique, nos résultats ont montré une large variété dans les vitesses de régénération des peroxyrédoxines étudiées :

- | Prx III et VI se régénèrent lentement et vraisemblablement seulement par néosynthèse.
- | 1 cys prx se régénère plus rapidement par ce même mécanisme.
- | De façon inattendue, nous avons clairement démontré que prx II se régénère grâce à un mécanisme combinant néosynthèse et rétro-réduction.

La lysine deutérée (d4 ou d8) s'est avérée être un marqueur de choix pour notre étude. En qualité d'acide aminé essentiel, elle nous a permis d'étudier les mécanismes biologiques impliqués dans le renouvellement des peroxyrédoxines avec une grande spécificité.

Par ailleurs, nous avons montré que ce marqueur ne modifie pas le comportement chromatographique des peptides analysés permettant leur quantification relative par nanoLC-MS/MS.

Les deux approches testées (MALDI-MS et nanoLC-MS/MS) ont offert des résultats satisfaisants démontrant la validité de ces techniques dans le cadre d'une étude quantitative.

Cependant, le traitement des données générées par les analyses nanoLC-MS/MS s'est avéré compliqué à défaut de logiciel permettant la détection des couples de peptides marqués et non marqués ainsi que le calcul automatique des ratios correspondant. Aujourd'hui, les récents développements du logiciel Biotools® (Bruker) permettent de répondre à cette attente en offrant la possibilité à l'utilisateur de définir lui-même le marqueur utilisé pour son étude quantitative.

4. Présentation de ce travail

Ce travail a fait l'objet d'un article joint ci-après paru dans The Journal of Biological Chemistry : « Regeneration of peroxiredoxins during recovery after oxidative stress : only some overoxidized peroxiredoxins can be reduced during recovery after oxidative stress ».

Nous avons également présenté ces résultats, sous forme de communications par affiche, au 16^{ème} congrès international de spectrométrie de masse (16th IMSC) qui s'est déroulé à Edimbourg (Ecosse) en Septembre 2003 ainsi qu'aux 20èmes journées françaises de spectrométrie de masse et d'analyse protéomique à Toulouse en Septembre 2003.

PUBLI 1

PUBLI 2

PUBLI 3

PUBLI 4

PUBLI 5

PUBLI 6

PUBLI 7

PUBLI 8

BIBLIOGRAPHIE

Chae H.Z., Chung S.J., Rhee S.G.

Thioredoxin-dependent peroxidase from yeast, *J. Biol. Chem.*, **1994**, 269, 27670-27678.

Choi H.J., Kang S.W., Yang C.H., Rhee S.G., Ryu S.E.

Crystal structure of a novel human peroxidase enzyme at 2.0 Å resolution, *Nat. Struct. Biol.*, **1998**, 5, 400-406.

Declercq J.P., Evrard C., Clippe A., Stricht D.V., Bernard A., Knoops B.

Crystal structure of human peroxiredoxin 5, a novel type of mammalian peroxiredoxin at 1.5Å resolution, *J. Mol. Biol.*, **2001**, 311, 751-759.

Drapeau G.R.

Substrate specificity of a proteolytic enzyme isolated from a mutant of *Pseudomonas fragi*, *J. Biol. Chem.*, **1980**, 255, 839-840.

Fournier I., Marinach C., Tabet J.C., Bolbach G.

Irradiation effects in MALDI, ablation, ion production, and surface modifications. Part II. 2,5-dihydroxybenzoic acid monocrystals., *J. Am. Soc. Mass Spectrom.*, **2003**, 14, 893-9.

Fridovich I., Freeman B.

Antioxidant defenses in the lung, *Annu. Rev. Physiol.*, **1986**, 48, 693-702.

Gorman J.J., Ferguson B.L., Nguyen T.B.

Use of 2,6-dihydroxyacetophenone for analysis of fragile peptides, disulphide bonding and small proteins by matrix-assisted laser desorption/ionization, *Rapid Commun. Mass Spectrom.*, **1996**, 10, 529-536.

Gu S., Pan S., Bradbury E.M., Chen X.

Use of deuterium-labeled lysine for efficient protein identification and peptide *de novo* sequencing, *Anal. Chem.*, **2002**, 74, 5774-5785.

Gu S., Pan S., Bradbury E.M., Chen X.

Precise peptide sequencing and protein quantification in the human proteome through *in vivo* lysine-specific mass tagging, *J. Am. Soc. Mass Spectrom.*, **2003**, 14, 1-7.

Hirotsu S., Abe Y., Okada K., Nagahara N., Hori H., Nishino T., Hakoshima T.

Crystal structure of a multifunctional peroxiredoxin heme-binding protein 23 kDa/ proliferation-associated gene product, *Proc. Natl. Acad. Sci. USA*, **1999**, 96, 12333-12338.

Oda Y., Huang K., Cross F.R., Cowburn D., Chait B.T.

Accurate quantification of protein expression and site-specific phosphorylation, *Proc. Natl. Acad. Sci. USA*, **1999**, 96, 6591-6596.

Phelan S.A., Johnson K.A., Beier D.R., Paigen B.

Characterization of the murine gene encoding Aop2 (antioxidant protein 2) and identification of two highly related genes, *Genomics*, **1998**, 54, 132-139.

Prosperi M.T., Ferbus D., Karczynski I., Goubin G.

A human cDNA corresponding to a gene overexpressed during cell proliferation encodes a product sharing homology with amoebic and bacterial proteins, *J. Biol. Chem.*, **1993**, 268, 11050-11056.

Rabilloud T., Berthier R., Vinçon M., Ferbus D., Goubin G., Lawrence J.J.

Early events in erythroid differentiation: accumulation of the acidic peroxidoxin (PRP/TSA/NKEF-B), *Biochem. J.*, **1995**, 312, 699-705.

Sies H.

Strategies of antioxidant defense, *Eur. J. Biochem.*, **1993**, 215, 213-219.

Watabe S., Kohno H., Kouyama H., Hiroi T., Yago N., Nakazawa T.

Purification and characterization of a substrate protein for mitochondrial ATP-dependent protease in bovine adrenal cortex, *J. Biochem. (Tokyo)*, **1994**, 115(4), 648-654.

Zhang R., Sioma C.S., Thompson R.A., Xiong L., Regnier F.E.

Controlling deuterium isotope effects in comparative proteomics, *Anal. Chem.*, **2002**, 74, 3662-3669.

PARTIE IV :

**CARACTERISATION D'ENZYMES PURIFIEES
CHEZ LES ALGUES BRUNES :**

**UNE APPROCHE PROTEOMIQUE DANS LE CAS OU LES
PROTEINES SONT INCONNUES DES BANQUES DE DONNEES**

CARACTERISATION D'ENZYMES PURIFIEES CHEZ LES ALGUES BRUNES :

UNE APPROCHE PROTEOMIQUE DANS LE CAS OU LES PROTEINES SONT INCONNUES DES BANQUES DE DONNEES

Cette étude a été réalisée en collaboration avec le Dr Catherine Leblanc et Carole Colin, de l'UMR 7139 (CNRS, Laboratoires Goëmar, UPMC) de la Station Biologique de Roscoff.

I. INTRODUCTION

1. Contexte biologique

Dans l'eau de mer, l'iode se trouve principalement sous forme d'iodate (IO_3^-) et, en plus faible proportion, sous forme d'iodure (I^-) et d'iode moléculaire (I_2 et HIO). La couche euphotique des océans (surface de l'eau pénétrée par le rayonnement solaire) contient $0,25 \mu\text{M}$ d'iode inorganique.

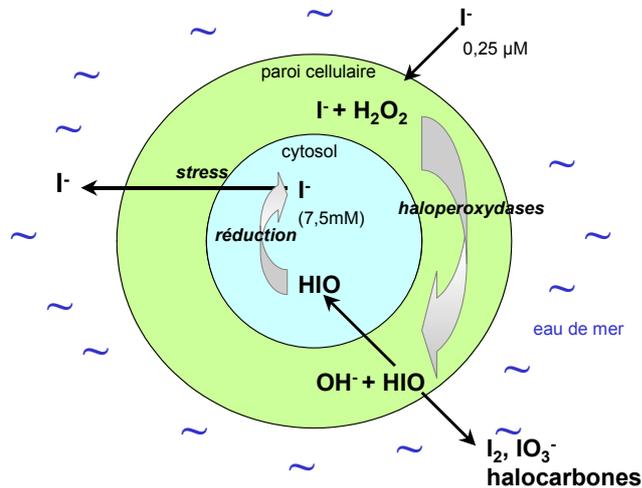
Par ailleurs, les Laminaires, des algues brunes très répandues sur le littoral breton (figure 1), sont de vrais réservoirs d'iode ; en particulier, les jeunes sporophytes de *Laminaria digitata* peuvent accumuler jusqu'à 150 000 fois la concentration en iode dans l'eau de mer. Ces algues sont également responsables de l'émission de composés volatils halogénés, majoritairement bromés et iodés, observée à marée basse. Cependant, le rôle physiologique de l'iode chez ces algues est encore mal connu.

Figure 1 : les Laminaires, des algues brunes très répandues sur les côtes bretonnes.



Des recherches, menées sur les sporophytes de *L. digitata*, montrent que l'absorption cellulaire d'iode sous forme d'iodure nécessite la présence de peroxyde d'hydrogène et d'une haloperoxydase dépendante du vanadate (v-HPO) [Küpper et coll., 1998]. L'iode diffuse ensuite dans le cytosol, sous forme d'acide hypo-iodeux qui, une fois réduit, est accumulé sous forme d'iodure (figure 2).

Figure 2 : métabolisme de l'iode chez les sporophytes de *L. digitata*.



Parallèlement, des travaux visant à mieux comprendre la signalisation par les oligosaccharides et les réactions de défense chez les algues [Potin et coll., 1999] ont mis en évidence que l'application de fragments de paroi, les oligoalginates, provoque un stress oxydant chez les sporophytes de *L. digitata*. Ce stress, en mimant l'attaque d'un pathogène, entraîne l'émission d'espèces activées de l'oxygène ou « burst oxydatif » [Küpper et coll., 2000] ainsi qu'un fort efflux d'iode et de composés volatils halogénés [Küpper et coll., 1998]. Il existe donc un lien entre le stress oxydant et l'émission massive d'iode et de composés halogénés.

Par conséquent, les haloperoxydases dépendantes du vanadate, enzymes clés du métabolisme halogéné, pourraient également être impliquées dans les réponses au stress oxydant et les mécanismes de défense. Ainsi, l'équipe de Roscoff a entrepris la caractérisation biochimique et moléculaire de différentes haloperoxydases afin de mieux comprendre leurs fonctions.

Des activités bromoperoxydases (BPO), capables d'oxyder à la fois le brome et l'iode, et une activité iodoperoxydase (IPO), spécifique de l'iode, ont été mises en évidence dans des extraits protéiques de *L. digitata*. Ces enzymes ont été purifiées à l'homogénéité par des techniques d'électroélution et d'électrophorèse semi-préparative.

Dans ce cadre, nous expliquerons dans les prochains paragraphes pourquoi nous avons mis en œuvre une approche protéomique par séquençage *de novo* afin d'obtenir des informations relatives à ces deux types d'haloperoxydases.

2. Particularités des haloperoxydases de *Laminaria digitata* pour l'approche protéomique

2.1. Quelles sont les informations dont nous disposons ?

Cette étude se place dans un contexte particulier puisque le génome de *L. digitata* n'est pas séquencé.

Un programme EST («Expressed Sequence Tag») chez *L. digitata* a permis d'identifier un ADNc (ADN complémentaire) partiel présentant des homologies de séquence avec des bromoperoxydases connues [Crépineau et coll., 2000].

| L'hybridation de l'ADN génomique du sporophyte de l'algue par une sonde intragénique préparée à partir de cet ADNc partiel (technique « southern blot ») met en évidence plusieurs bandes correspondant à plusieurs copies du gène ou/et de gènes homologues. Les haloperoxydases constituent donc une famille multigénique.

| Par ailleurs, le criblage d'une banque d'ADNc de *L. digitata* par une sonde préparée à partir de l'ADNc partiel d'haloperoxydase a permis de cloner deux nouveaux ADNc entiers très semblables, provenant probablement de localisations sub-cellulaires différentes. Le taux d'homologie de ces ADNc avec les bromoperoxydases rencontrées chez d'autres familles d'algues varie de 30% à 40%. Ces ADNc pourraient « coder » pour des bromoperoxydases.

2.2. Objectifs de l'étude et intérêt de la spectrométrie de masse dans ce contexte

Les haloperoxydases purifiées à partir des sporophytes de *L. digitata* sont préalablement séparées sur gel natif d'acrylamide (à une dimension) coloré au bleu de Coomassie. En parallèle, un test de détection (décrit dans la publication n°1 ci-jointe) spécifique de leur activité, bromo ou iodoperoxydase, permet de reconnaître les protéines sur le gel natif (figure 1, publication n°1).

Dans ce contexte, les protéines étant inconnues des banques de données, il s'agit d'obtenir, par spectrométrie de masse en tandem, des informations de séquences caractéristiques des protéines étudiées. Deux approches composent notre étude :

| La première approche consiste à vérifier l'homologie de séquence des protéines présentant une activité bromoperoxydase avec les deux ADNc dont nous disposons ainsi qu'avec des bromoperoxydases appartenant à d'autres familles d'algues (dont les séquences sont répertoriées dans les banques de données). Dans ce cadre, il s'agit d'une **approche par comparaison de données pré-existantes avec les listes de masses des fragments peptidiques obtenus par nanoLC-MS/MS**.

| La seconde approche, plus ciblée, concerne les protéines ayant une activité iodoperoxydase. Dans un premier temps, les séquences obtenues doivent permettre la **construction d'amorces nucléotidiques** (« primers dégénérés ») afin de cloner les nouveaux ADNc correspondants. Il

s'agit ensuite d'identifier l'ADNc cloné par comparaison directe de sa séquence protéique avec les celle des peptides analysés par nanoLC-MS/MS.

II. LA STRATEGIE PAR SEQUENÇAGE *DE NOVO*

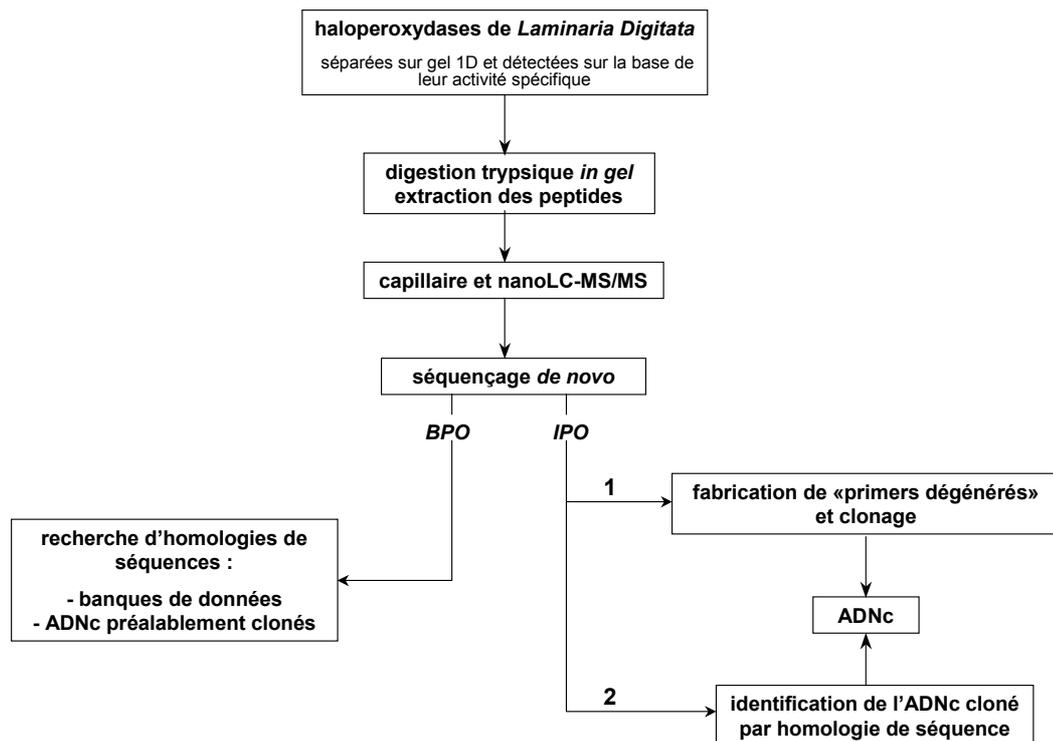
1. Stratégie utilisée pour notre étude

Cette étude consiste donc à séquencer *de novo* (cf. Introduction – Bibliographie, §II.4.) les haloperoxydases de *L. digitata*. Un schéma illustrant la stratégie que nous avons employée est présenté en figure 3.

Les protéines étudiées sont digérées *in gel* à la trypsine puis, les peptides extraits sont analysés par nanoLC-MS/MS et par capillaire LC-MS/MS (cf. parties expérimentales des publications jointes).

Dans les prochains paragraphes, nous nous proposons de résumer les étapes de l'approche par séquençage *de novo* et d'expliquer les principales difficultés rencontrées pour cette étude.

Figure 3 : stratégie employée pour l'analyse par séquençage *de novo* des haloperoxydases de *L. digitata*. BPO : bromoperoxydases, IPO : iodoperoxydases.



2. Mise en œuvre

Lorsque nous avons débuté ce travail, nous ne disposions pas de logiciels suffisamment performants pour interpréter automatiquement la séquence en acides aminés des peptides analysés par LC-MS/MS. Aussi, l'interprétation de nos résultats a principalement été réalisée de façon « manuelle », spectre après spectre. Nous verrons par la suite que des outils informatiques, très

récemment commercialisés, apportent aujourd'hui une aide précieuse pour l'approche systématique par séquençage *de novo* en LC-MS/MS.

2.1. Séquençage de novo « manuel »

L'analyse LC-MS/MS permet de générer un nombre plus ou moins important de spectres MS/MS selon la complexité du mélange et la concentration des peptides tryptiques présents dans l'échantillon. Devant l'ampleur du travail (jusqu'à 150 peptides fragmentés par analyse LC-MS/MS pour cette étude), nous nous sommes prioritairement intéressés aux spectres MS/MS les plus significatifs (en terme d'intensité et de nombre de fragments obtenus).

Des logiciels de séquençage (*De novo sequencing*® de Biotools®, Bruker ; *Peptide sequencing*® de BiolyNX®, Micromass) nous ont aidé pour l'interprétation des spectres MS/MS. Cependant, nous avons souhaité vérifier manuellement la cohérence des résultats obtenus avec ce type d'outil.

2.1.1. Déconvolution

Il peut être délicat d'interpréter les spectres de fragmentation d'ions multichargés. En effet, l'écart de masse entre deux fragments consécutifs n'a de signification que si l'on compare des ions possédant le même état de charge. Or, si l'on considère, par exemple, un ion parent doublement chargé, les fragments générés en MS/MS à partir de cet ion peuvent être, soit simplement (MH^+), soit doublement chargés (MH_2^{2+}).

C'est pourquoi, nous avons recouru à un algorithme de déconvolution permettant de transformer les spectres d'ions multichargés en spectres d'ions monochargés [Wehofsky et Hoffmann, 2002]. Selon le spectromètre de masse, nous avons utilisé le programme MaxEnt3® (Micromass) ou le programme de déconvolution du logiciel Data Analysis® (Bruker).

2.1.2. Séries b et y

La première difficulté rencontrée pour l'interprétation *de novo* des spectres de fragmentation de peptides tryptiques (cf. Introduction – Bibliographie, §II.3.3.) réside dans l'attribution des fragments à une série d'ions (principalement b et y pour les fragmentations à basse énergie). Cette distinction est très importante car elle définit le cadre de lecture de la séquence peptidique. Plusieurs indications sur les spectres MS/MS peuvent permettre de distinguer les séries de fragments b et y :

- | Les peptides tryptiques ont pour particularité de porter une charge permanente conférée par la lysine ou l'arginine en position C-terminale. C'est pourquoi, les spectres MS/MS correspondant présentent généralement une série de fragments y dominants et des signaux de plus faibles intensités pour la série des fragments b. L'information apportée par l'intensité des fragments peut d'ailleurs permettre une « caractérisation statistique » des spectres MS/MS utile au développement d'algorithmes servant à identifier la séquence en acides aminés des peptides analysés [Tabb et coll., 2003].
- | Le fragment y_1 peut servir de repère pour démarrer le séquençage. En effet, il possède un rapport m/z caractéristique : 147,1 ou 175,1 Da selon que le peptide tryptique se termine par une lysine ou une arginine respectivement.

Il existe souvent des ions, quoique de faible intensité, correspondant à d'autres séries de fragments et permettant de guider l'identification : les ions c et z voisins des fragments majoritaires b ou bien les ions a ou x vicinaux des ions y.

2.1.3. Informations complémentaires sur la composition en acides aminés

La différence de masse entre les ions consécutifs d'une même série, y ou b, permet donc de déduire la séquence en acides aminés du peptide analysé.

Deux autres types de fragments (fragments internes et ions immoniums), trouvés dans la plupart des spectres à basse énergie, résultent du clivage d'au moins deux liaisons internes à la chaîne peptidique et apportent des informations complémentaires sur la composition en acides aminés.

a. Les fragments internes

Le premier type est appelé fragment interne car il s'agit d'un fragment ayant perdu les côtés N et C-terminaux [Hunt et coll., 1986]. Ce type d'ions contient rarement plus de trois ou quatre acides aminés et se retrouve donc généralement dans le spectre au niveau des bas m/z.

Les fragments internes sont souvent en faible abondance, à l'exception des fragments contenant une Proline. En effet, le groupement imino de la Proline est inclus dans un cycle à cinq atomes et possède une densité électronique plus importante que les autres liaisons amides du peptide. Dès lors, la protonation et le clivage de la liaison amide de la Proline pour produire un fragment interne est un processus favorisé [Loo et coll., 1993].

Les fragments internes peuvent donc permettre de confirmer partiellement la séquence peptidique mais compliquent cependant l'interprétation du spectre MS/MS.

b. Les ions immoniums

Tableau 1 : rapports m/z des ions immoniums caractéristiques des acides aminés naturels les plus fréquemment rencontrés dans les spectres de fragmentation à basse énergie des peptides [Falik et coll., 1993].

Acides aminés	Masses caractéristiques (m/z)
Proline	70,05
Valine	72,07
Leucine	86,08
Isoleucine	86,08
Méthionine	104,04
Histidine	110,06
Phénylalanine	120,07
Tyrosine	136,06
Tryptophane	159,08

Le second type de fragment qui résulte de clivages multiples de la chaîne peptidique apparaît également dans les bas m/z du spectre : ce sont les ions immoniums des acides aminés.

Bien qu'ils soient rarement observés pour tous les acides aminés du peptide, ceux qui apparaissent donnent une information sur la composition de la séquence étudiée [Falick et coll., 1993]. Une liste des ions les plus couramment retrouvés dans les spectres est présentée dans le tableau 1.

2.1.4. Ambiguïtés pour la détermination de la séquence peptidique

L'interprétation des spectres de fragmentation est parfois ambiguë, notamment pour les acides aminés isobares (Leucines / Isoleucine) ou les acides aminés dont la différence de masse est très faible (cf. Introduction – Bibliographie, §II.3.3.).

Par ailleurs, il est fréquent d'observer des ions présentant une différence de -18 Da par rapport à un fragment b ou y. La présence de ces ions est difficilement interprétable :

- | Elle peut correspondre à la déshydratation des groupements hydroxy de résidus Sérine ou Thréonine : dans ce cas on observe un ion y ou b ayant perdu une molécule d'eau. Ces pertes de -18 Da sont alors généralement observées de façon récurrente sur le spectre pour tous les fragments y ou b contenant au moins un de ces deux types d'acide aminé.
- | La différence de masse observée (18 Da) correspond également à la différence de masse entre deux acides aminés : Phénylalanine / Acide glutamique (18,03 Da) ; Méthionine / Leucine ou Isoleucine (17,96 Da) ; Acide aspartique / Proline (17,97 Da).

2.2. Séquençage *de novo* « automatique » pour la LC-MS/MS

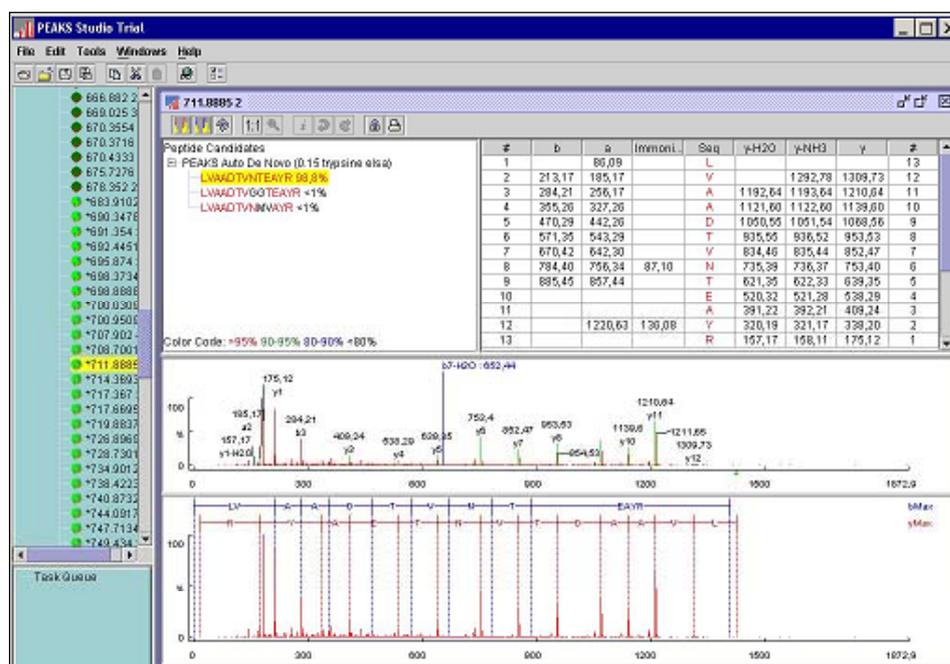
Nous avons récemment fait l'acquisition au laboratoire du logiciel PEAKS® (Bioinformatics Solutions Inc) permettant le séquençage *de novo* des peptides analysés par LC-MS/MS. Pour ce type d'analyse, où les spectres MS/MS sont générés de façon automatique et en grand nombre, cet outil constitue une aide précieuse à l'interprétation.

- | Dans un premier temps, l'utilisateur paramètre le logiciel avec les informations relatives à l'enzyme de digestion utilisée ainsi qu'aux modifications chimiques (ex. : carbamidométhylation des Cystéines) et post-traductionnelles (ex. : oxydation des Methionines) susceptibles d'affecter la masse des acides aminés des peptides.
- | Ensuite, la liste des m/z des parents sélectionnés et leurs fragments respectifs (« peak list », cf Le couplage nanoLC-MS/MS avec une trappe ionique § IV.2.) générée à partir de l'analyse LC-MS/MS est directement soumise au logiciel PEAKS. Cette liste doit contenir les informations relatives à l'état de charge des ions parents ce qui nécessite que les spectres MS aient été préalablement déconvolués.
- | Enfin, le logiciel traite l'ensemble des résultats de la « peak list » et propose pour chaque ion parent une ou plusieurs séquences dont la probabilité est définie par un score. Ainsi, la séquence peptidique peut être soit entièrement, soit partiellement interprétée. Le spectre MS/MS de chaque ion parent est reconstitué sur la base de l'intensité des ions fragments (issue de la « peak list ») (figure 4).

Nous avons réalisé au laboratoire des essais consistant à comparer les résultats issus d'un séquençage manuel avec ceux obtenus avec le logiciel PEAKS® ; ces tests, réalisés sur des séquences de natures variées, ont montré la fiabilité de l'algorithme. PEAKS® est donc aujourd'hui

utilisé au laboratoire en routine et permet le séquençage *de novo* des peptides analysés par LC-MS/MS de façon rapide et exhaustive.

Figure 4 : le logiciel PEAKS® (Bioinformatics Solutions Inc). A gauche, la liste des ions parents constitutifs de la « peak list ». A droite, exemple de résultat de séquençage *de novo* pour un peptide de la liste (en jaune) : proposition de séquences, fragments identifiés et spectres MS/MS reconstitués et interprétés.



2.3. Comparaison des résultats MS/MS obtenus avec la trappe ionique et le Q-TOF

Dans le cadre de cette étude, nous avons utilisé deux techniques différentes pour obtenir des informations de séquences des protéines étudiées : la nanoLC-MS/MS avec un Q-TOF et la capillaire LC-MS/MS avec une trappe ionique. Etant donnée la différence entre les deux systèmes HPLC utilisés pour ces couplages (débit et diamètre interne des colonnes de chromatographie), notre objectif n'est pas de comparer les deux techniques.

Cependant, de façon plus générale, nous avons cherché à comparer la nature des résultats obtenus en MS/MS avec les deux spectromètres de masse. Un exemple, présenté figure 5, illustre les différences observées.

