



**HAL**  
open science

# ETUDES STRUCTURALES DES PROTEINES PAR SPECTROMETRIE DE MASSE COUPLEE AUX ECHANGES HYDROGENE/DEUTERIUM ET A LA RETICULATION CHIMIQUE

Laetitia Cravello

► **To cite this version:**

Laetitia Cravello. ETUDES STRUCTURALES DES PROTEINES PAR SPECTROMETRIE DE MASSE COUPLEE AUX ECHANGES HYDROGENE/DEUTERIUM ET A LA RETICULATION CHIMIQUE. Sciences pharmaceutiques. Université Joseph-Fourier - Grenoble I, 2005. Français. NNT: . tel-00009698

**HAL Id: tel-00009698**

**<https://theses.hal.science/tel-00009698>**

Submitted on 7 Jul 2005

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

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

**THESE**

soutenue le 31 mars 2005 par

**Laetitia CRAVELLO**

pour l'obtention du titre de

Docteur de l'Université Joseph Fourier - Grenoble 1

Spécialité Biologie Structurale

**ETUDES STRUCTURALES DES PROTEINES PAR SPECTROMETRIE  
DE MASSE COUPLEE AUX ECHANGES HYDROGENE/DEUTERIUM  
ET A LA RETICULATION CHIMIQUE**

Composition du jury

*Rapporteurs* : Gérard BOLBACH

Bernard MONSARRAT

*Examineurs* : Andréa DESSEN

Eric FOREST

Eva PEBAY-PEYROULA

*A mes parents, à mon frère,*

Je tiens à exprimer ma profonde gratitude à :

Monsieur le Docteur Eric Forest (Laboratoire de Spectrométrie de Masse des Protéines, IBS, Grenoble),

*Pour m'avoir accueillie dans son laboratoire durant ces quatre années et pour avoir supervisé une partie de ce travail de thèse.*

Madame le Docteur Andréa Dessen (Laboratoire de Cristallographie Macromoléculaire, IBS, Grenoble),

*Pour avoir supervisé l'étude sur PcrV et PcrG, pour sa grande gentillesse et sa disponibilité.*

Monsieur le Docteur Gérard Bolbach (Université Pierre et Marie Curie, Paris) et Monsieur le Docteur Bernard Monsarrat (Institut de Pharmacologie et de Biologie Structurale, Toulouse),

*Pour l'honneur que vous me faites d'examiner mon travail de thèse.*

Madame le Professeur Eva Pebay-Peyroula (Directrice de l'Institut de Biologie Structurale, Grenoble),

*Pour avoir accepté de participer à ce jury de thèse.*

David Lascoux (Laboratoire de Spectrométrie de Masse des Protéines, IBS, Grenoble),

*Pour m'avoir appris ce qu'était la spectrométrie de masse, pour avoir participé à de nombreuses manipulations de ce travail de thèse. Pour ton amitié sincère et pour tous les bons moments passés ensemble, et il y en aura encore beaucoup d'autres, j'en suis sûre (merci aussi pour les cafés, l'hébergement, les remorquages...).*

Tous mes remerciements vont également à :

Markus Muller (Institute of Biotechnology, Zurich),

*Pour la création des logiciels de traitement des données obtenues par échange H/D et par réticulations chimiques.*

Christophe Geourjon et Sophie Misserey (Institut de Biologie et Chimie des Protéines, Lyon),

*Pour la partie bioinformatique du projet de détermination de la famille de repliement d'une protéine de structure inconnue.*

Gille Subra (Laboratoire des Aminoacides Peptides et Protéines, Montpellier),

*Pour la partie synthèse organique dans les réticulations chimiques.*

Manuelle Quinaud (Laboratoire de Cristallographie Macromoléculaire, IBS, Grenoble),

*Pour la structure modélisée de PcrV.*

Anne-Marie Di Guilmi (Laboratoire d'Ingénierie des Macromolécules, IBS, Grenoble),

*Pour m'avoir fourni la protéine PBP-2X*

Sébastien Brier et David Lemaire (Laboratoire de Spectrométrie de Masse des Protéines, IBS, Grenoble),

*Pour les discussions techniques, et surtout pour votre sympathie. (et pas merci de m'avoir martyrisé toutes ces années !!!)*

L'ensemble du Laboratoire de Spectrométrie de Masse des Protéines : Olivier Barré, Linda Berardi, Florestan Desmaris, Bernard Dublet, Armelle Hubert, Evelyne Laborie, Cédric Lastella, Hortense Mazon, Florence Sidaner, David et Jean Smith, Carol Teschke et Ali Tiss,  
*Pour tous les bons moments passés ensemble,*

Merci encore à :

Mes parents et mon frère, Florian,

*Pour tout l'amour qu'ils m'ont donné. Je n'en serais pas là sans vous.*

Ma famille,

*Pour leur soutien dans les bons comme dans les mauvais moments.*

Mes amis grenoblois, parisiens, toulousains, spinaliens, clermontois et autres expatriés,

*Pour m'avoir supporté malgré mes défauts (et c'est pas toujours facile) et avoir su apprécier mes qualités.*

## GLOSSAIRE

BS3 : disuccinimidyl suberate

Da : Dalton

DST : disuccinimidyl tartarate

EGS : ethyleneglycol bis-(succinimidylsuccinate)

ESI : ElectroSpray Ionization

GT : GlycosylTransferase

H/D : Hydrogène / Deutérium

LC-MS : chromatographie liquide couplée à la spectrométrie de masse

Lys-C : Endoprotéinase Lys-C

MALDI : Matrix Assisted Laser Desorption Ionization

MS : Mass Spectrometry (spectrométrie de masse)

MS-MS : Mass Spectrometry-Mass Spectrometry (spectrométrie de masse tandem)

P. aeruginosa : Pseudomonas aeruginosa

PBP : Penicilin Binding Protein

PDB : Protein Data Bank

RMN : Résonance Magnétique Nucléaire

TOF : Time Of Flight (temps de vol)

TP : TransPeptidase

Tris : Tris(hydroxyméthyl)méthylglycine

TTSS : Système de Sécrétion de Type III

**INTRODUCTION GENERALE ..... 1**

**CHAPITRE 1 : LA SPECTROMETRIE DE MASSE ..... 4**

INTRODUCTION..... 5

1 LES SOURCES ..... 7

**1.1 L'electrospray** ..... 7

1.1.1 La formation du spray. .... 8

1.1.2 La fission des gouttelettes ..... 9

1.1.3 La formation des ions en phase gazeuse ..... 9

*1.1.3.1 Le modèle de Dole*..... 9

*1.1.3.2 Le modèle d'Iribarne et Thomson* ..... 10

1.1.4 Le calcul de la masse..... 10

**1.2 Le MALDI**..... 12

2 LES ANALYSEURS ..... 14

**2.1 Les analyseurs de type filtre quadripolaire.** ..... 14

**2.2 Les analyseurs de type piège à ions (trappe ionique)**..... 17

2.2.1 Structure de la trappe..... 18

2.2.2 Le piégeage des ions..... 19

2.2.3 L'éjection des ions..... 20

**2.3 L'analyseur temps de vol**..... 21

2.3.1 Principe général ..... 21

2.3.2 L'extraction retardée ..... 22

2.3.3 Le réflectron ..... 23

2.3.4 L'injection orthogonale ..... 24

3 LA FRAGMENTATION DES IONS ..... 25

**3.1 Principe général**..... 25

**3.2 La fragmentation dans l'espace** ..... 26

3.2.1 Les triples quadripôles (QqQ) ..... 26

3.2.2 Les Qq-TOF ..... 26

3.2.3 Les TOF-TOF ..... 28

**3.3 La fragmentation dans le temps (cas des trappes ioniques)** ..... 30

**3.4 La fragmentation des peptides** ..... 31

## **CHAPITRE 2 : OPTIMISATION DE LA METHODE DES ECHANGES**

### **H/D COUPLES A LA SPECTROMETRIE DE MASSE..... 33**

INTRODUCTION.....	34
1 LE MODELE D'ETUDE : LA PROTEINE PBP-2X.....	36
<b>1.1 Contexte de l'étude : la bactérie Streptococcus pneumoniae.....</b>	<b>36</b>
<b>1.2 Les « Penicillin Binding Proteins ».....</b>	<b>37</b>
1.2.1 La description des PBPs.....	37
1.2.1.1 Les PBPs de haute masse moléculaire de classe A.....	38
1.2.1.2 Les PBPs de haute masse moléculaire de classe B.....	38
1.2.1.3 Les PBPs de basse masse moléculaire.....	39
1.2.2 Les rôles des PBPs.....	39
<b>1.3 Les PBPs et les <math>\beta</math>-lactamines.....</b>	<b>42</b>
1.3.1 Le mode d'action des $\beta$ -lactamines.....	42
1.3.2 La résistance des PBPs aux $\beta$ -lactamines.....	44
1.3.2.1 Les $\beta$ -lactamases.....	44
1.3.2.2 La modification des PBPs.....	45
<b>1.4 La protéine PBP-2X.....</b>	<b>45</b>
2 LES ECHANGES H/D.....	48
<b>2.1 La théorie des échanges H/D.....</b>	<b>48</b>
2.1.1 Les échanges H/D dans les protéines et dans les peptides.....	49
2.1.2 L'influence du pH.....	51
2.1.3 L'influence de la température.....	52
2.1.4 L'influence de la séquence locale en acides aminés.....	52
<b>2.2 Protocoles expérimentaux.....</b>	<b>53</b>
2.2.1 Les techniques de marquage.....	53
2.2.2 L'étape de blocage des rééchanges (ou « quenching »).....	54
2.2.3 La digestion enzymatique.....	54
<b>2.3 Applications des échanges H/D couplés à la spectrométrie de masse dans l'étude des protéines.....</b>	<b>56</b>
2.3.1 Etude du repliement des protéines.....	56
2.3.2 Etude des interactions protéine/protéine et protéine/ligand.....	57
2.3.3 Facilitation de la cristallogénèse.....	58
3 RESULTATS ET DISCUSSION.....	59
<b>3.1 Buts et principes de l'étude.....</b>	<b>59</b>

3.2	Recouvrement de la séquence.....	60
3.3	Deutération.....	62
3.4	Validation des résultats sur la structure tridimensionnelle.....	67
4	PERSPECTIVE D'OPTIMISATION DE LA METHODE .....	70
5	CONCLUSION .....	72

## **CHAPITRE 3 : MISE AU POINT DE LA TECHNIQUE DES RETICULATIONS CHIMIQUES COUPLEES A LA SPECTROMETRIE DE MASSE..... 73**

INTRODUCTION.....	74
1 LA RETICULATION CHIMIQUE .....	75
1.1 Les différents agents réticulants .....	75
1.1.1 Réactivité des agents réticulants.....	75
1.1.1.1 Réticulation sur les amines réactives .....	75
1.1.1.2 Réaction sur les thiols .....	79
1.1.1.3 Réticulation non sélectives .....	81
1.1.2 Les bras espaceurs .....	81
1.1.3 Le design des agents réticulants .....	83
1.1.3.1 Les agents homobifonctionnels .....	83
1.1.3.2 Les agents hétérobifonctionnels .....	83
1.1.3.3 Les agents trifonctionnels.....	84
1.2 Stratégie et applications .....	84
1.2.1 La réticulation intramoléculaire .....	84
1.2.1.1 Stratégie.....	84
1.2.1.2 Les produits obtenus.....	86
1.2.1.3 Applications.....	88
1.2.2 La réticulation intermoléculaire .....	89
1.2.2.1 Stratégie.....	89
1.2.2.2 Applications.....	91
1.3 L'analyse par spectrométrie de masse des produits réticulés .....	91
1.4 Amélioration de la détection des peptides réticulés .....	93
1.4.1 Les marquages isotopiques.....	93
1.4.1.1 Marquage isotopique de l'agent réticulant .....	93
1.4.1.2 Marquage isotopique des protéines .....	94

1.4.2	La fluorescence.....	95
1.4.3	Les agents réticulants clivables .....	97
1.4.4	Les systèmes de purification .....	98
2	BUTS ET PRINCIPES DE L'ETUDE.....	99
3	RESULTATS ET DISCUSSION.....	101
<b>3.1</b>	<b>La réticulation intramoléculaire .....</b>	<b>101</b>
3.1.1	Les conditions de réticulation.....	101
3.1.2	Les produits de réticulation .....	103
<b>3.2</b>	<b>Recherche des peptides pontés .....</b>	<b>105</b>
3.2.1	Identification des peptides pontés par calculs mathématiques.....	105
3.2.2	Identification des peptides pontés par MS-MS .....	106
3.2.3	Détermination des contraintes de distance .....	109
<b>3.3</b>	<b>Exploitation des données par bioinformatique.....</b>	<b>112</b>
4	PERSPECTIVES D'AMELIORATION DE LA METHODE.....	115
<b>4.1</b>	<b>L'augmentation du nombre de contraintes de distances .....</b>	<b>115</b>
4.1.1	Les agents réticulants immobilisés.....	115
4.1.2	Utilisation d'autres agents et d'autres enzymes .....	116
<b>4.2</b>	<b>L'amélioration de la précision des contraintes de distances .....</b>	<b>117</b>
<b>4.3</b>	<b>L'amélioration du logiciel de traitement bioinformatique.....</b>	<b>117</b>
5	CONCLUSION .....	118

## **CHAPITRE 4 : ETUDES STRUCTURALES DES PROTEINES PCRV ET PCRG ET CARACTERISATION DU COMPLEXE PCRV-PCRG..... 119**

	INTRODUCTION.....	120
1	LE CONTEXTE BIOLOGIQUE DE L'ETUDE .....	121
<b>1.1</b>	<b>La bactérie Pseudomonas aeruginosa.....</b>	<b>121</b>
<b>1.2</b>	<b>Le système de sécrétion de type III.....</b>	<b>121</b>
1.2.1	Les molécules effectrices .....	123
1.2.2	Le secreton .....	123
1.2.3	Le translocon .....	124
1.2.3.1	<i>L'étape de translocation.....</i>	<i>124</i>
1.2.3.2	<i>La régulation du TTSS.....</i>	<i>125</i>
<b>1.3</b>	<b>Les protéines PcrV et PcrG .....</b>	<b>125</b>
1.3.1	La protéine PcrV .....	125
1.3.2	Le complexe PcrV-PcrG .....	127

2	RESULTATS ET DISCUSSION .....	128
2.1	<b>Etudes structurales des protéines PcrV et PcrG seules .....</b>	<b>128</b>
2.1.1	La protéine PcrV .....	128
2.1.1.1	<i>Deutération de la protéine PcrV .....</i>	<i>128</i>
2.1.1.2	<i>Réticulation chimique intramoléculaire de la protéine PcrV .....</i>	<i>133</i>
2.1.2	Deutération de la protéine PcrG .....	136
2.2	<b>Caractérisation du complexe PcrV-PcrG par échange H/D.....</b>	<b>139</b>
3	CONCLUSIONS ET PERSPECTIVES .....	146

**CONCLUSION GENERALE .....** **147**

**MATERIELS ET METHODES .....** **150**

1	MATERIELS.....	151
1.1	<b>La protéine PBP-2X .....</b>	<b>151</b>
1.2	<b>Le cytochrome c.....</b>	<b>151</b>
1.3	<b>Les protéines PcrV et PcrG .....</b>	<b>151</b>
2	METHODES .....	153
2.1	<b>L'étude de la protéine PBP-2X.....</b>	<b>153</b>
2.1.1	Digestion de la protéine.....	153
2.1.2	Réalisation des cartes peptidiques par LC/MS/MS .....	153
2.1.3	Deutération des peptides et analyses LC/MS .....	154
2.1.4	ESI-MS and ESI-MS-MS .....	154
2.2	<b>Les réticulations chimiques sur le cytochrome c .....</b>	<b>155</b>
2.2.1	L'étape de réticulation.....	155
2.2.2	contrôle de la réticulation .....	155
2.2.3	Digestion peptidique.....	155
2.2.4	Mesures de la masse des produits de réticulation.....	156
2.2.4.1	<i>Mesures par MALDI-TOF .....</i>	<i>156</i>
2.2.4.2	<i>Mesures par LC-MS et LC-MS-MS .....</i>	<i>156</i>
2.2.4.3	<i>Mesures par MALDI-TOF-TOF .....</i>	<i>157</i>
2.3	<b>L'étude de PcrV, PcrG et du complexe PcrV/PcrG .....</b>	<b>157</b>
2.3.1	Cartes peptidiques des protéines .....	157
2.3.2	La deutération des protéines et du complexe .....	158
2.3.3	La réticulation chimique de PcrV .....	158
2.3.4	Détermination de la masse des produits de réticulation .....	159

2.3.4.1	<i>Mesures par LC-MS</i> .....	159
2.3.4.2	<i>Mesures par MALDI-TOF</i> .....	159

**REFERENCES BIBLIOGRAPHIQUES..... 160**

**ANNEXES..... 175**

# INTRODUCTION GENERALE

Les protéines sont des composants essentiels de tous les organismes vivants. Impliquées dans la plupart des processus biologiques, elles peuvent être des cibles thérapeutiques et aussi des médicaments potentiels.

Les projets de génomique et de protéomique ont permis de grandes avancées dans la compréhension de leur fonctionnement. De nombreux génomes sont maintenant séquencés et les gènes dont sont issues les protéines sont ainsi déterminés. Des nouvelles protéines sont identifiées chaque jour dans différents types cellulaires ou dans différents états de l'organisme. Cependant l'identification et la localisation des protéines ne suffisent pas à comprendre la totalité de leur mode d'action. Des études biochimiques et structurales sont alors indispensables.