1. La gamme de m/z sur la trappe ionique est limitée par la valeur du cut-off (cf. Introduction – Bibliographie, §1.1.4.2.) : avec cet instrument, nous perdons donc les informations relatives aux fragments y_1 et aux ions immoniums qui sont bien visibles sur le spectre obtenu avec le Q-TOF.

2. Les spectres MS/MS réalisés avec le Q-TOF ne présentent que peu d'ions de la série b, principalement de bas m/z. Bien que minoritaires, les fragments b sont généralement observés sur les spectres de fragmentation obtenus avec la trappe ionique. Leur présence permet de confirmer les informations de séquences déduites à partir des ions de la série y.

Nous essayons ici d'expliquer cette différence en nous référant aux voies de fragmentations peptidiques et au fonctionnement propre de la trappe ionique et du Q-TOF en MS/MS.

Parmi les travaux tentant de rationaliser les chemins de fragmentations sur le squelette peptidique, le « modèle du proton mobile » décrit par Dongré [Dongré et coll., 1996] semble offrir une approche cohérente des processus impliqués dans les collisions à basse énergie :

- | En l'absence d'un acide aminé fortement basique, le proton qui ionise le peptide n'est pas localisé sur un site spécifique. Il se situe généralement sur un site basique (par ex. azote du groupement N-terminal, hétéroatomes des chaînes latérales des acides aminés). Quelque soit le site de protonation initial du peptide considéré, par activation énergétique, le proton se délocalise en phase gazeuse et migre sur des sites moins basiques du peptide. Le transfert du proton sur les azotes des liaisons amides fragilise les liaisons peptidiques et favorise ainsi leurs fragmentations.
- | En présence d'acides aminés fortement basiques (Arginine, Histidine, Lysine, par ordre de basicité en phase gazeuse décroissante), le proton est d'autant plus localisé que la basicité du site en phase gazeuse est élevée. La dissociation d'un peptide contenant un acide aminé basique protonné sera plus énergétique que celle d'un peptide n'en contenant pas puisqu'elle nécessite de rendre mobile le proton « séquestré », capable d'induire la fragmentation des liaisons amides.

Les peptides trypsiques contiennent (au moins) un acide aminé basique en position C-terminale (Lysine ou Arginine). Les ions y générés par fragmentation d'un peptide trypsique emportent donc l'extrémité C-terminale contenant un acide aminé basique, site préférentiel de protonation. Les fragments trypsiques b ne contiennent, eux, généralement pas d'acide aminé basique ; la charge de ces ions n'est donc pas spécifiquement localisée. Aussi, si l'on considère le modèle du proton mobile précédemment décrit, les fragments trypsiques b sont plus fragiles que les fragments y protonnés en position C-terminale.

Considérons maintenant les différences entre les deux instruments utilisés pour fragmenter les peptides trypsiques en MS/MS :

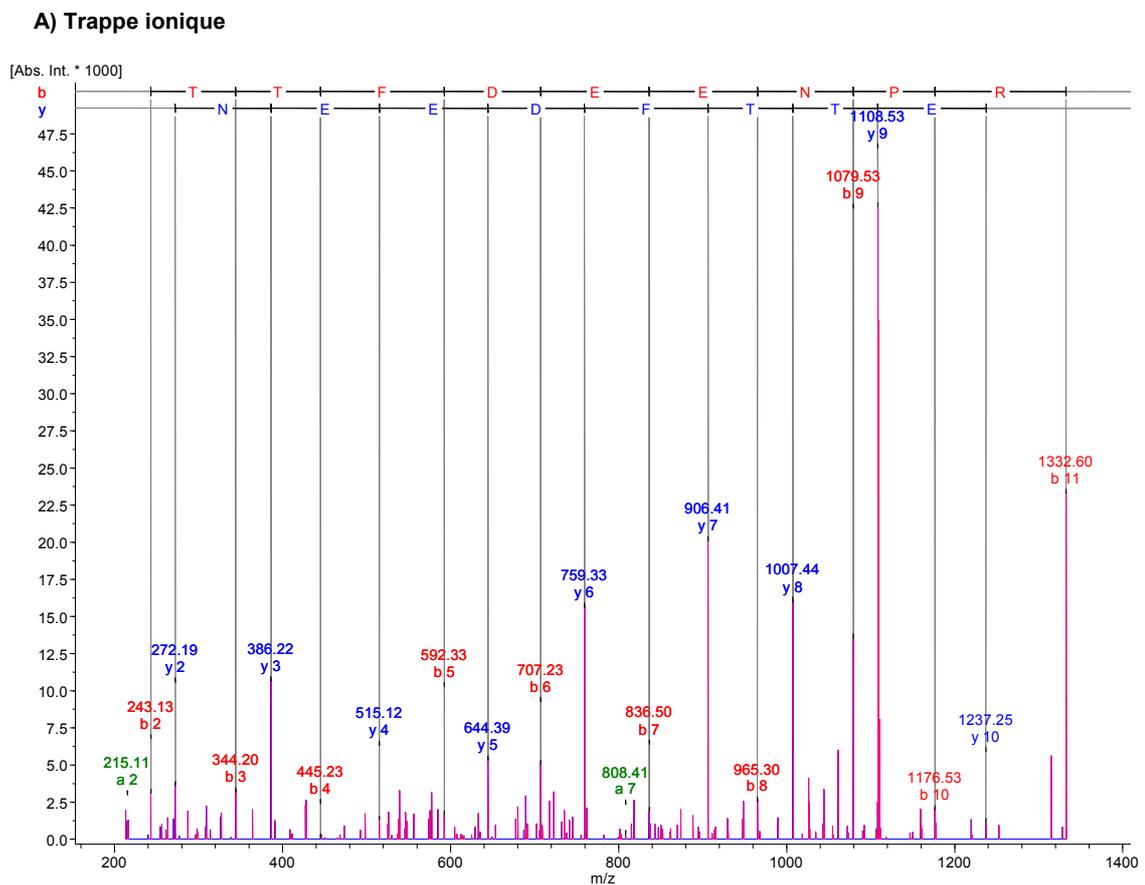
- | La MS/MS avec une trappe ionique permet une transmission exclusive de l'énergie de fragmentation aux ions précurseurs par application d'une fréquence de résonance qui leur est propre. Au fur à mesure des collisions avec les molécules d'hélium résiduelles, l'énergie cinétique de l'ion précurseur est progressivement transformée en énergie interne jusqu'à obtenir sa dissociation. Théoriquement, les ions fils issus d'une telle fragmentation ne peuvent pas se fragmenter simultanément en ions petits-fils.
- | Contrairement à la trappe ionique, la MS/MS effectuée dans une cellule de collision avec un spectromètre de masse hybride comme le Q-TOF peut générer des fragments multiples. En effet, l'énergie de fragmentation délivrée est transmise à l'ensemble des ions (parents et fragments) contenus dans la cellule de collision par application d'une différence de potentiel.

Par ailleurs, les collisions avec les molécules d'argon sont plus efficaces qu'avec de l'Hélium ; la fragmentation avec le Q-TOF est donc un processus plus rapide que dans la trappe. Ainsi, certains ions fils peuvent se re-fragmenter pour produire des ions petits fils (de 3^{ème} génération), et voire même de génération supérieure (ions immoniums).

Par conséquent, il nous semble probable que les ions b issus de la fragmentation des peptides tryptiques par MS/MS avec le Q-TOF soient re-fragmentés au cours de ce même processus en ions de plus faibles m/z comme, par exemple, des ions b de faibles masses ou des ions immoniums.

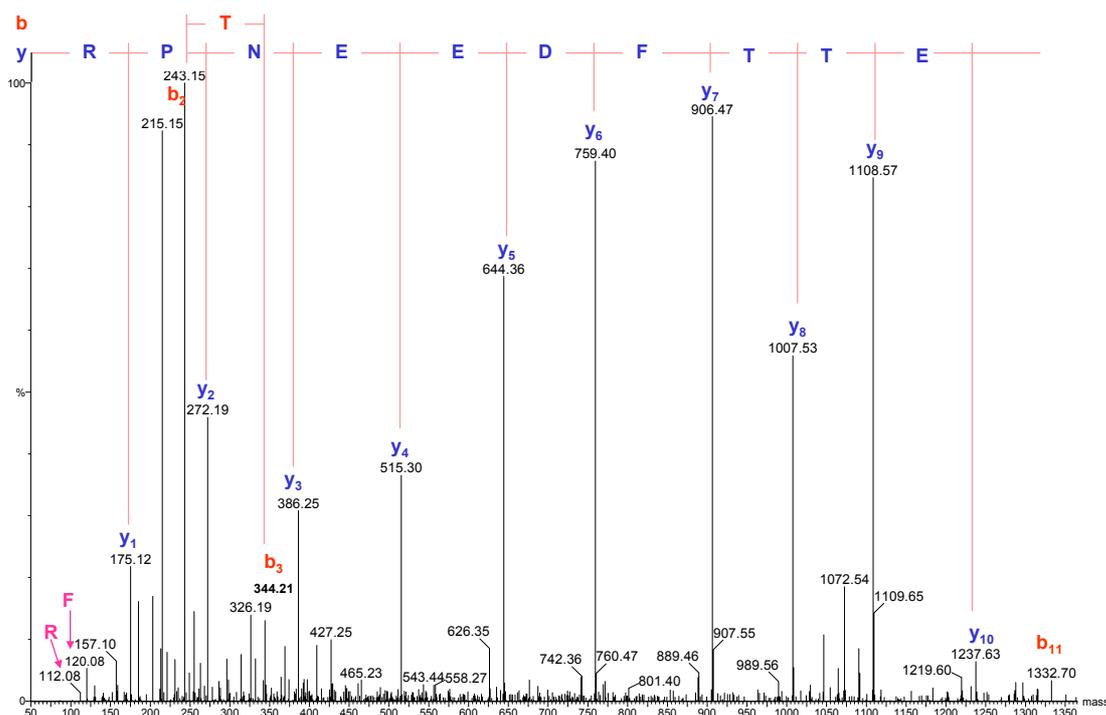
Les fragments tryptiques y, stabilisés par le proton porté par l'acide aminé basique en position C-terminale, seraient plus stables et donc majoritairement observés avec le Q-TOF.

Figure 5 : spectres MS/MS déconvolués du peptide 675,8 (2+) avec la trappe ionique (A) et le Q-TOF (B). Résultats issus de l'analyse LC-MS/MS d'un spot de bromoperoxydase. La séquence *de novo* est homologue à



une séquence peptidique obtenue à partir de l'ADNc de *L. digitata* : LETTFDEENPR.

B) Q-TOF



Finalement, la complémentarité des résultats obtenus avec la trappe ionique et le Q-TOF s'avère intéressante, voire indispensable, pour l'approche par séquençage *de novo*. Dans certains cas, elle peut permettre de lever les ambiguïtés soulevées, lors de l'interprétation des spectres, avec l'une ou l'autre de ces techniques.

III. RESULTATS

Nos travaux sont présentés dans les publications jointes à la fin de ce chapitre de thèse. Nous souhaitons ici rappeler les principaux résultats obtenus et détailler quelques résultats complémentaires.

1. Les bromoperoxydases

Trois « spots » protéiques (b, c, d) correspondant à une activité bromoperoxydase sur le gel natif ont été digérés à la trypsine et analysés par nanoLC-MS/MS. Nos résultats montrent que les trois bandes analysées sont des protéines homologues ou très similaires.

Les peptides caractérisés lors de l'approche par séquençage *de novo* des spots de bromoperoxydases sont, pour la majeure partie d'entre eux, identiques à ceux issus de la digestion trypsique *in silico* des séquences protéiques de vBPO1 et vBPO2 de *L. digitata* (cf. Table II, publication n°1).

Afin de confirmer plus largement cette homologie de séquence, nous avons utilisé le moteur de recherche MASCOT pour confronter les résultats de LC-MS/MS obtenus pour les bromoperoxydases avec les séquences protéiques de vBPO1 et vBPO2 (cf. Le couplage nanoLC-MS/MS avec une trappe ionique § IV.3.). Les séquences de vBPO1 et vBPO2 ont préalablement été introduites dans les bases de données du laboratoire : elles définissent à elles seules une banque. Nous avons donc interrogé cette banque à partir de nos résultats d'analyse et obtenu des informations complémentaires au

séquençage *de novo* : certains peptides ont été nouvellement identifiés, soit seulement sur la base de leur m/z, soit sur la base de leur m/z combiné à des informations de séquences. Les résultats obtenus sont présentés dans la figure 6 et le tableau 2.

Les peptides identifiés par le biais de l'interrogation avec MASCOT permettent de couvrir 40% de la séquence d'ADNc de vBPO1 (contre 22% avec les résultats initiaux). Par conséquent, ces résultats complémentaires confirment et renforcent nos conclusions : les gènes vBPO1 et vBPO2 clonés à partir des sporophytes de *L. digitata* codent pour des activités bromoperoxydases purifiées.

Tableau 2 : résultats obtenus avec le moteur de recherche MASCOT. Homologie entre les bromoperoxydases étudiées (d'après nanoLC-MS/MS avec le Q-TOF) et la séquence protéique vBPO1. Les peptides représentés en caractères normaux ont été identifiés sur la base de leur m/z. Ceux représentés en gras ont donné des informations de séquences significatives en MS/MS. Les peptides préalablement séquencés *de novo* sont en gris. CAM : carbamidométhylation des Cystéines ; Mox : Méthionine oxydée.

Résidus tryptiques vBPO1	MH ⁺ expérimentaux	MH ⁺ théoriques	Précision (ppm)	Séquences
284-291	872.58	872.51	80	DLATIALR
443-452	1083.51	1083.54	28	AFLGYEAGQK
191-201	1127.68	1127.61	62	GLNIPRSEGGK
292-300	1142.62	1142.54	70	DQLYTEAFR
79-88	1171.60	1171.52	68	FGQVDEDAYK
476-485	1176.62	1176.55	59	EIVGECVNEK (+CAM)
21-31	1273.70	1273.63	55	DSLDFLDPVPR
376-387	1338.84	1338.76	60	DVITPLPDSILR
267-277	1350.68	1350.61	52	LETTFDEENPR
535-546	1350.82	1350.73	67	LQQELPGLPEAR
178-190	1371.83	1371.76	51	LAAENLQGLPAFK
571-581	1381.74	1381.78	29	ILELYGRNLYK
488-500	1464.84	1464.75	61	LVEGLTYEGELNK
534-546	1506.96	1506.83	86	RLQQELPGLPEAR
19-31	1528.91	1528.80	72	VRDSLDFLDPVPR
197-210	1533.88	1533.78	65	SEGGKIDPVTDLFR
453-466	1563.81	1563.70	70	CFPNPVVSNDEGTK (+CAM)
453-467	1719.91	1719.80	64	CFPNPVVSNDEGTKR (+CAM)
90-104	1742.88	1742.79	52	LLECVFTSDINECEK
149-163	1779.03	1778.92	62	LAAQQAQEVYWMALLR (+Mox)
90-104	1856.96	1856.83	70	LLECVFTSDINECEK (+CAM)
359-375	2004.13	2004.01	60	MLRPEAYGALVHNTLMR (+Mox)
90-111	2481.31	2481.17	56	LLECVFTSDINECEKVPSGAGR (+CAM)
117-148	3326.81	3336.64	51	LTNPLGGTAHQVTGADSDNVFITTPDSLLSER

Figure 6 : couverture de séquence de vBPO1 (40%) d'après les résultats présentés dans le tableau 2.

```

1   YEEPPEPTQP  LLSGNVCRVR  DSLDFLDPVP  RAKVTLLKRL  AIAKDEISVG
51  PTCHINNGDE  ENVPLFAGQY  HKTLPHDKFG  QVDEDAYKKL  LECVFTSDIN
101 ECEKVPAG  RRGGAKLTNP  LGGTAHQVTG  ADSDNVFITT  PDSLLSERLA
151 AQQAEVYWMA  LLRDIPFGEF  AKNDYVRLAA  ENLQGLPAFK  GLNIPRSEGG
201 KIDPVTDLFR  TTWPGVTTGP  VVSQFMLSDF  LIDSIKVTPK  ADPLTPGVYD
251 MTAFQPWLDV  QNGASKLETT  FDEENPRFIR  NGRDLATIAL  RDQLYTEAFR
301 AALILFTEGA  LGGEVGPYAE  AERQQGFATF  GEPHILTAMA  SASSVTRHAW
351 YAKWQVHRML  RPEAYGALVH  NTLMRDVITP  LPDSILRNTE  LLNRVEVHNQ
401 RMNPDGEKTF  LLPMAAAQGS  PTHPAYPSGH  AINNGAYITA  LKAFLGYEAG
451 QKCFPNPVVS  NDEGTRIKY  KPSGREIVGE  CVNEKGKLV  GLTYEGELNK
501 ISANVLLGRS  HIGVHWRMDG  VYGALMGETS  CVRRLQOELP  GLPEAREVEG
551 KRRRGDIPPA  TYKFRLYSGK  IILELYGRNLY  KLDGKLCEGA  FTGDDFCDPI
601 DEDDYSSFDD  IVEEHAQFSL  HGHTEL

```

2. Les iodoperoxydases

Contrairement au cas des bromoperoxydases nous ne disposons d'aucun ADNc préalablement cloné. Aussi notre approche, plus ciblée, se décompose en deux étapes :

- | Dans un premier temps, le séquençage *de novo* à partir des données issues de l'analyse LC-MS/MS du spot protéique d'iodoperoxydase nous a permis d'obtenir quelques informations de séquences suffisamment longues et spécifiques pour permettre à l'équipe de Roscoff de synthétiser des « primers dégénérés » puis d'obtenir une sonde nucléotidique par amplification PCR. Grâce à cette sonde, deux ADNc d'iodoperoxydases (vIPO1 et vIPO2) ont été nouvellement clonés et séquencés. **C'est la première fois qu'une séquence entière d'iodoperoxydase est caractérisée.**
- | Ensuite, en utilisant une approche identique à celle décrite précédemment, nous avons exploité plus finement nos données et vérifié avec MASCOT l'homologie de la séquence protéique des vIPO avec les peptides analysés par spectrométrie de masse (cf. table II, publication n°2).

Nous avons ainsi mis en évidence un mélange d'au moins deux formes protéiques dont les séquences diffèrent par quelques acides aminés seulement.

IV. CONCLUSION ET PERSPECTIVES

L'utilisation de la nanoLC-MS/MS s'est avérée déterminante pour la caractérisation des deux types d'haloperoxydases purifiées chez les Laminaires et ayant des activités distinctes, bromo et iodoperoxydases.

- | Nous disposons d'informations préalables concernant les bromoperoxydases (ADNc et données protéiques pour une activité similaire chez d'autres familles d'algues). Nous avons donc recherché des homologies de séquence avec ces données pré-existantes.
- | Par contre, ne disposant d'aucune information concernant les iodoperoxydases, nous avons développé une approche plus ciblée consistant, dans un premier temps, à obtenir par

séquençage *de novo* des séquences suffisamment longues et spécifiques pour permettre la construction de « primers dégénérés » afin de cloner des ADNc correspondant à l'activité iodoperoxydase. Dans un second temps, nous avons vérifié l'homologie de séquence entre l'ADNc cloné et la protéine étudiée sur le gel.

La mise en œuvre d'une stratégie par séquençage *de novo* a permis, d'une part, d'attribuer une fonction biochimique aux deux ADNc de bromoperoxydases déjà caractérisés et, d'autre part, de cloner de nouveaux ADNc correspondant à l'activité iodoperoxydase. Nous avons ainsi pu établir un lien entre les activités enzymatiques des protéines étudiées et les données moléculaires obtenues à partir des sporophytes de *L. digitata*. Les gènes identifiés pourront ensuite être sur-exprimés afin d'approfondir la caractérisation biochimique des haloperoxydases permettant ainsi de progresser dans la compréhension de leurs fonctions biologiques.

La modélisation informatique de la structure tri-dimensionnelle des protéines de bromo et d'iodoperoxydase et leur comparaison avec des bromoperoxydases présentes chez d'autres algues ou une chloroperoxydase de champignon a déjà permis d'émettre des hypothèses quant à la spécificité des haloperoxydases pour les halogénures.

Par ailleurs, les bromoperoxydases semblent être constituées de dimères qui s'associent spontanément en solution en oligomères de plus haute masse. Il semblerait probable également, que des modifications post-traductionnelles sur la chaîne protéique joue un rôle dans l'état d'agrégation de ces molécules. Des analyses complémentaires de spectrométrie de masse supramoléculaire, en conditions natives, permettraient de caractériser plus finement ces édifices non covalents.

V. PRESENTATION DE CE TRAVAIL

Ce travail a fait l'objet de deux articles joints ci-après :

- | Publication n°1 : parue dans The Journal of Biological Chemistry : « The brown algal kelp *Laminaria Digitata* features distinct bromoperoxidase and iodoperoxidase activities ».
- | Publication n°2 (version provisoire) : « Vanadium-dependent iodoperoxidase in the brown algal kelp *Laminaria Digitata* ».

Nous avons également présenté ces résultats, sous forme de communication par affiche, au 3^{ème} congrès européen de phycologie qui s'est déroulé à Belfast (Irlande) en Aout 2003 (deuxième prix du meilleur poster).

PUBLI 1

PUBLI 2

PUBLI 3

PUBLI 4

PUBLI 5

PUBLI 6

PUBLI 7

PUBLI 8

Vanadium-dependent iodoperoxidase in the brown algal kelp

Laminaria digitata

Carole COLIN¹, Catherine LEBLANC^{1*}, Gurvan Michel¹, Elsa WAGNER², Emmanuelle LEIZE-WAGNER², Alain Van DORSSELAER², Philippe POTIN¹

¹ UMR 7139, CNRS-Laboratoires Goëmar-UPMC, Station Biologique, BP 74, F-29682 Roscoff Cedex, France.

² Laboratoire de Spectrométrie de Masse Bio-Organique, UMR 7509, CNRS-ULP, 25 rue Becquerel, F-67087 Strasbourg Cedex 2, France

* To whom correspondence should be addressed: Catherine Leblanc, Tel.: 33-2-98292332; Fax: 33-2-98292324; E-mail: leblanc@sb-roscoff.fr.

INTRODUCTION

Haloperoxidases (HPOs) catalyze, in the presence of hydrogen peroxide, the oxidation of halides (X^- : iodide, bromide or chloride) to their corresponding hypohalous acids or a related electron oxidized halogenating intermediate such as OX^- , X_3^- and X^+ . A variety of halogenated compounds can subsequently be generated if the appropriate nucleophilic acceptors are present (1). Myeloperoxidase, a bactericidal enzyme secreted by activated human phagocytes, specifically catalyses the production of hypohalous acid from chloride and bromide utilizing hydrogen peroxide released from oxidative burst (2). In fungus, hypohalous acids has been identified as strong bactericidal agents (3), while halogenated organic compounds are suggested to play a defensive role in marine organisms preventing bacterial fouling and biological surface colonization in natural habitats (4). In the marine environment, HPOs are thought to function in the biosynthesis of the numerous halogenated natural products, which range from relatively simple volatile halogenated hydrocarbons (5) to halogenated phenols (6), terpenes and indoles (7,8). In many cases the halogenated compounds are of pharmacological interest due to their biological activities (9-11).

HPOs are classified as chloro- (CPO), bromo- (BPO) and iodoperoxidase (IPO) and this nomenclature is based on the most electronegative halide which can be oxidized by hydrogen peroxide. On the basis of the cofactor requirement, three major classes of HPOs are known: HPOs containing no prosthetic group, the heme-containing HPOs and the vanadium-containing HPOs (v-HPOs) that bind a vanadate ion (HVO_4^{2-}) as a prosthetic group (9). Enzymes representing these two last-mentioned classes differ in at least two aspects: catalytic mechanism and stability. Heme-containing HPOs catalyze the formation of hypohalous acid by a redox mechanism whereas in v-HPOs the vanadate group does not change its redox state but may function as a Lewis acid in a substrate-inhibited bi-bi ping-pong mechanism (12,13). In contrary with heme-containing HPOs, v-HPOs possess a high level of thermostability as well as of tolerance to organic solvents and to high concentration of their substrates and products (14). These properties make v-HPOs good candidates as industrial catalysis in biotransformation (15,16).

Among v-HPOs, v-CPO activities have been characterized in terrestrial fungi (17,18), v-BPOs have been mainly described in marine algae (14), but also identified from lichens (19) and a v-IPO activity has been recently purified from the brown alga *Laminaria digitata* (20). However, at the molecular and structural levels, studies have essentially been focused on *Curvularia inaequalis* v-CPO, on the brown alga *Ascophyllum nodosum* v-BPO and the red alga *Corallina sp.* v-BPO, which native X-ray crystallographic structures have been solved (21). Moreover, kinetic and structural studies of peroxide-bound forms and mutants from v-CPO have been developed, as well as mutants from *Corallina pilulifera* v-BPO (21-23). All these studies have shown the importance

of several residues involved in hydrogen and covalent bond interactions with vanadate in the active site. The vanadium resides in a trigonal bipyramidal coordination geometry by four oxygen atoms and N^{ε2} of an histidine residue of the protein. Another histidine residue may participate as a proton donor/acceptor during the enzymatic reaction (21). The mechanism proposed by (24) for the v-CPO halogenation reaction and developed by (25) have been reinforced by structural studies on algal v-BPOs (21,26). Upon coordination of peroxide, a peroxovanadium intermediate of v-HPO may be formed during catalysis prior oxidation of the halide. (27) reported the regioselective halogen-dependent oxidation of indoles catalyzed by v-BPO. The halide binding site at the vanadium center is less probable (28). Recently, (22) demonstrated in v-CPO that neither the phenylalanine nor the serine residues are directly involved in chlorination or bromination in contrary with the hypothesis formulated by (29) and (30), respectively. Thus the origin of the halide selectivity remains one of the interesting, yet unanswered question.

In this context, in an effort to elucidate the difference of specificity of the v-HPOs toward the halides, we have undertaken the characterization of the v-IPO from the brown alga *L. digitata*. In this report, we described for the first time a v-IPO full-length cDNA and analyzed its three-dimensional structure model as well as its active site by comparison with others v-HPOs. We show that in v-IPO the presence of specific residues may be one structural explanation for the observed differences in iodine specificity and phylogenetic relationships among v-HPO enzymes are discussed.

EXPERIMENTAL PROCEDURES

Plant Material-

L. digitata sporophytes were collected from the shore in the vicinity of Roscoff (Brittany, France), and frozen in liquid nitrogen for DNA, RNA or protein extractions.

v-HPO Extraction and Enzyme Activity assays-

v-HPOs from *L. digitata* sporophytes were extracted and purified from adult thalli as described in (20). Haloperoxidase activities were detected on non-denaturing gels in a staining solution consisting of 100 mM potassium phosphate buffer, pH 7.4, in the presence of 0.1 mM o-dianisidine, 10 mM potassium iodide and 0.45 mM H₂O₂. Iodoperoxidase specific activities were measured as described in (20). Steady state kinetic reactions were performed in the presence of different iodide, bromide and chloride concentrations. Sodium sulfate was used to maintain a constant ionic strength of 550 mM in the final solution.

Mass Spectrometry Analyses-

The strategy previously described by (20) was used for the *L. digitata* v-IPO purification. To remove completely polyethylene glycol, purified protein sample was applied to a DEAE-Sepharose anionic exchange column (Amersham Pharmacia) equilibrated with 50 mM NaCl, 50 mM Tris-HCl, pH 9.0. Protein was eluted by increasing steps of 50, 100, 500 mM, 1 and 2 M of NaCl in 50 mM Tris-HCl, pH 9.0. The active fraction (see below for haloperoxidase activity assay) was concentrated by filtration (YM-30, Millipore, Saint-Quentin, France) and equilibrated at 10 mM aqueous ammonium acetate, pH 6.5.

cDNA Isolation, Cloning and Sequencing-

The *L. digitata* sporophytes cDNA λZAP II library (31) was used as a template polymerase chain reaction (PCR) to amplify a v-IPO cDNA fragment using two degenerated primers (forward; 5'-GACACSGTBAACACSGAGGCBTAT-3' and reverse; 5'-GGCVACGTTVACVGC-3')

designed from LC-MS/MS data (20). A 600-pb PCR fragment was cloned in the pCR 2.1-TOPO vector (Invitrogen) and sequenced on both strands using the ABI prism 3100 genetic analyser (Applied Biosystems). The *L. digitata* sporophytes cDNA library was used a second time to obtain a 5' end fragment of the cDNA using a cDNA λZAP II specific primer (forward) and a specific primer from the cloned cDNA fragment (reverse; 5'-ACGCCGTAGTTGACGAAGC-3'). The 220-pb PCR fragment obtained was cloned in the pCR 2.1-TOPO vector and labelled with [α -³²P] dCTP using the Megaprime labelling kit (Amersham Biosciences) to screen the *L. digitata* sporophytes cDNA library. After *in vivo* excision (Stratagene), cDNA clones were sequenced on both strands.

The 5' end cDNA was obtained using the 5' rapid amplification of cDNA ends (RACE) system. Total RNA was extracted from *L. digitata* sporophytes according to (32). First-strand cDNA synthesis **and RACE PCRs were performed** recommendations using the SMART RACE cDNA amplification kit (Clontech) and v-IPO following specific primers (primerRACE1, 5'-CTTGTGGAACGTGGCGGCGGATGA-3' or primerRACE2, 5'-TCGGAGTTGATCGCTGAGGCCGGA-3') according to the manufacturer's. The RACE products were purified using a QIAquick gel extraction kit (QIAGEN), cloned into the pCR 2.1-TOPO vector (Invitrogen) and fully sequenced.