L'établissement de la relation structure/fonction des protéines permet la compréhension de leur rôle dans les processus biologiques. De plus, les protéines n'agissent que rarement seules et sont souvent impliquées dans des complexes avec d'autres protéines, avec des acides nucléiques ou avec des petites molécules. Différentes méthodes sont actuellement disponibles pour étudier la structure des protéines, comme par exemple la cristallographie ou la RMN. Même si ces méthodes sont extrêmement performantes, elles présentent tout de même des limites. La cristallographie nécessite le passage de la protéine sous la forme d'un cristal, la RMN peut difficilement étudier des molécules dont la taille est supérieure à 40 kDa, et les deux méthodes requièrent des quantités importantes d'échantillon. De plus, au regard du nombre croissant de protéines identifiées pour lesquelles l'obtention d'une structure est désirée, l'analyse systématique de toutes ces protéines nécessiterait un temps considérable. Il a donc été nécessaire de développer d'autres méthodes alternatives complémentaires des méthodes physico-chimiques classiques. C'est dans ce cadre que la spectrométrie de masse a montré tout son potentiel. Associée aux échanges hydrogène/deutérium (H/D) ou la réticulation chimique, elle permet l'obtention de données structurales intéressantes.

Le travail réalisé durant cette thèse a consisté à développer des méthodes innovantes utilisant la spectrométrie de masse pour étudier la structure des protéines et à appliquer ces méthodes à une problématique biologique. Après une description de la spectrométrie de masse, qui permettra d'apprécier toute la potentialité de cet outil analytique, nous décrivons la méthode des échanges H/D associés à la spectrométrie de masse. Nous développerons une nouvelle technique, basée sur la méthode classique, à partir d'une protéine modèle : la protéine PBP-2X. Puis nous nous intéresserons aux réticulations chimiques. Cette méthode récemment associée à la spectrométrie de masse sera testée et optimisée sur le cytochrome c.

Enfin, nous pourrons appliquer ces deux nouvelles méthodes à une problématique biologique. La structure de deux protéines, PcrV et PcrG, impliquées dans le système de sécrétion de type III de *Pseudomonas aeruginosa*, sera étudiée et le complexe qu'elles peuvent former sera caractérisé.

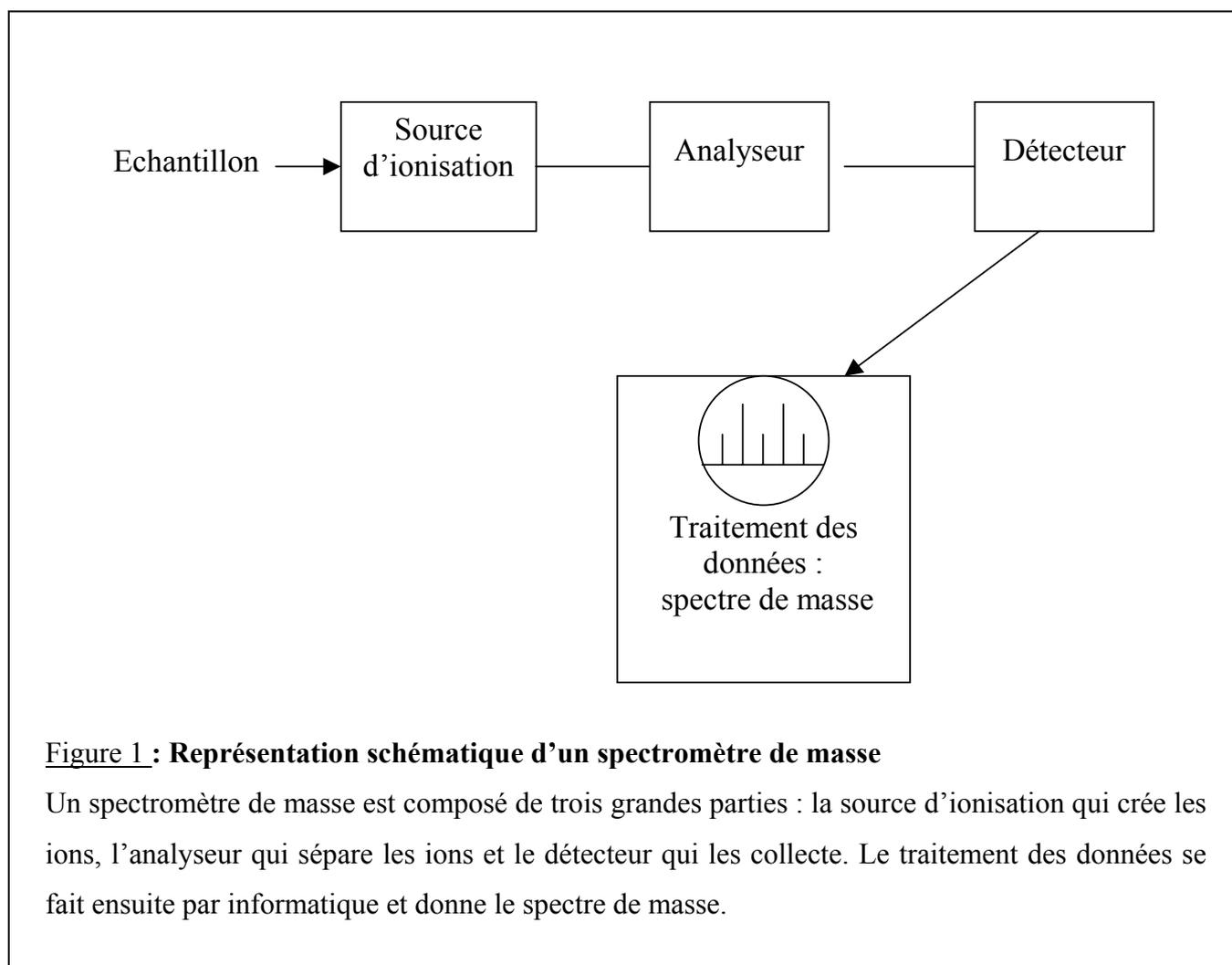
CHAPITRE 1 :  
LA SPECTROMETRIE DE  
MASSE

## INTRODUCTION

La spectrométrie de masse est une méthode analytique qui permet de déterminer la masse moléculaire d'un composé chimique ou biologique.

Cette technique a une place privilégiée dans le domaine de l'analyse grâce à sa faible consommation d'échantillon (parfois de l'ordre de la femtomole) et donc sa sensibilité, à sa grande résolution, et à la relative rapidité des analyses.

Un spectromètre de masse est classiquement composé de trois grandes parties : la source d'ionisation, l'analyseur et le détecteur (figure 1). Une molécule subit une ionisation au niveau de la source. Au niveau de l'analyseur, les ions sont séparés en fonction de leur rapport masse sur charge ( $m/z$ ). Enfin le détecteur collecte ces ions, quantifie leur intensité et amplifie le signal. Ces dernières étapes se déroulent dans un vide poussé pour éviter toute collision entre les ions et les molécules de gaz.



**Figure 1 : Représentation schématique d'un spectromètre de masse**

Un spectromètre de masse est composé de trois grandes parties : la source d'ionisation qui crée les ions, l'analyseur qui sépare les ions et le détecteur qui les collecte. Le traitement des données se fait ensuite par informatique et donne le spectre de masse.

Après le détecteur, un système informatique permet le traitement des données et génère un spectre de masse qui précise la variation du courant ionique observé en fonction du rapport  $m/z$  et permet de déterminer la masse moléculaire de l'espèce analysée.

Depuis le développement des sources d'ionisation dites douces, le MALDI et l'electrospray, par Tanaka et Fenn au cours des années 1980 (Prix Nobel 2002) (Fenn et al., 1989 ; Tabet et Rebuffat, 2003), la spectrométrie de masse a connu un essor considérable dans le monde de la biologie. Les applications dans ce domaine sont très nombreuses et notamment dans l'étude des protéines.

La spectrométrie de masse peut en effet intervenir en contrôle-qualité pour déterminer la pureté et l'homogénéité d'une protéine, ou être utilisée en protéomique pour identifier les protéines (Gevaert et Vanderkerckhove, 2000). Elle est également un outil de choix dans les études structurales pour la vérification fine d'une séquence, la localisation de certains motifs, les études de changement de conformation et de repliement, la caractérisation de complexes covalents ou non et de leurs zones d'interaction.

Les sources, analyseurs et détecteurs existants peuvent être associés de manière différente et ainsi créer une grande variété d'appareil. Le choix d'un type de source ou d'un type d'analyseur dépendra de la nature de l'échantillon à analyser et du type de donnée souhaité (sensibilité, résolution et précision de mesure).

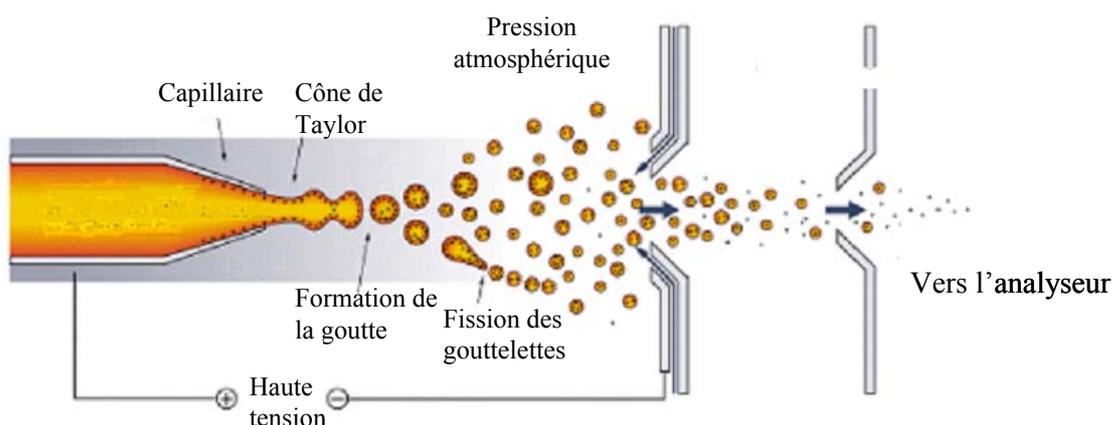
Dans cette étude nous avons utilisé essentiellement quatre types d'appareils : un spectromètre de masse de type electrospray-triple quadripôle (API III, Perkin Elmer-Sciex), un spectromètre de masse de type electrospray-piège ionique (Esquire 3000+, Bruker Daltonics), un spectromètre de masse de type electrospray-Q-TOF (Q-TOF Micro, Waters), et un spectromètre de masse de type MALDI-TOF (Voyager Elite XL, Applied Biosystems).

# 1 LES SOURCES

## 1.1 *L'electrospray*

Les sources de type electrospray permettent la production d'ions en phase gazeuse à partir d'un échantillon en solution. Le fait de partir d'un échantillon en solution est le grand avantage de ce type de source. En effet la majeure partie des réactions chimiques et biologiques se déroulant en phase liquide, l'analyse de ces réactions peut se faire de façon simple (après seulement un dessalage ou une séparation par exemple). De plus, les sources electrospray peuvent facilement être couplées à des techniques de séparation de type chromatographie liquide ou électrophorèse capillaire.

Dans le mode d'ionisation electrospray, l'échantillon est solubilisé dans un solvant organique à pH acide (ou basique) et amené par un capillaire dans la source à pression atmosphérique. L'extrémité du capillaire étant portée à un haut potentiel, il en résulte un champ électrostatique entre l'extrémité du capillaire et la contre-électrode. Les ions migrent à la surface du liquide, un spray est formé. Le processus d'electrospray peut être assisté d'un gaz nébuliseur. Les gouttelettes chargées sont ensuite désolvatées, puis séchées par un flux d'azote chauffé circulant à contre courant. Ceci entraîne une série d'éclatement de la goutte et des ions en phase gazeuse sont finalement produits (Kearle et Tang, 1993).



**Figure 2 : Représentation schématique d'une source electrospray**

La haute tension fait migrer les charges à l'extrémité du capillaire. Des gouttelettes chargées sont créées. Suite à leur désolvatation, elles subissent des explosions coulombiennes, se transforment en microgouttelettes puis en nanogouttelettes et enfin en ions en phase gazeuse. Les ions sont ensuite transférés vers l'analyseur.

Le processus de création d'ion par électrospray peut se décomposer en trois grandes étapes : la formation du spray, la fission des gouttelettes et enfin la formation des ions en phase gazeuse. (figure 2)

Les ions ainsi formés sont ensuite transférés vers l'analyseur.

### 1.1.1 La formation du spray.

Le spray est formé par la différence de potentiel entre l'extrémité du capillaire et la contre-électrode (de 2 à 6kV). Le champ électrostatique qui en résulte induit la migration des ions positifs et négatifs. Si le capillaire est considéré comme étant la cathode, les ions positifs vont migrer vers l'extrémité du capillaire. Les forces de répulsion des charges vont alors augmenter et dépasser la tension de surface du solvant. Il y a formation d'un ménisque qui va s'étendre en suivant la ligne des champs électrique, c'est le cône de Taylor (Figure 3). Lorsque le champ électrique à l'extrémité de ce cône sera assez important, des fines gouttelettes seront émises. La source électrospray peut être apparentée à une cellule électrolytique dans laquelle se passe une réaction d'oxydoréduction entre le capillaire et le solvant. Ce mode de création de gouttelette est appelé « mécanisme électrophorétique ».

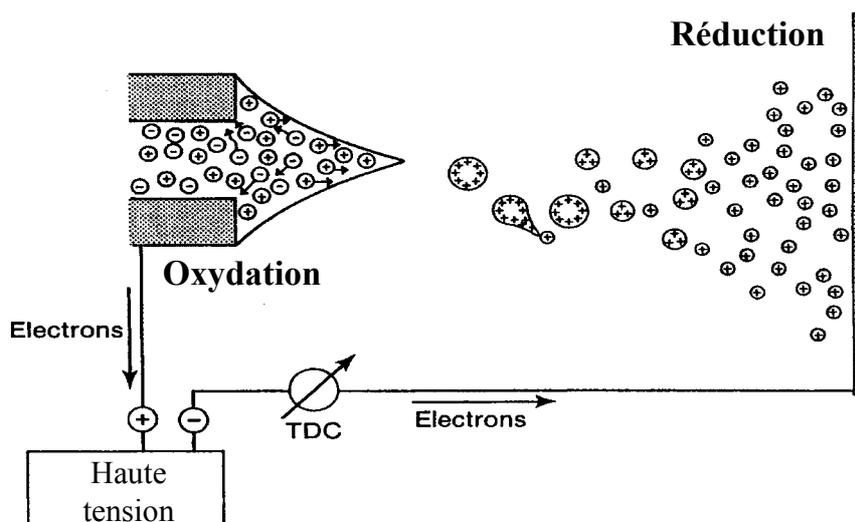


Figure 3 : Représentation schématique de la création du spray

La source est assimilée à une cellule électrolytique  
D'après Wong et al., 1988

## 1.1.2 La fission des gouttelettes

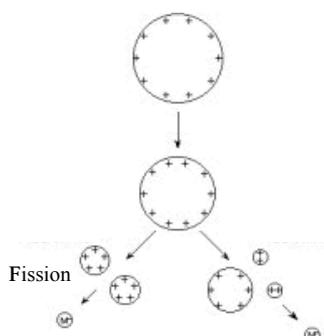
Les gouttelettes du spray sont progressivement évaporées avec l'aide du gaz rideau d'azote à contre-courant et de la température à l'interface. Les charges se concentrent alors dans la goutte dont le rayon diminue. Les forces de répulsion coulombiennes augmentent jusqu'à devenir supérieures aux forces de cohésion de la goutte. On atteint alors la limite de Rayleigh, et une fission coulombienne a lieu. Ce phénomène se répète plusieurs fois, les gouttelettes deviennent des micro-gouttelettes puis des nano-gouttelettes.

## 1.1.3 La formation des ions en phase gazeuse

Lorsque la goutte atteint un diamètre de 10 nm environ, on ne connaît pas les phénomènes exacts qui conduisent à l'ion en phase gazeuse. Plusieurs modèles ont été proposés, les deux principaux sont le modèle de Dole et le modèle d'Iribarne et Thomson.

### 1.1.3.1 Le modèle de Dole.

Il estime que si le processus de division par fission coulombienne va assez loin, et si la solution est suffisamment diluée, on finit par obtenir une goutte qui contient une seule molécule avec une ou plusieurs charges. Cette molécule passe alors sous phase gazeuse après évaporation du solvant (Dole et al., 1968). C'est le modèle de la goutte sèche. (figure 4)

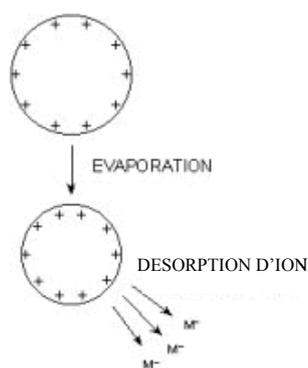


**Figure 4 : Le modèle de Dole**

L'ion en phase gazeuse est produit suite à une ultime fission coulombienne

### 1.1.3.2 Le modèle d'Iribarne et Thomson

Ils ne penchent pas pour un éclatement de la gouttelette mais pour une évaporation ionique. L'ion serait alors extrait par désorption de champ (Iribarne et Thomson, 1976), (figure 5).



**Figure 5 : Le modèle d'Iribarne et Thomson**

Le champ électrique à la surface de la goutte est si intense que l'ion est extrait par désorption de champ.

Il n'est pas possible de conclure sur un mécanisme précis quant à la création de l'ion en phase gazeuse, chaque modèle présentant des limites. Le modèle de Dole paraît compliqué à appliquer à des solutions très concentrées, quant au modèle d'Iribarne et Thomson, il est difficilement concevable pour des macromolécules biologiques pour des raisons cinétiques et thermodynamiques. Dans la pratique, le processus de création des ions pourrait dépendre de l'analyte et se rapprocher des deux modèles à la fois selon les cas.

### 1.1.4 Le calcul de la masse

L'ionisation de type electrospray est caractérisée par la formation d'ions multichargés. Les ions sont déjà présents en solution sous forme d'anions ou de cations. En mode positif, les charges sont apportées par des protons  $H^+$  ou par des cations  $Na^+$ ,  $K^+$ ,  $NH_4^+$ , le plus souvent issus du solvant. Dans le cas de l'analyse des peptides, on considère que les charges positives

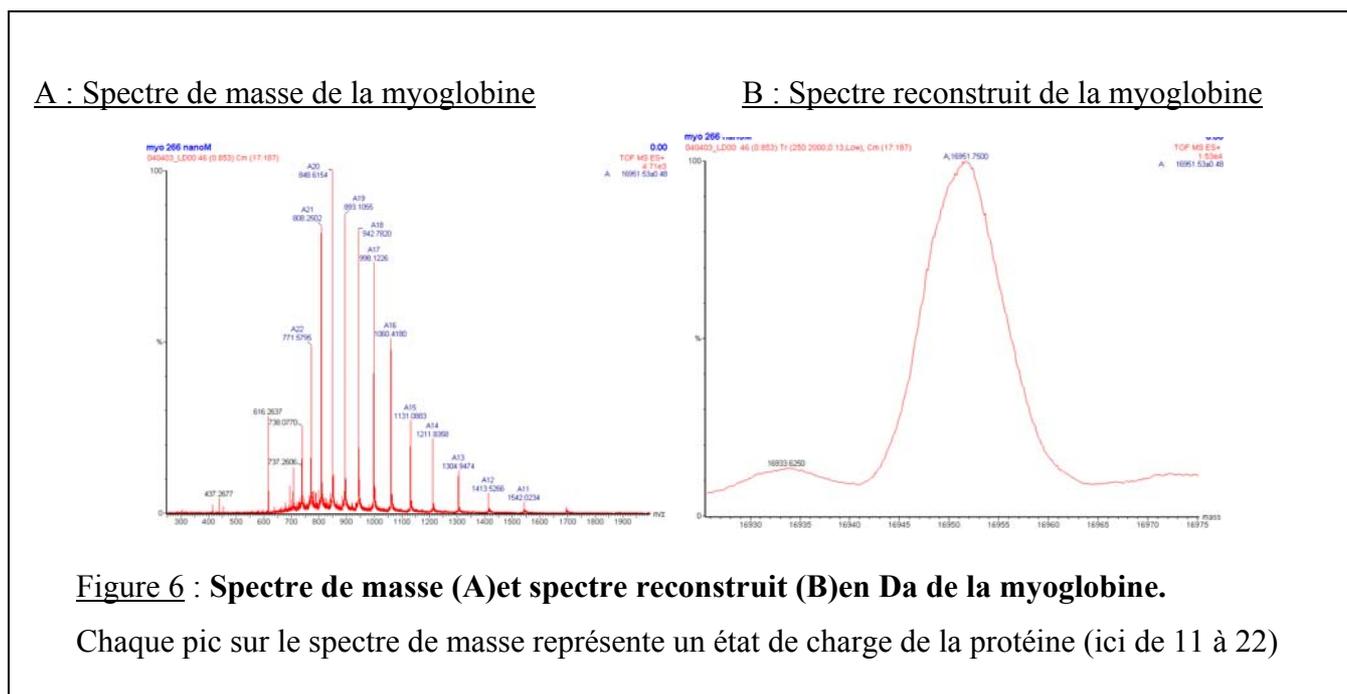
sont issues de la protonation du groupement NH<sub>2</sub> terminal et des résidus basiques (lysine, arginine et histidine) ou de la neutralisation des charges négatives. Les charges négatives dans le cas des analyses en mode négatif sont liées à la déprotonation des résidus acides (acide aspartique, acide glutamique) et du groupement COOH terminal ou à des contre-ions (Loo et al., 1988).

Le spectre de masse d'une protéine correspond aux intensités relatives des ions multichargés en fonction de leur rapport m/z. Cette répartition statistique de tous les états de charge de l'échantillon a généralement l'allure d'une courbe de Gauss. La masse moléculaire est calculée à partir de deux états de charge consécutifs grâce à la résolution d'un système d'équations.

$$\begin{cases} m1=(M+n)/n \\ m2=(M+n+1)/(n+1) \end{cases}$$

Avec m<sub>1</sub> et m<sub>2</sub> les valeurs de m/z de deux pics consécutifs (avec m<sub>1</sub>>m<sub>2</sub>), M la masse du produit analysé et n l'état de charge du pic m<sub>1</sub>. Cette équation n'est valable que dans le cas où la charge est apportée par un proton.

Des logiciels de déconvolution permettent de réaliser ce calcul et d'obtenir ainsi un spectre déconvolué (ou spectre reconstruit) de la protéine qui représentera cette fois l'intensité en fonction du poids moléculaire et non plus en fonction du rapport m/z (figure 6).



## 1.2 Le MALDI

L'ionisation de type MALDI (Matrix-assisted laser desorption/ionisation) est un autre type d'ionisation douce très utilisé dans l'analyse des protéines.

La technique MALDI permet le passage de molécules très polaires et/ ou thermiquement fragiles à l'état solide sous forme d'ions intacts à l'état gazeux.

Les produits à analyser sont dissous dans une matrice qui possède une forte absorption à la longueur d'onde du laser utilisée, et vont ainsi pouvoir transmettre l'énergie du laser aux analytes à ioniser. L'irradiation du mélange par le laser va provoquer le dépôt d'énergie dans la phase condensée par excitation électronique des molécules de la matrice. Il y a formation d'un plasma et les ions sont désorbés par transfert de protons entre la matrice photoexcitée et la substance analysée (figure 7). Ce mécanisme de transfert est encore aujourd'hui mal défini mais pourrait résulter de réactions acide/base entre la matrice et l'analyte (Busch et al., 1995). Lorsque les ions sont formés, ils peuvent être accélérés par l'application d'un champ électrostatique.

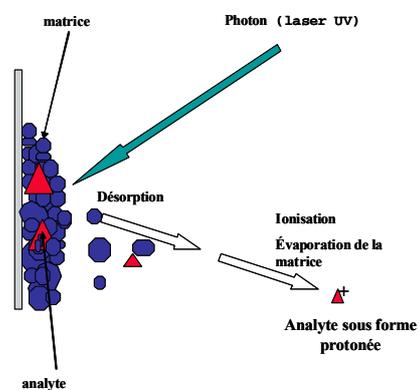


Figure 7 : Schéma du processus de désorption/ionisation

Le mélange matrice/échantillon (avec un large excès de matrice) est déposé sur une plaque en métal. Après évaporation du solvant, la matrice va alors co-cristalliser avec l'analyte. Le diamètre de la surface irradiée est de l'ordre de 100 à 500  $\mu\text{m}$ . Ainsi, l'impact faible du laser permet de transmettre une grande quantité d'énergie par unité de matrice, en minimisant

toutefois la dégradation potentielle de l'échantillon que le faisceau laser incident aurait pu générer. Le type de laser le plus souvent utilisé est le laser à azote qui a une longueur d'onde de 337 nm. Le choix de la matrice est très important, il va dépendre du laser utilisé, mais aussi du composé à analyser (et notamment de la gamme de masse dans laquelle il se situe). Les plus couramment utilisées pour l'analyse des protéines et des peptides sont l'acide 2,5-dihydroxybenzoïque (2,5-DHB) (Karas et al., 1990) et les dérivés de l'acide cinnamique (Beavis et al., 1989 et 1992), comme l'acide férulique, l'acide sinapinique et l'acide  $\alpha$ -cyano-4-hydroxycinnamique.

L'intérêt majeur de la technique d'ionisation MALDI est de créer essentiellement des ions monochargés (ou faiblement chargés), et d'ainsi faciliter l'interprétation des spectres de masse obtenus.

La technique d'ionisation de type MALDI est constamment améliorée, notamment par l'émergence de techniques d'ionisation dérivées. Les nouvelles sources de type AP-MALDI (MALDI à pression atmosphérique) (Laiko et al., 2000), DIOS (Desorption/ Ionization On Silicon), (Wei et al., 1999) et AP-DIOS (Laiko et al., 2002), qui n'utilisent pas de matrice, ont permis d'élargir les possibilités et les applications des sources de type MALDI en biologie.

Quel que soit le processus d'ionisation, une fois l'ion créé, il est transféré vers l'analyseur. Ce transfert doit s'accompagner d'un gradient de pression dans l'interface source/analyseur, réalisé par différentes étapes de pompage. La transmission du faisceau d'ion est assistée par des quadripôles, hexapôles ou octopôles qui servent de guides et par des lentilles qui refocalisent le faisceau.

## **2 LES ANALYSEURS**

Les analyseurs ont pour fonction de séparer les ions précédemment créés en fonction de leur rapport  $m/z$ . Il existe une grande variété d'analyseurs. Les plus courants sont des analyseurs dits à balayage qui détectent les ions successivement au cours du temps. De plus en plus d'appareils sont équipés d'analyseurs hybrides, associant plusieurs types d'analyseurs, et combinant ainsi les spécificités intéressantes de chacun. Le choix d'un analyseur va dépendre principalement :

de sa gamme de masse (la valeur limite du rapport  $m/z$  mesurable)

de sa précision dans la mesure de la masse

de sa résolution, c'est la capacité à différencier des ions de rapport  $m/z$  voisins.

Dans cette étude, les analyseurs que nous avons utilisés sont de type quadripôle, trappe ionique, temps de vol, couplage quadripôle/temps de vol et temps de vol/temps de vol.

### ***2.1 Les analyseurs de type filtre quadripolaire.***

Le principe de cet analyseur est d'utiliser la stabilité des trajectoires des ions dans le quadripôle en fonction des tensions qui y sont appliquées pour les séparer selon leur rapport  $m/z$ .

Il est composé de quatre cylindres parallèles reliés électriquement deux à deux (Dawson et al., 1986). Les deux paires de barres sont de polarité opposée, ce qui forme un champ électrique à deux dimensions (figure 8). L'application simultanée d'une tension continue et d'une tension radiofréquence impose aux ions qui passent entre les cylindres un chemin oscillant. Pour des valeurs précises de tension et de radio fréquence, seuls les ions avec un rapport  $m/z$  donné peuvent atteindre l'extrémité des quadripôles et se diriger vers le détecteur.

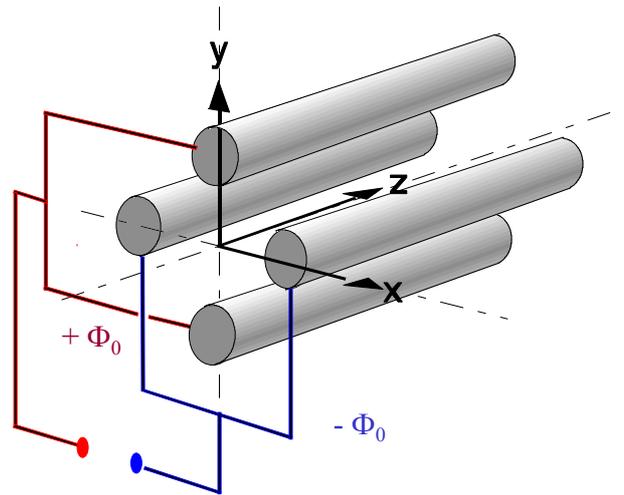


Figure 8 : Schéma d'un analyseur quadripolaire

Lorsque les ions entrent dans l'analyseur, ils subissent donc les forces résultantes du champ électrique total composé du champ alternatif quadripolaire superposé à un champ constant selon l'équation :

$$\Phi_0 = \pm(U \pm V \cos \omega t)$$

ou  $\Phi_0$  est la tension appliquée aux barres,  $\omega$  la fréquence angulaire,  $U$  la tension continue et  $V$  l'amplitude du voltage RF.

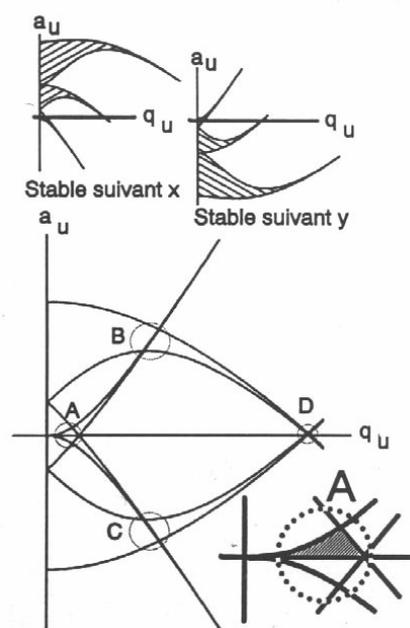
Ces forces vont alors définir les zones de stabilité des ions pour qu'ils soient transmis. Leur trajectoire va obéir à l'équation de Mathieu (résolue en 1866).

$$\frac{d^2 u}{d\xi^2} + (a_u - 2q_u \cos 2\xi) \cdot u = 0$$

$$u = x \text{ ou } y, \quad \xi = \omega t / 2,$$

$$a_u = \frac{8zeU}{mr_0^2 \omega^2}; \quad q_u = \frac{4zeV}{mr_0^2 \omega^2}$$

avec  $r_0$  : le rayon du cercle inscrit entre les 4 barres ;  $\omega$  : la fréquence angulaire telle que  $\omega = 2\pi f$  où  $f$  est la fréquence du champ alternatif ;  $m$  : la masse de l'ion ;  $ze$  : la charge de l'ion ;  $a_u$  et  $q_u$  : les solutions de l'équation différentielle de Mathieu.



**Figure 9 : Zones de stabilité dans un quadripôle selon x ou y,**  
D'après De Hoffmann et al., 1994

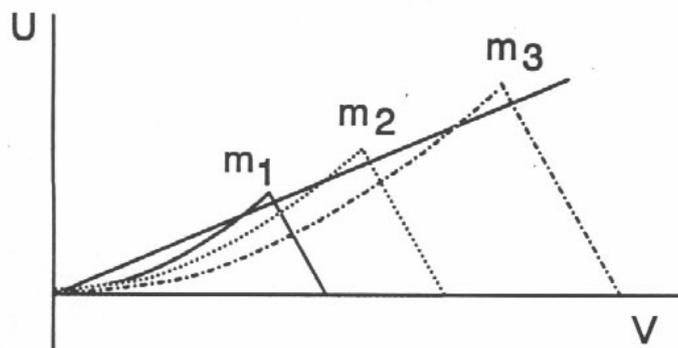
Les ions vont donc être séparés en fonction de leur rapport  $m/z$ , par la variation de  $U$  et de  $V$ . Le champ électrique qu'elles vont créer va imposer aux ions une trajectoire oscillante selon  $x$  et  $y$ , tout au long de l'axe  $z$ . Pour que l'ion arrive jusqu'au détecteur, il suffira que les oscillations selon  $x$  et  $y$  ne dépassent pas  $r_0$ , le rayon du cercle inscrit entre les barres.

Dans un diagramme,  $a_u = f(q_u)$ , on peut représenter les zones de stabilité, ou zones de valeurs de  $U$  et de  $V$  telles que  $x$  et  $y$  n'atteignent pas au cours du temps des valeurs supérieures ou égales à  $r_0$  (figure 9).

En superposant les diagrammes de stabilité selon  $x$  et  $y$ , quatre zones de travail possibles, marquées de  $A$  à  $D$ , sont obtenues. De manière générale, les analyseurs utilisent la zone  $A$  où  $U$  est positif.

D'après les équations, le déplacement d'une masse à l'autre se fait en divisant de manière proportionnelle  $a_u$  et  $q_u$ . Le triangle initial défini comme la zone  $A$  subit comme une homothétie. Ainsi un balayage de  $U$  et  $V$  avec  $U/V$  constant permet de détecter les différentes masses (figure 10). Plus la pente de la droite sera élevée, meilleure sera la résolution. Par

contre, dans ce type d'appareil, on peut noter que la résolution et la sensibilité sont inversement liées.



**Figure 10 : droites de fonctionnement d'un analyseur quadripolaire sur la première zone de stabilité. D'après Hoffmann et al., 1994.**

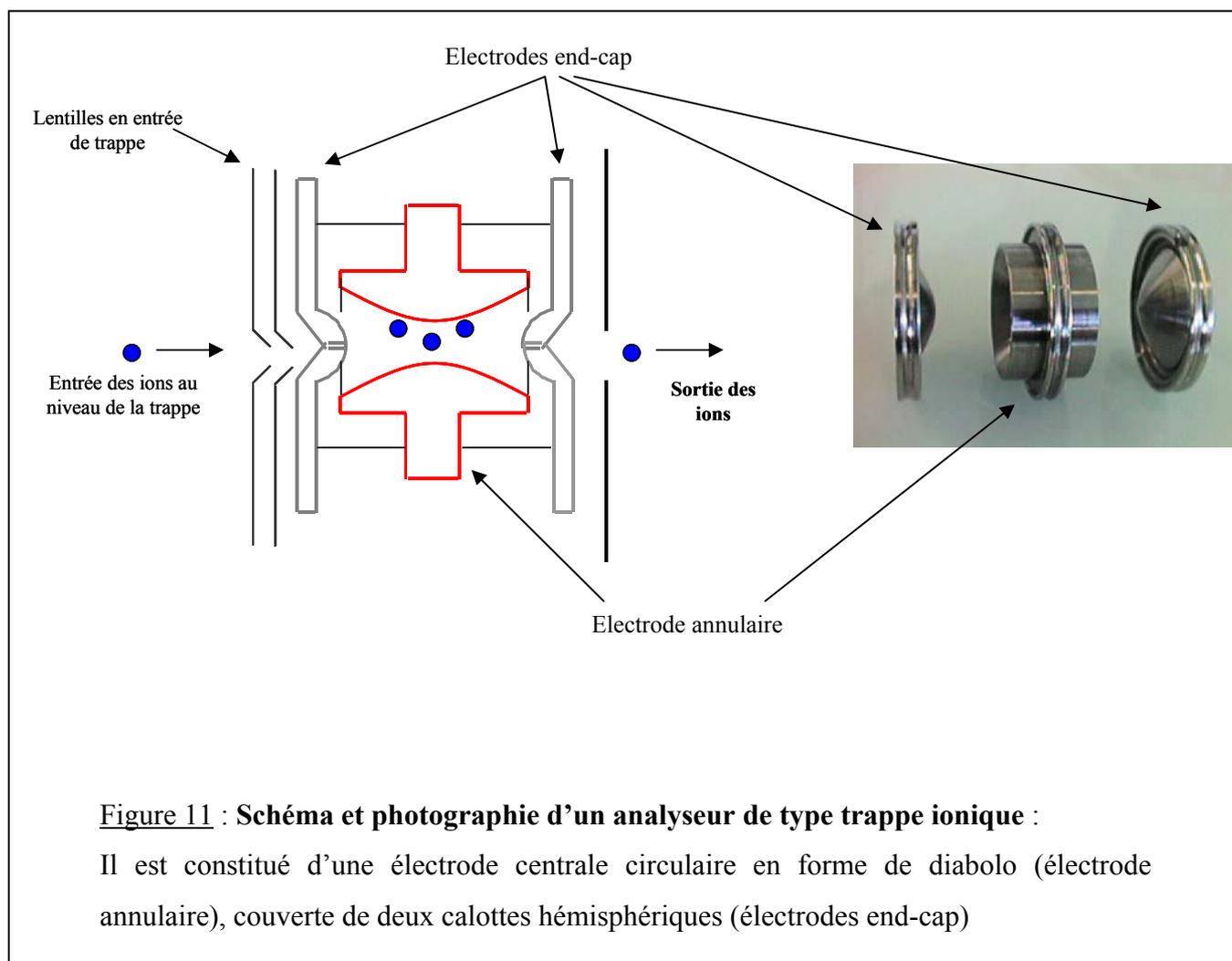
A pente nulle, la résolution sera nulle, mais tous les ions seront transmis dès lors que la bonne tension  $V$  est appliquée, c'est le mode RF only très utilisé notamment dans le cadre du couplage de plusieurs analyseurs.