Primary Sequence, Structural Analysis and Phylogenetic Analysis-

The v-IPO sequences were obtained from trace data using Sequencher™ program version 4.0 (Gene Codes Corporation) and analysed by DNAMAN program version 4.15. Peptides signal were predicted by different programs, SignalP, PSORT and TargetP, and *in silico* digestion by trypsin was performed using the PeptideCutter program (programs available on the www.expasy.org server). The amino acid sequences of v-CPO, v-BPO and v-IPO were aligned first using MAFFT (available on the <http://bioinformatics.uams.edu/mafft/>, (33)) and then manually using the GENEDOC software (free access on www.psc.edu/biomed/genedoc). The three-dimensional structures of *C. inaequalis* v-CPO (NCBI Accession # 1VNC) and *A. nodosum* v-BPO (NCBI Accession # 1QI9) were used as a guide to ensure that homologous regions were compared. Protein sequence alignments were displayed using ESPript (34).

For phylogenetic analyses, when all the sequences were analyzed, only the most conserved portion of the alignment was taken into account, in the region of the active site (see FIG. 4), leading to a total of 136 amino acids used for distance-based methods and of 103 amino acids (89 amino acid informative positions) for parsimony analyses. When excluding v-CPO sequences, 541 amino acid sites were used in distance-based methods and 492 amino acids (373 informative) for parsimony analyses. Distance matrix, maximum-parsimony methods (35) were applied using the Phylo-Win software package (36). The neighbor-joining method (37) was performed with PAM distance corrected for multiple substitutions according to Dayhoff's matrix and bootstrap analyses were used to provide confidence estimates for the phylogenetic tree topologies (38).

Three-dimensional v-IPO model was generated by the swiss-model homology-modelling server using the structure of *A. nodosum* v-BPO as template (NCBI accession number # 1QI9). The structure was analysed using Swiss-Pdb Viewer application (free access on the www.expasy.org server).

Southern and Northern Blot Analysis-

Southern analysis was performed as already described (20). For Northern analyses, total RNA (10 μ g/lane) was separated by agarose gel electrophoresis, blotted onto nylon filters and hybridized with ³²P-labelled random-primed probes prepared according to the manufacturer's recommendations (Amersham). Northern hybridizations were performed at 42°C overnight with 50% formamide, 7% SDS, 250 mM NaCl, 120 mM Na(PO₄) pH 7.2. Membranes were washed for 10 min at 42°C successively in 2x SSC, 0.1% SDS and in 1x SSC, 0.1% SDS. A 531-bp 3' end specific probe of v-BPO1 was prepared by PCR (AAATGGTTGAGACGCCTTGG forward primer;

GCCTTACGCCACGCTTATAA reverse primer) from full-length cDNA (20). Two different v-IPO DNA probes were obtained by PCR using the following primers : a 416-pb 5' end v-IPO specific probe (in FIG. 2 referred to as # 1, CCAGAGGAGGACACCATACAC forward, GTTAAGGTGGAGCGTGCCT reverse) and a 773-bp v-IPO intragenic probe (# 2, CCGGAGCAGACTACATGAC forward, GTTGACGGCCAGCTTGTT reverse).

RESULTS

Mass Spectrometry Analysis of purified v-IPO Activity from L. digitata-

Mass spectrometry analyses were carried out for the purified iodoperoxidase activity (20). Among peptides sequenced by LC-MS/MS, eleven presented sequence homologies with *A. nodosum* and *F. distichus* v-BPOs (TABLE I). Moreover, three pairs of v-IPO peptides with significantly distinct masses were homologous to each other (Pept3 and 4; Pept6 and 7; Pept10 and 11; TABLE I).

Specific activity and Competition Studies on v-IPO Specific Activities-

The IPO specific activity showed typical Michaelis-Menten kinetics as function of iodide in the initial part of the curve, but over 20 mM, the IPO specific activity decreased rapidly (FIG. 1A). At fixed concentration of iodide (6.1 mM), the IPO specific activity decreased over a concentration of 50 mM bromide, whereas it was not affected by the presence of chloride in the same range of concentrations (FIG. 1B). The Lineweaver-Burke analysis at various fixed levels of bromide showed that Vmax of the IPO specific activity was independent of bromide whereas Km for iodide increased with the bromide concentration (FIG. 1C).

Molecular Characterization and Transcription of L. digitata v-IPO gene-

After screening of the *L. digitata* sporophyte cDNA library by a 5' end specific v-IPO probe, nine positive clones were mapped by restriction analyses and sequenced in both directions. As they corresponded to the same cDNA, the longest one was referred to as v-IPO (3733 bp). v-IPO was a full-length cDNA with a short 5'UTR (137 bp) and a large 3' UTR (1724 bp) (FIG. 2A). Clones obtained from 5' **RACE analyses** did not more extend its 5' end. v-IPO encoded a putative protein of 624 amino acids. No putative signal peptide was predicted and the molecular mass of the mature v-IPO calculated from the amino acid sequence was 66,326 Da.

Fourteen peptide masses characterized by LC-MS on the purified IPO activity corresponded with a good accuracy to those identified by *in silico* tryptic digestion of v-IPO protein sequence (TABLE II) and they homogeneously covered 34% of the v-IPO protein. Nine of these peptides, partially or fully sequenced by LC-MS/MS, were identical to the corresponding peptides of v-IPO and seven of them were first identified based on their homologies with other v-BPO sequences (see TABLE I).

The expression of iodoperoxidase gene was examined by Northern-blot analysis in sporophytes of *L. digitata*. Based on hybridization with intragenic probe of v-IPO a major transcript around 4 kb and a minor one at 2.5 kb were observed, whereas, by comparison, only one strong band at 3.5 kb was detected for *L. digitata* v-BPO transcripts. **With** a 5' end v-IPO specific probe, a single lower signal appeared at 4 kb (FIG. 2B).

In a Southern blot analysis with v-IPO intragenic probe (FIG. 3), several hybridization bands were detected in *L. digitata*, but presented less bands than those hybridized by *L. digitata* v-BPO intragenic probe (20). Moreover, the two patterns were significantly different.

Comparisons and Phylogenetic Analyses of v-HPOs-

The deduced amino acid sequence of *L. digitata* v-IPO was aligned with available protein sequences of v-HPOs (FIG 4). It shared 37.1 % and 27.2 % identities with *L. digitata* v-BPO1 and *C. officinalis* v-BPO, respectively, while it was closer from *A. nodosum* v-BPO and *F. distichus* v-BPO with 49.4 % and 47.0 % sequence identities respectively. As the overall sequence homologies between v-CPO of fungi and others v-HPO from brown and red algae were very low, a phylogenetic tree was built by Neighbor-joining analysis based on comparisons of amino acid sequences around the conserved active site region (FIG. 5). A tree based on parsimony presented the same topology and the corresponding bootstrap values are thus also included in the tree. The exclusion of fungal proteins from the phylogenetic analyses resulted in identical topology of the neighbor-joining and parsimony subtrees, with higher bootstrap values (data not shown). Fungal v-CPOs formed a monophyletic cluster. Interestingly, an hypothetical protein predicted on the genome of ascomycetous fungi *Magnaporthe grisea* merged at the basis of the already identified fungal v-CPOs. Most distantly related, the algal v-HPOs formed another monophyletic cluster, subdivided into two different subgroups, the brown and the red algal proteins. In the *Corallina sp.* cluster, v-BPO1 and v-BPO2 from *C. pilulifera* did not emerge together, but branched each other with their orthologues in *C. officinalis*. As expected from sequence comparisons, *L. digitata* v-IPO is more closely related to v-BPO from fuclean algae than to v-BPO from the same species and this branching is also well supported by bootstrap.

In fact, the overall secondary structure organization was highly conserved between *A. nodosum* v-BPO and *L. digitata* v-IPO. v-IPO was dominated by helical secondary structure elements and presented the same α -helices and β -strands arrangement than *A. nodosum* v-BPO monomer (FIG. 4). The six cysteine residues involved in intra-molecular disulphide bridges within the *A. nodosum* v-BPO monomer (26) were present at similar position in v-IPO. In the N-terminal helice of v-IPO, the two cysteine residues, involved in dimer interface v-BPO, were not observed. Almost all residues involved in active site and vanadate binding sites were also conserved (FIG. 4 and see FIG. 6).

Three-dimensional Structure of *L. digitata* v-IPO-

The overall folds from the regions of the active site of the crystallized *A. nodosum* v-BPO (26), *C. officinalis* v-BPO (39), *C. inaequalis* v-CPO (29) and the modellized *L. digitata* v-IPO were surimposable (FIG. 6). In the v-IPO protein structure, the vanadium site residues were at the top of the four-helix bundles in a broad channel that was lined on one half with predominantly polar residues, including Arg549^{vIPO} and His483^{vIPO}. The other half of the channel presented hydrophobic residues : Trp400^{vIPO}, Pro474^{vIPO}, His476^{vIPO} and Thr475^{vIPO} (FIG. 6). The putative His483^{vIPO} catalytic residue formed a strong putative hydrogen bond (2,9Å) to Asp333^{vIPO} residue, like in the v-BPOs structure, while in v-CPO this catalytic residue was hydrogen bonded to Trp289^{vCPO} and Ala290^{vCPO} residues. In the same way, v-IPO like the v-BPOs showed the additional histidine residue near the active site (CoHis478^{vBPO}, AnHis411^{vBPO} and His476^{vIPO}) (FIG. 5) while in v-CPO, Phe397 was localized at this position. Among residues involved in hydrogen bound to the vanadate cofactor, the main difference between v-IPO, v-BPO and v-CPO concerned a serine residue, which is involved within the strongest hydrogen bond to the vanadate cofactor in v-CPO and in v-BPO, and which is replaced by an alanine residue in v-IPO (FIG.6).

Other differences appeared when the structures of active site funnel entrances are compared between the homodimeric crystallized *A. nodosum* v-BPO and a modellized homodimer *L. digitata* v-IPO (FIG. 7). As established for the structure of *A. nodosum* v-BPO (26), in the homodimer v-IPO the residues of one subunit formed the bottom of the cleft and whereas the top was formed by the residues of the other subunit of the dimer. Two major residues differed between v-IPO and v-BPO (FIG. 7), Ala331^{vIPO} instead of AnPhe276^{vBPO} (bottom) and Ala357^{vIPO} instead of Arg299^{vBPO} (top), which leading to a narrower cavity funnel entrance in v-BPO than in v-IPO.

DISCUSSION

Characterization of the L. digitata v-IPO-

v-IPO purified from *L. digitata* yielded only one band on polyacrylamide gels under denaturing conditions with molecular weight around 80 kDa and gel filtration analysis supposed an dimeric structure of the enzyme (20). The LC-MS/MS results suggested that *L. digitata* v-IPO exists as a heterodimer of monomers which present quite identical molecular masses and link by no covalent bond. If we suggest posttranslational regulations, the molecular mass of the translated v-IPO cDNA was agreed with this of the purified v-IPO (20). LC-MS/MS data suggest the mature v-IPO protein sequence correspond to one of the v-IPO monomer.

Evolution of v-IPO Genes

Southern and northern analysis suggest that v-IPO is a large multigenic family but smaller than the v-BPO one (20). Although the *L. digitata* v-BPO and v-IPO diverged in their primary sequences, they were highly conserved in their overall folding. Primary sequence of the v-IPO presented high homologies with other brown algae v-BPO like *A. nodosum* or *F. distichus* v-BPO which confirm first data on LC-MS/MS peptides sequences (20). The phylogenetic analysis proposed that a gene duplication occurred before speciation. In *L. digitata*, two forms of haloperoxidases were present, v-BPO and v-IPO, which possessed a multicopy of genes. The driving force of evolution is gene copy.

Active Site Structure and Specificity for Halides-

Whereas specific activities are highly different toward iodide and bromide, three-dimensional structure comparison of the v-BPO and the v-IPO model didn't revealed major changes to the overall structure of the proteins. Looking the active site region, we can suggest that the presented model structure of v-IPO agrees with the proposed mechanisms of halogenation (25,26,29). All the amino acid residues involved in hydrogen-bonding contacts with vanadate are nearly conserved between v-CPO, v-BPO and v-IPO. Nevertheless, in v-IPO, a crucial difference can be observed in the region of the active site near the vanadium : a serine residue in v-CPO and v-BPOs, replaced by an alanine residue in v-IPO. While, in *A. nodosum* v-BPO, this residue was suggested as a possible candidate to form a carbon-bromide covalent bond (30), the solvent accessibility of this serine in *C. officinalis* v-BPO seems to be too low in the phosphate-bound form of the enzyme (39). Recently, (22) investigated the site-directed mutagenesis of this serine in *C. inaequalis* v-CPO. It resulted in tuning down the rate of halogenation of the enzyme but the mutant is still able to catalyse the oxidation of halides. Nevertheless, the mutant present a significant decrease of specificity toward chloride and at low pH the mutation apparently affects the binding of bromide. These recent data permits to throw back in the v-IPO the direct and single role of the alanine residue in the iodide specificity.

The presence and/or absence of some residues completely eliminated the ability of v-IPO to oxidize bromide. Two hypothesis can be proposed to explain this v-IPO specificity to iodide : (1) The enzyme recognizes discriminately halogens. The substrate specificity is changed by binding pocket modification and only iodide can be bound. (2) The enzyme doesn't recognize discriminately halogens. Halogen binding is still present for the different halogens but conformational changes involve the impossible oxidation of some halogens. The present results show that the reaction between native v-IPO and bromide (as an inhibitor) is competitive with the binding of iodide. In v-IPO, the iodide site can be occupied by both iodide and bromide that is in favor of the hypothesis (2).

The active site funnel entrance of the v-HPO, a region rich on hydrophobic patches and polar residues, could provide binding sites for the halogen ions. *A. nodosum* v-BPO present a cavity funnel entrance narrower than the v-IPO one. This is principally due to some alanine residues in the

v-IPO instead main residues in v-BPO. Halides are known to bind at hydrophobic sites in other enzymes such as in haloalkane dehalogenase (40), certain amylases (41), peroxidases (42) and myeloperoxidase (43).

In the active site region, the presence of an alanine residue in v-IPO instead the serine residue in v-CPO and v-BPO could induce changes in the charge neutralisation at the active site center and it would bring about a putative swing round of the vanadium. But this mutation doesn't seem to result directly to the specificity of the v-IPO for iodide but certainly to its oxidation. Moreover, the difference at the active site funnel entrance between the different v-HPO could affect the substrate specificity and stereoselectivity by inhibiting the oxidation of some halides during the reaction but not their fixations to the enzyme which seem to be the same for both bromide and iodide. So, in the v-IPO, the punctual mutation of certain residues and particularly the presence of some alanine residues instead major residues may be one explanation for the observed differences in iodide specificity between v-IPO and the other v-HPO. These results form the basis for future investigation on the analysis of recombinant and mutants of v-IPO and v-BPO (A481S^{vIPO} and AnS416A^{vBPO} / A357R^{vIPO} and AnR299A^{vBPO} / A331F^{vIPO} and AnF276A^{vBPO}) with substrates and reaction intermediates.

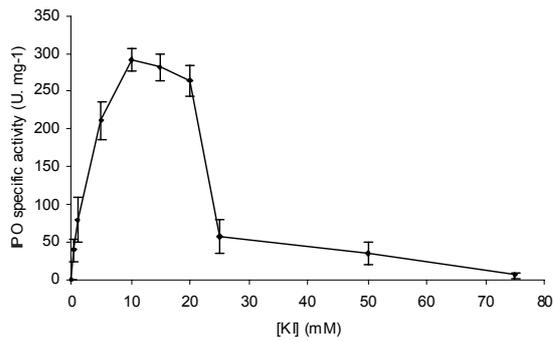
The difference in oxidative ability between the different v-HPO is not due to a single residue but apparently many factors are involved. In terms of specificity constant, no mutant with a higher K_{cat}/K_m values than that of the recombinant v-HPO has been found yet. In fact, all mutations and activity studies could show that differences in the catalytic site between the v-HPO seem to involve only the catalytic activity but not the specificity toward halides. The hydrogen bonding network in the active site and bound water molecules seem to play an important role in the tuning of the reactivity of these v-HPO but not directly on the halide specificity.

FIGURES

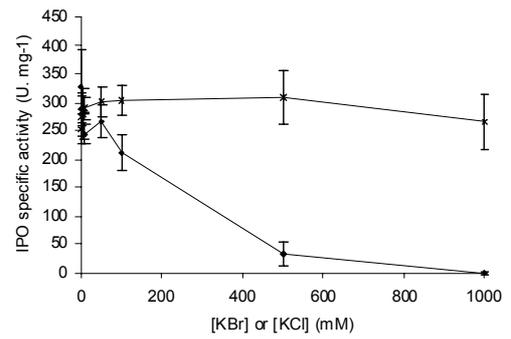
FIG. 1. Iodoperoxidase specific activity of the purified *L. digitata* v-IPO and effect of bromide or chloride.

(A) Iodoperoxidase specific activity of the purified v-IPO as a function of iodide or (B) as a function of bromide (◆-◆) or chloride (*-*) concentrations (the iodide concentration is fixed at 6.1 mM). (C) Kinetic analysis (Lineweaver-Burke) at various iodide concentration and 500 mM KBr (●-●), 100 mM KBr (○-○), 50 mM KBr (*-*), 10 mM KBr (□-□), 3mM KBr (Δ-Δ) or 0 mM KBr (◇-◇).

A



B



C

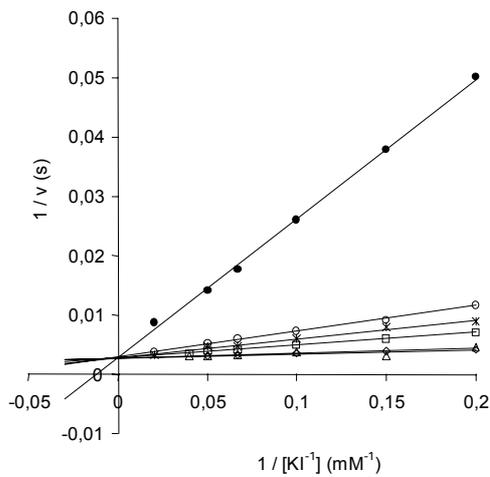


FIG. 2. Transcription analysis of v-IPO and v-BPO1 genes from *L. digitata*.

(A) Physical map of v-IPO cDNA showing Open Reading Frames (black arrows) and the position of the two different DNA probes (gray squares), used for hybridizations; probe #1 is specific of the 5' end region and probe #2 corresponds to conserved active site region of v-IPO. (B) Northern blot hybridizations of total RNA, with a 3'end v-BPO1 specific probe (first lane) and with the #1 and #2 v-IPO probes (see above).

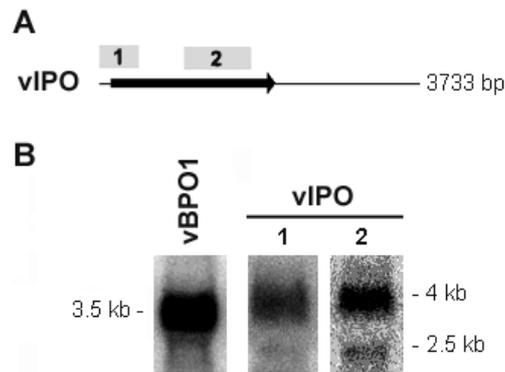


FIG. 3. Southern blot hybridization of *L. digitata* genomic DNA with an intragenic probe from v-IPO cDNA. *L. digitata* genomic DNA was digested with *SalI*, *SalI/XhoI*, *XhoI* (lanes 1 to 3 respectively).

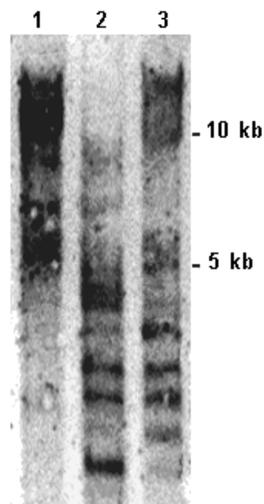


FIG. 4. Multiple amino acid sequence alignment of algal v-BPOs and v-IPO.

An-vBPO, v-BPO from *A. nodosum* (NCBI Accession # P81701); Ld-vBPO1, v-BPO1 from *L. digitata* (NCBI Accession # CAD37191); Ld-IPO, v-IPO from *L. digitata* (This study); Co-vBPO, v-BPO from *C. officinalis* (NCBI Accession # 1QHBA). Conserved residues in all four sequences are in white capitals on a black background, residues that occur in at least three sequences are in bold black capitals and surrounded by black boxes. Secondary structures (α -helices and β strands) of the An-vBPO and Co-vBPO are indicated on top and bottom lines, respectively, according to their resolved crystal structures (26,39). Below the alignment, the vanadate H-bonding residues are indicated with a *black triangle up*, the covalent link of the vanadate group with a *black solid circle*, the extra-histidine residue of the v-BPO class with a *black triangle down* and the catalytic histidine residue with a *star*. The portion of the alignment used for phylogenetic analyses is flanked by white triangles.

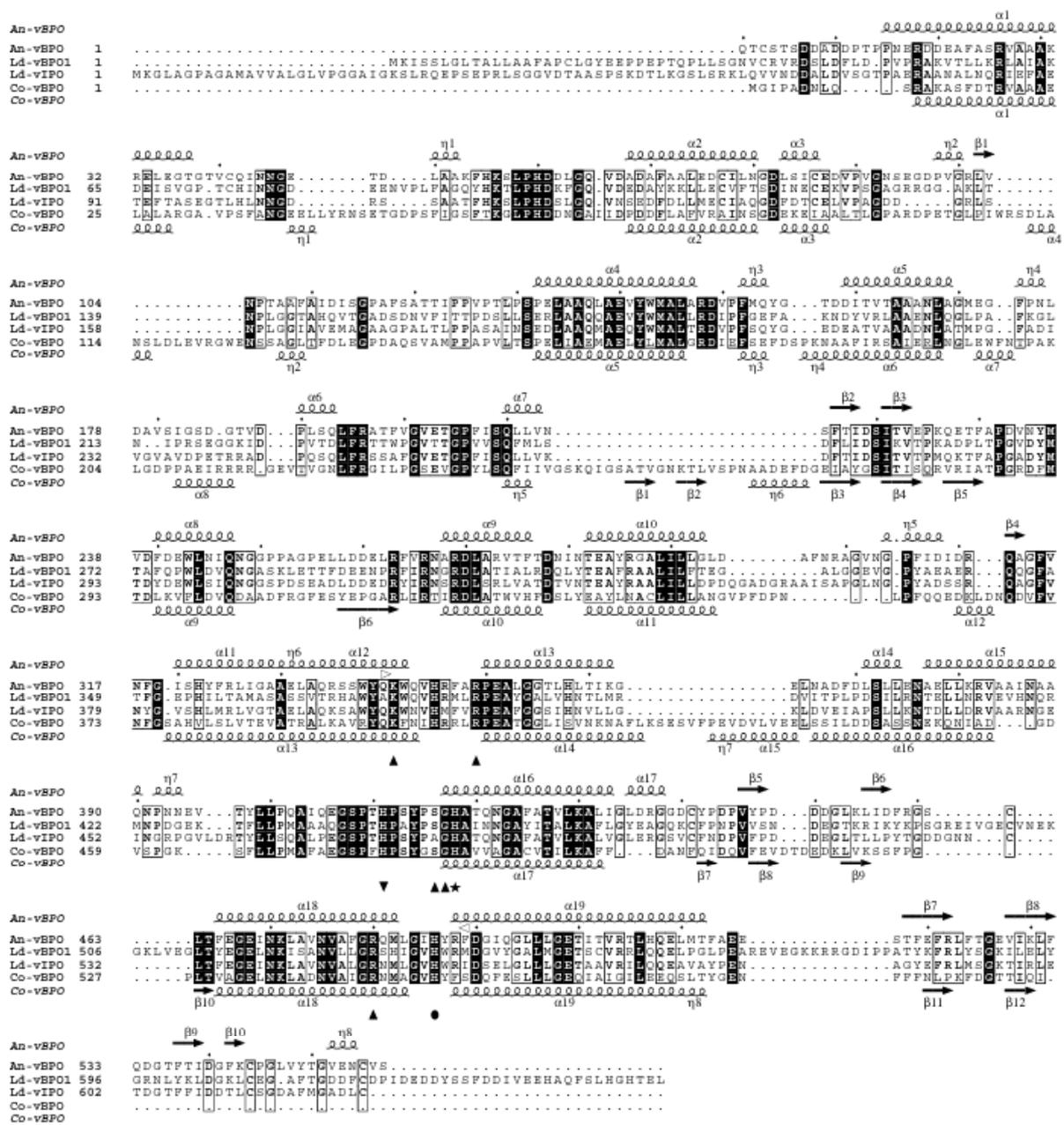


FIG. 5. Phylogeny of vanadium-dependent haloperoxidases.

The topology shown here is the tree obtained using the Neighbor-Joining method (with the Dayhoff's PAM matrix), with 137 amino acids used in the analyses, covering the active site region (see Fig. 4). The scale bar represents the expected number of changes per sequence position. Numbers at the nodes refer to the bootstrap values (100 replicates) in distance and maximum-parsimony analyses, respectively. Absence of data indicates that the node was not reliable. Ed-vCPO, v-CPO from *Embellisia didymospora* (NCBI Accession # CAA72344) ; Db-vCPO, v-CPO from *Drechslera biseptata* (NCBI Accession # CAA72008) ; Ci-vCPO, v-CPO from *C. inaequalis* (NCBI Accession # 1VNC) ; MgHypProt, hypothetical protein from *Magnaporthe grisea* (NCBI Accession # EAA54225) ; Co-vBPOp, v-BPO from *C. officinalis* obtained by protein sequencing (NCBI Accession # 1QHBA) ; Co-vBPOc, cloned v-BPO from *C. officinalis* (NCBI Accession # AAM46061) ; Cp-vBPO1, v-BPO1 from *C. pilulifera* (NCBI Accession # BAA31261) ; Cp-vBPO2, v-BPO2 from *C. pilulifera* (NCBI Accession # BAA31262) ; Ld-vBPO1, v-BPO1 from *L. digitata* (NCBI Accession # CAD37191) ; Ld-vBPO2, v-BPO2 from *L. digitata* (NCBI Accession # CAD37192) ; Ld-vIPO, v-IPO from *L. digitata* (this study) ; Fd-vBPO, v-BPO from *Fucus distichus* (NCBI Accession # AAC35279) ; An-vBPO, v-BPO from *A. nodosum* (NCBI Accession # P81701).

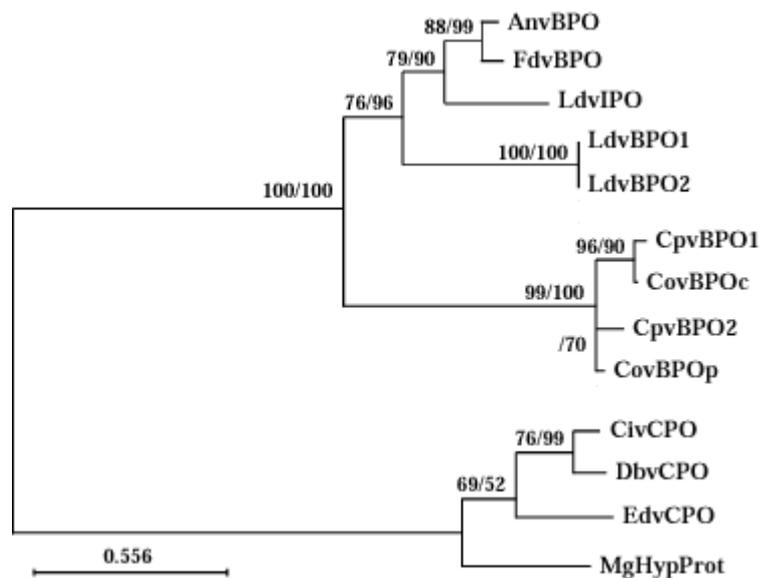


FIG. 6. Comparison of the active site structure of different algal v-HPOs.

(A) *C. inaequalis* v-CPO (PDB # 1VNC), (B) *C. officinalis* v-BPO (PDB # 1QHBA), (C) *A. nodosum* v-BPO (PDB # 1QI9A) and (D) the modellized *L. digitata* v-IPO. In the center of the picture, are represented vanadate (A, C and D) or the inorganic phosphate (B). Residues are drawn in atom colors and hydrogen-bonding contacts with vanadate are indicated by dashed lines.