## ***2.2 Les analyseurs de type piège à ions (trappe ionique)***

La trappe ionique est en fait l'homologue en trois dimensions des analyseurs de type quadripolaires. A l'inverse du quadripôle, qui ne laisse passer que les ions d'un rapport  $m/z$  donné, le piège ionique accumule à l'intérieur de sa zone de champ tous les ions sélectionnés. La capacité de confinement des ions est liée à la formation d'un pseudo puits de potentiel. Les ions piégés sont ensuite libérés successivement suivant leur rapport  $m/z$ , cette libération intervient lorsqu'ils deviennent instables (McLucky et al., 1994).

## 2.2.1 Structure de la trappe

Elle est composée d'une électrode centrale circulaire en forme de diabolo (l'électrode annulaire), couverte de deux calottes hémisphériques (les électrodes end-cap) (De Hoffmann et al. ,1994) (figure 11).



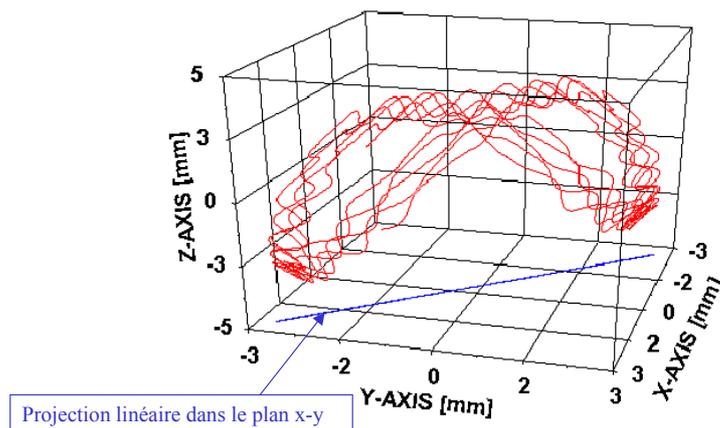
La géométrie de ces électrodes est choisie pour se rapprocher au maximum d'un champ quadripolaire idéal lorsqu'un seul potentiel de type RF est appliqué sur l'électrode annulaire ; dans ce cas les électrodes end-cap restent à la masse et ne servent que dans les expériences de MS-MS pour l'isolation des ions. Dans la réalité, il semble qu'une tension soit tout de même appliquée sur les électrodes end-cap pour compenser la non-perfection géométrique du piège. Comme dans le cas du quadripôle, la trajectoire tridimensionnelle des ions à l'intérieur du piège est imposée par un champ quadripolaire et elle peut être définie par l'équation de

Mathieu. Cependant il convient d'ajouter une nouvelle dimension : z. Ainsi, u=x, y ou z, et une nouvelle composante, r, peut être définie telle que  $r^2=x^2+y^2$ .

### 2.2.2 Le piégeage des ions

Le piégeage peut se faire par l'application d'un potentiel RF sur l'électrode annulaire.

Les ions ne vont pouvoir rester piégés que si leur trajectoire est stable simultanément axialement et radialement. Lorsque le diagramme de stabilité axiale est superposé au diagramme de stabilité radiale, deux zones, A et B, où les ions sont stables dans les deux dimensions sont identifiées. Dans les appareils classiques, seule la zone A est utilisée (figure 13).



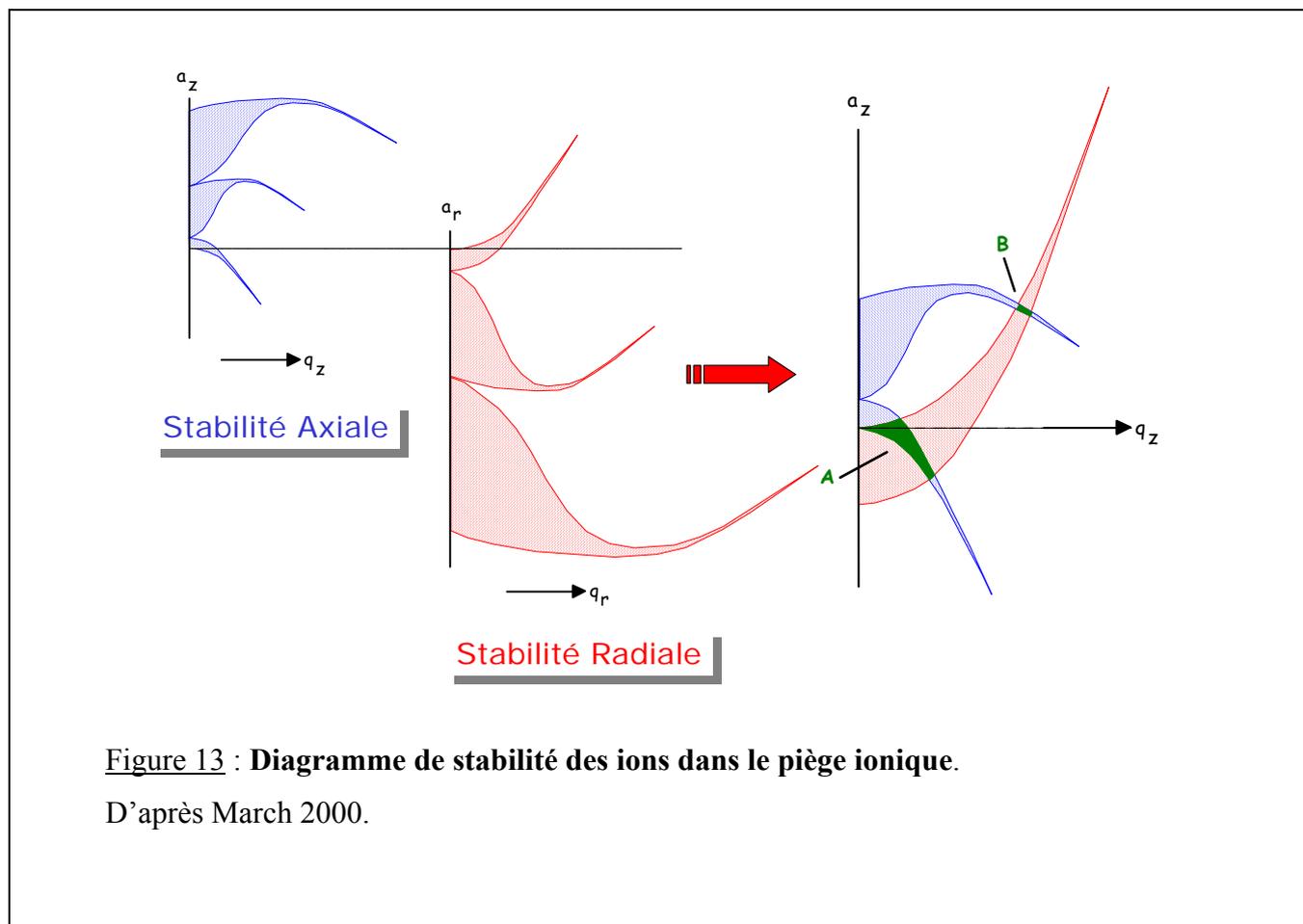
**Figure 12 : Figure de Lissajous**

Elle représente la trajectoire des ions piégés sous forme d'un huit tridimensionnel à laquelle se superpose une oscillation.

D'après March, 2000

La répulsion mutuelle des ions induisant des augmentations du rayon de leur trajectoire, il est nécessaire d'utiliser un gaz neutre, généralement de l'hélium, pour les confiner (McLucky et

al. 1994). Cet hélium va également servir à refocaliser les ions lors de leur entrée dans la trappe en diminuant leur énergie cinétique.



### 2.2.3 L'éjection des ions

L'éjection des espèces ioniques de la trappe se fait selon cinq processus principaux :

Leur trajectoire peut se situer en dehors d'une zone de stabilité.

Leur trajectoire peut se situer à l'intérieur d'une zone de stabilité, mais l'amplitude d'oscillations des ions peut dépasser les limites géométriques de la trappe.

Les ions peuvent subir des collisions défocalisantes avec les particules contenues dans le piège.

Les effets d'espace de charge peuvent contribuer à la défocalisation des ions s'ils sont en trop grand nombre dans le piège.

Dans le cas le plus simple, l'éjection sélective des ions se fait par une déstabilisation axiale selon l'axe z. Il se produit donc un balayage de l'amplitude du potentiel RF défini par  $q_z$ . Les ions atteignent le  $q_{eject}$  et sont ainsi éjectés séquentiellement selon des rapports  $m/z$  croissants (figure 14).

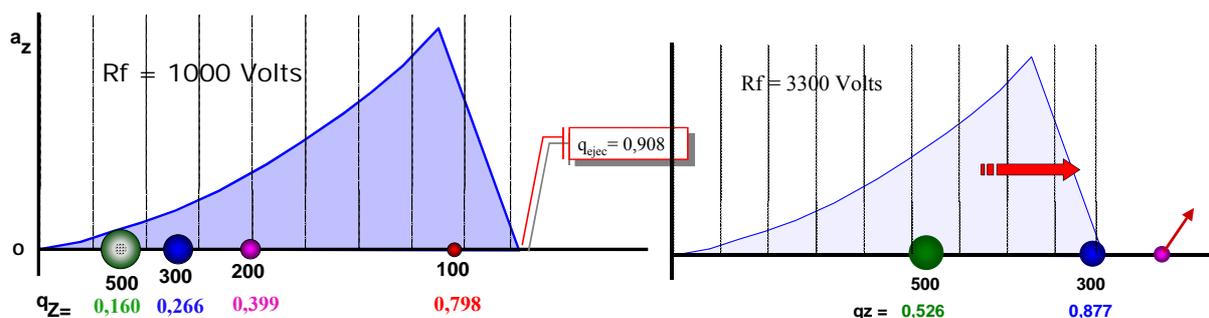


Figure 14 : Ejection sélective des ions

D'après Bécue, 2004 (communication personnelle).

## 2.3 L'analyseur temps de vol

### 2.3.1 Principe général

L'analyseur de type temps de vol ou TOF (pour Time Of Flight) est par son concept de base le plus simple des analyseurs.

Si dans la source, les ions sont accélérés par une tension  $V_s$  et parcourent une distance  $d$  pour atteindre le détecteur sans aucune autre accélération. Un ion de masse  $m$  et de charge totale  $q=ze$  aura une énergie cinétique  $E_c$  en sortie de source telle que :

$$E_c = \frac{1}{2} mv^2 = qV_s$$

Si les ions parcourent ensuite une distance  $d$  en un temps  $t$  dans un tube où ne règne aucune différence de potentiel et où le vide permet d'éviter d'éventuelles collisions.

Comme

$$t=d/v$$

Alors :

$$t=d*\sqrt{(m/(2zeV_s))}$$

Le temps de vol des ions permettra donc de déterminer leur rapport  $m/z$ . Le facteur de proportionnalité entre  $t$  et  $m/z$  sera donc fonction de l'appareillage et des conditions expérimentales. Le principal problème de cette technique est le manque de résolution. En effet, la dispersion énergétique des ions à l'entrée de l'analyseur entraîne une grande perte de résolution. Cette dispersion a pu être compensée par l'introduction de systèmes qui permettent la refocalisation des ions à la fois sur le plan énergétique et sur le plan temporel. Ces systèmes sont l'extraction retardée (Vestal et al., 1995) et le réflectron (Cotter 1992).

### 2.3.2 L'extraction retardée

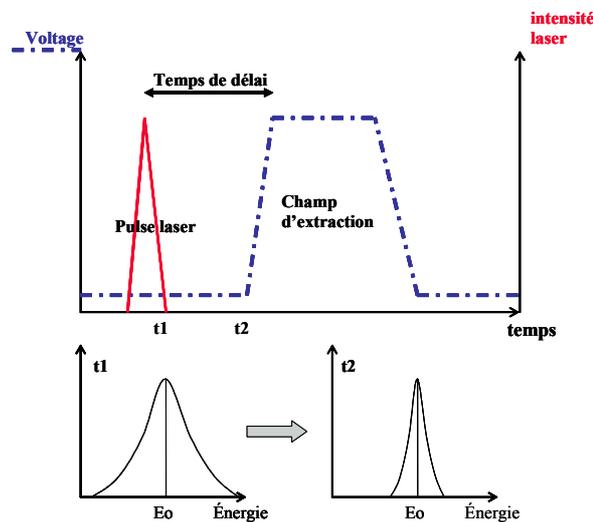


Figure 15 : L'extraction retardée :

Refocalisation énergétique des ions dans la source du fait d'un temps de délai avant l'extraction des ions

Source : Voyager Biospectrometry Workstation, User Guide (Applied Biosystem, 2001)

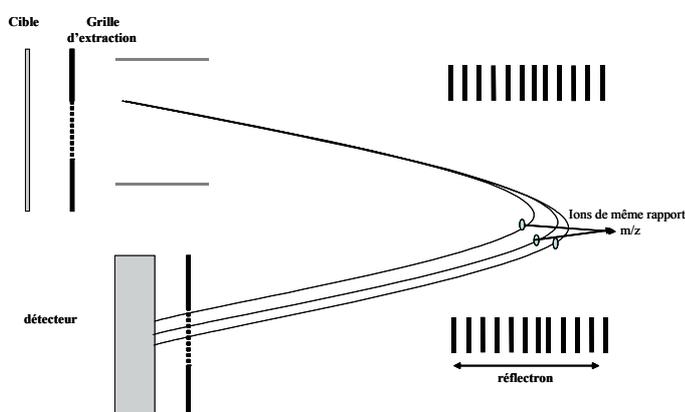
Le laser permet de donner une grande énergie aux ions, puis ces ions rentrent dans une phase de relaxation, avant d'être réaccélérés. Ce délai permet aux ions une refocalisation énergétique (figure 15). Ainsi, l'analyse gagnera en résolution et par conséquent en précision dans la mesure de la masse et en sensibilité. De plus l'appareil sera alors moins sensible aux fortes fluences laser.

### 2.3.3 Le réflectron

Le système réflectron va permettre une focalisation des ions d'un même rapport  $m/z$  sur le plan temporel. Le réflecteur est composé d'une série d'électrodes annulaires portées à des potentiels croissants qui définissent un champ électrique repousseur homogène, la trajectoire est ainsi inversée (figure 16).

Une tension fixe est appliquée au niveau du réflectron. Les ions les plus énergétiques sont les plus rapides et ils pénètrent donc plus dans la zone de champs. Leur trajectoire jusqu'au détecteur sera alors plus longue, et ainsi tous les ions d'un même rapport  $m/z$  parviendront au même moment au détecteur, quelle que soit leur énergie de départ. Cette méthode permet de gagner en résolution.

Le grand inconvénient de cette technique est la transmission des ions de plus haut rapport  $m/z$ . Ainsi, l'analyse des molécules de faible masse sera plutôt réalisée en mode réflectron, tandis que l'analyse des grosses molécules se fera préférentiellement en mode linéaire. Le  $m/z$  limite dépendant beaucoup du type d'appareillage utilisé.



**Figure 16 : Schéma de fonctionnement d'un TOF avec réflectron.**

Ce dernier permet de refocaliser les temps de vol des ions ayant un même rapport  $m/z$ .

### 2.3.4 L'injection orthogonale

Les analyseurs de type temps de vol sont préférentiellement adaptés aux méthodes d'ionisation pulsées comme le MALDI, le tir laser donnera alors le temps de départ pour la mesure du temps. Les méthodes d'ionisation continue, comme l'electrospray, ne sont donc à priori pas adaptées au TOF, cependant il est intéressant de combiner les propriétés de l'electrospray à celle du TOF. Pour cela il est possible d'injecter les ions dans le TOF, orthogonalement suivant l'axe z. Avec un tel dispositif, une faible dispersion en vitesse initiale suivant l'axe z est observée, ce qui permet d'obtenir une bonne résolution.

L'injection se déroule alors selon deux processus (figure 17), l'un continu, l'autre pulsé. Les ions arrivent de manière continue dans la source selon l'axe y, avec une vitesse  $V_y$ . Ils sont accélérés de manière pulsée selon l'axe z. Le « pusher » va donc accélérer les ions à un même niveau d'énergie et donner le départ de la mesure du temps de vol (Verentchikov.et al., 1994).

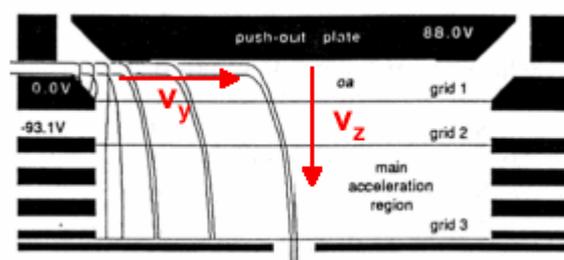


Figure 17 : Schéma du système d'injection orthogonale

Le système d'injection orthogonal a permis de coupler les sources electrospray avec les analyseur temps de vol et d'ainsi créer les ES-TOF.

### **3 LA FRAGMENTATION DES IONS**

#### ***3.1 Principe général***

La fragmentation des ions permet d'obtenir des informations sur leur structure. Cette fragmentation est le résultat d'une augmentation de l'énergie interne, l'ion parent (ou précurseur va alors se dissocier et former des ions fils (ou fragments). La dissociation des ions précurseurs peut résulter d'événements de collision (CID pour collision-induced dissociation). La collision peut avoir lieu en source avec des molécules de gaz ou de solvant. La multiplicité des ions précurseurs alors concernés par la fragmentation, va générer de nombreux ions fragments. La filiation des ions sera alors difficilement établie et l'analyse des spectres d'ions fragments sera fastidieuse. Pour obtenir avec certitude l'identité des ions fragments, il est nécessaire de les isoler physiquement. Les spectromètres de masse tandem permettent cette isolation. Dans ces appareils, l'analyse se fait en deux phases : dans la première, les ions parents (ou précurseurs) sont sélectionnés, et dans la seconde les ions issus de sa fragmentation, appelés ions fils ou ions produits, sont analysés. Entre ces deux étapes a lieu la fragmentation, qui peut se réaliser dans des cellules de collision grâce à un gaz inerte avec lequel les ions entreront en collision.

Les différentes étapes d'analyse de la MS-MS peuvent se faire dans des analyseurs distincts, l'un sélectionnant l'ion parent, l'autre analysant les ions fils, on parle alors de MS-MS dans l'espace. C'est le principe par exemple utilisé dans les triples quadripôles, dans les analyseurs hybrides (de type Q-TOF par exemple) ou dans les TOF-TOF.

Mais les deux étapes peuvent également se dérouler dans un même analyseur, c'est la MS-MS dans le temps, mise en oeuvre dans les trappes à ion.

Il est théoriquement possible de coupler un nombre infini d'analyseurs et d'ainsi réaliser des expériences de MS/MS/MS/MS..... ou  $MS^n$ , cependant l'intensité des ions diminuant à chaque étape du fait de la longueur de leur trajet et de la présence de gaz de collision, le nombre de MS consécutives reste dans la pratique limité.

## ***3.2 La fragmentation dans l'espace***

### **3.2.1 Les triples quadripôles (QqQ)**

Ce type d'appareil est constitué de trois quadripôles. Le premier quadripôle va sélectionner l'ion parents, et le troisième va analyser les ions fragments. Le second quadripôle servira de cellule de collision. Il travaillera alors en mode « RF only ». De ce fait, les ions seront focalisés et transmis avec une grande efficacité.

Selon les modes de balayages des quadripôles Q1 et Q3, plusieurs types d'expériences pourront être réalisés.

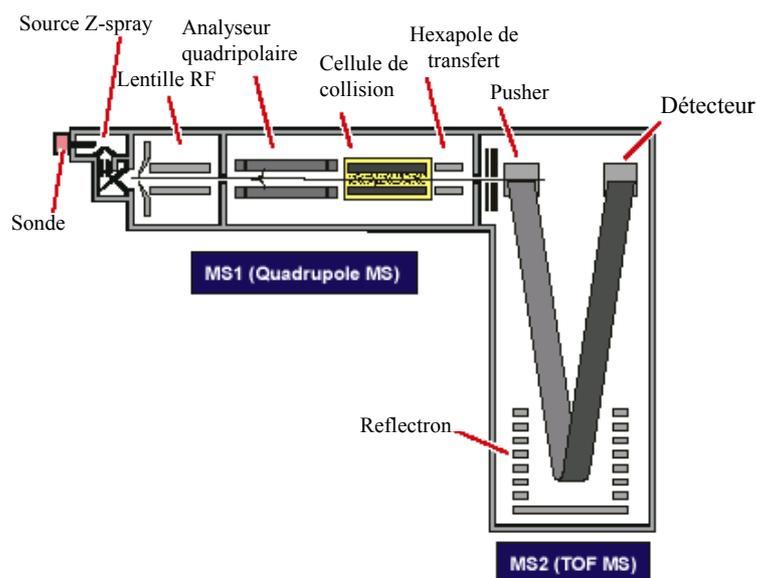
Si une tension fixe est appliquée en Q1, et un balayage de tension en Q3, le spectre obtenu sera un spectre d'ions produits. Pour un précurseur particulier choisi en Q1, les ions fragments sont recherchés.

Si un balayage de tension est réalisé en Q1 et une tension fixe appliquée en Q3, un spectre d'ions précurseurs sera obtenu. Les différents ions parents qui donnent le même ion fixe seront identifiés par cette méthode.

Si un balayage de tension est effectué en Q1 et en Q3, un spectre de perte de neutre sera obtenu. Les deux quadripôles balayent simultanément la gamme de masse en laissant une différence de masse constante entre eux. Le spectre de masse permettra alors de déterminer tous les ions fragments qui ont perdu une molécule, par rapport à leur précurseur, d'une masse égale à la différence de balayage.

### **3.2.2 Les Qq-TOF**

Si l'on compare cet instrument hybride à un triple quadripôle, le TOF remplacera le troisième quadripôle (figure 18).



**Figure 18 : Schéma de fonctionnement du Qq-TOF**

D'après le guide d'utilisation du Q-TOF micro, waters

Cet appareil est constitué d'un premier quadripôle, puis d'un hexapôle qui tiendra lieu de cellule de collision. Le « pusher » accélérera les ions dans la direction axiale grâce à des « pulses » de tensions de fréquences élevées et les ions atteindront le temps de vol ici équipé d'un réflectron.

En MS simple, le premier quadripôle sera utilisé en « RF only », tous les ions seront donc transmis au temps de vol, qui les séparera selon leur rapport  $m/z$ . Lors des expériences MS-MS, les ions précurseurs seront sélectionnés dans le premier quadripôle et fragmentés dans la cellule de collision. Le TOF réalisera l'analyse des ions fragments (Chernushevich et al., 2001).

### 3.2.3 Les TOF-TOF

Les instruments TOF-TOF, qui sont le couplage de deux analyseurs de type temps de vol vont permettre d'obtenir des fragmentations riches et efficaces, tout en gardant un temps d'analyse rapide.

Classiquement dans les appareils possédant un temps de vol, la fragmentation est réalisée par un processus appelé PSD (Post Source Decay). Ce procédé de fragmentation utilise la fragmentation spontanée des ions métastables ; seules les informations données par les ions fragmentant après la source, dans la première zone-champs, seront interprétables. Lorsque l'ion précurseur se dissocie, l'ion parent, les ions fils et les neutres ont la même vitesse ; ils arriveront donc en même temps au niveau du réflecteur. Le neutre ne sera pas réfléchi. L'ion parent pénétrant plus loin dans le réflecteur (en raison de sa plus grande énergie cinétique) que l'ion fils, il arrivera plus tard au niveau du détecteur, la masse de l'ion fragment sera alors déterminée. Ce type d'analyse nécessite différentes acquisitions à différentes tensions du réflecteur, afin d'obtenir une résolution suffisante sur une large gamme de masse. Les spectres obtenus par PSD seront donc des spectres composites (Kaufmann et al., 1994).

Dans les analyses PSD, l'ion parent est sélectionné par une porte temporelle, qui va dévier la trajectoire de tous les ions jusqu'à la valeur du temps de vol de l'ion parent désiré.

Les analyseurs de type TOF-TOF sont nés de la volonté d'obtenir des fragmentations de haute énergie (impossibles dans les appareillages de type MALDI-QTOF), tout en s'affranchissant du balayage au niveau du réflecteur imposé par le PSD.

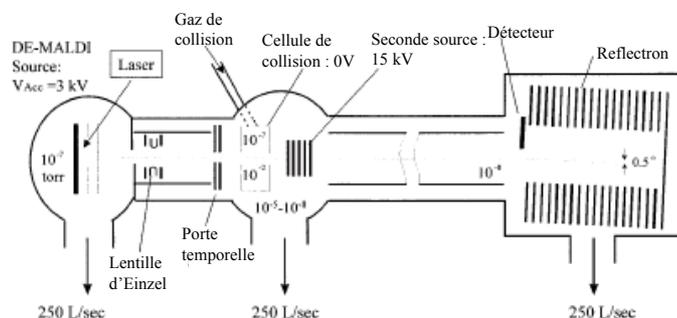
Le principe de ces instruments est la ré-accélération des ions après la sélection par la porte temporelle. Ainsi, la gamme d'énergie cinétique dans laquelle se trouvent les ions va diminuer, il ne sera pas nécessaire de balayer la tension du réflecteur et le spectre d'ions fragments pourra être obtenu en une seule étape.

Selon les constructeurs, la géométrie et le principe des appareils de type TOF-TOF peut varier.

L'appareil 4700 Proteomics d'Applied Biosystems peut être apparenté à un triple quadripôle dans lequel le premier et le dernier quadripôles sont remplacés par des TOF.

Dans cet appareil, après l'accélération de tous les ions, la sélection de l'ion parent donné se fait par une double porte temporelle, les ions sont ensuite décélérés dans la cellule de collision flottante et fragmentés par CID de haute énergie. S'en suit une post-accélération de l'ion parent et de ses fragments, comme s'ils arrivaient dans une nouvelle source. Enfin, les ions

fragments sont analysés par le réflecteur en une seule étape, contrairement aux analyses réalisées par PSD classique. (figure 19) (Medzihradzky et al., 2000).



**Figure 19 : Schéma de fonctionnement du MALDI-TOF-TOF d'Applied Biosystems**

Dans l'appareil Ultraflex TOF/TOF de Bruker, la fragmentation est réalisée avant la porte temporelle par CID (dans une cellule de collision placée en source) ou par LID (Laser Induced-Dissociation). Dans le cas du LID, de fortes fluences laser vont créer des ions métastables. Ces ions vont fragmenter dans la première zone de temps de vol. Les fragments et leurs précurseurs respectifs ayant la même vitesse, ils atteindront la porte temporelle au même moment. La « famille d'ions » désirée est alors sélectionnée au niveau de la porte temporelle et entre dans le système LIFT, qui consistera en une sorte de deuxième source. Dans la zone LIFT, tous les ions sont refocalisés en énergie et ré-accélérés. L'analyse se fait ensuite dans le deuxième temps de vol. Seule une fraction de l'énergie des ions est dépendante de leur masse, l'énergie apportée par le LIFT étant la même pour tous les ions. Les ions seront donc dans une gamme énergétique restreinte et l'analyse pourra se faire sans balayage au niveau du réflecteur. Le "Post Lift Metastable Suppressor" permet la suppression des ions parents afin d'éviter d'autres fragmentations métastables dans le deuxième tube de vol (figure 20) (Suckau et al., 2003).

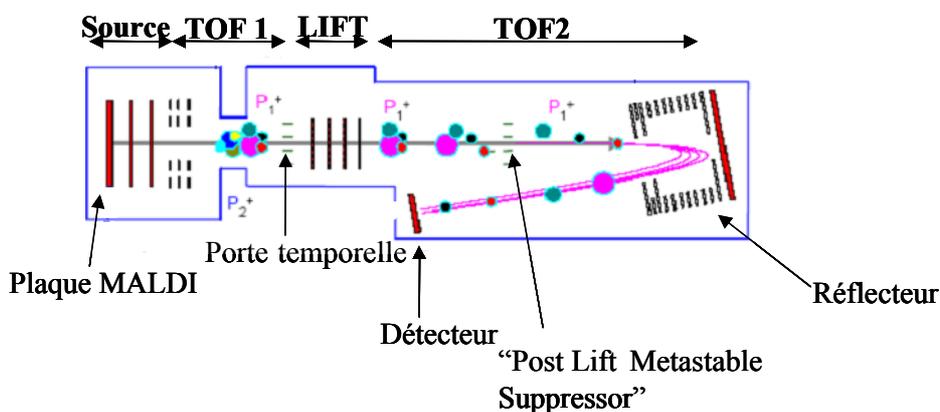


Figure 20 : Schéma de fonctionnement de l'Ultraflex TOF/TOF de Bruker

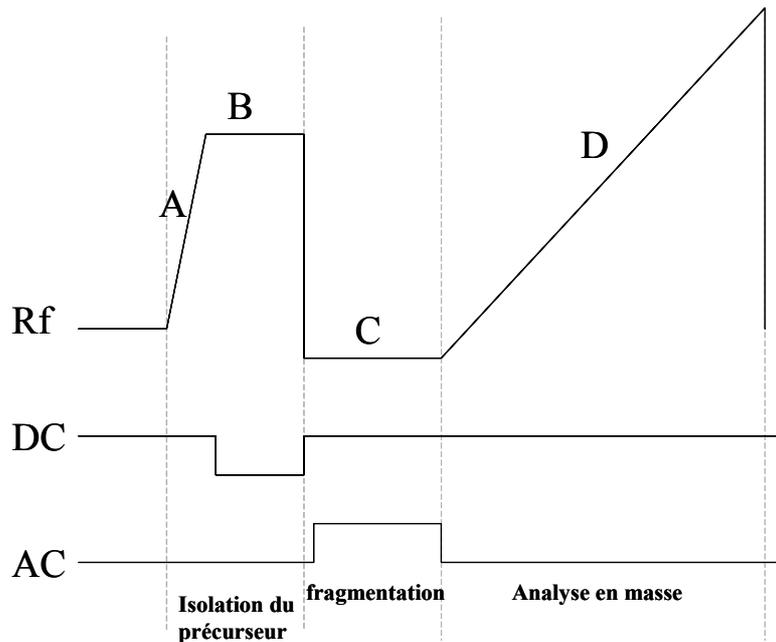
### 3.3 La fragmentation dans le temps (cas des trappes ioniques)

Dans ce type d'appareils, un seul analyseur est utilisé à la fois pour l'isolation des précurseurs, leur fragmentation et l'analyse des ions fragments.

Plusieurs processus concernant l'isolation de l'ion précurseur sont décrits (Bonner 1977, Strife et al., 1988).

L'isolation du précurseur peut se faire par l'augmentation linéaire de l'amplitude RF et de la tension auxiliaire jusqu'au seuil de coupure juste inférieur à celui correspondant au rapport  $m/z$  de l'ion précurseur. Les ions de rapport  $m/z$  inférieurs sont donc ainsi déstabilisés et éliminés (phase A figure 21). Les ions de rapport  $m/z$  supérieurs au  $m/z$  de l'ion précurseur sont éliminés par l'addition d'une tension continue au potentiel RF (phase B figure 21).

La fragmentation de l'ion précurseur se fait par l'application d'une tension alternative et à l'aide de l'hélium. La trappe ionique contiendra alors les ions fils (phase C figure 21), qui seront expulsés séquentiellement de la même manière que lors d'une expérience de MS simple (phase D figure 21).

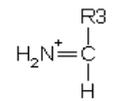
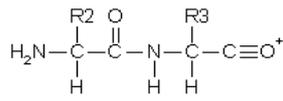
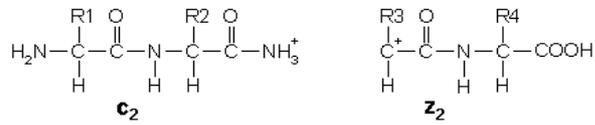
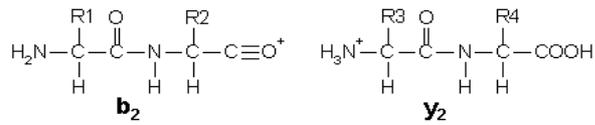
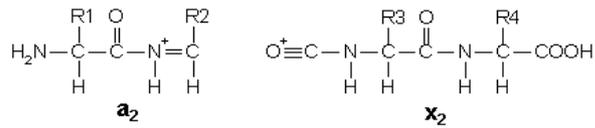
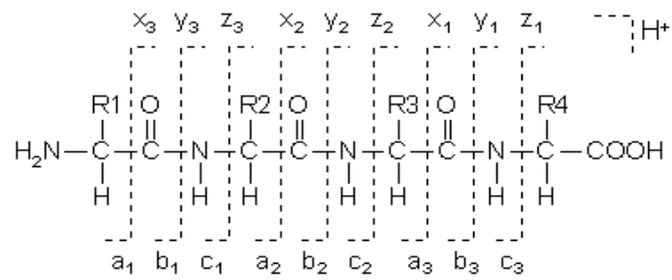


**Figure 21 :** Représentation des différentes phases lors des expériences de MS-MS dans une trappe ionique.

### 3.4 La fragmentation des peptides

La fragmentation des peptides est largement utilisée en biologie, car elle permet de les séquencer et donc de les identifier. Les peptides se fragmentent toujours de la même façon, essentiellement au niveau des liaisons peptidiques. La nomenclature qui leur est appliquée (Roepstorff et Fohlman, 1984) différencie les fragments pour lesquels la charge est portée par la partie N-terminale du peptide (fragments  $a_n$ ,  $b_n$ ,  $c_n$ ) et ceux pour lesquels elle est portée par la partie C-terminale ( $x_n$ ,  $y_n$ ,  $z_n$ ) (figure 22).

L'analyse de peptides par MS-MS permet ainsi de détecter des séries de fragments. La différence de masse entre des ions consécutifs d'une même série permet de déterminer l'identité des acides aminés et d'en déduire la séquence du peptide.



**Fragment interne acylium**

**Ion immonium**

**Figure 22: Nomenclature des ions issus de la fragmentation peptidique par spectrométrie de masse en tandem**

CHAPITRE 2 :  
OPTIMISATION DE LA  
METHODE DES ECHANGES  
H/D COUPLES A LA  
SPECTROMETRIE DE MASSE

## INTRODUCTION

Les protéines jouent un grand rôle dans de nombreux processus biologiques. Leur simple identification ne suffit souvent pas à comprendre toutes les fonctions dans lesquelles elles sont impliquées. En effet, la caractérisation d'une protéine requiert la compréhension de trois caractéristiques : fonction, structure et dynamique, puisqu'elles agissent souvent en formant des complexes avec différents partenaires, qu'elles ne sont actives que dans une conformation spécifique et que cette conformation peut varier au cours du temps, selon sa localisation et son activité.