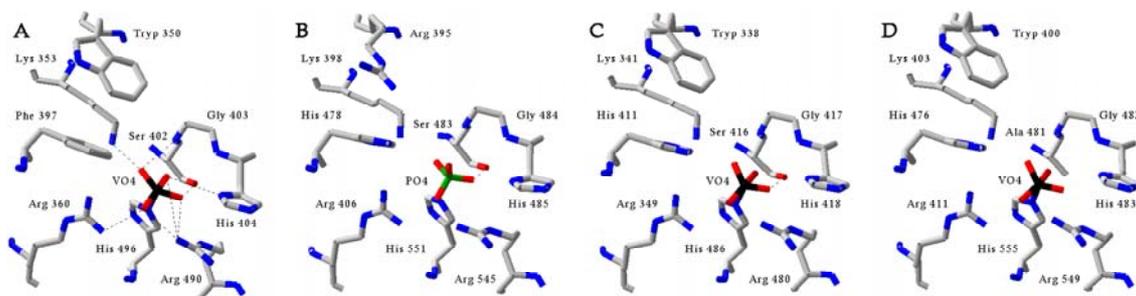


FIG. 7. Structures of active site funnel entrances of the crystallized *A. nodosum* v-BPO (A, B) and the modellized *L. digitata* v-IPO (C, D).

(A, C) Residues are drawn in atom colors excepted vanadate that appear in red and the main different residues between v-BPO and v-IPO that are colored in green and blue. (B, D) Stereo view representations of the funnel entrance surface. For color codes residues, see above.

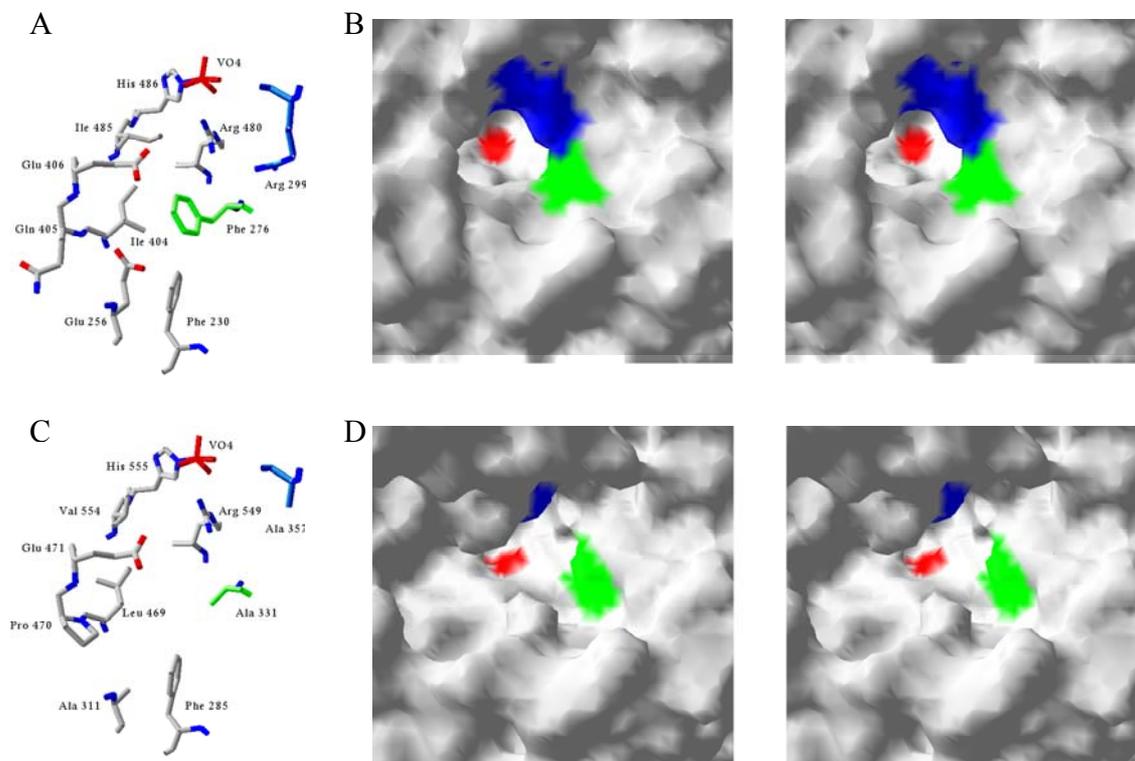


TABLE I

LC-MS/MS analyses of purified v-IPO tryptic digests and comparison with *A. nodosum* and *F. distichus* v-BPO protein sequences

Precursor ions obtained by LC-MS/MS are reported in the 1st and 3rd columns. The 2nd and 4th columns correspond to partial peptide sequence alignments of *de novo* sequenced v-IPO peptides (Ld-vIPO_Pept) and v-BPOs. Conserved residues in all sequences are in white capitals on a black background, residues that occur in at least one v-IPO peptide sequences and at least one v-BPO are in black capitals on a gray background. An-vBPO, v-BPO from *A. nodosum* (NCBI Accession # P81701) ; Fd-vBPO, v-BPO from *F. distichus* (NCBI Accession # AAC35279). Peptide positions are indicated before each sequence.

v-IPO (obs) MH ⁺	Partial peptide alignment	v-IPO (obs) MH ⁺	Partial peptide alignment
1217.70	Ld-vIPO_Pept1 ...D P Q S K L F R An-vBPO 150 D P L S Q L F R Fd-vBPO 269 D P F S Q L F R	1096.66	Ld-vIPO_Pept6 ...GRPGV L D R 1396.78 Ld-vIPO_Pept7 N G E I N G R P G V L D R An-vBPO 383 V A A I N A A Q N P N N E Fd-vBPO 502 V A E I N A A Q N P N N E
1611.86	Ld-vIPO_Pept2 D F T I D S I T V T ... An-vBPO 198 S F T I D S I T V E Fd-vBPO 317 S F T I D A I T V E	946.6	Ld-vIPO_Pept8 L A V N V A F G R An-vBPO 472 L A V N V A F G R Fd-vBPO 591 L A V N V A F G R
1422.76	Ld-vIPO_Pept3 L V A A D T V N T E A Y R	1656.9	Ld-vIPO_Pept9 I D S E I G L L L G E T A A V R
1452.78	Ld-vIPO_Pept4 L V A I D T V N T E A Y R An-vBPO 274 V T F T D N I N T E A Y R Fd-vBPO 393 V S F V D N I N T E A Y R		An-vBPO 489 F D G I Q G L L L G E T I T V R Fd-vBPO 608 F D G I Q G L L L G E T I T V R
1646.88	Ld-vIPO_Pept5 A I S A P C L N G P Y A D S S R An-vBPO 295 D A F N R A C V N G P F I D I D R Fd-vBPO 414 G A F S R E C I N G P F I D S D R	2098.10	Ld-vIPO_Pept10 I L Q Q E A V A Y P E N A... R
		2150.05	Ld-vIPO_Pept11 ... P E N A G Y E F R An-vBPO 505 T L H Q E L M T F A E S T F E F R Fd-vBPO 624 T L H Q E L M T F A E E A T E F R

TABLE II

Correspondences between observed LC-MS data from the purified v-IPO tryptic digests and the theoretical digests of v-IPO from L. digitata

The monoisotopic masses of purified v-IPO observed by LC-MS (1st column) were compared with the calculated v-IPO tryptic digest masses (2nd column); accuracies are indicated in the 4th column. Residues positions are reported in the 3rd column. The peptides, partially or fully sequenced by LC-MS/MS, are shown in bold, and those already mentioned in Table I were marked with an asterisk. Abbreviations: MSO, oxidation of methionine to sulfoxide; MC, miss cleavage.

v-IPO (obs.)		v-IPO (calc.)		
MH+		MH+	Tryptic peptide residues	
<i>Da</i>		<i>Da</i>	ΔM	
			ppm	
2451,13		2451,14	86-107	4
1217,70*		1217,66 (1 MC)	242-251	33
1061,58		1061,54	243-251	38
1880,08		1880,02	252-269	32
1611,86*		1611,79 (MSO)	270-283	43
1452,78*		1452,73	325-341	34
1524,84		1524,80	342-356	26
1646,68*		1646,81	357-373	79
1594,84		1594,78 (MSO)	374-386	38
1197,78		1197,71	427-437	58
2025,19		2025,14 (1 MC)	427-444	25
1396,78*		1396,73	449-461	36
1656,96*		1656,91	558-573	30
2098,10*		2098,02	574-591	38

REFERENCES

- Gribble, G. W. (2003) *Chemosphere* **52**, 289-297
- Albert, C. J., Crowley, J. R., Hsu, F.-F., Thukkani, A. K., and Ford, D. A. (2001) *J. Biol. Chem.* **276**, 23733-23741
- Barnett, P., Kruitbosh, D. L., Hemrika, W., Dekker, H. L., and Wever, R. (1997) *B. B. A.* **1352**, 73-84
- Potin, P., Bouarab, K., Salaün, J.-P., Pohnert, G., and Kloareg, B. (2002) *Curr. Opin. Plant Biol.* **5**, 308-317
- Gschwend, P. M., MacFarlane, J. K., and Newman, K. A. (1985) *Science* **227**, 1033-1035
- Fenical, W. (1975) *J. Phycol.* **11**, 245-259
- Faulkner, D. (2002) *J. Nat. Prod. Rep.* **19**, 1-48
- Polzin, J. J., Rorrer, G. L., and Cheney, D. P. (2003) *Biomol. Eng.* **20**, 205-215
- Littlechild, J. A. (1999) *Curr. Opin. Chem. Biol.* **3**, 28-34
- Hansen, E. H., Albertsen, L., Schafer, T., Johansen, C., Frisvad, J. C., Molin, S., and Gram, L. (2003) *Appl. Env. Microbiol.* **69**, 4611-4617
- Vairappan, C. S. (2003) *Biomol. Eng.* **20**, 255-259
- de Boer, E., and Wever, R. (1988) *J. Biol. Chem.* **263**, 12326-12332
- Arber, J. M., de Boer, E., Garner, C. D., Hasnain, S. S., and Wever, R. (1989) *Biochem.* **28**, 7968-7973
- Vilter, H. (1995) in *Metal ions in biological systems* (Sigel, H., and Sigel, A., eds) Vol. 31, pp. 325-362, Marcel Dekker, Inc., New York, Basel, Hong Kong
- Soedjak, H. S., and Butler, A. (1990) *Biochem.* **29**, 7974-7981
- Neidleman, S. L., and Geigert, J. (1987) *Endeavour* **11**, 5-15
- Vollenbroek, E. G. M., Simons, L. H., van Schijndel, J. W., Barnett, P., Balzar, M., Dekker, H. L., van der Linden, C., and Wever, R. (1995) *Biochem. Soc. Trans.* **23**, 267-271
- Barnett, P., Hemrika, W., Dekker, H. L., Muijsers, A. O., Renirie, R., and Wever, R. (1998) *J. Biol. Chem.* **273**, 23381-23387
- Plat, H., Krenn, B. E., and Wever, R. (1987) *Biochem. J.* **248**, 277-279

20. Colin, C., Leblanc, C., Wagner, E., Delage, L., Leize-Wagner, E., van Dorsselaer, A., Kloareg, B., and Potin, P. (2003) *J. Biol. Chem.* **278**, 23545-23552
21. Butler, A., Carter, J. N., and Simpson, M. T. (2001) in *Handbook on Metalloproteins* (Bertini, I., Sigel, A., and Sigel, H., eds), pp. 153-179, Marcel Dekker, Inc., New York, Basel
22. Tanaka, N., Hasan, Z., and Wever, R. (2003) *Inorg. Chim. Acta* **In Press**
23. Carter, J. N., Beatty, K. E., Simpson, M. T., and Butler, A. (2002) *J. Inorg. Biochem.* **91**, 59-69
24. Messerschmidt, A., Prade, L., and Wever, R. (1997) *Biol. Chem.* **378**, 309-315
25. Hemrika, W., Renirie, R., Macedo-Ribeiro, S., Messerschmidt, A., and Wever, R. (1999) *J. Biol. Chem.* **274**, 23820-23827
26. Weyand, M., Hecht, H.-J., Kiess, M., Liaud, M.-F., Vilter, H., and Schomburg, D. (1999) *J. Mol. Biol.* **293**, 595-611
27. Martinez, J. S., Carroll, G. L., Tschirret-Guth, R. A., Altenhoff, G., Little, D. R., and Butler, A. (2001) *J. Am. Chem. Soc.* **123**, 3289-3294
28. Butler, A. (1999) *Coord. Chem. Rev.* **187**, 17-35
29. Messerschmidt, A., and Wever, R. (1996) *Proc. Natl. Acad. Sci. USA* **93**, 392-396
30. Dau, H., Dittmer, J., Epple, M., Hanss, J., Kiss, E., Rehder, D., Schulzke, C., and Vilter, H. (1999) *FEBS Lett.* **457**, 237-240
31. Crépineau, F., Roscoe, T., Kaas, R., Kloareg, B., and Boyen, C. (2000) *Plant Mol. Biol.* **43**, 503-513
32. Apt, K. E., Clendennen, S. K., Powers, D. A., and Grossman, A. R. (1995) *Mol. Gen. Genet.* **246**, 455-464
33. Katoh, K., Misawa, K., Kuma, K., and Miyata, T. (2002) *Nucleic Acids Res.* **30**, 3059-3066
34. Gouet, P., Robert, X., and Courcelle, E. (2003) *Nucleic Acids Res.* **31**, 3320-3323
35. Fitch, W. M. (1971) *Syst. Zool.* **20**, 406-416
36. Galtier, N., Gouy, M., and Gautier, C. (1996) *Comput. Appl. Biosci.* **12**, 543-548
37. Saitou, N., and Nei, M. (1987) *Mol. Biol. Evol.* **4**, 406-425
38. Felsenstein, F. (1985) *Evolution* **39**, 783-791
39. Isupov, M. N., Dalby, A. R., Brindley, A. A., Izumi, Y., Tanabe, T., Murshudov, G. N., and Littlechild, J. A. (2000) *J. Mol. Biol.* **299**, 1035-1049
40. Verschuere, K. H., Kingma, J., Rozeboom, H. J., Kalk, K. H., Janssen, D. B., and Dijkstra, B. W. (1993) *Biochem.* **32**, 9031-9037
41. Machius, M., Wiegand, G., and Huber, R. (1995) *J. Mol. Biol.* **246**, 545-559
42. Fukuyama, K., Sato, K., Itakura, H., Takahashi, S., and Hosoya, T. (1997) *J. Biol. Chem.* **272**, 5752-5756
43. Fiedler, T. J., Davey, C. A., and Fenna, R. E. (2000) *J. Biol. Chem.* **275**, 11964-11971

BIBLIOGRAPHIE

Crépineau F., Roscoe T., Kaas R., Kloareg B., Boyen C.

Characterisation of complementary DNAs from the expressed sequence tag analysis of life cycle stages of *Laminaria Digitata* (Phaeophyceae), *Plant Mol. Biol.*, **2000**, 43 (4), 503-513.

Falick A.M., Hines W.M., Medzihradzsky K.F., Baldwin M.A., Gibson B.W.

Low-mass ions produced from peptides by high-energy collision-induced dissociation in tandem mass spectrometry, *J. Am. Soc. Mass Spectrom.*, **1993**, 4 (11), 882-93.

Dongré A.R., Jones J.L., Somogyi A., Wysocki V.H.

Influence of peptide composition, gas phase basicity, and chemical modification on fragmentation efficiency : evidence for the mobile proton model, *J. Am. Chem. Soc.*, **1996**, 118, 8365-74.

Hunt D.F., Yates J.R., Shabanowitz J., Winston S., Hauer C.R.

Protein sequencing by tandem mass spectrometry, *Proc. Natl. Acad. Sci. USA*, **1986**, 83 (17), 6233-7.

Küpper F.C., Schweigert N., Ar Gall E., Legendre J.M., Vilter H., Kloareg B.

Iodine uptake in laminariales involves extracellular, haloperoxidase-mediated oxidation of iodide., *Planta*, **1998**, 207, 163-171.

Küpper F.C., Kloareg B., Guern J., Potin P.

Oligoglucuronates elicit an oxidative burst in the brown algal kelp *Laminaria Digitata*., *Plant Physiol.*, **2000**, 125, 278-291.

Loo J.A., Edmonds C.G., Smith R.D.

Tandem mass spectrometry of very large molecules. Dissociation of multiply charged proline-containing proteins from electrospray ionization, *Anal. Chem.*, **1993**, 65, 425-438.

Potin P., Bouarab K., Küpper F.C., Kloareg B.

Oligosaccharide recognition signals and defense reaction in marine plant-microbe interaction., *Curr. Opin. Microbiol.*, **1999**, 2, 276-283.

Tabb D.L., Smith L.L., Brezi L.A., Wysocki V.H., Lin D., Yates J.R.

Statistical characterization of ion trap tandem mass spectra from doubly charged tryptic peptides, *Anal. Chem.*, **2003**, 75, 1155-1163.

Wehofsky M., Hoffmann R.

Automated deconvolution and deisotoping of electrospray mass spectra, *J. Mass Spectrom.*, **2002**, 37, 223-229.

PARTIE V :

CARACTERISATION DE GLYCANNES PAR
SPECTROMETRIE DE MASSE MS/MS

CARACTERISATION DE GLYCANNES PAR SPECTROMETRIE DE MASSE MS/MS

Ce travail a débuté en collaboration avec la société SERONO (Serono Biotech Center, Fenil-Sur-Corsier, Suisse) pour l'analyse, par MS/MS, des glycosylations de gonadotropines recombinantes. La stratégie de recherche de cette industrie s'étant ensuite focalisée sur d'autres projets prioritaires, nous nous sommes alors intéressés à des N-glycannes très similaires provenant de protéines recombinantes commercialisées par Dextra (Dextra Laboratories, Reading, U.K.).

Dans un second temps, nous avons initié le développement informatique d'un logiciel aidant à l'interprétation des spectres de fragmentation des glycannes. Pour la programmation de ce logiciel, nous avons travaillé en collaboration avec le Dr Luc Moulinier (Laboratoire de Génomique et Biologie Structurale, IGBMC, CNRS, UMR 7104, Strasbourg).

I. CONTEXTE BIOLOGIQUE : LES GONADOTROPINES

Les gonadotropines, la lutropine ou Luteinizing Hormone (LH), la follitropine ou Follicle Stimulating Hormone (FSH), et la choriogonadotropine placentaire (CG), sont associées à la famille des hormones glycoprotéiques. Chez les mammifères, leur rôle est d'assurer la production de stéroïdes sexuels et de réguler les fonctions de reproduction.

Ces hormones sont des hétérodimères non covalents constitués de sous-unités α et β . La sous-unité α est commune aux trois hormones et possède deux sites de N-glycosylation. La sous-unité β , propre à chaque hormone, confère au dimère sa spécificité biologique [Pierce et Parsons, 1981].

L'hétérogénéité des hormones glycoprotéiques varie suivant, le sexe et l'âge de l'individu mais également, suivant ses caractéristiques endocriniennes [Ronin, 1995]. L'importance des glycoformes des hormones glycoprotéiques a été largement étudiée ces dernières années afin d'établir des relations entre la structure des glycannes et l'activité biologique des hormones glycosylées [Sairam, 1983 ; Hård et coll., 1990 ; Thotakura et Blithe, 1995 ; Amoresano et coll., 1996]. Les glycannes des gonadotropines affectent directement ou indirectement l'activité biologique de ces hormones alors que la partie protéique reste le facteur déterminant de la spécificité hormonale [Ronin, 1995].

Aujourd'hui, certains cas de stérilité masculine ou féminine peuvent être soignés grâce à la production de gonadotropines recombinantes. La société SERONO est l'une des sociétés pharmaceutiques qui produit des gonadotropines recombinantes (r-hFSH, r-hLH et r-hCG) exprimées dans des cellules CHO, une lignée de cellules issue des ovaires de hamster chinois. L'importance de la

glycosylation des hormones est telle que, pour s'assurer de l'efficacité biologique des gonadotropines recombinantes, il est indispensable de vérifier leur structure au cours des différentes étapes de la production.

Cette problématique requiert donc une stratégie analytique permettant de caractériser les glycosylations protéiques. Dans ce contexte, nous avons entrepris :

- | d'évaluer les capacités de deux instruments, une trappe ionique et un MALDI TOF-TOF, pour l'élucidation structurale par MS/MS de N-glycannes connus
- | de mettre au point un algorithme facilitant l'interprétation des spectres de fragmentation générés lors de ces analyses.

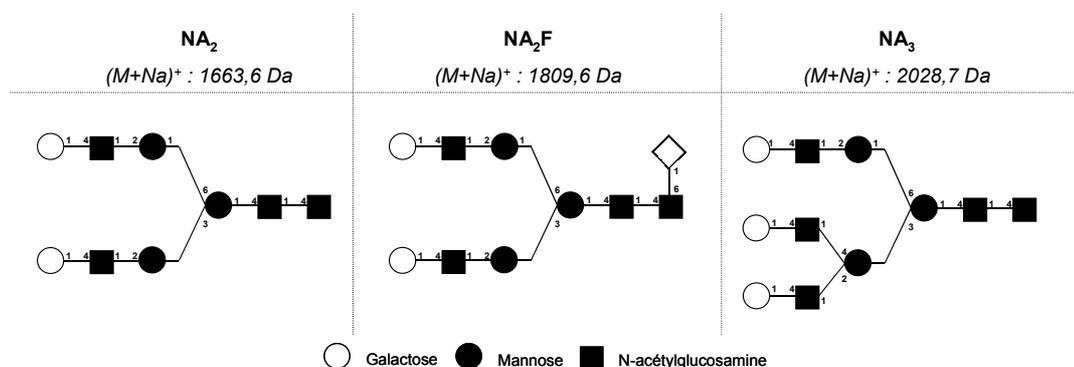
II. CARACTERISATION PAR MS/MS DE N-GLYCANNES

1. Mise en oeuvre

Les N-glycannes choisis pour notre étude de caractérisation par MS/MS sont représentatifs des motifs glycanniques couramment rencontrés chez les glycoprotéines (figure 1) :

- | une structure biantennée, asialylée, galactosylée (NA_2)
- | une structure biantennée, asialylée, galactosylée et substituée par un fucose (NA_2F)
- | une structure triantennée, asialylée, galactosylée (NA_3).

Figure 1 : structure des N-glycannes analysés.



Nous avons solubilisé les sucres dans une solution [H_2O/CH_3CN 1 :1 + 0,1% $HCOOH$] à une concentration de 100 pmol/ μL .

- | Cette solution a été directement analysée par nanoES-MS/MS sur la trappe ionique (ESQUIRE 3000+, Bruker).
- | Pour l'analyse MALDI-MS/MS (ULTRAFLEX TOF-TOF, Bruker), nous avons utilisé une matrice DHB et déposé 50 picomoles de glycannes sur la cible. Les deux modes de fonctionnement du TOF-TOF, LID et CID (cf. Introduction - Bibliographie, §1.1.4.2.b.), ont été

testés afin d'évaluer l'effet de la présence de gaz de collision dans la première partie du tube de vol sur la fragmentation des glycanes.

N.B. : La matrice DHB possède une affinité protonique supérieure à celle de la matrice CHCA, elle est donc moins « chaude » et moins favorable aux fragmentations métastables [cf. Introduction – Bibliographie, §1.2.4.]. Cependant, après comparaison des résultats obtenus avec les deux matrices, nous avons choisi d'utiliser la matrice DHB donnant des signaux plus intenses en mode MS ce qui facilite la sélection de l'ion parent pour l'analyse MS/MS.

Le mode d'ionisation négatif, quoique moins sensible, permet l'analyse des glycanes par MS/MS avec la trappe ionique [Carte, thèse 2001], cependant cette possibilité n'était pas encore envisageable avec le TOF-TOF lorsque nous avons réalisé ces expériences. Aussi, nous avons travaillé en mode d'ionisation positif.

Compte tenu de l'affinité des sucres pour les cations alcalins, il s'est avéré que nous détectons préférentiellement les glycanes sous forme d'adduits de sodium ($M+Na$)⁺. Nous avons donc fragmenté ces espèces cationisées par MS/MS et comparé les résultats obtenus avec les deux instruments.

2. La caractérisation du N-glycane NA₂

2.1. MS/MS avec la trappe ionique

Le N-glycane NA₂, (MNa ⁺, 1603,6 m/z), a été isolé et fragmenté avec la trappe ionique (cf. spectre de la figure 2). Les fragments identifiés à partir de ce spectre sont présentés dans le tableau 1 et la figure 3.

Figure 2 : spectre MS/MS avec la trappe ionique de l'ion ($M+Na$)⁺ du N-glycane NA₂.

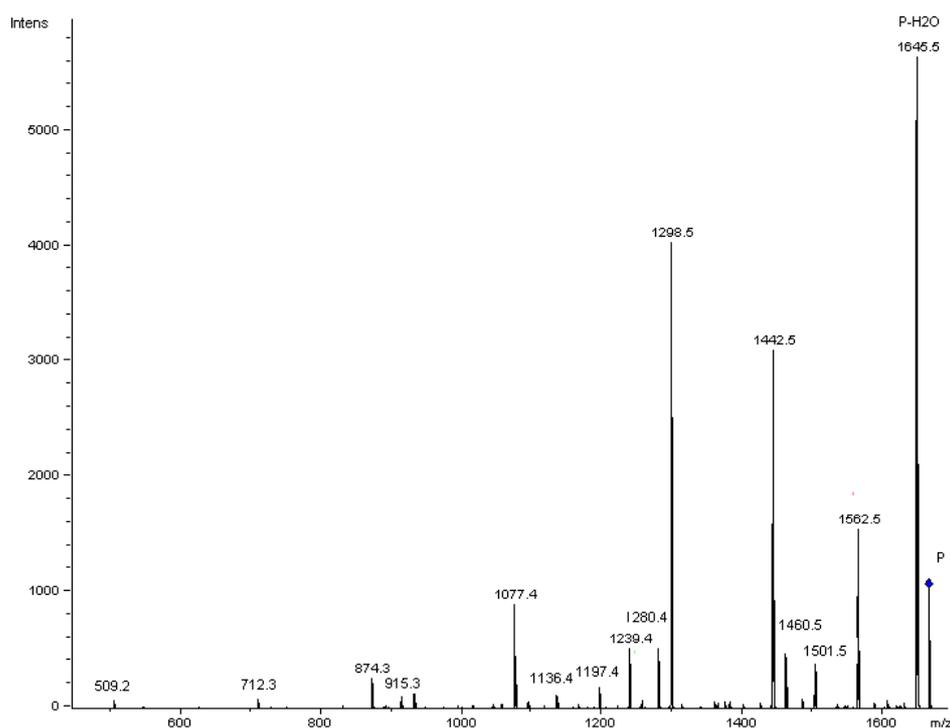


Tableau 1 : fragments obtenus par MS/MS de l'ion $(M+Na)^+$ du N-glycane NA_2 avec la trappe ionique. Les fragments intracycliques sont indiqués en rouge.

m/z	fragments $(M+Na)^+$	m/z	fragments $(M+Na)^+$
1645,5	$NA_2 - H_2O$	1197,4	$^{0,2}A_4$
1562,5	$^{1,3}A_6$; $^{1,3}A_{2\alpha/\beta}$	1136,4	$Y_{3\alpha/\beta}$
1501,5	$Y_{5\alpha/\beta}$	1077,4	$Y_{5\alpha/\beta}/B_4$; $Y_{4\alpha/\beta}/B_5$
1460,5	C_5	915,3	$Y_{5\alpha}/Y_{5\beta}/B_4$; $Y_{3\alpha/\beta}/B_5$; $Y_{4\alpha}/Y_{4\beta}$
1442,5	B_5	874,3	$Y_{4\alpha/\beta}/B_4$
1298,5	$Y_{4\alpha/\beta}$	712,3	$Y_{5\alpha}/Y_{4\beta}/B_4$; $Y_{4\alpha}/Y_{5\beta}/B_4$; $Y_{3\alpha/\beta}/B_4$; $Y_{4\alpha}/Y_{4\beta}/B_5$
1280,4	$Z_{4\alpha/\beta}$	509,3	$Y_{4\alpha}/Y_{4\beta}/B_4$
1239,4	C_4		

Nous remarquons sur le spectre la présence d'une espèce majoritaire correspondant probablement à la déshydratation du glycanne NA_2 (-18 Da). Ce type de fragmentation, étudié par l'équipe de Hofmeister en utilisant un marquage avec ^{18}O , proviendrait de la perte spécifique du groupement hydroxy en C-1 du monosaccharide situé à l'extrémité réductrice du glycanne [Hofmeister et coll., 1991].

NB : Pour notre étude, une étape de MS^n de degré supérieur (MS^3) est nécessaire afin de fragmenter cet ion relativement stable. Notre intérêt étant tourné vers la comparaison des résultats de MS/MS obtenus avec deux types d'instruments, nous ne présenterons pas ici les résultats de MS^3 qui n'apportent, d'ailleurs, pas d'informations supplémentaires quant à la nature des fragments observés.