Il existe un large éventail de techniques pour étudier les protéines. Leur fonction, par exemple, est le plus souvent caractérisée par des méthodes biochimiques. Leur structure et leur dynamique sont plutôt étudiées par des techniques biophysiques.

Alors que beaucoup de méthodes biophysiques donnent des informations structurales sur les protéines, la plupart des techniques, comme le dichroïsme circulaire, l'ultracentrifugation ou encore la spectrométrie infrarouge, ne donnent que des informations globales. Seules, la cristallographie par diffraction des rayons X et la RMN peuvent donner des informations locales et avec une grande résolution sur la structure des protéines. Cependant ces deux techniques ont leurs limites. En ce qui concerne la cristallographie, l'étape de cristallogénèse reste le principal obstacle, et certaines protéines, comme les protéines membranaires ou les protéines très désordonnées sont très difficilement cristallisables. Le fait même que la protéine soit étudiée sous forme solide conduit au piégeage d'une conformation donnée qui ne reflétera pas la dynamique de la protéine et qui pourrait, dans certains cas, ne pas être la conformation active. La RMN, même avec les récents développements au niveau des appareillages, est difficilement applicable à des protéines de plus de 40 kDa. De plus, ces deux méthodes requièrent de grandes quantités de protéines pures

Ainsi il est nécessaire de continuer à développer de nouvelles méthodes d'études structurales et dynamiques des protéines. C'est dans ce contexte que la technique des échanges hydrogène/deutérium, apparue il y a plus de 40 ans, est devenue une technique des plus intéressantes. La mesure du taux d'échange peut en effet fournir des informations sur la protéine analysée soit d'une manière globale, soit très localement.

Dans ce chapitre, nous allons appliquer cette méthode sur une protéine de taille relativement importante, la protéine PBP-2X de *Streptococcus pneumoniae*. L'utilisation de cette protéine modèle permettra l'optimisation de la technique, qui sera alors plus généralisable, quelle que

soit la taille de la protéine étudiée, et plus résolutive dans la définition des zones d'intérêt. Ce chapitre sera composé de plusieurs parties. Nous commencerons par une étude bibliographique de la protéine PBP-2X que nous avons utilisée comme modèle, puis la théorie des échanges H/D sera développée. La partie expérimentale présentera les résultats obtenus lors des essais des différentes approches pour l'optimisation de la technique des échanges H/D. Enfin, ces différentes approches seront discutées afin d'en chercher la validation.

# **1 LE MODELE D'ETUDE : LA PROTEINE PBP-2X**

## ***1.1 Contexte de l'étude : la bactérie Streptococcus pneumoniae***

Les streptocoques sont des bactéries Gram positives ; disposées le plus souvent en chaînettes ou paires. Ils sont entourés d'une capsule polysaccharidique qui empêche la liaison des anticorps à la paroi de la cellule et inhibe donc la phagocytose. Leur pouvoir pathogène est donc typiquement invasif. Ils sont responsables de pneumonies, de méningites et d'otites (Klein, 1999).

Les infections pneumococciques constituent un important problème de santé publique. La fréquence, la gravité et la mortalité de ces infections restent très élevées malgré une antibiothérapie disponible. Dans les pays développés, elles affectent surtout les nourrissons et les personnes âgées. Elles sont responsables du décès d'environ 1,1 millions de personnes à travers le monde chaque année (Klein, 1999), ce qui représente jusqu'à 9% des décès dans certains pays en voie de développement.

Au fil du temps et sur l'ensemble de la planète, ces bactéries ont développé des résistances aux antibiotiques utilisés pour les combattre. En France, pour l'année 1999, 44% des souches isolées présentaient une sensibilité diminuée à la pénicilline G (Laurans et al., 2001). Toutefois l'importance de ce problème ainsi que le degré de résistance varient beaucoup selon les régions et selon les classes d'âge. La prise d'antibiotiques exerce une pression de sélection sur les agents infectieux. Ce mécanisme est considéré comme le moteur principal dans le développement des résistances.

La résistance de *Streptococcus pneumoniae* aux antibiotiques de la classe des  $\beta$ -lactamines est due à la modification des cibles primaires de ces antibiotiques : les Penicillin Binding Proteins (PBP) (Hakenbeck et Coyette, 1998).

## **1.2 Les « *Penicillin Binding Proteins* »**

### **1.2.1 La description des PBPs**

Les « Penicillin Binding Proteins » (PBPs) sont des protéines bactériennes localisées dans le périplasma et ancrées à la membrane cytoplasmique. Elles interviennent dans la construction et le maintien de l'intégrité du peptidoglycane. Le nombre et la taille des PBPs varient selon les espèces bactériennes.

Elles peuvent être divisées en trois classes selon leur séquence en acides aminés et leur fonction : les PBPs de haute masse moléculaire des classes A et B et les PBPs de basse masse moléculaire.

Les PBPs de haute masse moléculaire de classe A sont des enzymes bifonctionnelles. Elles possèdent une activité glycosyltransférase (GT) et une activité D,D-transpeptidase (TP) (Goffin et al., 1998).

Les PBPs de haute masse moléculaire de la classe B sont des enzymes monofonctionnelles qui possèdent une activité D,D-transpeptidase (Goffin et al., 1998).

Les PBPs de basse masse moléculaire ont une activité D,D-carboxypeptidase (Hakenbeck et al., 1982).

Les domaines TP et D,D-carboxypeptidase des PBPs appartiennent, tout comme les  $\beta$ -lactamases à serine, à une superfamille de protéines appelée Active-site Serine Penicillin Recognizing Enzymes (ASPRES) (Joris et al. 1988 ; Frère et al., 1988). Cette famille de protéines est caractérisée par trois motifs d'acides aminés que différentes études structurales ont permis de localiser (Dideberg et al. 1987 ; Jelsch et al. 1993 ; Kelly et Kuzin 1995 ; Parès et al. 1996). Le premier motif est une tétrade S\*-X-X-K, où S\* représente une serine active. Ce motif est situé du côté N-terminal d'une longue hélice. Le deuxième motif forme une triade, S/Y-X-N et est localisé d'un côté du site actif. De l'autre côté, sur la face opposée, se trouve le troisième motif K/H-S/T-G.

### 1.2.1.1 Les PBPs de haute masse moléculaire de classe A

*Streptococcus pneumoniae* possède trois PBPs de ce type : PBP1a, PBP1b et PBP2a. Elles sont organisées de façon similaire. Cette organisation moléculaire est très proche chez différentes bactéries et notamment chez *E.coli*.

La région N-terminale est composée d'une partie cytoplasmique et d'une partie plus hydrophobe qui permet l'ancrage dans la membrane. Du côté N-terminal de la région périplasmique se trouve le domaine GT.

Du côté C-terminal, les motifs caractéristiques de la famille ASPRE sont retrouvés indiquant la présence du domaine TP (Di Guilmi et al., 1998, 1999, 2003).

Certaines données expérimentales suggèrent que les domaines des PBPs de classe A qui lient la pénicilline ont une activité TP, mais aucune activité in vitro n'a été démontrée. Par contre elles possèdent des activités estérase et thiolestérase (Jamin et al., 1993).

Chez *E.coli*, la délétion de PBP1a ou PBP1b seule est tolérée, mais au moins l'une des deux protéines doit être présente. Ceci indique une fonction redondante dans la synthèse du peptidoglycane, les deux protéines se compensent (Yousif et al., 1985).

De la même façon, chez *Streptococcus pneumoniae*, la délétion individuelle d'une des PBPs de classe A n'entrave pas la croissance de l'organisme, mais la présence de PBP1a ou de PBP2a est essentielle (Kell et al. 1993).

### 1.2.1.2 Les PBPs de haute masse moléculaire de classe B

Deux PBPs de cette classe ont été identifiées chez *S. pneumoniae* : PBP2x et PBP2b. Ces protéines sont ancrées dans la membrane par la région N-terminale hydrophobe. Les motifs caractéristiques de la famille ASPRE et donc des domaines TP, sont localisés dans la partie C-terminale de PBP2b et dans la partie centrale de PBP2x (Dowson et al., 1989 ; Laible et al., 1989). Ces protéines sont essentielles à la viabilité de la bactérie puisque leur suppression ne peut être tolérée (Kell et al., 1993).

### 1.2.1.3 Les PBPs de basse masse moléculaire

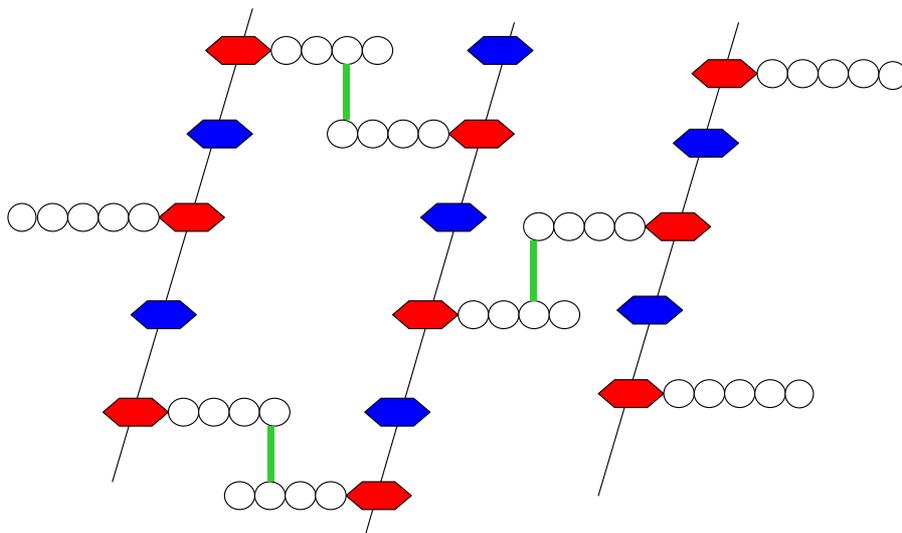
Une seule PBP de ce type a été identifiée chez *Streptococcus pneumoniae* : PBP3. Son activité D,D-carboxypeptidase est sensible à la pénicilline. Elle est ancrée à la membrane par son extrémité C-terminale.

## 1.2.2 Les rôles des PBPs

Les PBPs interviennent dans la synthèse et le maintien de l'intégrité du peptidoglycane.

Chez les eubactéries, le peptidoglycane est le principal constituant de la paroi. Il contribue au maintien de la forme de la bactérie et de son intégrité cellulaire et peut servir de support pour d'autres macromolécules pariétales.

Le peptidoglycane est un hétéropolymère qui est composé de chaînes osidiques et de peptides (figure 23).



**Figure 23 : Structure chimique du peptidoglycane :**

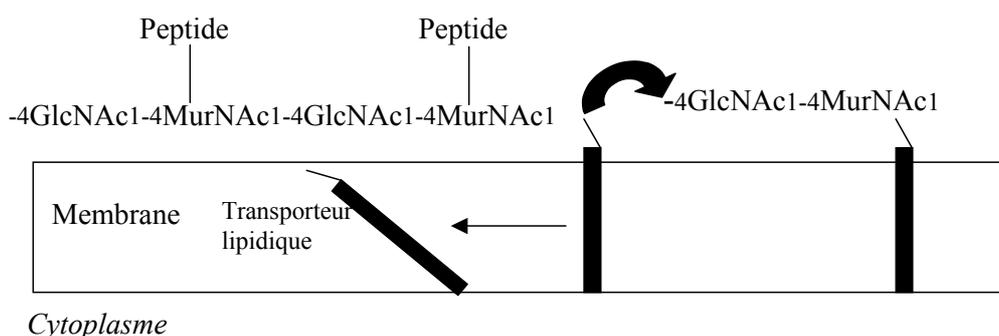
alternance de N-acétylglucosamine (◀) et d'acide N-acétylmuramique (▶) qui sont substitués par des chaînes peptidiques (○○○○). Sur ces chaînes se trouve un pont interpeptidique qui relie le 4ème acide aminé d'une chaîne peptidique (un D-Ala) au 3ème acide aminé d'une autre chaîne (une Lys chez *S. pneumoniae*) (—)

Les chaînes osidiques sont constituées par une alternance de N-acétylglucosamine et d'acide N-acétylmuramique sur lesquels les chaînes peptidiques sont substituées. Les ponts interpeptidiques donnent au peptidoglycane une structure en réseau. Cette structure est commune à la plupart des bactéries même s'il existe plus d'une centaine de peptidoglycans chimiquement différents. De plus, à l'intérieur d'une même bactérie sa composition chimique varie selon sa localisation au sein de la paroi cellulaire (Ghuysen et al., 1968 ; Schleifer et Kandler, 1972).

La rigidité du peptidoglycane est conférée par les chaînes de sucres puisque seules des petites rotations sur l'axe  $\beta$ 1-4 sont possibles. L'élasticité, quant à elle, est donnée par la flexibilité des chaînes peptidiques interconnectées. Le peptidoglycane peut alors accepter des variations de volume de la bactérie dues à des éventuels changements de pression osmotique (Holtje, 1998).

Les PBPs interviennent dans les étapes finales de la synthèse du peptidoglycane. Elles utilisent comme précurseur le lipide II issu des étapes membranaires de synthèse.

La polymérisation est réalisée par les PBPs de haute masse moléculaire de classe A grâce à une réaction de glycosyltransférase qui permet le transfert du disaccharide pentapeptidique à l'extrémité d'une chaîne (figure 24). Un brin de muréine naissant attaché à un lipide est transféré sur le C4 d'un résidu glucosamine d'un précurseur de la muréine lui aussi associé avec un lipide (Holtje, 1998).

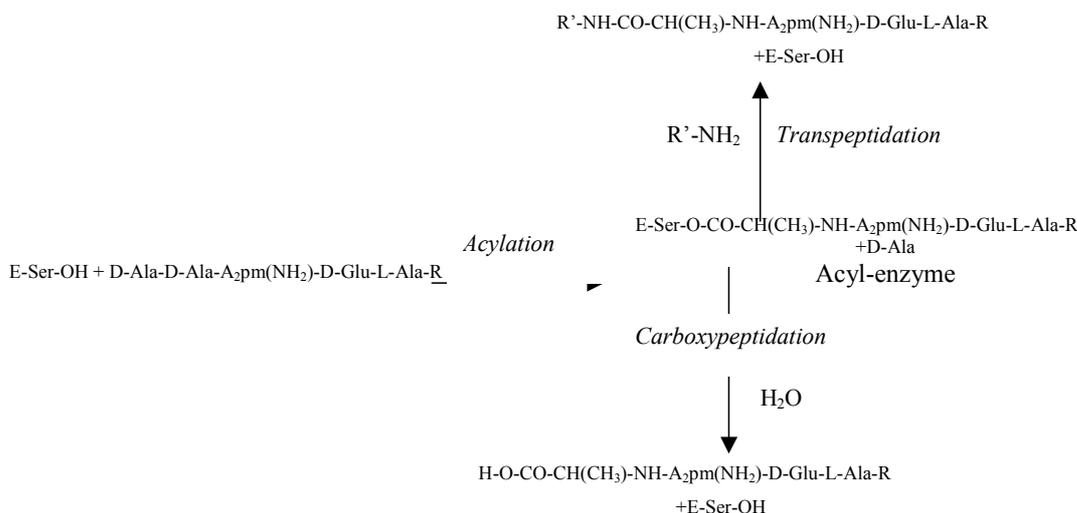


**Figure 24 : La polymérisation du peptidoglycane :**

La réaction de glycosyltransfert catalyse le transfert d'un brin naissant de muréine, attaché à un lipide, sur le carbone 4 d'un résidu glucosamine d'un précurseur attaché lui aussi à un lipide.

Les fonctions transpeptidases des PBPs de haute masse moléculaire réalisent ensuite la réticulation de la chaîne néoformée sur une chaîne adjacente. La rupture de la liaison D-Ala-D-Ala du substrat donneur fournit l'énergie nécessaire pour former un lien covalent avec le groupement nucléophile de la serine active du domaine TP. Puis cet acyl-enzyme subit une aminolyse qui entraîne la formation d'un pont peptidique entre l'avant-dernier D-Ala du pentapeptide donneur et le groupe ε-amine (une lysine chez *Streptococcus pneumoniae*) du résidu accepteur (Ghuysen, 1994). Enfin les D,D-carboxypeptidases portées par les PBPs de basse masse moléculaire éliminent le dernier D-Ala du pentapeptide donneur pour qu'il n'y ait pas de nouvelle réaction de transpeptidation (figure 25) (Hakenbeck et al., 1982).

Les β-lactamines peuvent bloquer chacune des trois étapes de synthèse décrites ci-dessus en inhibant chacune des classes de PBPs.



**Figure 25 : Représentation schématique des mécanismes de transpeptidation et de carboxypeptidation :**

L'hydroxyle de la serine active de l'enzyme (E) réalise une attaque nucléophile sur le D-Ala préterminal, ce qui forme un acyl-enzyme. L'acyl-enzyme subit soit une aminolyse dans le cas de la transpeptidation, soit une hydrolyse dans le cas de la carboxypeptidation.

Les PBPs semblent avoir également un rôle dans la division cellulaire et interviennent dans la synthèse du peptidoglycane pendant la septation (Spratt et al., 1975 ; Botta et Park, 1981). Elles pourraient faire partie d'un complexe multi-enzymatique de la division cellulaire

(Bramhill, 1997). En particulier, l'extrémité N-terminale de PBP2x pourrait intervenir dans le recrutement et la mise en contact des autres partenaires du divisome (Pares et al. 1996 ; Gordon et al. 2000).

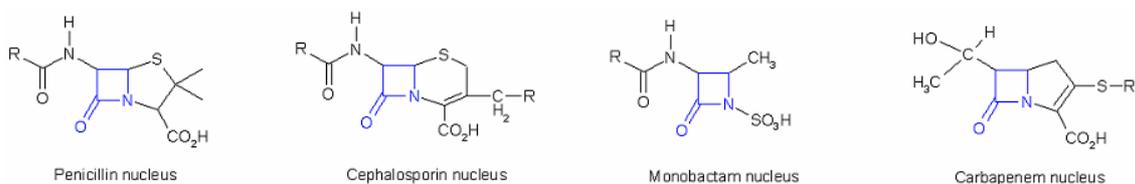
Selon la séquence et la localisation dans la bactérie de chacune des PBPs, il est possible de leur attribuer un rôle plus ou moins spécifique.

Lors de la division cellulaire PBP2x, PBP1a, PBP2a et PBP2b colocalisent initialement sur l'équateur. Durant le cycle PBP2x et PBP1a restent au milieu de la cellule en division sur le septum. PBP2a et PBP2b se dupliquent et se séparent lors du cycle cellulaire. PBP1b se localise à la fois dans le septum et à la périphérie (Morlot et al. 2003). Ainsi, PBP2x est considérée comme agissant plus particulièrement dans la division cellulaire, le gène *pbp2x* étant en outre localisé dans un « cluster » dédié à la division cellulaire (Massidda et al. 1998). A l'inverse, PBP2b agirait plus spécifiquement dans la phase de croissance cellulaire.

### 1.3 Les PBPs et les $\beta$ -lactamines

#### 1.3.1 Le mode d'action des $\beta$ -lactamines

Les  $\beta$ -lactamines, dont fait partie la pénicilline, sont des antibiotiques très utilisés en médecine et représentent près de 60% des antibiotiques utilisés dans les thérapies anti-infectieuses. Ces molécules sont caractérisées par un cycle  $\beta$ -lactame (amide cyclisée en  $\beta$ ), et leurs cibles primaires sont les PBPs. Elles peuvent être classées selon leur structure chimique (figure 26).

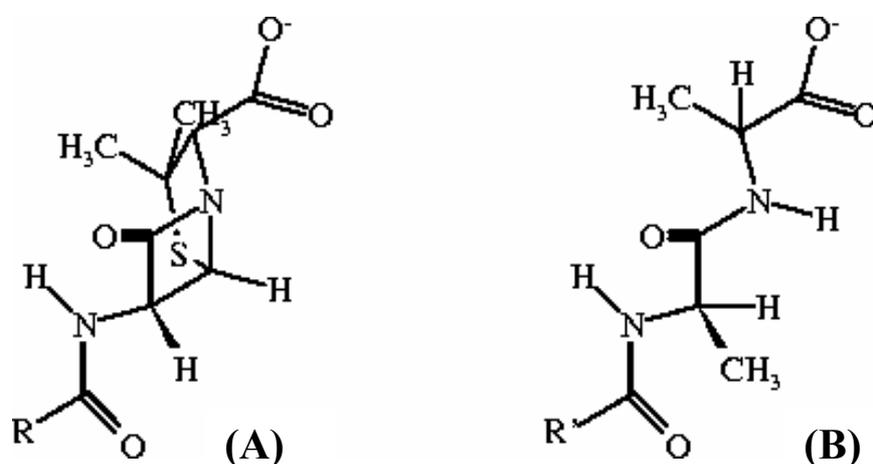


**Figure 26 : Structures chimiques des principaux noyaux des  $\beta$ -lactamines**

Depuis leur apparition, différentes générations d'antibiotiques naturels ou semi-synthétiques ont été développées. Mais l'utilisation massive de ces composés a conduit à l'émergence de souches de bactéries résistantes. L'impossibilité de combattre efficacement ces souches est devenue un véritable problème de santé publique.

En 1965 Tipper et Strominger montrent que ces antibiotiques bloquent l'étape finale de la biosynthèse du peptidoglycane : la transpeptidation (Tipper et Strominger, 1965)

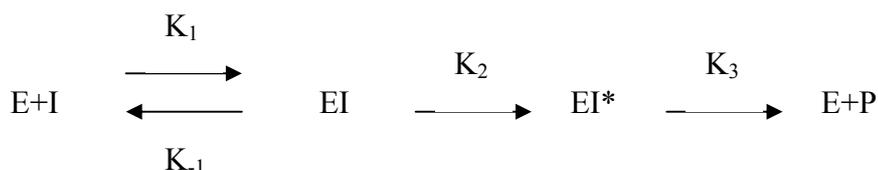
Les propriétés antibiotiques des  $\beta$ -lactamines sont dues à l'analogie de structure entre la fonction amide du cycle  $\beta$ -lactame et les liaisons D-Ala-D-Ala des substrats naturels des PBPs. (figure 27)



**Figure 27 : Homologie structurale entre les  $\beta$ -lactamines (A) et le peptide D-Ala-D-Ala (B) substrat naturel des PBPs.**

Ces inhibiteurs suicides ont une action bactériostatique due à l'arrêt de la synthèse du peptidoglycane, mais aussi un effet bactéricide.

Le mécanisme d'interaction entre les PBPs et les  $\beta$ -lactamines se décompose en trois étapes (Frère et al., 1975) représentées par la formulation ci-dessous :



Où E représente l'enzyme active, I la  $\beta$ -lactamine active et P la  $\beta$ -lactamine inactive.

Un complexe non covalent (EI) de Michaelis est formé de façon réversible. Puis, l'anneau  $\beta$ -lactame ouvert et le site de la serine active se lient de façon covalente et forment ainsi un complexe acyl-enzyme (EI\*). Enfin ce complexe est déacylé, l'enzyme est régénérée et l'antibiotique est relargué, temporairement inactif (Jamin et al., 1993).

### 1.3.2 La résistance des PBPs aux $\beta$ -lactamines

Dès le début de l'utilisation des  $\beta$ -lactamines, des organismes résistants sont apparus. La résistance aux  $\beta$ -lactamines peut résulter de deux mécanismes différents : la production de  $\beta$ -lactamases ou la modification des PBPs.

#### 1.3.2.1 Les $\beta$ -lactamases

Dans la majorité des cas, les bactéries produisent des  $\beta$ -lactamases qui inactivent l'antibiotique avant qu'il atteigne sa cible en hydrolysant le noyau  $\beta$ -lactame. Les  $\beta$ -lactamases sont sécrétées dans le milieu extérieur pour les bactéries Gram positives et dans le périplasme pour les bactéries Gram négatives. Les  $\beta$ -lactamases peuvent être divisées en quatre groupes selon leur taille, la nature de leur site actif et leur alignement de structure (Massova et Mobashery, 1998 ; Matagne et al., 1999). Les motifs caractéristiques de la super famille ASPRE sont retrouvés dans les  $\beta$ -lactamases de type A, C, et D (Joris et al., 1988). Elles possèdent toutes trois une serine au sein de leur site actif. Les  $\beta$ -lactamases de type A sont codées par un plasmide, tandis que celles des types C et D sont codées par un

chromosome. Les  $\beta$ -lactamases de type B sont quant à elles des métalloenzymes à  $Zn^{2+}$  et possèdent un large spectre d'activité. Leur expression peut être constitutive ou inductible.

### 1.3.2.2 La modification des PBPs

Certaines bactéries, comme *Streptococcus pneumoniae*, ne possèdent pas de  $\beta$ -lactamases, elles peuvent alors développer des PBPs de haute masse moléculaire qui ont perdu leur affinité pour les  $\beta$ -lactamines ou en acquérir de nouvelles qui sont de faible affinité pour les antibiotiques (Hakenbeck et al., 1999 ; Dowson et al., 1994). Ainsi, la mutation d'un seul acide aminé modifie l'affinité des antibiotiques pour les PBPs et constitue un premier niveau de résistance aux  $\beta$ -lactamines (Grebe et Hakenbeck, 1996). En fait, dans ces nouvelles protéines des dizaines de substitutions peuvent être observées. Ceci résulte de la structure mosaïque des gènes codants pour les PBPs acquise grâce à des événements de recombinaison moléculaire intra et inter-espèce. Mais la situation est plus complexe. En effet, dans les résistants isolés cliniquement, plus de quatre PBPs peuvent être altérées (Laible et al., 1991) et les effets de ces mutations semblent être coopératifs. Les PBPs des organismes résistants ont une faible affinité pour les  $\beta$ -lactamines et leur activité hydrolytique sur des substrats analogues diminue (Mouz et al., 1998). Cependant l'efficacité de leur turn-over catalytique n'est pas affectée. Il en résulte une paroi plus hydrophobe et composée de peptides chimiquement différents chez les organismes résistants (Garcia-Bustos et al., 1988, 1990).

## 1.4 La protéine PBP-2X

La protéine PBP2X est donc une cible primaire des antibiotiques de la classe des  $\beta$ -lactamines, et elle intervient dans les phénomènes de résistance à ces antibiotiques. Elle apparaît, en effet, très souvent mutée dans des souches de *Streptococcus pneumoniae* résistantes (Grebe et Hakenbeck, 1996).

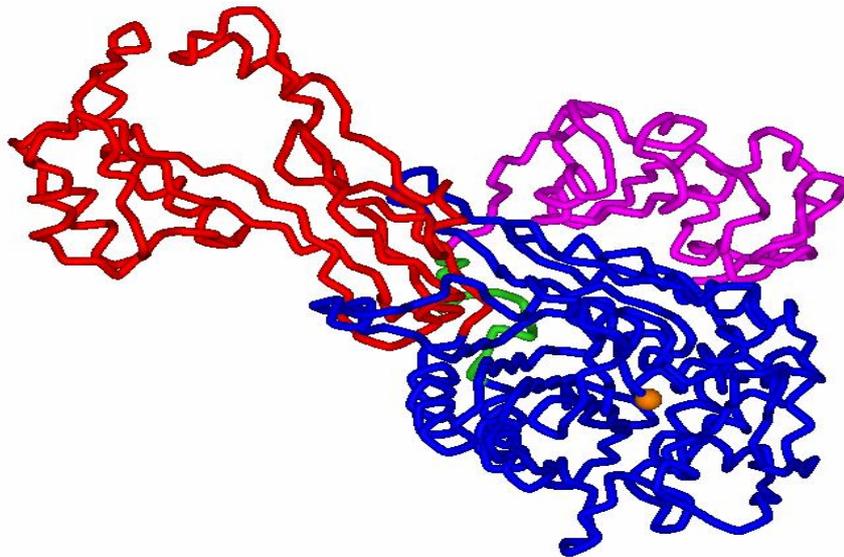
La structure de PBP2X a été obtenue par cristallographie d'abord à 3.5Å de résolution (Parès et al. 1996), puis à 2.4 Å de résolution (Gordon et al. 2000). Cette structure n'a pu être obtenue sur la protéine PBP2X entière mais sur une partie de celle-ci : PBP2X\*. PBP2X\* est constituée de la région periplasmique soluble de PBP2X déletée de la partie cytoplasmique et

de la région hydrophobe d'ancrage à la membrane. Ainsi constituée d'une PBP2X sans les acides aminés 1 à 48, cette protéine est soluble.

Les différentes structures révèlent une organisation en trois domaines. Le premier domaine, le domaine N-terminal (1-265), est un long domaine en forme de « pince à sucre ». Sa fonction n'est pas encore connue, mais il pourrait intervenir dans le recrutement des protéines du divisome.

Le domaine central de PBP2X (266-316) est le domaine TP et fait partie de la famille ASPRE. Il possède une organisation similaire à celle des  $\beta$ -lactamases de classe A. La serine active, qui fait partie des motifs ASPRE, est en position 337, au centre d'un long sillon. Ce domaine possède 8 hélices  $\alpha$  et 5 brins  $\beta$ .

Après le « linker » (617-629), se trouve le domaine C-terminal (630-750). Sa structure suggère une duplication de gène codant pour un domaine  $\alpha/\beta$  aboutissant à un domaine  $\alpha/\beta/\alpha/\beta$ . La fonction de ce domaine n'est pas connue, mais intervient probablement dans des interactions (figure 28).



**Figure 28 : Structure de PBP2X :**

Le domaine N-terminal est coloré en rouge, le domaine central TP en bleu, le linker en vert, et le domaine C-terminal en rose. Dans la cavité du domaine TP, la serine du site actif est représentée en orange.

D'après Parès et al., 1996 et Gordon et al., 2000.

Les résistants cliniques de *Streptococcus pneumoniae* présentent près de 100 mutations localisées à des endroits divers sur PBP2X (Hakenbeck et al. 1998 ; Asahi et al. 1999). Il existe cependant des points chauds de mutation (Dessen et al. 2001). Ainsi, on observe beaucoup de mutations dans le site actif, dans le linker et dans la partie C-terminale, par contre la partie N-terminale en contient très peu. Les acides amines les plus fréquemment mutés sont Thr 338, Gln 552 et Met 339 (Mouz et al. 1998 ; Mouz et al. 1999 ; Chesnel 2003). La mutation de la Thr 338 induit l'absence d'un groupement hydroxyl, ce qui déstabilise le site actif. La mutation de la Met 552 crée un effet électrostatique avec l'antibiotique ce qui diminue également l'efficacité d'acylation. La faible affinité de liaison des  $\beta$ -lactam aux PBP2X des organismes résistants est en grande partie attribuée à la diminution de l'efficacité de formation de la liaison ester et dans une moindre mesure à la faible reconnaissance du site actif, ce qui expliquerait le large spectre de résistance des PBP2X mutées (Lu et al., 2001).

## **2 LES ECHANGES H/D**

Les échanges hydrogène/deutérium sont utilisés depuis plusieurs décennies pour étudier la structure des protéines et leurs interactions avec différents ligands. Cependant la quantification des taux de deutération a longtemps été problématique. Différentes techniques ont été testées, comme par exemple la spectroscopie infrarouge ou la RMN. La spectrométrie de masse s'est aussi révélée un outil de choix de par sa rapidité de mise en œuvre et sa faible consommation d'échantillons.

### ***2.1 La théorie des échanges H/D***

Lorsqu'une protéine est incubée dans un solvant deutéré, ses hydrogènes peuvent s'échanger avec les deutériums du solvant. Cependant tous les hydrogènes d'une protéine n'ont pas le même comportement vis à vis de la deutération. Trois groupes d'hydrogène peuvent être distingués.

Le premier groupe est constitué des hydrogènes aliphatiques et aromatiques. Ils ne s'échangent pas dans les conditions standard. Ces hydrogènes ne seront concernés par les substitutions isotopiques que suite à des catalyses chimiques importantes.

Le deuxième groupe subit des réactions d'échange très rapides. Il est constitué des hydrogènes labiles des chaînes latérales des acides aminés et des hydrogènes des fonctions carboxy et amino terminales. Dans l'échelle de temps de nos expérimentations, les vitesses d'échange entre ces hydrogènes et le solvant sont trop rapides pour que nous puissions les observer.

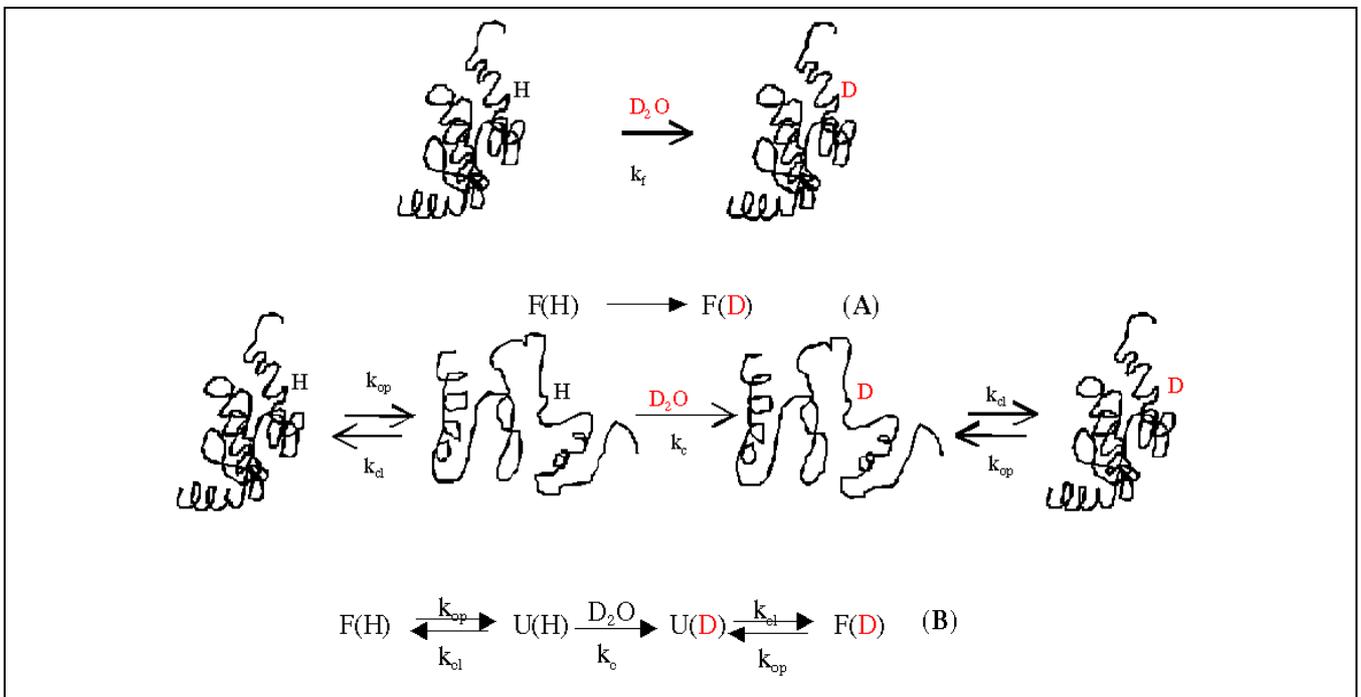
Le troisième groupe est constitué des hydrogènes amidiques des liaisons peptidiques. La vitesse d'échange de ces hydrogènes est très variable dans une protéine structurée et va refléter l'environnement local de chaque acide aminé dans la structure tridimensionnelle. La vitesse d'échange isotopique d'un proton amidique d'une liaison peptidique est fortement dépendante de son implication dans des liaisons hydrogènes intramoléculaires et de son accessibilité au solvant. Ainsi, des protons participant à des liaisons hydrogènes dans des hélices  $\alpha$  ou des feuillets  $\beta$  ou enfouis au cœur de la protéine auront des taux d'échange beaucoup plus faibles que ceux situés dans des boucles flexibles à la surface de la protéine.