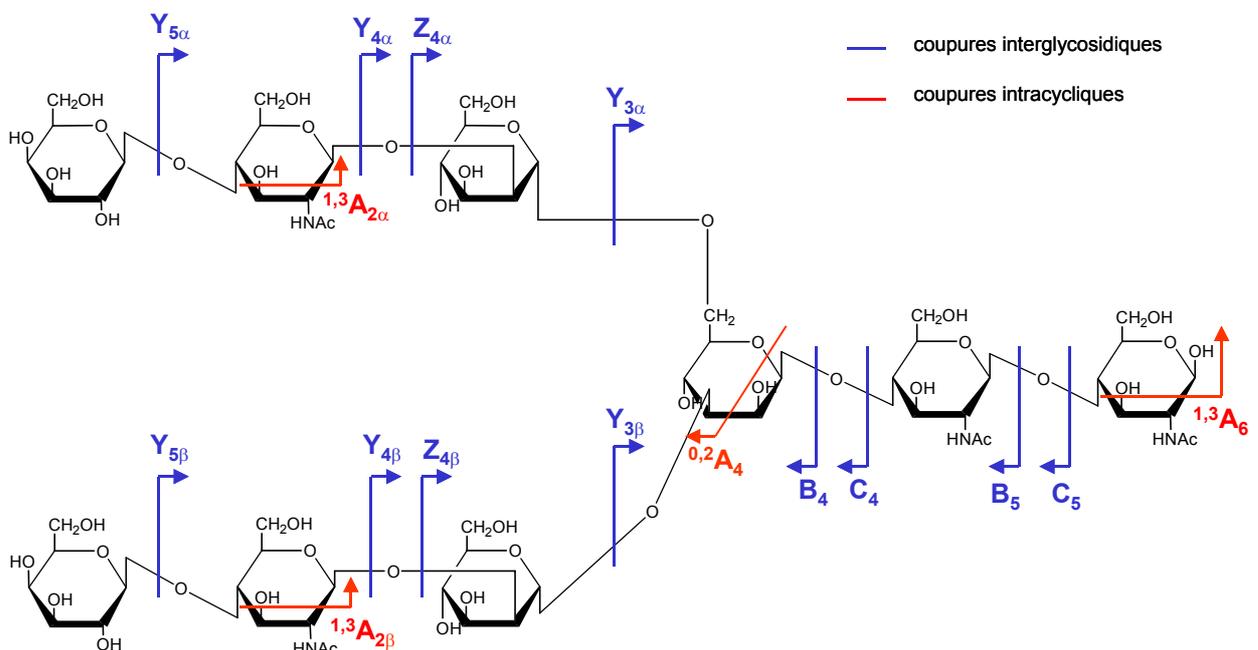
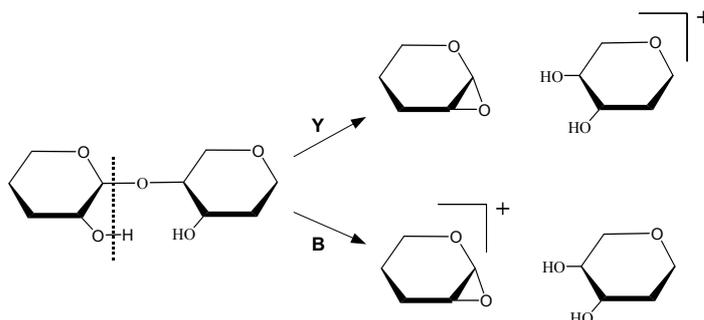


Figure 3 : structure du N-glycane NA₂ et fragments observés en MS/MS avec la trappe ionique.

2.1.1. Fragments Y et B

Les principaux fragments observés, de type Y et B, permettent de couvrir de façon significative la séquence du polysaccharide NA₂.

Figure 4 : voie de fragmentation des liaisons glycosidiques conduisant aux fragments B et Y. La position de l'atome d'hydrogène hydroxylique participant à ce mécanisme est indéterminée. Modèle proposé par [Harvey et coll., 1997].



Le mécanisme de fragmentation des polysaccharides conduisant à l'obtention de fragments de type Y et B est celui d'une fragmentation d'ion à charge piégée. Le processus de décomposition comprend l'élimination d'un proton et son transfert sur l'oxygène de la liaison glycosidique (figure 4). Le type d'ion obtenu, B ou Y, n'est alors fonction que de la localisation de la charge. Les travaux de Harvey [Harvey et coll., 1997] ont montré que, dans les spectres de fragmentation d'oligosaccharides $[M+Na]^+$, l'atome d'hydrogène provient d'un groupement hydroxy et non du C-2 comme cela semble être le cas pour la fragmentation de dérivés perméthylés [Gillece-Castro et Burlingame, 1990].

2.1.2. Fragments internes

Nous observons la présence de nombreux fragments internes, de types Y/B, ou Y/Y/B qui résultent de fragmentations consécutives du glycane.

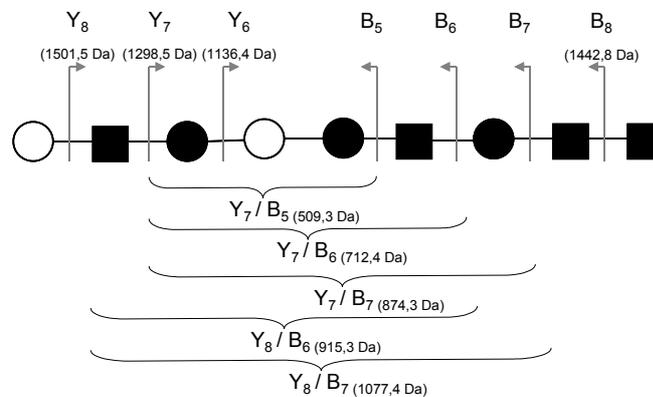
Il est vrai que, si les fragmentations multiples pour les peptides sont rarement observées par spectrométrie de masse en tandem avec la trappe ionique (cf. Introduction - Bibliographie, §I.1.4.2.), il n'en est pas de même pour la fragmentation des polysaccharides [Weiskopf et coll., 1997].

2.1.3. Le motif de branchement

Les fragmentations interglycosydiqes (Bi, Ci, Yj et Zj) peuvent parfois renseigner sur le motif de branchement du polysaccharide (cf. Introduction - Bibliographie, §III.3.4.3.). Cependant, cette distinction n'est pas toujours évidente : dans le cas présent, nos résultats ne permettent pas de vérifier que nous avons affaire à une structure biantennée.

En effet, un enchaînement linéaire des unités monosaccharidiques constituant notre N-glycane peut théoriquement générer des fragments isobares à ceux observés sur notre spectre. A titre d'exemple, nous proposons figure 5, une structure qui permet de vérifier cette affirmation.

Figure 5 : exemple d'oligosaccharide linéaire permettant de générer des fragments isobares à ceux observés en MS/MS avec la trappe ionique pour le N-glycane NA₂. Les symboles utilisés font référence à la figure 1.



Il est cependant important de rappeler que la plupart des techniques utilisées pour séparer les glycanes de leur glycoconjugué sont spécifiques d'un type de glycosylation : N ou O-glycosylation.

- | Les N-glycanes sont caractérisés par un motif commun biantenné appelé core. Aussi, les informations pertinentes tirées de l'analyse structurale par MS/MS des N-glycanes vont essentiellement concerner les branchements additifs au core sur le squelette du polysaccharide.
- | Pour les O-glycanes, qui n'ont pas de structure de base prédéfinie, toutes les informations de branchement obtenues par MS/MS sont intéressantes.

2.1.4. Clivages intracycliques

Les clivages intracycliques sont peu fréquents à basse énergie (cf. Introduction - Bibliographie, §III.3.4.2.). Nous observons deux pics sur notre spectre pouvant correspondre à ce type de fragment ; ils ne permettent pas de déduire d'informations quant à l'isomérisation de position des liaisons glycosidiques.

2.2. MS/MS avec le TOF-TOF

Le N-glycane NA_2 , $(M+Na)^+$, a été isolé et fragmenté avec le TOF-TOF (cf. spectre de la figure 6). Les fragments identifiés à partir de ce spectre sont présentés dans le tableau 2 et la figure 7.

Figure 6 : spectre MS/MS avec le TOF-TOF, en mode CID, de l'ion $(M+Na)^+$ du N-glycane NA_2 . Les fragments correspondant aux clivages de liaisons interglycosidiques sont annotés en noir, ceux correspondant à des ruptures intracycliques sont présentés en rouge. Pour plus de clarté, les fragments internes ne sont pas annotés.

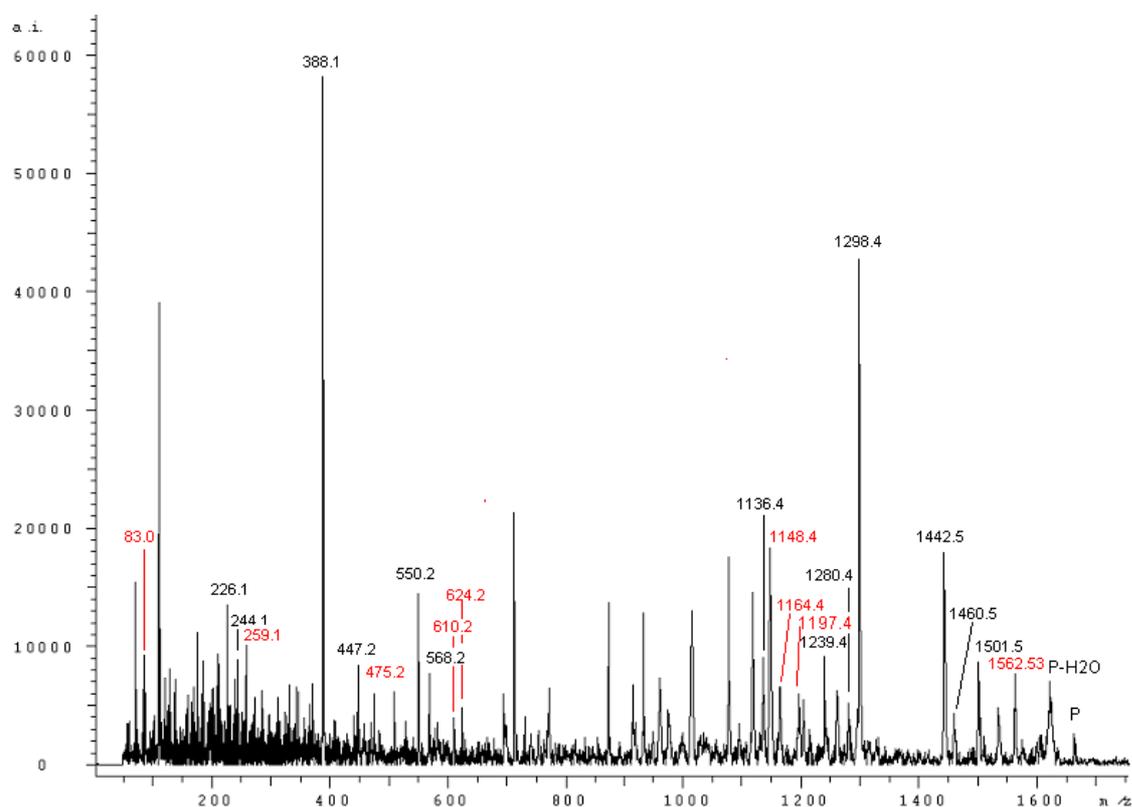
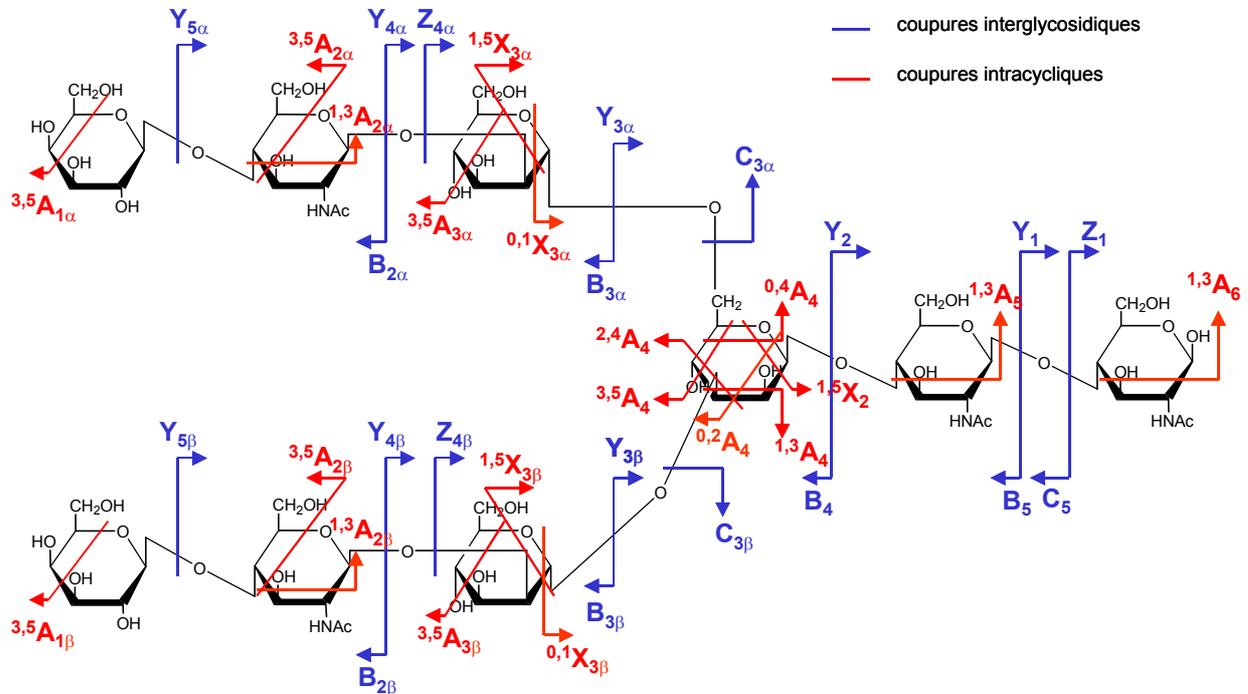


Tableau 2 : fragments obtenus par MS/MS de l'ion $(M+Na)^+$ du N-glycane NA_2 avec le TOF-TOF en mode CID. Les fragments intracycliques sont indiqués en rouge.

m/z	fragments $(M+Na)^+$	m/z	fragments $(M+Na)^+$
75,0	$^{3,5}A_{1\alpha/\beta}$; $^{3,5}A_{3\alpha/\beta}$	1136,4	$Y_{3\alpha/\beta}$
226,1	Z_1	1148,4	$^{0,1}X_{3\alpha/\beta}$
244,1	Y_1	1164,4	$^{1,5}X_{3\alpha/\beta}$
259,1	$^{3,5}A_{2\alpha/\beta}$	1197,4	$^{0,2}A_4$
388,1	$B_{2\alpha/\beta}$	1239,4	B_4
447,2	Y_2	1280,4	$Z_{4\alpha/\beta}$
475,2	$^{1,5}X_2$	1298,4	$Y_{4\alpha/\beta}$
550,2	$B_{3\alpha/\beta}$	1442,5	B_5
568,2	$C_{3\alpha/\beta}$	1460,5	C_5

610,2	$^{0,4}A_4; ^{2,4}A_4; ^{1,3}A_4$	1501,5	$Y_{5\alpha/\beta}$
624,2	$^{3,5}A_4$	1562,5	$^{1,3}A_{2\alpha/\beta}; ^{1,3}A_5; ^{1,3}A_6$

Figure 7 : structure du N-glycane NA₂ et fragments observés en MS/MS avec le TOF/TOF en mode CID.



2.2.1. Analyse des résultats et comparaison avec la MS/MS sur la trappe ionique

Le spectre MS/MS obtenu avec le TOF-TOF est relativement complexe à interpréter du fait de la quantité de fragments que nous avons générés au cours de cette analyse.

Deux ions, majoritaires sur le spectre, montrent que les fragments Y₄ et B₂, issus du clivage de la liaison interglycosidique entre le mannose du core et le résidu N-acétylglucosamine des branches du glycanne, sont des espèces relativement stables.

La comparaison des résultats obtenus à basse énergie avec la trappe ionique avec ceux obtenus à haute énergie avec le TOF-TOF nous conduisent aux observations suivantes :

- 1. Les fragmentations intracycliques**, observées lors de collisions à haute énergie, **sont favorisées avec le TOF-TOF**. Ces fragments, quoique d'intensité relative assez faible, sont nombreux et peuvent renseigner sur l'isomérisation de certaines liaisons comme nous le verrons ensuite.
- 2. Les fragments internes, déjà abondants avec la trappe, sont très fréquemment observés avec le TOF-TOF** : ils résultent alors de clivages multiples combinant des ruptures interglycosidiques et intracycliques et sont, par conséquent, difficiles à interpréter.
- Avec les deux techniques, **les fragments résultant de clivages interglycosidiques (Y, Z, B et C) couvrent de façon significative le squelette du polysaccharide NA₂**. Avec le TOF-TOF, la série des

ions Y suffit à reconstituer la séquence entière du sucre, les fragments B, C et Z confirment cet enchaînement. Avec la trappe ionique, la complémentarité des informations apportées par les fragments Y, Z d'une part et, B, C d'autre part, renseignent sur la séquence complète du glycanne.

4. Par ailleurs, avec la trappe ionique, le clivage des liaisons situées sur les antennes du glycanne génère essentiellement des fragments portant une charge sur le côté réducteur (Y et Z), alors que les clivages se produisant sur « le tronc » du core génèrent, eux, des fragments chargés du côté non réducteur (B et C). Il semblerait donc que la charge portée par le glycanne NA₂ et ses fragments, lors de l'analyse MS/MS avec la trappe ionique, soit préférentiellement localisée sur le résidu mannose situé à l'embranchement du core.

Au contraire, la technique MALDI-TOF-TOF nous permet d'observer simultanément les fragments Y et B correspondant à la rupture d'une même liaison glycosidique. Il ne semble donc pas qu'il y ait de localisation préférentielle de la charge sur le glycanne avec cette technique.

Nous proposons deux hypothèses pour expliquer cette différence :

| Le processus d'ionisation, MALDI ou ES, utilisé selon la technique peut influencer la position de la charge sur l'ion parent.

| D'autre part, comme nous l'avons expliqué précédemment (cf. Introduction - Bibliographie, §I.1.4.2.), le transfert d'énergie dans une trappe ionique se fait étape par étape et favorise la formation des espèces les plus stables sur le plan énergétique. Contrairement, le TOF-TOF permet d'étudier les fragmentations métastables issues de la décomposition rapide des ions dans la première partie du tube de vol.

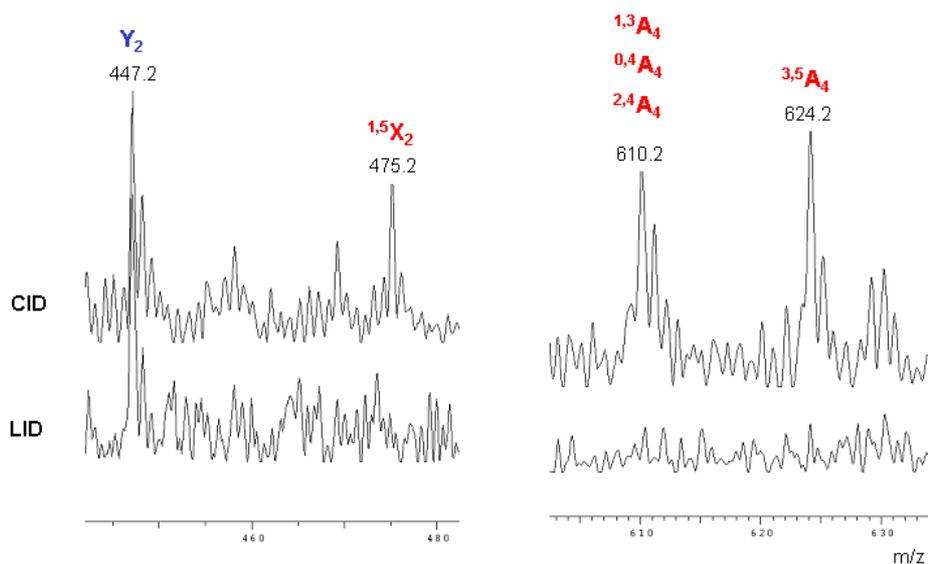
Aussi, les résultats suggèrent que les fragments observés avec la trappe ionique correspondent à des espèces stables sur lesquelles la charge ionique se situe sur le mannose branché du core. Les ions correspondant à la localisation de la charge sur d'autres résidus du glycanne seraient moins stables et par conséquent, non observés en MS/MS avec cette technique.

2.2.2. Comparaison des modes LID et CID

Les spectres obtenus en mode LID et CID (cf. Introduction – Bibliographie, §I.2.4.2.) pour la MS/MS du glycanne NA₂ avec le TOF-TOF sont d'allure relativement similaire. Les différences observées concernent principalement les fragments intracycliques obtenus spécifiquement à haute énergie et dont la présence est accentuée avec le gaz de collision. Ainsi, certains fragments A ou X ne sont observés que sur les spectres CID (cf. figure 8) ; d'autres, observés avec les deux modes de fonctionnement, sont généralement plus intenses en mode CID.

Par conséquent, la présence de gaz dans la première partie de l'analyseur favorise les fragmentations métastables en rendant les collisions efficaces entre les analytes plus fréquentes. L'utilisation du mode CID peut donc s'avérer très intéressante pour la caractérisation structurale des glycanes et notamment la détermination des isoméries de liaison.

Figure 8 : comparaison des résultats obtenus avec le TOF-TOF pour la MS/MS du glycanne NA₂ en modes LID et CID.



2.2.3. Les isoméries de liaison

Malgré l'abondance des fragments intracycliques observés, il est difficile de déduire du spectre des informations sur les liaisons qui soient à la fois intéressantes et non ambiguës.

En effet, les fragments $^{0,1}X_3$, $^{1,5}X_3$ et $^{1,5}X_2$ sont peu informatifs puisqu'ils nous renseignent sur l'implication du carbone 1 des résidus mannose du core dans les liaisons osidiques.

Des informations pertinentes proviennent de la fragmentation du cycle mannose tri-substitué : la présence simultanée des fragments $^{3,5}A_4$ et $^{0,2}A_4$ nous permet de caractériser sans ambiguïté le branchement 1-3. Cependant, étant donnée la symétrie des antennes, les fragments $^{0,4}A_4$, $^{2,4}A_4$ et $^{1,3}A_4$ sont isobares et ne peuvent être distingués ; il n'est donc pas possible de discriminer le branchement 1-6.

3. La caractérisation du N-glycane NA₂F

3.1. MS/MS avec la trappe ionique

Le N-glycane NA₂F, (MNa⁺, 1609,6 m/z), a été isolé et fragmenté avec la trappe ionique (cf. spectre de la figure 9). Les fragments identifiés à partir de ce spectre sont présentés dans le tableau 3 et la figure 10.

Figure 9 : spectre MS/MS avec la trappe ionique de l'ion (M+Na)⁺ du N-glycane NA₂F.

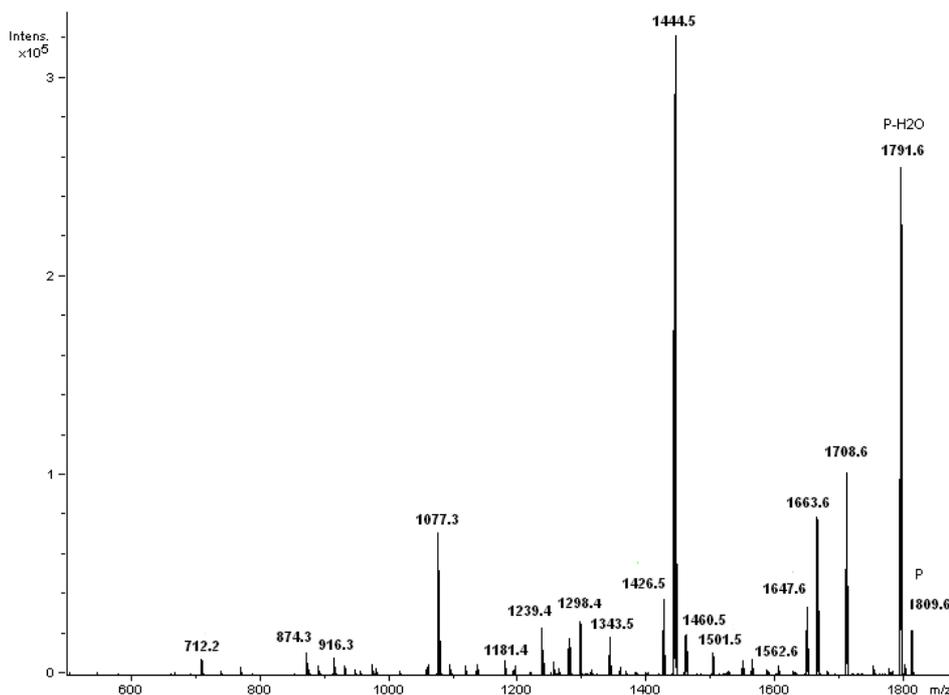


Tableau 3 : fragments obtenus par MS/MS de l'ion (M+Na)⁺ du N-glycane NA₂F avec la trappe ionique.

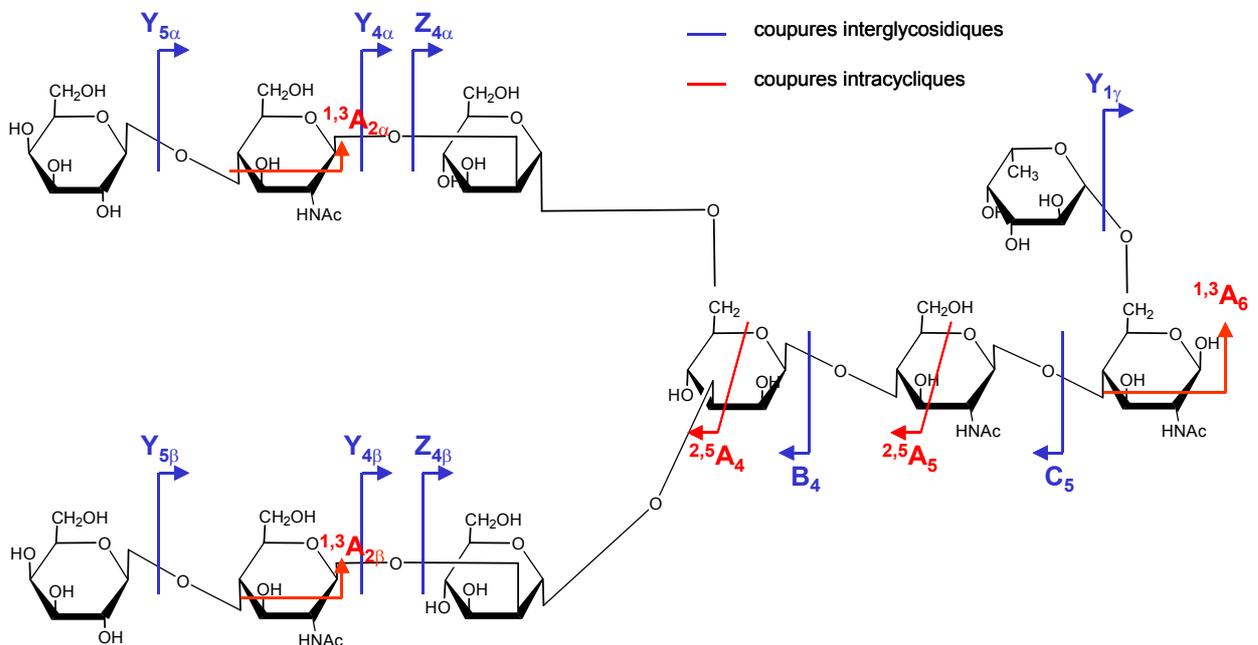
m/z	fragments (M+Na) ⁺	m/z	fragments (M+Na) ⁺
1791,6	NA ₂ F – H ₂ O	1343,5	^{2,5} A ₅
1708,6	^{1,3} A ₆ ; ^{1,3} A _{2α/β}	1298,4	Y _{4α/β} /Y _{1γ} ; Y _{5α/β} /B ₅
1663,6	Y _{1γ}	1239,4	B ₄
1647,6	Y _{5α/β}	1181,4	^{2,5} A ₄
1562,6	^{1,3} A ₆ /Y _{1γ} ; ^{1,3} A _{2α/β} /Y _{1γ}	1077,3	Y _{5α/β} /B ₄ ; Y _{4α/β} /B ₅
1501,5	Y _{5α/β} /Y _{1γ}	916,3	Y _{5α} /Y _{5β} /B ₄ ; Y _{3α/β} /B ₅ ; Y _{4α/β} /Y _{5α/5β} /B ₅
1460,5	C ₅	874,3	Y _{4α/β} /B ₄
1444,5	Y _{4α/β}	712,2	Y _{4α} /Y _{4β} /B ₅ ; Y _{4α/β} /Y _{5α/5β} /B ₄
1426,5	Z _{4α/β}		

Les résultats sont assez similaires à ceux obtenus pour le glycanne NA₂ : nous observons principalement des fragmentations des liaisons interglycosidiques couvrant la séquence du polysaccharide.

Le fucose est un groupement relativement labile caractérisé sur le spectre par la présence du fragment Y_{1γ}, cependant l'analyse MS/MS ne permet pas de localiser cet ose sur le squelette glycanne.

De récents travaux au laboratoire montrent que des étapes de spectrométrie de masse en tandem de degré supérieur (jusqu'à MS⁵) sur le glycanne NA₂ marqué à son extrémité réductrice par un groupement 2-aminobenzamide s'avèrent déterminantes pour caractériser la position du fucose sur le glycanne [Carte, thèse, 2001].

Figure 10 : structure du N-glycanne NA₂F et fragments observés en MS/MS avec la trappe ionique.



3.2. MS/MS avec le TOF-TOF

Le N-glycanne NA₂F, (M+Na)⁺, a été isolé et fragmenté avec le TOF-TOF (cf. figure 11). Les fragments identifiés à partir de ce spectre sont présentés dans le tableau 4 et la figure 12.

Tableau 4 : fragments obtenus par MS/MS de l'ion (M+Na)⁺ du N-glycanne NA₂F avec le TOF-TOF en mode CID.

m/z	fragments (M+Na) ⁺	m/z	fragments (M+Na) ⁺
169,0	B _{1γ}	1257,4	C ₄
259,1	^{3,5} A _{2αβ}	1310,4	^{1,5} X _{3αβ}
388,1	B _{2αβ}	1426,5	Z _{4αβ}
550,2	B _{3αβ}	1444,5	Y _{4αβ}

Caractérisation de glycannes par MS/MS

568,2	C _{3αβ}	1460,5	C ₅
593,2	Y ₂	1647,6	Y _{5αβ}
610,2	^{0,4} A ₄ ; ^{2,4} A ₄ ; ^{1,3} A ₄	1663,6	Y _{1γ}
621,2	^{1,5} X ₂	1675,6	^{1,5} X _{5αβ} ; ^{0,1} X _{1γ}
624,2	^{3,5} A ₄	1708,6	^{1,3} A _{2αβ} ; ^{1,3} A ₅ ; ^{1,3} A ₆
1239,4	B ₄		

Figure 11 : spectre MS/MS avec le TOF-TOF, en mode CID, de l'ion $(M+Na)^+$ du N-glycane NA₂F. Les couleurs font référence à la figure 6.

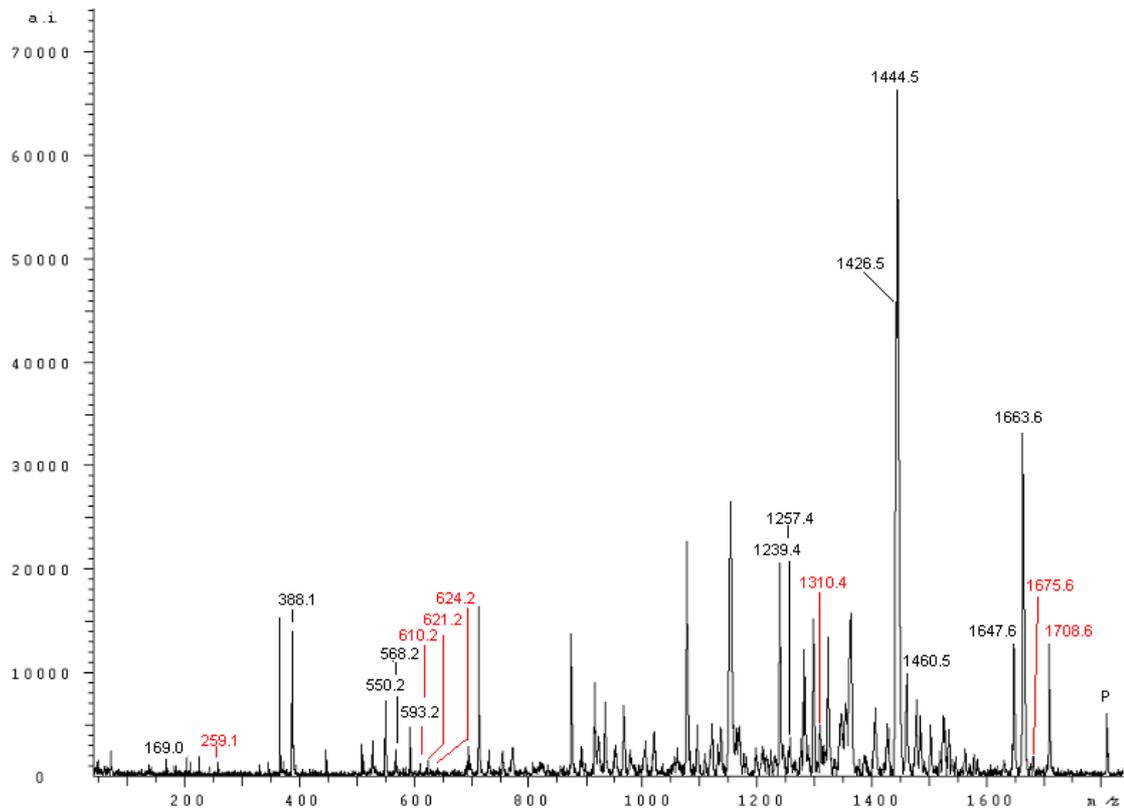
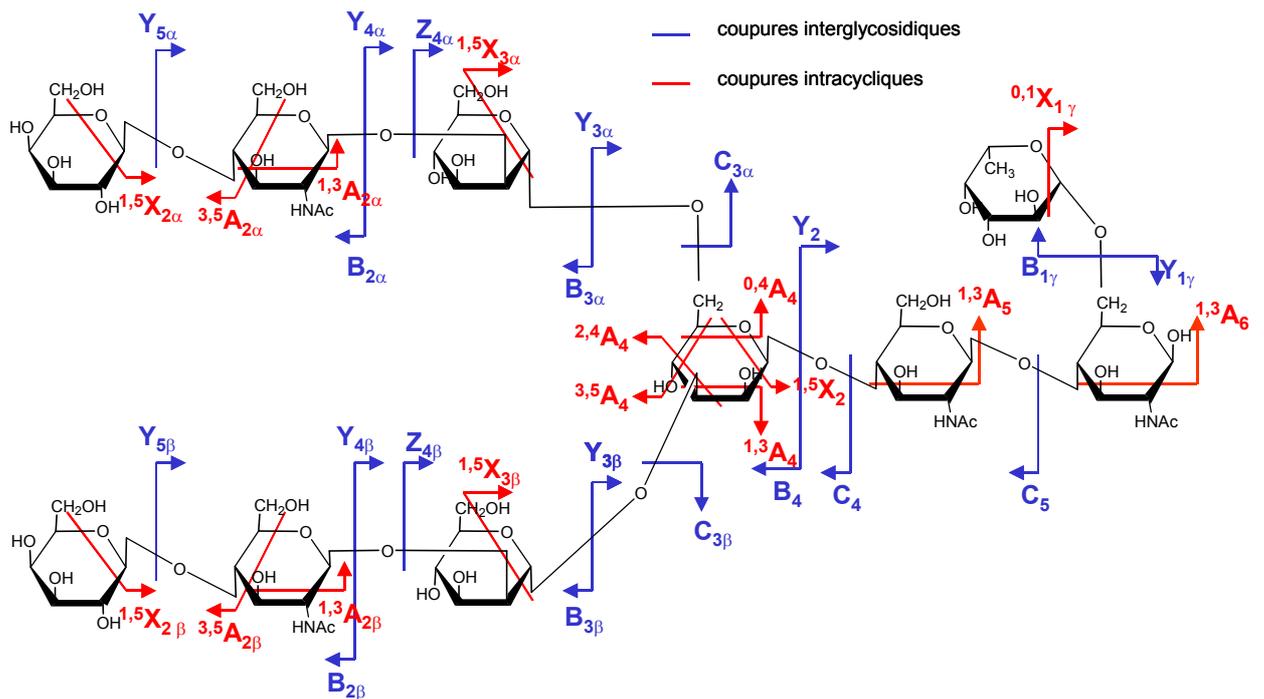


Figure 12 : structure du N-glycane NA₂F et fragments observés en MS/MS avec le TOF/TOF en mode CID.



Comme précédemment avec la structure NA₂, le spectre obtenu avec le TOF-TOF est très complexe du fait de l'abondance des fragments internes.

La complémentarité des informations issues des clivages interglycosidiques et intracycliques reflète la séquence du polysaccharide NA₂F sans toutefois permettre de discriminer clairement les branchements 1-3 et 1-6. Les fragments B_{1γ} et Y_{1γ} sont, eux, caractéristiques de la présence de fucose dans le glycanne.

4. La caractérisation du N-glycanne NA₃

4.1. MS/MS avec la trappe ionique

Le N-glycanne NA₃, (MNa⁺, 2028,7 m/z), a été isolé et fragmenté avec la trappe ionique (cf. figure 13). Les fragments identifiés à partir de ce spectre sont présentés dans le tableau 5 et la figure 14.

Figure 13 : spectre MS/MS avec la trappe ionique de l'ion (M+Na)⁺ du N-glycanne NA₃.

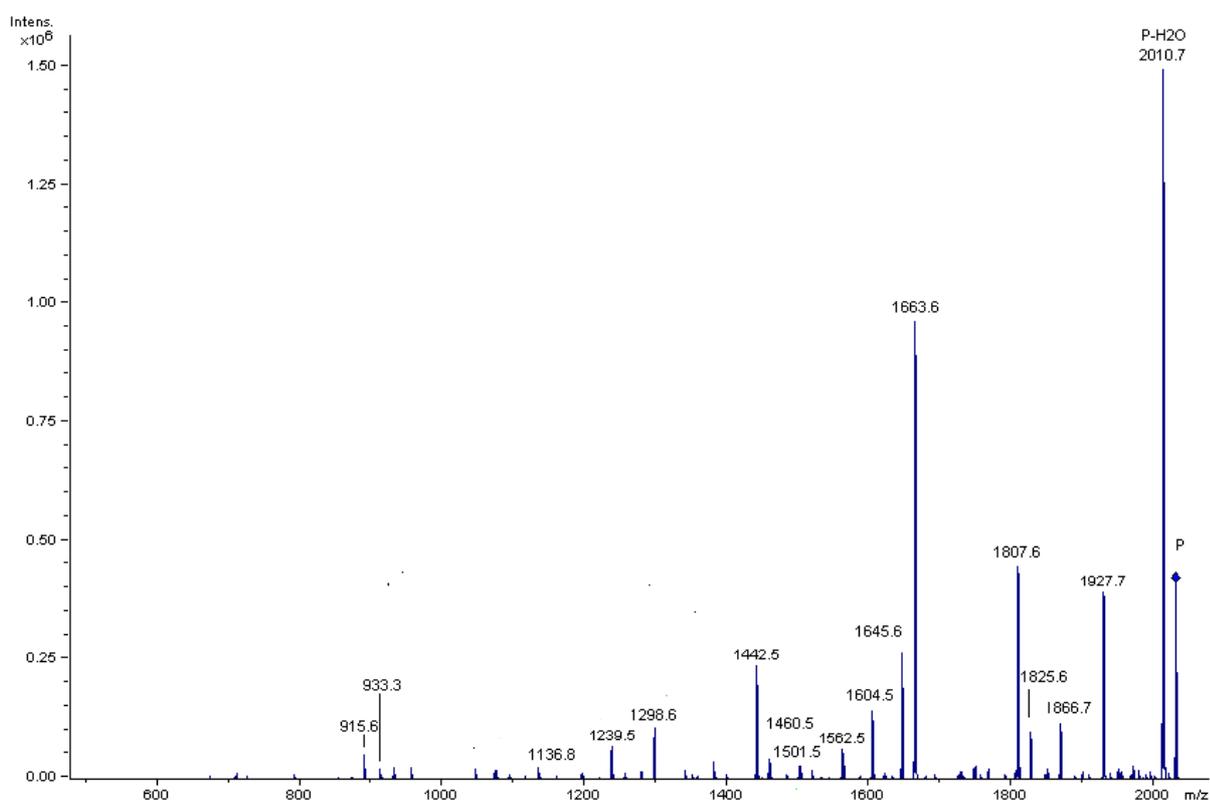


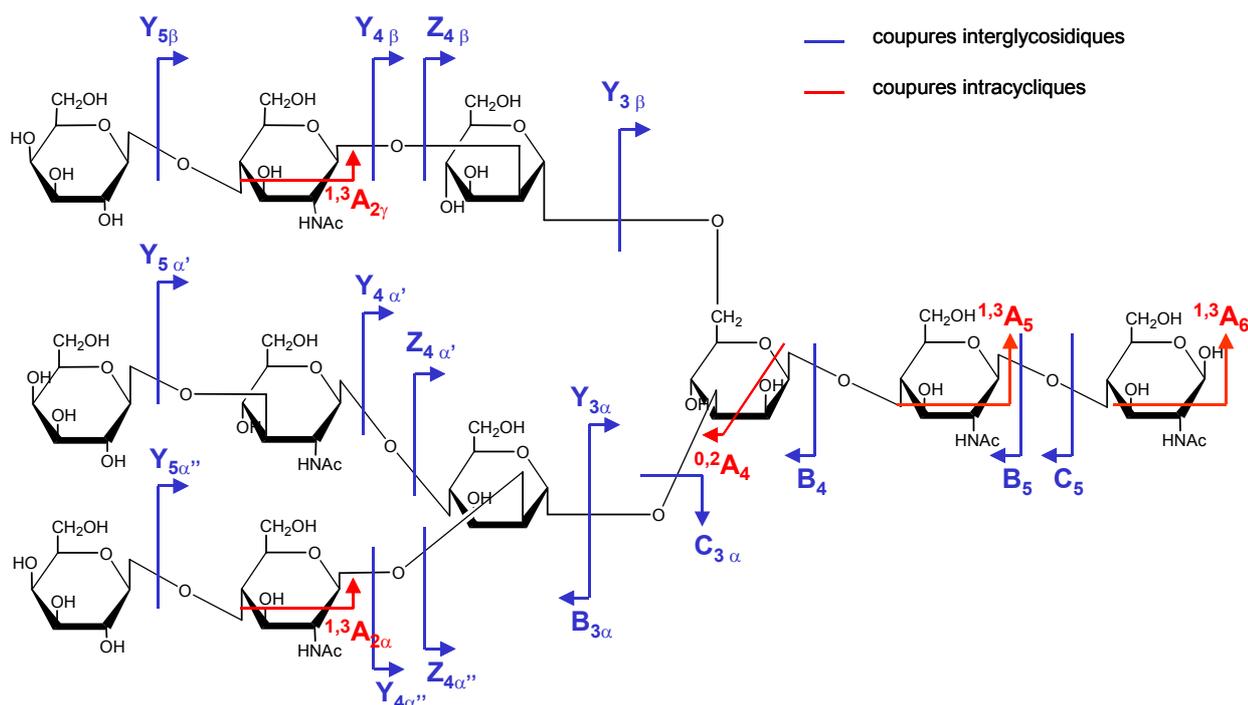
Tableau 5 : fragments obtenus par MS/MS de l'ion (M+Na)⁺ du N-glycanne NA₃ avec la trappe ionique.

m/z	fragments (M+Na) ⁺	m/z	fragments (M+Na) ⁺
915,6	B _{3α}	1562,5	^{0,2} A ₄
933,32	C _{3α}	1604,5	B ₄
1136,8	Y _{3α}	1645,6	Z _{4α'/α''/β}
1239,5	Z _{4α'/α''/β} /B ₄	1663,6	Y _{4α'/α''/β}
1298,6	Y _{3β} /B ₅ ; Y _{4α'/α''/β} /Y _{5α'/α''/β} /B ₅ ; Y _{5α'/α''/β} /Y _{5α'/α''/β} /B ₄ ; Y _{4α'/α''/β} /Y _{4α'/α''/β}	1807,6	B ₅
1442,5	Z _{4α'/α''/β} /B ₅	1825,6	C ₅

Caractérisation de glycanes par MS/MS

1460,5 | $Y_{4\alpha/\alpha''/\beta}/B_5$
1501,5 | $Y_{3\beta}$

1866,7 | $Y_{5\alpha/\alpha''/\beta}$
1927,7 | $^{1,3}A_{2\alpha\beta}$; $^{1,3}A_5$; $^{1,3}A_6$

Figure 14 : structure du N-glycane NA₃ et fragments observés en MS/MS avec la trappe ionique.


Contrairement aux deux structures étudiées précédemment, la fragmentation du glycane NA₃ avec la trappe ionique permet de générer des fragments caractéristiques de la branche 1-3 : B_{3α} et C_{3α}. Cependant, nous ne visualisons pas de fragments de type B_{3β} ou C_{3β} correspondant à la branche 1-6.

4.2. MS/MS avec le TOF-TOF

Le N-glycane NA₃, (M+Na)⁺, a été isolé et fragmenté avec le TOF-TOF (cf. spectre de la figure 15). Les fragments identifiés à partir de ce spectre sont présentés dans le tableau 5 et la figure 16.

Tableau 6 : fragments obtenus par MS/MS de l'ion (M+Na)⁺ du N-glycane NA₃ avec le TOF-TOF en mode CID.

m/z	fragments (M+Na) ⁺	m/z	fragments (M+Na) ⁺
185.0	B _{1β/α/α'}	568.2	C _{3β}
203.1	C _{1β/α/α'}	610.2	^{0,4} A ₄
226.1	Z ₁	624.2	^{3,5} A ₄
244.1	Y ₁	640.2	^{0,3} A ₄
259.1	^{3,5} A _{2β/α/α'}	915.3	B _{3α}
272.1	^{1,5} X ₁	933.3	C _{3α}
305.1	^{0,2} A _{2β/α/α'}	1076.4	^{2,4} X ₂ ; ^{1,3} A ₄
388.1	B _{2β/α/α'}	1136.4	Y _{3α}
406.1	C _{2β/α/α'}	1164.4	^{1,5} X _{3α}
429.1	Z ₂	1529.5	^{1,5} X _{3β}
447.2	Y ₂	1645.6	B _{4β/α/α'}

462.2	$3,5A_{3\alpha}$	1664.6	$2,4A_5$
475.2	$1,5X_2$	1807.6	B_5
550.2	$B_{3\beta}$	1866.7	$Y_{5\beta/\alpha/\alpha'}$

Figure 15 : spectre MS/MS avec le TOF-TOF, en mode CID, de l'ion $(M+Na)^+$ du N-glycane NA_3 . Les couleurs font référence à la figure 6.

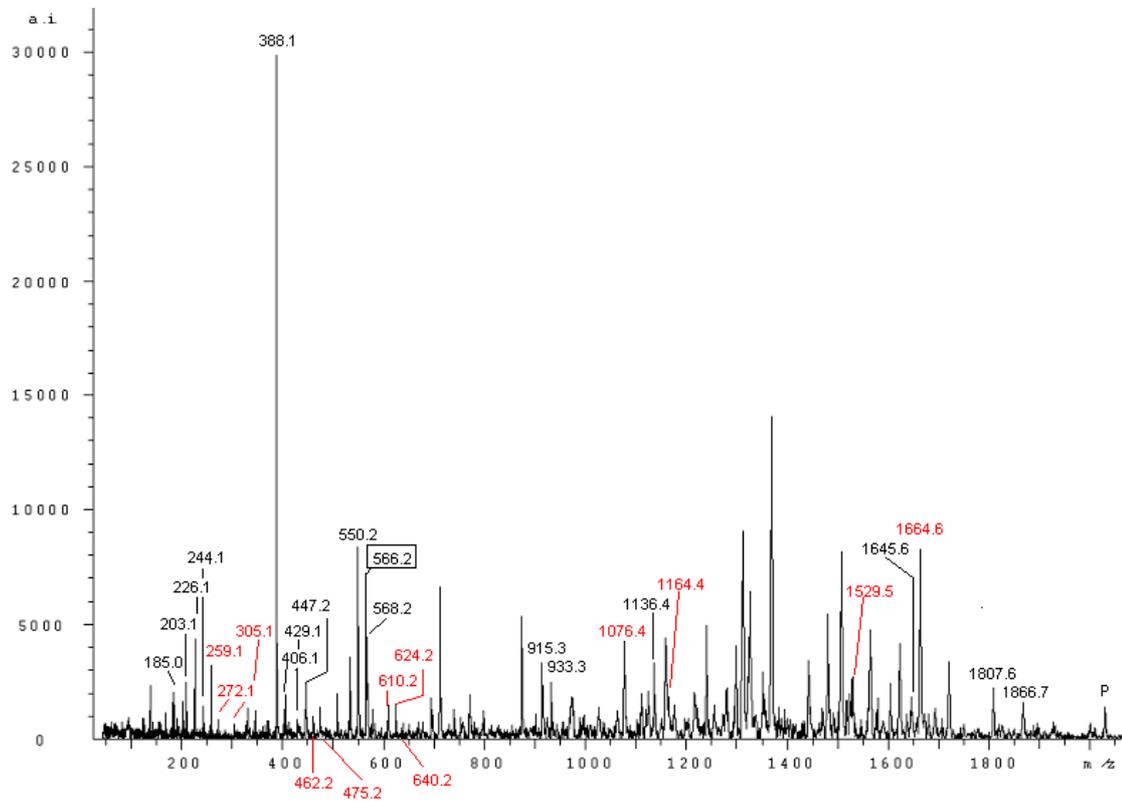
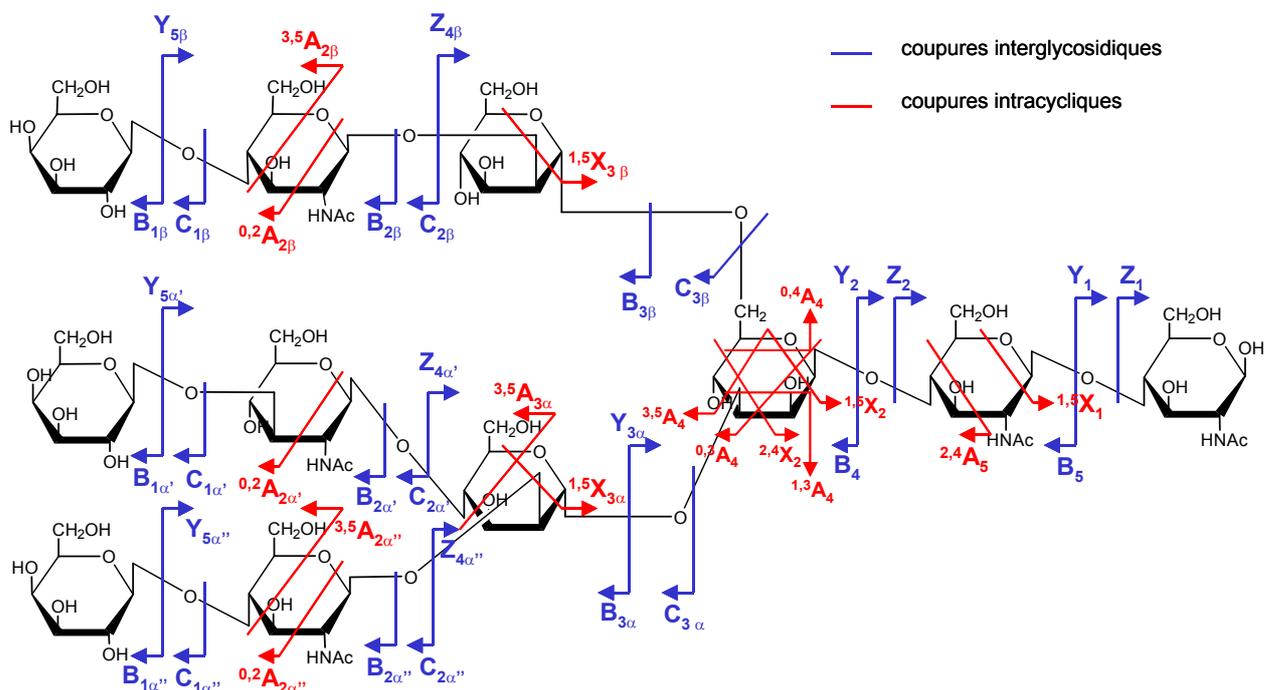


Figure 16 : structure du N-glycane NA_3 et fragments observés en MS/MS avec le TOF/TOF en mode CID.

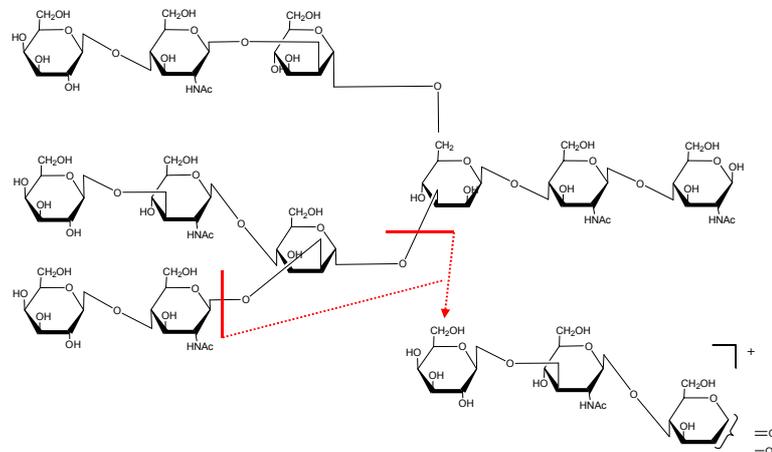


La plupart des clivages interglycosidiques et intracycliques trouvés dans les spectres des oligosaccharides biantennés sont également présents dans le spectre du glycanne NA₃.

Les fragments ^{0,4}A₄ d'une part, ainsi que ^{3,5}A₄ et ^{0,3}A₄ d'autre part, nous permettent de discriminer sans ambiguïté les branchements 1-3 et 1-6 sur le mannose en position centrale.

Certains ions apparaissent cependant n'avoir rien en commun avec les spectres des N-glycannes biantennés. Parmi eux, l'ion de m/z 566 (encadré sur la figure 15) a préalablement été identifié par Harvey lors de travaux sur la CID à haute énergie d'oligosaccharides complexes ionisés par MALDI [Harvey et coll., 1997]. Cet ion, caractéristique des structures tri et tetraantennées, correspondrait à un fragment interne dont le mécanisme de formation est présenté figure 17.

Figure 17 : Mécanisme de formation de l'ion de m/z 566 caractéristique du motif triantenné du glycanne NA₃. D'après Harvey [Harvey et coll., 1997].



On peut également remarquer que la différence de masse entre les fragments B_{3α} et B_{3β} ainsi qu'entre les fragments ^{1,5}X_{3β} et ^{1,5}X_{3α} rend compte de l'asymétrie des antennes. De plus, le fragment ^{3,5}A_{3α} établit clairement la présence du branchement 1-4 de l'antenne 3.

5. Conclusion

Nous avons donc testé les capacités de deux types d'instruments pour l'étude structurale par MS/MS de N-glycannes connus.

- | L'ES trappe ionique a permis de générer des spectres relativement simples avec principalement des fragments issus de coupures des liaisons interglycosidiques, donnant des informations de séquence des structures étudiées.
- | Le MALDI TOF-TOF a donné des spectres beaucoup plus complexes avec des fragments issus du clivage des liaisons interglycosidiques mais également des fragments intracycliques résultant de mécanismes de dissociation à haute énergie. Ces derniers se sont avérés particulièrement intéressants pour la détermination de certaines isoméries de liaison et, parfois, du motif de branchement du glycanne.

Avec les deux techniques, nous avons aussi observé des fragments internes résultant de clivages multiples du glycanne. Ces fragments, particulièrement abondants avec le TOF-TOF, présentent l'inconvénient de rendre l'interprétation des spectres très complexe.

Leur élucidation pourrait cependant permettre d'approfondir la caractérisation structurale du glycanne, comme nous l'avons montré par comparaison de nos résultats avec ceux présentés par l'équipe d'Harvey [Harvey et coll., 1997].

Dans ce but, il serait intéressant de réaliser une étude sur un plus grand nombre de structures : la comparaison des spectres MS/MS permettrait de corréler la présence de certains fragments internes à une particularité du motif glycanne (isomérisation de liaisons, branchement etc...).

D'autre part, la comparaison des deux techniques montre que l'analyse MS/MS avec la trappe ionique est plus sensible qu'avec le TOF-TOF. Elle permet également d'isoler, de façon plus fine que le TOF-TOF, un ion précurseur issu d'un mélange :

| Pour les glycanes étudiés, la trappe offre une sensibilité adaptée à l'analyse MS/MS de solutions de concentrations comprises entre 10 à 50 picomoles / μ L. Nous avons analysé des glycanes purs mais la trappe permet également d'isoler finement un précurseur dans un mélange (± 1 m/z). De plus, cet instrument présente l'avantage de permettre la réalisation d'étapes de spectrométrie de masse de degré supérieur qui peuvent s'avérer utiles pour approfondir la caractérisation structurale des glycanes.

| Avec le MALDI-TOF-TOF, nous avons déposé 50 picomoles de glycanes purs sur la cible. Pour notre étude et dans nos conditions, des dépôts de quantités inférieures sont critiques pour la sensibilité de l'analyse MS/MS. L'isolement du précurseur avec cette technique n'est pas très fine (fenêtre de 10 m/z environ) ce qui peut être gênant pour l'analyse de glycanes issus d'un mélange. Ainsi, les étapes préalables de fractionnement et de purification des glycanes (cf. Introduction – Bibliographie, §III.3.3.1) sont d'autant plus importantes pour l'analyse MS/MS avec le TOF-TOF.

A ce jour, aucune technique analytique ne peut, à elle seule, permettre de caractériser l'hétérogénéité structurale de la famille des glycanes. La multiplicité des problèmes à résoudre pour l'analyse des glucides nécessite de mettre en œuvre plusieurs techniques. Nous souhaitons illustrer ici les capacités de la nouvelle génération d'instrument TOF-TOF pour l'analyse structurale de glycanes par MS/MS. Cette technique ouvre une voie nouvelle pour l'analyse des fragmentations induites à haute énergie grâce à sa rapidité de mise en œuvre et sa sensibilité comparativement à des techniques pré-existantes comme le MALDI-PSD.

La dérivation des glycanes, par exemple par peracétylation, perméthylation ou bien l'amination réductive, permettent, dans certains cas, d'obtenir des informations structurales pertinentes et peuvent faciliter l'interprétation des spectres obtenus par MALDI PSD [Lemoine et coll., 1996]. Il serait par conséquent intéressant d'exploiter les performances de l'analyse MALDI-TOF-TOF pour l'analyse de glycanes dérivatisés.

III. DEVELOPPEMENT D'UN OUTIL BIO-INFORMATIQUE POUR ASSISTER L'INTERPRETATION DES SPECTRES DE FRAGMENTATION DES GLYCANNES

1. Introduction

Un certain nombre d'approches par spectrométrie de masse en tandem (MS/MS), utilisant principalement l'ionisation FAB [Egge et coll., 1987], ES [Andrew et coll., 1988] ou MALDI [Spengler et coll., 1995 ; Harvey, 1999], ont été décrites pour l'étude structurale d'oligosaccharides. La MS/MS est aujourd'hui une méthode sensible et rapide à mettre en œuvre pour obtenir des informations sur la composition d'un glycanne préalablement isolé.

Contrairement à d'autres types de biomolécules, comme les peptides ou les oligonucléotides, l'analyse des carbohydrates s'avère généralement complexe car elle requiert la distinction des isomères résultant des combinaisons possibles des branchements et des liaisons entre les différentes unités monosaccharides. Par conséquent, comme nous l'avons montré précédemment, l'interprétation des spectres de fragmentation des glycanes, sans données de référence, peut être difficile et parfois ambiguë.

A l'heure actuelle, il n'existe pas de bibliothèques de composés standards relatifs aux N et O-glycanes rencontrés chez les diverses espèces animales ou végétales. C'est probablement la raison principale pour laquelle aucune procédure automatique pour l'identification des carbohydrates, comparable à celles existant pour les protéines, n'a encore été établie.

Lorsque nous avons démarré ce travail de thèse, les spectres MS/MS de glycanes devaient donc être interprétés manuellement sur la base de règles empiriques. C'est pourquoi nous avons souhaité initier le développement d'un programme informatique permettant d'assister l'assignation manuelle des pics contenus dans les spectres MS/MS de glycanes complexes.

2. Stratégies possibles

Deux stratégies, résumées figure 18, ont préalablement été envisagées afin d'élaborer un outil informatique visant à faciliter l'interprétation des données de MS/MS issues de l'analyse de glycanes complexes.

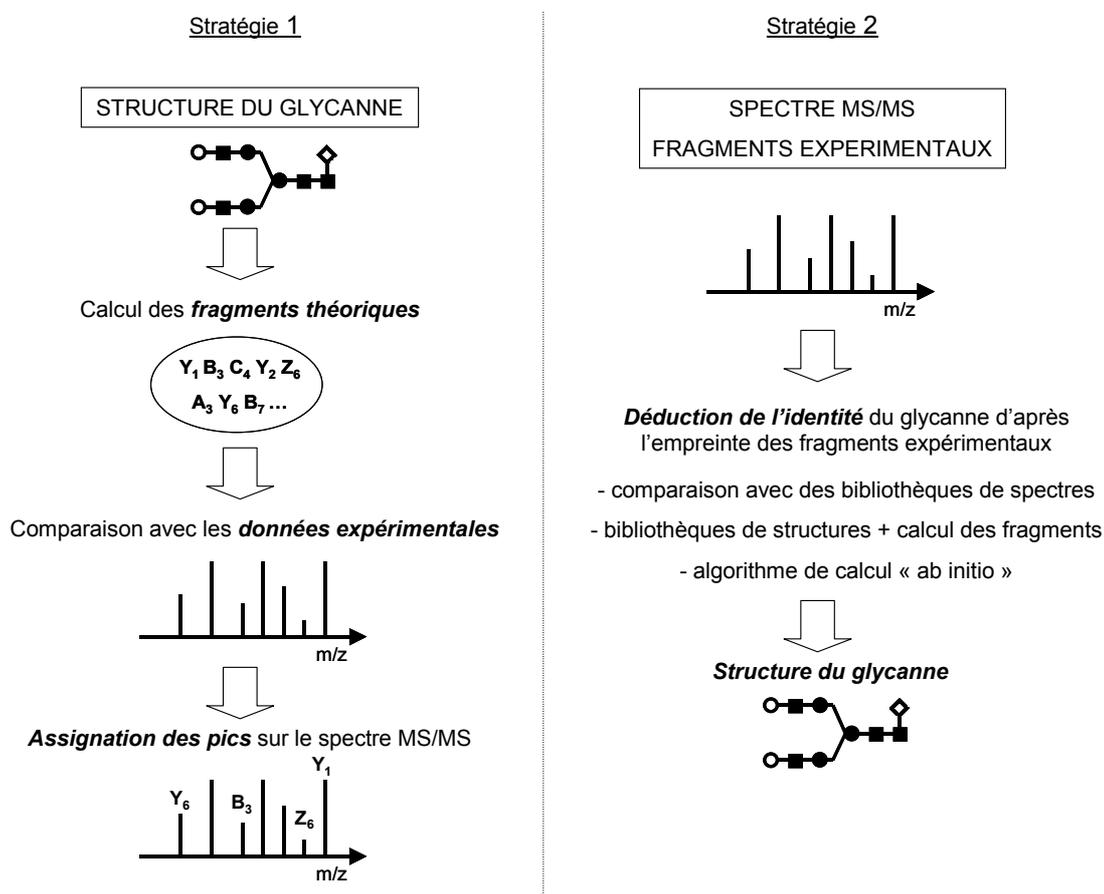
2.1. Stratégie 1

La première stratégie repose sur l'étude d'un motif glycanique connu et consiste à élaborer un algorithme permettant de calculer la masse de tous les fragments théoriques pouvant être générés à partir de cette structure. La comparaison des fragments expérimentaux obtenus par MS/MS avec les fragments calculés permet alors de faciliter l'assignation des pics observés sur les spectres.

L'inconvénient majeur de cette stratégie est de ne pas permettre l'identification de composés inconnus mais d'aider seulement à vérifier l'identité de structures prédéfinies ou probables. Il faut cependant souligner que, bien souvent, le contexte biologique et les résultats d'analyses

complémentaires sont des sources d'informations préliminaires concernant la nature du polysaccharide étudié. Cette stratégie peut s'avérer notamment utile dans le cadre d'analyses de contrôle structural, comme, par exemple, pour les glycanes issus de protéines recombinantes.

Figure 18 : schéma illustrant les stratégies envisagées dans le cadre du développement d'un outil informatique facilitant l'interprétation des données de MS/MS issues de l'analyse de glycanes complexes.



2.2. Stratégie 2

La seconde stratégie est plus puissante puisqu'elle propose l'identification du glycanne analysé à partir de son spectre de fragmentation. Plusieurs solutions peuvent être envisagées pour sa mise en œuvre :

- | Elle peut reposer sur l'utilisation d'une banque de données de spectres MS/MS de glycanes connus. L'homologie entre les spectres expérimentaux et ceux de la banque permettrait alors l'identification structurale. Toutefois, la nature des fragmentations observées en MS/MS étant étroitement liée à l'instrument, aux conditions d'analyse et au type d'ion fragmenté (MH⁺, MNa⁺, ...), cette stratégie requiert, pour être performante, une bibliothèque exhaustive de spectres réalisés dans diverses conditions analytiques.
- | Nous pouvons également envisager une banque de données de structures glycaniques combinée à un algorithme permettant le calcul de leurs fragments théoriques (solution analogue aux moteurs de recherche utilisés pour interpréter les données de MS/MS sur les peptides).

| Une autre solution, que nous appellerons calcul « ab initio », consiste à envisager toutes les combinaisons possibles de résidus monosaccharides ayant la masse du glycanne analysé, puis à calculer les fragments théoriques susceptibles d'être générés par chacune de ces combinaisons afin de les comparer avec les ions fragments expérimentaux. Cette approche présente l'avantage de ne pas recourir à l'utilisation de banques de données. Cependant, étant donnée la variété des possibilités de branchement et de liaisons des oses entre eux, le nombre de combinaisons à envisager [Laine, 1994] peut être un facteur limitant. Aussi, cette solution nécessite un système informatique très puissant.

Quelque soit la solution proposée pour mettre en œuvre la stratégie 2, reste à savoir si les informations tirées des spectres de fragmentation sont suffisamment spécifiques pour permettre une identification univoque d'un motif glycanne. Il est certain que les fragmentations produites à haute énergie permettant des clivages intracycliques sont un atout non négligeable pour la réussite éventuelle de ce type d'approche.

2.3. Stratégie choisie

Compte tenu du contexte dans lequel nous avons démarré notre étude, la caractérisation de N-glycanes de gonadotropines recombinantes, notre objectif n'était pas d'identifier ces glycosylations mais d'en vérifier la structure. De plus, comme mentionné précédemment, nous ne disposons d'aucune banque de données concernant les glycanes. Enfin, nous avons cherché à développer un outil simple dans la mesure de nos moyens informatiques et de nos connaissances, nouvellement acquises, sur les glycosylations.

Pour ces raisons, nous avons opté pour la première des deux stratégies décrites précédemment. En collaboration avec un informaticien, nous avons donc tenté de développer un algorithme de calcul des fragments pouvant être générés à partir d'un motif glycanne connu.

3. Mise en oeuvre

Pour expliquer brièvement la programmation de notre outil, le motif du glycanne peut être comparé à un arbre défini par l'enchaînement des résidus osidiques : l'extrémité réductrice du glycanne correspond à la racine de l'arbre et les antennes aux branches. Les résidus monosaccharides sont caractérisés par la masse qui leur est propre.

L'algorithme est conçu pour parcourir l'arbre depuis la racine jusqu'à l'extrémité des branches en reconnaissant les embranchements (définis comme étant des nœuds) et en numérotant les liaisons entre chaque unité monosaccharide sur la base de la nomenclature définie par Domon et Costello [Domon et Costello, 1988].

Chaque liaison est alors définie comme un site potentiel de clivage permettant de générer deux fragments complémentaires : l'un emportant l'extrémité réductrice, la racine, l'autre emportant les branches. Le calcul de la masse des fragments, Y ou B, est basé sur la somme de la masse des résidus constitutifs du fragment. Nous avons, pour ces calculs, envisagé les cas où le parent et les fragments sont, soit sous forme protonnée (+H), soit sous forme d'adduit de sodium (+Na).

L'algorithme permet également le calcul de la masse des fragments internes générés suite à un double clivage de type Y/B. La difficulté réside ici dans le discernement, parmi toutes les combinaisons possibles, des fragments Y/B ayant un sens : on ne peut, par exemple, combiner un fragment Y du tronc de l'arbre avec un fragment B issu de l'une des branches.

4. Présentation du logiciel

Le logiciel que nous avons développé ainsi qu'un exemple d'utilisation sont présentés figure 19. Les différentes étapes du fonctionnement de ce logiciel peuvent être résumées de la façon suivante :

La première étape consiste à construire la structure étudiée dans la fenêtre de dessin. Nous utilisons à cet effet les symboles de couleurs proposés qui représentent un panel de résidus monosaccharides, dérivatisés ou non, parmi ceux les plus fréquemment rencontrés. Dans le menu « setup », l'utilisateur peut également définir de nouveaux résidus et leur associer un symbole permettant de les discerner. Un bouton de liaison sert à unir les oses entre eux. En dernier lieu, il est nécessaire de désigner l'extrémité réductrice du sucre (en jaune sur la figure) à l'aide de la fonction « Reroot ».

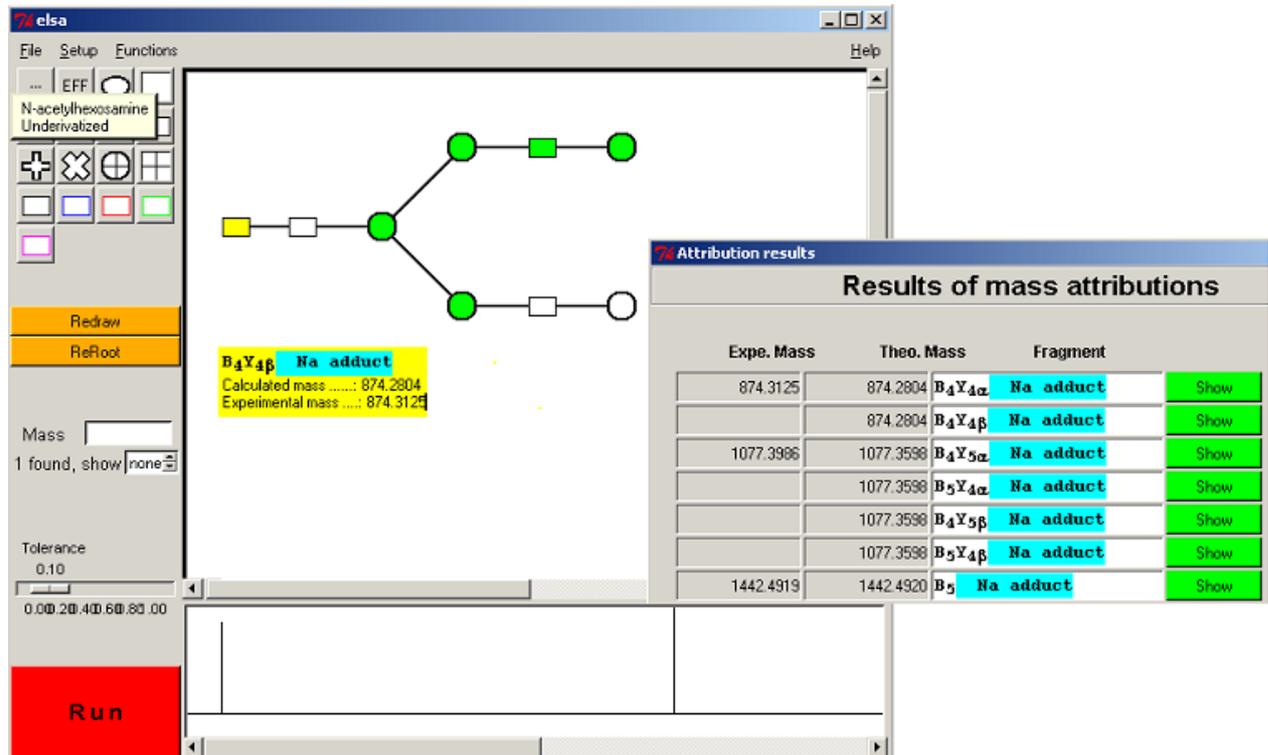
Dans un second temps, l'utilisateur charge la liste de masses des ions fragments qui ont été détectés sur le spectre MS/MS de la structure étudiée. Le format informatique de cette liste dépend de l'instrument utilisé pour l'analyse ; nous avons adapté le logiciel à l'utilisation des fichiers générés avec le TOF-TOF et la trappe ionique Bruker (cf. expériences décrites précédemment). Cette liste contient également des informations relatives à l'intensité des ions qui seront utilisées par la suite.

Nous ajustons ensuite la tolérance : exprimée en Da, elle correspond à l'écart de masse toléré entre les fragments expérimentaux et théoriques pour l'assignation des pics du spectre MS/MS.

Le bouton « Run » permet de lancer le programme :

- | Tous les fragments théoriques, Y, B et Y/B relatifs à la structure étudiée sont calculés.
- | La comparaison des masses théoriques et des masses expérimentales permet de générer une liste de résultats correspondant aux fragments identifiés.
- | Un spectre MS/MS représentatif des ions issus de la liste de masses expérimentales est reconstitué ; les pics qui ont été assignés se distinguent alors par leur couleur. L'intensité relative des fragments est restituée selon les informations contenues dans la liste de masses préalablement soumise au logiciel.
- | Dans la liste des résultats, un bouton « show » permet de visualiser, sur la fenêtre de dessin, le fragment identifié sur la structure du glycanne de départ (en vert sur la figure).

Figure 19 : présentation du logiciel développé et exemple de résultats obtenus pour l'assignation de quelques fragments issus de l'analyse MS/MS avec la trappe ionique du composé NA₂.



5. Conclusion

Nous avons par conséquent initié le développement d'un outil informatique permettant d'assister l'assignation manuelle des fragments observés sur les spectres MS/MS de glycanes.

Ce travail a consisté à envisager, parmi les stratégies possibles, celle qui semblait à la fois réalisable et adaptée à nos besoins. Nous avons également conçu cet outil de manière à ce qu'il soit convivial et ré-utilisable par la suite. Nous souhaitons en effet, à l'issue des dernières améliorations encore nécessaires, le rendre disponible sur internet sous forme d'un programme téléchargeable.

Le point fort de notre approche réside dans le fait que l'algorithme développé permet de s'intéresser à n'importe quel glycanne, aussi complexe soit-il, puisque sa structure et les unités glucidiques qui le composent sont définis par l'utilisateur. Son principal inconvénient est de ne pas permettre la prédiction des fragments intracycliques limitant ainsi l'interprétation des spectres aux fragments de type « basse énergie ». Cependant, le calcul des fragments internes de type Y/B est un sérieux atout puisque ceux-ci sont souvent abondants, difficiles à interpréter et recèlent parfois des informations structurales.

Depuis que nous avons démarré ce travail, une équipe d'informaticiens du département de spectroscopie de l'institut allemand de recherche sur le cancer à Heidelberg a rendu accessible, sur internet, un outil à vocation très similaire, dénommé GLYCO-FRAGMENT [Lohmann et Von der Lieth,

2003]. L'algorithme développé est très puissant puisqu'il calcule à la fois les fragments interglycosidiques et intracycliques d'une structure prédéfinie, rendant possible l'interprétation des spectres de fragmentation à haute énergie. Il ne permet cependant pas de calculer les fragments internes résultant de clivages successifs. Par ailleurs, leur approche s'inscrit dans une stratégie qui vise à développer parallèlement des bibliothèques de glycannes ; elle présente donc l'inconvénient de ne pas permettre la prédiction des fragments de structures absentes de ces banques.

Finalement, il est intéressant d'observer la complémentarité des approches envisagées pour le développement d'un tel programme informatique. Aucun outil aujourd'hui ne permet encore de remplacer complètement l'interprétation manuelle faite par l'analyste. Cependant, avec l'intérêt croissant qui est porté à l'étude des modifications post-traductionnelles en protéomique, et plus spécialement aux glycosylations, de nouveaux programmes devraient certainement voir le jour.

BIBLIOGRAPHIE

Amoresano A., Siciliano R., Orru S., Napoleoni R., Altarocca V., De Luca E., Sirna A., Pucci P.

Structural characterization of human recombinant glycohormones follitropin, lutropin, and choriogonadotropin expressed in Chinese hamster ovary cells, *Eur. J. Biochem.*, **1996**, 242, 608-618.

Carte N.

La trappe ionique et l'ionisation électrospray : un nouveau potentiel pour la caractérisation de biomolécules, *Thèse de l'université Louis Pasteur de Strasbourg*, **2001**.

Domon B., Costello C.E.

A systematic nomenclature for carbohydrate fragmentations in FAB-MS/MS spectra of glycoconjugates, *Glycoconj. J.*, **1988**, 5, 397-409.

Egge H., Peter-Katalinic J.

Fast atom bombardment mass spectrometry for structural elucidation of glycoconjugates, *Mass Spectrom. Rev.*, **1987**, 6, 331-93.

Gillece-Castro B., Burlingame A.L.

Meth. Enzymol., **1990**, 193, 689.

Härd K., Mekking A., Damm J.B.L., Kamerling J.P., De Boer W., Wijands R.A., Vliengenhart J.F.G.

Isolation and structure determination of the intact sialylated N-linked carbohydrate chains of recombinant human follitropin expressed in Chinese hamster ovary cells, *Eur. J. Biochem.*, **1990**, 193, 263-271.

Harvey D.J., Bateman R.H., Green M.R.

High-energy collision-induced fragmentation of complex oligosaccharides ionized by matrix-assisted laser desorption/ionization mass spectrometry, *J. Mass Spectrom.*, **1997**, 32, 167-187.

Harvey D.J.

Matrix-assisted laser desorption/ionization mass spectrometry of carbohydrates, *Mass Spectrom. Rev.*, **1999**, 18, 349-451.

Hofmeister G.E., Zhou Z., Leary J.A.

Linkage position determination in lithium-cationized disaccharides : tandem mass spectrometry and semiempirical calculations, *J. Am. Chem. Soc.*, **1991**, 113, 5964.

Laine R.A.

A calculation of all possible oligosaccharides isomers both branched and linear yields 1.05×10^{12} structures for a reducing hexasaccharide : the *Isomer Barrier* to development of single-method saccharide sequencing or synthesis system, *Glycobiol.*, **1994**, 4, 759-767.

Lemoine J., Chirat F., Domon B.

Structural analysis of derivatized oligosaccharides using post-source decay matrix assisted laser desorption / ionization mass spectrometry, *J. Mass Spectrom.*, **1996**, 31, 908-912.

Lohmann K., Von der Lieth C.W.

GLYCO-FRAGMENT : a web tool to support the interpretation of mass spectra of complex carbohydrates, *Proteomics*, **2003**, 3, 2028-2035.

Pierce J.G., Parsons T.F.

Glycoprotein hormones : structure and functions, *Ann. Rev. Biochem.*, **1981**, 50, 465-495.

Ronin C.

Glycosylation des hormones hypophysaires : polymorphisme structural et biologique en clinique humaine, *Regard sur la biochimie*, **1995**, 3, 12-21.

Sairam M.R.

Gonadotropic hormones : relationship between structure and function with emphasis on antagonists, *Hormonal proteins and peptides*, ed. Li, C.H., Academic press, New York., **1983**, 11, 1-79.

Spengler B., Kirsch D., Kaufmann R., Lemoine J.

Structure analysis of branched oligosaccharides using post-source decay in matrix assisted laser desorption ionization mass spectrometry, *J. Mass Spectrom.*, **1995**, 30, 782-787.

Thotakura N.R., Blithe D.L.

Glycoprotein hormones : glycobiology of gonadotrophins, tyrotrophin and free alpha subunit, *Glycobiol.*, **1995**, 5, 3-10.

Weiskopf A.S., Vouros P., Harvey D.J.

Electrospray ionization-ion trap mass spectrometry for structural analysis of complex N-linked glycoprotein oligosaccharides, *Anal. Chem.*, **1988**, 70, 4441-47.

Weiskopf A.S., Vouros P., Harvey D.J.

Characterization of oligosaccharide composition and structure by quadrupole ion trap mass spectrometry, *Rapid Commun. Mass Spectrom.*, **1997**, 11, 1493-1504.

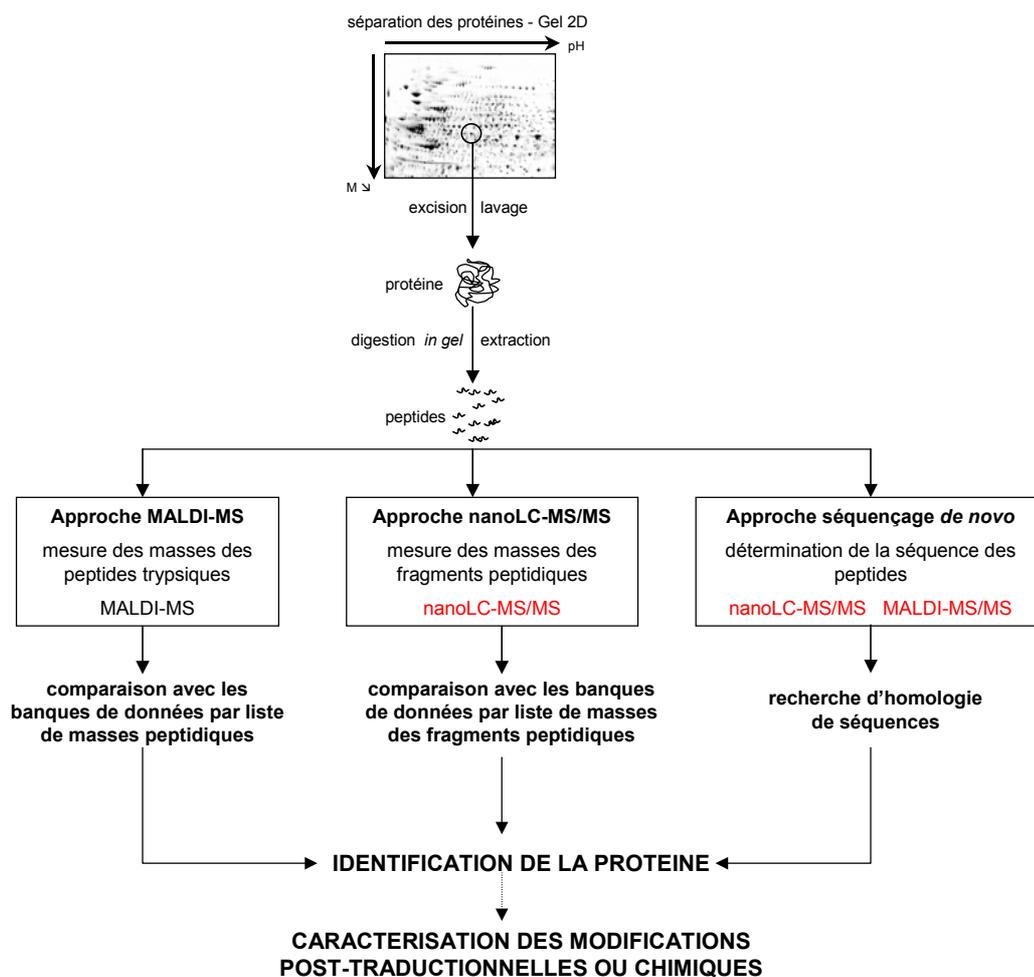
CONCLUSION GENERALE

CONCLUSION GENERALE

Au cours de ce travail de thèse, nous avons pu, grâce à l'intégration de la MS/MS dans l'analyse protéomique, développer de nouvelles approches analytiques permettant la résolution de certains problèmes biologiques pour lesquels la seule approche par MALDI-MS était inefficace.

L'ensemble de ces approches intégrant l'analyse par MS/MS sont les suivantes (figure ci-dessous) :

- | **approche classique MALDI-MS**
- | **approche nanoLC-MS/MS** consistant en la mesure des masses des fragments peptidiques et leur comparaison avec les banques de données
- | **approche par séquençage *de novo*** permettant la détermination de séquences peptidiques pour rechercher des homologues de séquence (dans le cas où les protéines étudiées sont inconnues des banques de données).



De l'identification à la caractérisation des protéines, les différentes stratégies protéomiques. Les techniques développées au cours de ce travail de thèse sont en rouge.

C'est la combinaison de toutes ces approches qui permet aujourd'hui de passer du stade d'identification des protéines à leur caractérisation (modifications post-traductionnelles ou chimiques, séquençage *de novo*) dans le cadre de l'analyse protéomique.

Grâce à ces nouvelles approches de la spectrométrie de masse en tandem, nous avons pu caractériser :

- | des modifications chimiques (oxydation de cystéines) sur des peroxyrédoxines
- | des enzymes purifiées chez les algues brunes, inconnues des banques de données
- | des glycosylations de type N-glycannes.

► Sur le plan instrumental, l'intégration de la MS/MS dans l'analyse protéomique a nécessité **l'optimisation des paramètres du couplage nanoLC-MS/MS avec une trappe ionique** (ESQUIRE 3000+, Bruker Daltonics). Ce type d'analyse, entièrement automatisée, permet l'obtention d'informations de séquences pour des peptides présents dans des quantités allant de 10 femtomoles à 1 picomole de peptides, ce qui correspond à la quantité de peptides statistiquement extraits des gels 2D colorés au bleu, ou, de façon intense, à l'argent.

L'approche développée est aujourd'hui incontournable en protéomique : il s'agit d'une stratégie complémentaire ou alternative à la stratégie associant gel 2D et MALDI-MS pour l'identification des protéines.

De plus, nos efforts ont porté sur la mise au point de méthodes d'analyse permettant une caractérisation fine des protéines (étude des modifications post-traductionnelles ou chimiques) ainsi que le séquençage *de novo* pour l'identification par homologie de séquence dans le cas où les protéines étudiées sont absentes des banques de données.

► Dans le cadre de la caractérisation des protéines et de leurs modifications post-traductionnelles, **nous avons également développé une approche MS/MS avec un instrument de nouvelle génération, un MALDI-TOF-TOF (ULTRAFLEX, Bruker Daltonics), pour l'étude structurale des glycosylations.**

Dans ce contexte, nous avons évalué et comparé les capacités du MALDI-TOF-TOF et d'une ES-trappe ionique (ESQUIRE 3000+, Bruker Daltonics) pour l'étude, par MS/MS, de N-glycannes connus.

- | La trappe ionique a permis de générer (à partir de solutions de glycannes purs de concentrations comprises entre 10 et 50 picomoles / μ L) des spectres relativement simples à interpréter contenant principalement des fragments issus de coupures des liaisons interglycosidiques, donnant des informations de séquences des structures étudiées.
- | Le TOF-TOF, moins sensible (50 picomoles déposées), a quant à lui, donné des spectres plus complexes avec des fragments interglycosidiques et intracycliques provenant de mécanismes de dissociation à haute énergie. Ces derniers se sont avérés particulièrement intéressants pour la détermination de certaines isoméries de liaison, et parfois du motif de branchement du glycanne.

| Les deux techniques présentent l'inconvénient de générer également des fragments internes résultant de clivages multiples du glycanne et compliquant l'interprétation des spectres.

Cette étude nous a conduit à **initier la conception d'un outil informatique permettant d'assister l'assignation manuelle des fragments observés sur les spectres MS/MS de glycannes**. Nous sommes ainsi intervenus sur le plan du développement bio-informatique en protéomique qui est un aspect indispensable contribuant à l'évolution des techniques de spectrométrie de masse. En effet, la quantité et la complexité des données générées peuvent rapidement devenir un facteur limitant de l'interprétation analytique.

Ces avancées instrumentales nous ont ensuite permis de résoudre plusieurs problèmes biologiques.

► **Caractérisation par approche protéomique de protéines impliquées dans le système de défense des cellules contre le stress oxydant.**

1. Dans le cadre de l'étude de peroxyrédoxines (prxs), nous avons mis au point une stratégie permettant de détecter, de localiser et de caractériser la sur-oxydation d'une cystéine (en acide cystéine sulfonique ou sulfonique) contenue dans une séquence protéique. Cette approche, reposant sur l'utilisation d'une endoprotéase spécifique, l'AspN, permet de vérifier sans ambiguïté l'état d'oxydation de cystéines appartenant à des isoformes protéiques se distinguant par leur point isoélectrique lors de l'analyse par gel 2D.

La complémentarité de plusieurs techniques de spectrométrie de masse a été nécessaire à l'aboutissement de cette étude : MALDI-MS en modes d'ionisation positif et négatif, nanoES-MS/MS. Dans ce contexte, l'utilisation d'une matrice originale (DHAP + DAHC) pour l'analyse MALDI-MS en mode d'ionisation négatif a été particulièrement intéressante.

2. Nous avons ensuite étudié la régénération des prxs suite à un stress oxydant en utilisant un marquage *in vivo* des protéines suivi d'une approche protéomique quantitative.

La lysine deutérée (d4 ou d8) s'est avérée être un marqueur de choix pour notre étude. En qualité d'acide aminé essentiel, elle nous a permis d'étudier les mécanismes biologiques impliqués dans le renouvellement des prxs avec une grande spécificité.

Nous avons testé deux approches pour la quantification par spectrométrie de masse : MALDI-MS et nanoLC-MS/MS. Elles ont toutes deux permis d'aboutir aux mêmes conclusions sur le plan biologique : nous avons observé une large variété dans les vitesses de régénération des prx étudiées ; certaines prx se régénèrent lentement (prx III et VI) tandis que d'autres se régénèrent rapidement (1 cys prx). Par ailleurs, cette étude a également permis de mettre en évidence un phénomène inattendu : alors que la plupart des prxs étudiées se renouvèlent uniquement par néosynthèse, prxII se régénère également, en partie, grâce à un mécanisme de rétro-réduction.

► Caractérisation par approche séquençage *de novo* d'enzymes purifiées chez les algues brunes

L'utilisation de la nanoLC-MS/MS s'est avérée déterminante pour la caractérisation par séquençage *de novo* de deux types d'haloperoxydases purifiées chez les Laminaires et ayant des activités distinctes, bromo et iodoperoxydases.

- | Nous disposions d'informations préalables concernant les bromoperoxydases (ADNc et données protéiques pour une activité similaire chez d'autres familles d'algues). Nous avons donc recherché des homologues de séquence avec ces données pré-existantes.
- | Par contre, ne disposant d'aucune information concernant les iodoperoxydases, nous avons développé une approche plus ciblée consistant, dans un premier temps, à obtenir par séquençage *de novo* des séquences suffisamment longues et spécifiques pour permettre la construction de « primers dégénérés » afin de cloner des ADNc correspondant à l'activité iodoperoxydase. Dans un second temps, nous avons vérifié l'homologie de séquence entre l'ADNc cloné et la protéine étudiée sur le gel.

La mise en œuvre d'une stratégie par séquençage *de novo* a permis, d'une part, d'attribuer une fonction biochimique aux deux ADNc de bromoperoxydases déjà caractérisés et, d'autre part, de cloner de nouveaux ADNc correspondant à l'activité iodoperoxydase. Nous avons ainsi pu établir un lien entre les activités enzymatiques des protéines étudiées et les données moléculaires obtenues à partir des sporophytes de *Laminaria digitata*.

PARTIE EXPERIMENTALE

PARTIE EXPERIMENTALE

I. LES SPECTROMETRES DE MASSE

1. Spectrométrie de masse MALDI

1.1. Appareillage utilisé

Les mesures de masses par MALDI-TOF ont été réalisées avec des instruments et des logiciels de la marque Bruker (Bruker Daltonics GmbH, Brême, Allemagne) :

- | le BIFLEX, équipé du logiciel d'acquisition XACQ (version 5.0)
- | l'ULTRAFLEX TOF-TOF, permettant également les analyses MS/MS d'ions métastables, équipé du logiciel FLEXcontrol (version 1.2) pour l'acquisition des spectres.

Le logiciel d'exploitation des données équipant ces spectromètres est XMASS (version 5.0). L'interprétation des spectres MS/MS obtenus avec le TOF-TOF peut être réalisée à l'aide du logiciel Biotools 2.2.

Ces instruments possèdent un système de déplacement de la cible selon les axes x-y par rapport au laser. Une caméra permet de visualiser le dépôt et l'impact des tirs laser.

La désorption/ionisation de l'échantillon est réalisée avec un laser à azote ($\lambda = 337$ nm) dont la fréquence est ajustable entre 1 et 20 Hz. La fluence du faisceau peut être adaptée selon les applications. Les deux instruments sont pourvus d'un système à extraction retardée.

Un spectre de masse est généré à chaque tir laser. Selon les échantillons, il peut être nécessaire d'accumuler plusieurs spectres (en général de 20 à 100) afin d'obtenir un bon rapport signal sur bruit.

1.2. Les principaux paramètres de l'analyse MALDI-TOF

Pour les analyses de peptides et d'oligosaccharides, nous avons principalement travaillé en mode réflecteur sur la gamme des m/z inférieurs à 5 kDa. Les tensions appliquées aux lentilles d'accélération et de focalisation ont été ajustées en conséquence :

- | pour le BIFLEX, IS1 = 19 kV, IS2 = 15,65 kV, lens = 9,75 kV et réflecteur = 20 kV
- | pour l'ULTRAFLEX, IS1 = 20 kV, IS2 = 17,5 kV, lens = 5 kV et réflecteur = 21 kV.

1.3. Les principaux paramètres de l'analyse MALDI-TOF-TOF

La technologie LIFT permet l'analyse d'ions fragments résultant de décompositions métastables. Les ions formés dans la source MALDI sont accélérés à l'aide d'une différence de potentiel de 8 kV et se fragmentent dans la première partie du tube de vol. L'ion parent et ses ions fils sont ré-accelérés dans la cellule LIFT avec une tension de 19 kV. L'ion parent est ensuite éliminé et les fragments sont analysés dans la seconde partie du tube de vol, en mode réflecteur. Un gaz de collision, de l'argon, peut être introduit à la sortie de la source pour accentuer la fréquence des fragmentations métastables.

1.4. L'étalonnage du spectromètre de masse MALDI-TOF

Deux types d'étalonnage sont possibles :

| **L'étalonnage interne** est celui qui permet d'obtenir la meilleure précision de mesure. Un mélange étalon est incorporé à l'échantillon et l'ensemble est co-cristallisé avec la matrice. L'assignation des composés de référence permet le calcul automatique des constantes de la courbe d'étalonnage qui sert ensuite au calcul de la masse des composés analysés. Le principal inconvénient de cette méthode est le risque de suppression du signal propre aux analytes, lié à une trop forte concentration en étalons.

| **L'étalonnage externe** consiste à déposer la solution étalon en un endroit distinct de l'échantillon sur la cible. Dans un premier temps, l'acquisition d'un spectre de masse du mélange standard permet le calcul des constantes d'étalonnage. Les échantillons sont ensuite analysés, en ayant soin de reproduire les mêmes conditions analytiques que celles utilisées pour le mélange étalon, et leurs masses sont calculées d'après les constantes préalablement définies. Cette méthode présente l'avantage d'éviter les phénomènes de suppression du signal liés au mélange de l'étalon avec l'échantillon. Cependant la mesure de masse est moins précise qu'avec un étalonnage interne.

Pour l'analyse de peptides tryptiques en protéomique, deux des fragments issus de l'autolyse de la trypsine ($m/z = 842,51$ et $2211,11$) ont servi d'étalons internes.

Pour les autres analyses, nous avons utilisé un mélange de peptides standards constitué de Leu-enképhaline-arginine, Angiotensine, Substance P, Bombésine et ACTH. La masse monoisotopique de chacun de ces composés, sous forme $(M+H)^+$, a alors servi de référence (cf. tableau 2).

1.5. La préparation de l'échantillon

Les matrices les plus couramment employées pour nos analyses sont :

- | l'acide α -cyano-4-hydroxycinnamique (CHCA) pour l'analyse de peptides
- | l'acide 2,5 dihydroxybenzoïque (DHB) pour l'analyse d'oligosaccharides.

Les modes de dépôt utilisés sont le dépôt en sandwich ou la goutte séchée.

2. Spectrométrie de masse à source électrospray

Deux types de spectromètres de masse à source électrospray ont été utilisés lors de ce travail de thèse : un instrument hybride, le Q-TOF et une trappe ionique.

2.1. Le spectromètre de masse Q-TOF

Les analyses ont été effectuées sur un appareil de type électrospray Q-TOF à géométrie orthogonale, le Q-TOF II (Micromass Ltd, Altricham, UK). L'acquisition et le traitement des données ont été réalisés avec le logiciel MassLynx (version 3.5).

Les ions sont produits au niveau d'une source à géométrie en Z et introduits dans le spectromètre de masse par un orifice (« cône d'échantillonnage » ou « skimmer ») de 0,4 mm de diamètre.

2.1.1. Introduction des échantillons pour les analyses nano-ESI et nanoLC-ESI

Parmi les différents types de sources adaptables au Q-TOF II, nous avons utilisé les sources nanospray et LC nanospray.

En mode nano-ES, l'échantillon (1 à 5 μL) est placé dans un capillaire en quartz métallisé de 1 à 3 μm de diamètre interne (Protana, Odense, Danemark). De l'azote, introduit à l'entrée du capillaire, permet de pousser le liquide et d'amorcer la formation du spray.

En mode LC nano-ES, la nano-colonne analytique est connectée directement à l'entrée d'un capillaire en silice de 20 μm de diamètre interne. Un gaz coaxial (N_2) facilite la nébulisation de l'échantillon.

Dans les deux cas, le capillaire est ajusté perpendiculairement à l'entrée du spectromètre de masse devant le cône d'échantillonnage.

2.1.2. Etalonnage du Q-TOF

Pour étalonner le Q-TOF, nous utilisons du Glu-fibrinopeptide B (F3261 ; Sigma-Aldrich, St Quentin Fallavier, France). L'espèce doublement chargée correspondant à ce peptide (m/z 785,8) est sélectionnée par le quadripôle et fragmentée à 30 eV dans la cellule de collision. Les fragments ainsi générés sont analysés par le TOF ; l'instrument est alors étalonné sur la base de leurs masses sur une gamme de 50 à 1300 m/z .

2.2. La trappe ionique

Nous avons travaillé avec une trappe ionique ESQUIRE 3000+ (Bruker Daltonics GmbH, Brême, Allemagne) équipée soit :

- | avec une source microspray orthogonale Hewlett Packard
- | avec une source nanospray linéaire dite « off-line » (Bruker Daltonics GmbH, Brême, Allemagne)
- | avec une source nanospray orthogonale « on-line » (Bruker Daltonics GmbH, Brême, Allemagne).

L'acquisition des spectres de masse et le traitement des données sont réalisés, respectivement, par les logiciels Bruker ESQUIRE Control (version 5.1) et Data Analysis (Version 3.1).

Les spectres sont enregistrés en moyennant 10 à 100 balayages (« scans ») de la gamme de m/z . Ce nombre de balayages augmente significativement lorsque le degré de MS^n croît.

2.2.1. Les performances de l'ESQUIRE 3000+

L'ESQUIRE 3000+ permet l'analyse sur trois gammes de m/z distinctes qui diffèrent par la vitesse de balayage. La résolution de l'analyseur est étroitement liée à ces paramètres ; la résolution diminue lorsque la gamme de m/z est étendue et lorsque la vitesse de balayage augmente (tableau 1).

Tableau 1 : résolution de l'ESQUIRE 3000+ selon la vitesse de balayage et la gamme de m/z .

gamme de m/z	résolution		vitesse de balayage ($m/z / s$)
	largeur des pics à mi-hauteur	résolution maximale	
standard (50 – 3000)			
normal	0,6	5000	13000
enhanced	0,45	6600	5500
maximum	0,3	10000	1650
extended (200 – 6000)	4	1500	27000

2.2.2. Introduction des échantillons

La source microspray comprend un capillaire métallique, appelé aiguille de nébulisation, de 500 μm de diamètre interne. Elle permet d'introduire les échantillons dans l'ESQUIRE 3000+ à un débit d'environ 2 à 4 $\mu\text{L} / \text{min}$, soit en mode de couplage avec un système HPLC capillaire, soit avec un pousse seringue (Harvard Apparatus Inc., South Natick, U.S.A.).

En mode nanoES off-line, 1 à 5 μL d'échantillon sont introduits dans un capillaire en quartz métallisé de 1 à 3 μm de diamètre interne (Protana, Odense, Danemark). Le capillaire est ajusté face à l'entrée de l'interface à l'aide d'un système de vis micrométriques. Un gaz de nébulisation (N_2) permet d'amorcer la formation du spray.

La configuration de la source nanopray on-line est décrite dans le chapitre « Mise en œuvre expérimentale – couplage avec une trappe ionique ».

2.2.3. Etalonnage de la trappe ionique

L'étalonnage de la trappe ionique sur la gamme de masse de 50 à 1700 m/z est réalisée à partir des ions multi-chargés issus de l'analyse d'un mélange de cinq peptides (dilués dans une solution H₂O / ACN, 1 :1 v/v, + 0,1% HCOOH) dont les concentrations sont mentionnées dans le tableau 2.

Tableau 2 : états de charge et rapports m/z des peptides du mélange utilisé pour l'étalonnage de la trappe ionique.

Peptides	Concentrations (pmol / μ L)	Masses monoisotopiques		
		(1+)	(2+)	(3+)
Leu-enképhaline-arginine	4	712,38	356,69	/
Angiotensine	2,5	1046,54	523,77	/
Substance P	5	1347,74	674,37	449,91
Bombésine	4	1620,81	810,90	/
ACTH	3,5	/	1233,10	822,40

2.3. La préparation des échantillons pour l'analyse nanoES

Le dessalage des échantillons est une étape incontournable de l'analyse nanoES-MS off-line. Nous utilisons à cet effet un Zip Tip C18 (Millipore Corporation, Bedford, U.S.A.) qui est un cône contenant 0,6 μ L de phase C18 avec des particules de silice sphériques (15 μ m, 200Å). La quantité de matériel pouvant être retenue sur le Zip Tip est de l'ordre de 2 μ g. Il peut être utilisé sur une gamme de pH allant de 2 à 13. Le protocole de dessalage sur Zip Tip peut être résumé de la façon suivante :

- | La première étape correspond à l'activation de la colonne. Le Zip Tip est rincé avec une solution ACN/H₂O (1/1, v/v) ; il s'agit de chasser l'air contenu dans la phase.
- | Le Zip Tip est ensuite conditionné pour permettre par ailleurs la fixation des peptides. Cet étape de conditionnement consiste à rincer la phase avec une solution d'eau acidifiée (H₂O + 0.1% HCOOH) afin d'éliminer toute trace d'acétonitrile.
- | Le pipetage de l'échantillon, comprenant une série d'étapes d'aspiration et de refoulement, permet ensuite aux peptides de se fixer sur la colonne.
- | Une solution (H₂O + 0.1% HCOOH) est utilisée pour permettre l'élution des sels.
- | La dernière étape correspond à l'élution des peptides : ils sont décrochés de la phase par une solution d'acétonitrile (ACN/H₂O, 40/60, v/v + 0.1% HCOOH).

II. LES SYSTEMES CHROMATOGRAPHIQUES

1. Le MicroBlotter

Les éléments constitutifs du système HPLC sont les suivants :

- | une pompe ABI 140D comprenant deux seringues en céramique permettant de délivrer un mélange binaire
- | un injecteur et un four ABI 112A
- | un détecteur UV ABI 785A comprenant une cellule capillaire (1 μ L).

Nous avons équipé le système d'une colonne adaptée à l'analyse des peptides (LC Packings, PepMap C18, 0.3 \times 150 mm, 3 μ m, 100 Å) et travaillé à un débit de 4 μ L/min.

2. Le système Agilent

2.1. L'HPLC

Les différents modules constituant le système chromatographique utilisé pour les montages LC capillaire et nanoLC sont les suivants :

- | une pompe HP 1100 Series Capillary LC System ou HP 1100 Series Nano LC System
- | une pompe binaire travaillant en mode isochratique à un débit de 20 μ L/min, G1312A
- | un dégazeur, G1379A
- | un injecteur automatique thermostaté pouvant contenir des plaques ou des flacons et équipé d'une vanne à 6 voies, G1377A
- | un compartiment colonne thermostaté doté d'une vanne à 6 voies, G1316A
- | un détecteur à barrettes de diodes, G1315A, équipé d'une cellule de 500 nL et de trajet optique 10 mm (utilisé uniquement pour le montage microLC).

Nous avons utilisé les colonnes analytiques suivantes :

- | en LC capillaire: Agilent ZORBAX SB-C18, 0,3 \times 150 mm, 5 μ m, 300 Å
- | en nanoLC : Agilent ZORBAX SB-C18, 0,075 \times 150 mm, 3,5 μ m, 300 Å.

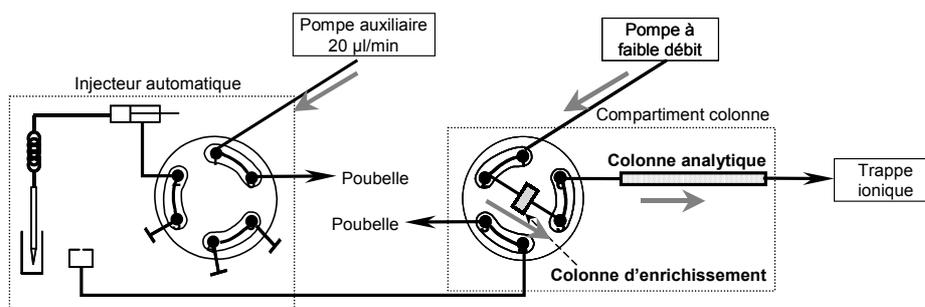
Nous avons, par ailleurs, utilisé une colonne dite d'enrichissement ou de pré-concentration pour notre montage : Agilent ZORBAX SB-C18, 0,3 \times 5 mm, 5 μ m, 300 Å.

2.2. Le système de pré-concentration

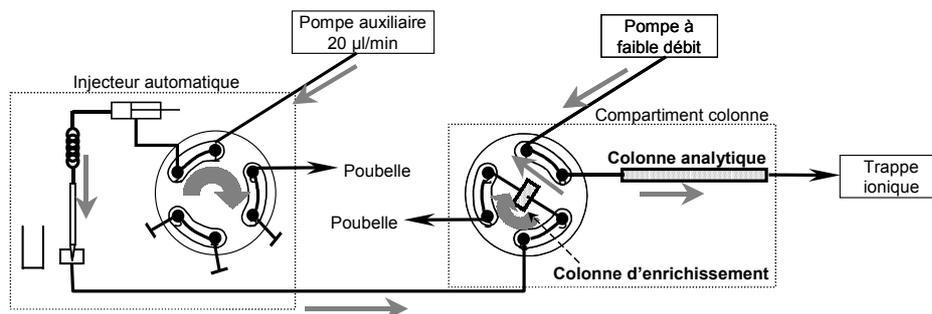
La configuration du montage de pré-concentration est décrite figure 1.

Figure 1 : configuration du montage de pré-concentration et étapes de l'analyse avec le système Agilent.

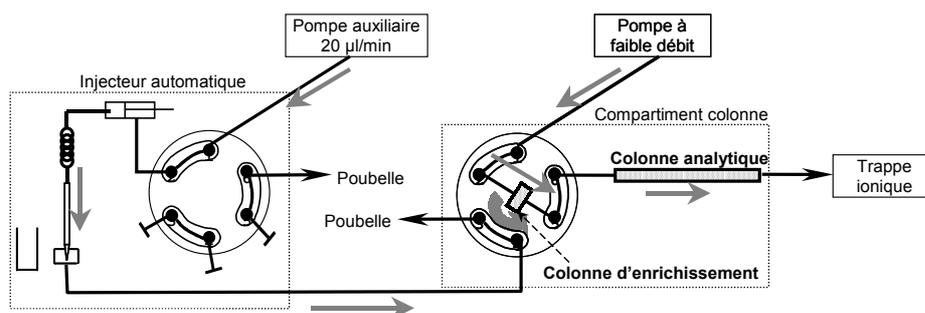
Etape 1 : chargement. L'échantillon est prélevé dans un vial ou une plaque 96 puits (deux boucles montées en série permettent d'injecter jusqu'à 16 μL). Les deux colonnes sont balayées par le solvant délivré par la pompe capillaire ou la nanoLC (5% B pour l'équilibration).



Etape 2 : rinçage. Les deux vannes ont basculé. La pompe auxiliaire ($\text{H}_2\text{O} + 0.1\% \text{HCOOH}$) entraîne l'échantillon sur la colonne d'enrichissement. Les analytes sont ainsi piégés sur la phase en tête de pré-colonne et le flux de solvant aqueux permet de rincer l'échantillon : les sels sont éliminés vers la poubelle pendant 5 minutes. La colonne analytique est toujours dans le circuit de la pompe à faible débit (5% B).



Etape 3 : analyse. La vanne du compartiment colonne bascule à nouveau. La colonne analytique et la colonne d'enrichissement sont dans le circuit de la pompe à faible débit. Les peptides sont dirigés vers la colonne analytique dans le sens inverse de leur chargement. Au fur et à mesure de leur élution de la colonne analytique, ils sont analysés par la trappe ionique.



2.3. Les gradients testés

Les solvants utilisés pour le gradient sont les suivants :

- | A : H₂O + 0.1% HCOOH
- | B : ACN + 0.1% HCOOH.

Tableau 3 : gradients de chromatographie testés dans le cadre de l'approche protéomique à haut débit par nanoLC-MS/MS.

Gradient LONG			Gradient COURT		
Temps (min)	%B	Débit (nL / min)	Temps (min)	%B	Débit (nL / min)
5.0	5.0	200	5.0	5.0	200
65.0	1.5%B/m	200	37.0	45.0	200
70.0	90.0	200	38.0	95.0	200
80.0	5.0	200	47.0	95.0	200
100.0	5.0	200	48.0	5.0	200
			63.0	5.0	200

1%B/min

3. La CapLC

La CapLC (Micromass Ltd, Manchester, UK) comprend un injecteur automatique et un système de pré-concentration (figure 2).

Tout d'abord, l'échantillon est rincé et concentré sur une colonne d'enrichissement (LC-Packings, PepMap C18, 0,3 × 1mm, 5 µm, 100 Å) durant 3 minutes à un débit de 30 µL / min avec une solution H₂O + 0,1% HCOOH (solvant C).

Les peptides retenus sur la pré-colonne sont ensuite élués en sens inverse vers la colonne analytique (LC-Packings, PepMap C18, 0,075 × 150 mm, 3 µm, 100 Å).

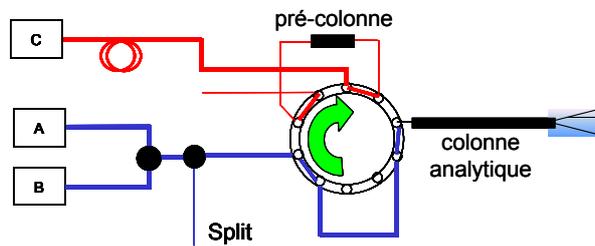
L'éluion des peptides est réalisée à un débit de 200 nL / min. Les solvants A et B sont les mêmes que ceux utilisés avec le système Agilent (cf. § II.2.3.).

Le gradient analytique comprend les étapes suivantes :

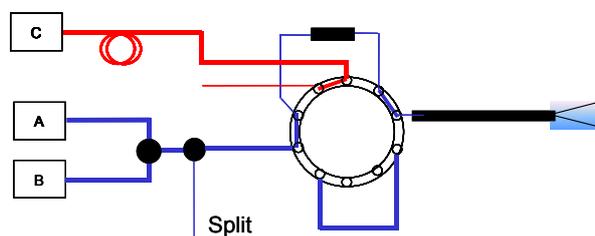
- | passage de 5 à 45% de solvant B en 35 minutes
- | 95% de solvant B durant 5 minutes
- | équilibration de la colonne pendant 20 minutes avec 95% de solvant A.

Figure 2 : configuration du montage de pré-concentration et analyse avec la CapLC.

1. Chargement de l'échantillon sur la colonne de pré-concentration



2. Elution des peptides de la pré-colonne vers la colonne analytique



III. L'ANALYSE PROTEOMIQUE

1. Réduction / alkylation / digestion

Pour l'analyse protéomique des protéines séparées sur gels mono et bi-dimensionnels, nous avons mis au point un protocole de réduction / alkylation / digestion *in gel*.

L'acquisition d'un robot MassPREP Station (Micromass, Manchester, UK) nous a permis d'automatiser chacune de ces étapes. Les deux avantages principaux du robot sont :

- | le gain de temps (environ 5h avec le robot – hors temps de digestion - pour une plaque de 96 spots, contre 2 jours pour l'exécution manuelle du protocole)
- | la diminution du risque de contamination des spots (notamment par de la kératine) grâce à l'automatisation de toutes les étapes ; le risque de pollution lié à l'expérimentateur reste cependant possible lors de la préparation des divers réactifs et de la manipulation du gel.

1.1. Lavage des spots

Les spots d'intérêt sont excisés du gel et placés dans une plaque à 96 puits. Ils sont ensuite lavés avec 100 μ L d'un mélange ACN / 25mM NH_4HCO_3 (1/1, v/v) pendant 10 minutes. Cette opération est répétée 2 fois. Les spots sont ensuite déshydratés avec 50 μ L d'ACN (5 minutes).

1.2. La réduction et l'alkylation

La solution permettant la réduction des ponts disulfures contient 10 mM de dithiothreitol (D9779 ; Sigma-Aldrich, St Quentin Fallavier, France) dilué dans un tampon NH_4HCO_3 25mM. 50 μL de cette solution sont déposés sur chacun des morceaux de gel qui sont ensuite laissés à incuber pendant 1h à 57°C. L'alkylation est réalisée avec 50 μL d'une solution d'iodoacétamide 55mM (I1149 ; Sigma-Aldrich, St Quentin Fallavier, France) préparée dans NH_4HCO_3 25mM. Après 20 minutes, l'excédent de réactif est éliminé et les morceaux de gels sont finalement lavés et déshydratés.

1.3. La digestion trypsique

Les protéines sont digérées à l'aide d'une solution contenant 12,5 ng/ μL de trypsine (Promega Corporation, Madison, U.S.A.) dans NH_4HCO_3 25mM. Le volume de la solution d'enzyme déposé par puits est fonction de la taille des spots (5 à 10 μL). La réaction de digestion se déroule à température ambiante pendant une nuit.

2. L'extraction des peptides

L'étape d'extraction doit permettre la diffusion passive des peptides du gel vers la solution d'extraction. Sa qualité, déterminante pour l'analyse par spectrométrie de masse, dépend principalement de deux paramètres : le temps d'extraction et le solvant utilisé.

Pour nos études, deux protocoles d'extraction ont été utilisés :

- | Pour l'analyse MALDI-MS, 5 μL d'une solution ACN/ H_2O (60/35, v/v) + 5% HCOOH sont ajoutés dans chaque puits. La diffusion des peptides hors du gel dure 60 minutes. La solution d'extraction ainsi obtenue est directement déposée sur la cible MALDI.
- | Pour l'analyse LC-MS/MS, 30 μL d'une solution ACN/ H_2O (60/35, v/v) + 5% HCOOH sont ajoutés dans chaque puits pendant 60 minutes. Le volume de la solution d'extraction est ensuite réduit à environ 8 ou 16 μL , selon le système HPLC utilisé, par évaporation de l'acétonitrile à température ambiante ou sous un flux d'azote. Cette solution est ainsi injectée sur la colonne analytique.