### 2.1.1 Les échanges H/D dans les protéines et dans les peptides.

La vitesse d'échange isotopique des hydrogènes amidiques d'un peptide sont extrêmement différentes selon si ce peptide est structuré ou non. Pour chacun des hydrogènes concernés, son implication dans des liaisons hydrogène intramoléculaires (Hilser et Freire, 1996), sa distance par rapport à la surface de la protéine (Resing et al., 1999), la flexibilité de la région peptidique à laquelle il appartient (Zhang et al., 1996), sont différents facteurs connus dont dépend la vitesse d'échange isotopique.

Différents modèles cinétiques ont été développés pour décrire les échanges H/D dans une protéine native. Deux composantes sont intégrées dans ses modèles : le facteur structural qui représentera les dynamiques de dépliement/repliement et le facteur chimique ( $k_c$ ) qui représentera la vitesse d'échange des protons totalement exposés dans des zones non structurées. Différentes hypothèses expliquant la levée des barrières structurales ralentissant l'échange, ont été proposées. Elles impliquent ou la pénétration du solvant dans la structure, ou des dépliements localisés (Woodward et al., 1982 ; Englander et Kallenbach, 1984 ; Miller et Dill, 1995).

Actuellement, il est admis que les échanges peuvent se réaliser soit directement sans dépliement, soit dans des formes de dépliement transitoires, la structure d'une protéine étant non pas figée mais dynamique (figure 29).



**Figure 29 : Modèles d'échanges H/D**

L'échange a lieu directement (A), ou sur des formes transitoirement dépliées (B).  
D'après Engen et Smith, 2000.

L'échange sans dépliement serait prédominant pour les hydrogènes amidiques exposés au solvant et non impliqués dans des liaisons hydrogènes. A l'inverse, la protéine doit se trouver dans un état transitoirement déplié pour que se produisent des échanges au sein des hydrogènes assurant la stabilité des structures secondaires et tertiaires, c'est à dire peu accessibles au solvant et impliqués dans des liaisons hydrogènes. Même dans des conditions physiologiques qui favorisent la forme native, certaines molécules peuvent se déplier pendant un temps très court. Ce dépliement peut être localisé ou impliquer la protéine entière (global). Les constantes de vitesse du dépliement et du repliement sont désignées sur la figure par  $k_{op}$  et  $k_{cl}$  respectivement. En accord avec ce modèle, l'échange isotopique ne peut avoir lieu qu'après une rupture transitoire des liaisons intramoléculaires rendant les hydrogènes amidiques échangeables.

La vitesse d'échange peut être exprimée comme la somme des contributions des échanges des formes pliées et dépliées (Woodward et al., 1982) selon cette équation :

$$k_{obs} = k_f + k_u$$

Où  $k_f$  est la constante de vitesse d'échange résultant de la dynamique conformationnelle de l'état natif,  $k_u$  est la constante de vitesse pour l'échange nécessitant un dépliement et  $k_{obs}$  la constante de vitesse globale observée.

La constante  $k_f$  peut être exprimée comme suit (Kim et Woodward, 1993) :

$$k_f = \beta \cdot k_c$$

Où  $\beta$  est la probabilité qu'un hydrogène particulier soit accessible au solvant (et aux catalyseurs) et  $k_c$  la constante de vitesse d'échange de l'hydrogène concerné si celui-ci se trouvait dans un peptide non structuré.

$\beta$  est alors dépendant de l'accessibilité au solvant et de la participation de l'hydrogène concerné à des liaisons hydrogènes. Le facteur  $k_c$  est quant à lui lié aux conditions expérimentales (pH, température) et à la nature des chaînes latérales des acides aminés voisins.

## 2.1.2 L'influence du pH

Les échanges isotopiques sont fortement dépendants du pH. La réaction peut s'effectuer par des mécanismes de catalyse acide ou basique.

La constante de vitesse d'échange H/D,  $k_E$ , s'exprime comme suit :

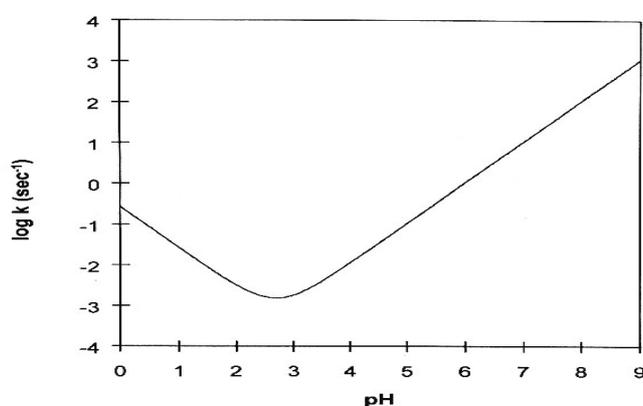
$$k_E = k_H [H^+] + k_{OH} [OH^-]$$

Où  $k_H$  et  $k_{OH}$  sont les constantes d'échange en catalyse acide et en catalyse basique respectivement.

Un troisième mécanisme d'échange direct avec l'eau a été également observé, mais sa constante de vitesse est généralement négligeable (Roder et al., 1985).

Des études sur des peptides modèles, des polyalanines, ont montré que  $k_{OH}$  est supérieure d'un facteur  $10^8$  à  $k_H$  (Englander et Kallenbach, 1984). Le pH influence donc les vitesses d'échange qui sont principalement conditionnées par les ions  $OH^-$  pour des pH supérieurs à 3 et par les ions  $H^+$  pour des pH inférieurs à 3. Ainsi, la constante de vitesse  $k_E$  peut être reliée au pH (figure 30).

Le minimum d'échange se produit pour un pH voisin de 2-3, pour lequel les échanges dus à la catalyse acide sont équivalents à ceux dus à la catalyse basique. En dehors de cette zone de pH, la constante de vitesse d'échange augmente d'un facteur 10 par unité de pH (Bai et al., 1993). C'est pour cette raison qu'il est impératif de contrôler très précisément le pH lors des expériences d'échange H/D.



**Figure 30 : Influence du pH sur la vitesse des échanges H/D**

Cas d'un peptide polyalanine.

D'après Englander et Kallenbach, 1984

### **2.1.3 L'influence de la température**

Les échanges H/D sont également, dans une moindre mesure, dépendant de la température. En effet, une augmentation de la température de 10°C se traduit par une augmentation de la vitesse d'échange d'un facteur 3 (Zhang et Smith, 1993).

La forte dépendance de la constante de vitesse des échanges vis à vis de la température et du pH, sera utilisée dans les protocoles expérimentaux. En effet, en dehors des périodes de marquages, le pH et la température seront diminués afin de limiter au maximum les ré-échanges. C'est l'étape de « quenching » (cette étape sera développée plus bas).

### **2.1.4 L'influence de la séquence locale en acides aminés.**

Les vitesses d'échange des hydrogènes amidiques sont également affectées par les chaînes latérales des acides aminés voisins. Cette sensibilité a été étudiée avec des di-peptides modèles (Bai et al., 1993). Ces études ont montré que les résidus d'acides aminés polaires (serine, thréonine, cystéine et méthionine) avaient tendance à délocaliser la densité d'électrons de la liaison peptidique. L'effet inductif ainsi créé va augmenter la constante de catalyse basique et donc réduire la fixation de deutérium sur la liaison peptidique.

Les acides aminés tels que la leucine, l'isoleucine, la valine, la phénylalanine, la tyrosine et le tryptophane réduisent quant à eux la constante de vitesse d'échange du fait de leur faible activité catalytique acide et basique.

La constante d'échange des acides aminés basiques et acides n'est que très faiblement modifiée.

Certaines chaînes latérales des acides aminés apolaires peuvent également avoir une influence sur la constante de vitesse d'échange de par l'encombrement stérique qu'ils peuvent générer.

## 2.2 *Protocoles expérimentaux*

### 2.2.1 Les techniques de marquage

Deux méthodes de marquage existent (Deng et al., 1999). La plus couramment utilisée est le marquage continu et certaines études utilisent également le marquage pulsé.

Le marquage continu consiste à incuber une protéine avec du D<sub>2</sub>O en présence ou non de dénaturant. Lors du marquage continu, les populations des états natifs et dépliés changent durant le temps d'exposition de la protéine au D<sub>2</sub>O. Les protéines qui sont, ou deviennent, dépliées pendant le marquage se deutèrent entièrement, et les protéines qui ne se déplient pas pendant ce temps ne sont deutérées que partiellement. Les niveaux de deutération des protéines marquées de cette manière rendent compte du nombre de molécules qui se déplient durant le marquage. Le spectre de masse obtenu dans ces conditions représentera alors le cumul des populations ayant existé dans un état donné durant la période de marquage. Ce type de marquage est utilisé préférentiellement lors de l'analyse des protéines dans des conditions où l'état natif est favorisé. Il peut être utilisé pour identifier des régions de la protéine qui se déplient à des vitesses différentes (Deng et al. 1999). Il permet la détermination directe des vitesses de repliement dans les peptides. Cette approche est particulièrement efficace lorsque l'échange isotopique dans la forme native de la protéine est lent.

Le marquage pulsé consiste à laisser la protéine un temps donné dans les conditions spécifiques pour lesquelles on veut des informations (conditions dénaturantes ou renaturantes). Le marquage au D<sub>2</sub>O est ensuite réalisé pendant un temps très bref. L'exposition de la protéine au D<sub>2</sub>O est alors très courte par rapport à l'échelle de temps des dynamiques de dépliement/repliement. Les niveaux de deutération résultant d'un marquage pulsé sont alors représentatifs des populations instantanées des molécules pliées et dépliées. Le spectre de masse sera une représentation instantanée des populations de protéine présentes à la fin de l'étape de deutération. Les deutériums détectés refléteront seulement les molécules qui se seront dépliées pendant le temps bref de marquage et ne représenteront pas les molécules qui auraient pu se déplier ou se replier avant l'ajout de D<sub>2</sub>O. Le marquage pulsé peut être utilisé pour étudier le repliement et le dépliement des protéines dans les conditions où la population de protéines dépliées est artificiellement augmentée par l'ajout de dénaturant, le chauffage ou la variation du pH. Un aperçu de la séquence des événements de dépliement de la protéine pourra être ainsi obtenu.

Ces protocoles ont été notamment utilisés pour étudier la dénaturation de l'aldolase par l'urée (Deng et al., 1999) ou encore les cinétiques de dépliement et de repliement du lysozyme (Canet et al., 1999). Pour ce type de marquage très bref, un système spécifique de cinétique rapide doit être utilisé, c'est le système « quenched flow ».

### **2.2.2 L'étape de blocage des rééchanges (ou « quenching »)**

Une étape clé de la méthode d'échange H/D consiste à bloquer les échanges isotopiques après avoir exposé la protéine à un tampon deutéré. C'est l'étape de « quenching ». Cette étape est primordiale, car pour être interprétables, les taux de deutération observés par spectrométrie de masse devront être les plus proches possibles de ceux réellement présents à la fin du temps de deutération choisi. Il nous faudra obtenir une « photographie » de l'état de deutération de la protéine à la fin de la période de marquage.

La plupart des étapes suivant la deutération (chromatographie, digestion enzymatique, analyse par spectrométrie de masse) se déroulent en milieu hydrogéné et dans des conditions dénaturantes. Ces conditions peuvent favoriser le rééchange des deutériums incorporés lors du marquage. Les échanges, et donc les rééchanges, étant, comme nous l'avons déjà vu, fortement liés aux conditions expérimentales, le « quenching » consistera à diminuer le pH et la température. Ainsi, pour limiter un maximum les rééchanges, toutes les étapes suivant la deutération seront réalisées à pH 2,5 et à 0°C. Ces conditions expérimentales pourront être parfois contraignantes et imposer des changements dans les protocoles habituels, mais seront indispensables pour l'obtention de résultats significatifs.

### **2.2.3 La digestion enzymatique**

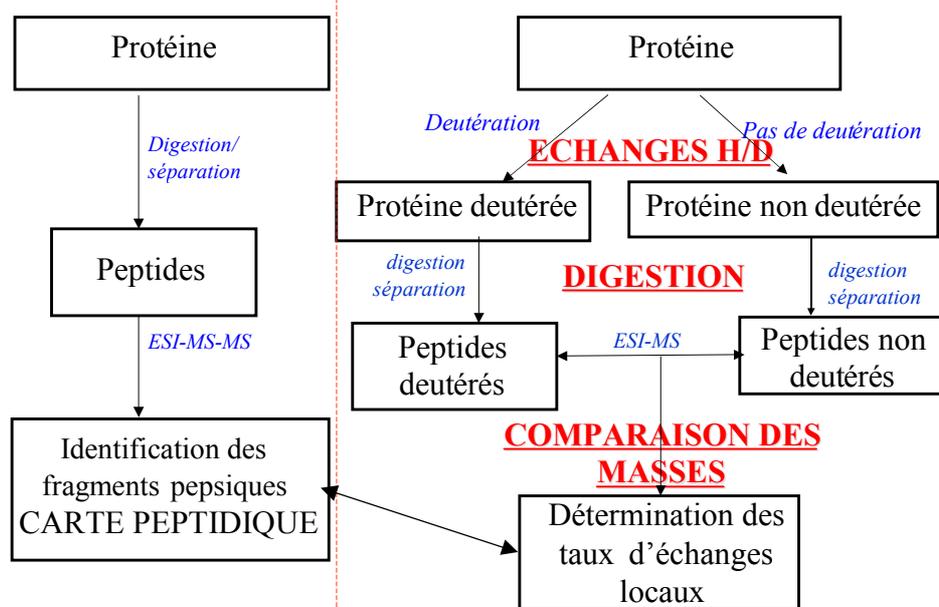
L'analyse des protéines entières après échanges H/D permet d'obtenir des informations mais ne renseigne pas sur la localisation des changements structuraux éventuels.

Des techniques de protéolyse donneront des informations de deutération plus locales.

En 1993, Zhang et Smith ont développé une technique de protéolyse en adaptant des travaux plus anciens (Rosa et Richards, 1979 ; Englander et Kallenbach, 1984). Le tritium a été remplacé par le deutérium comme sonde conformationnelle et la dégradation d'Edman a été

remplacée par la spectrométrie de masse pour l'identification des fragments. La protéolyse devant se réaliser dans les conditions de blocage du rééchange (à pH 2,5 et à 0°C), l'enzyme la plus couramment utilisée est la pepsine. Cette enzyme, non spécifique, conduit à la formation de nombreux peptides parfois chevauchant, ce qui permet l'obtention d'une meilleure résolution dans la localisation des zones d'intérêt, cependant cette faible spécificité complique beaucoup l'identification des peptides obtenus.

Classiquement, les expériences d'échange H/D associées à la spectrométrie de masse suivent un protocole identique (figure 31). La protéine d'intérêt est digérée par la pepsine, les fragments sont identifiés par MS-MS. La protéine est incubée parallèlement dans un solvant deutéré et dans un solvant non deutéré pendant le temps voulu, la réaction d'échange est bloquée par diminution du pH et de la température. La protéine est ensuite digérée par la pepsine, les peptides obtenus sont séparés par HPLC et analysés en ligne par spectrométrie de masse. La comparaison de la masse des peptides issus de la protéine non deutérée avec la masse de ceux issus de la protéine deutérée, permettra de déterminer un taux de deutération pour chacun des peptides. Les peptides issus de la digestion pepsique ayant été préalablement identifiés par des expériences de MS-MS, il sera ainsi possible de localiser les zones d'incorporation de deutériums et donc les régions d'intérêt.



**Figure 31 : Protocole expérimental pour la détermination d'échanges H/D locaux**

Il est également possible d'étudier l'incorporation de deutériums sur une protéine par spectrométrie de masse MALDI (Nazabal et al., 2003). Les échantillons deutérés et digérés sont alors directement déposés sur la plaque. L'avantage du MALDI-MS par rapport à ESI-MS est qu'un grand nombre d'échantillons peuvent être analysés dans une courte période. De plus, cette technique est plus tolérante aux sels que l'électrospray. Enfin, elle produit essentiellement des ions monochargés, ce qui facilite l'interprétation. Cependant, contrairement à l'électrospray, elle ne peut être couplée en ligne avec des systèmes de chromatographie liquide, qui sont très utiles pour l'analyse de mélanges complexes.

Même dans les conditions de « quenching », quelques atomes d'hydrogène de la protéine peuvent s'échanger avec des deutériums présents dans le milieu lors de l'étape de digestion, et inversement, certains de ces deutériums peuvent se rééchanger avec des hydrogènes lors de l'étape de séparation des peptides par HPLC. L'analyse après l'étape de deutération doit donc être la plus rapide possible. Il est nécessaire de faire un compromis entre le temps nécessaire pour une séparation optimale des peptides, afin de simplifier l'analyse, et la minimisation du temps entre la fin de la deutération et l'arrivée dans le spectromètre. Il est également possible de corriger les gain et perte de deutération des liaisons peptidiques. Cette correction se fait grâce à la présence de témoins, le contrôle 0% et le contrôle 100% de deutération (Zhang et Smith, 1993).

Cependant, si les temps d'analyse restent raisonnables et si les différentes analyses sont réalisées dans des conditions identiques, la correction n'est pas nécessaire et les taux de deutération obtenus seront considérés comme d'une exactitude suffisante pour l'application. Tout va alors dépendre de la précision que l'on souhaite obtenir dans la détermination des taux d'incorporation.

## ***2.3 Applications des échanges H/D couplés à la spectrométrie de masse dans l'étude des protéines***

### **2.3.1 Etude du repliement des protéines**

La technique des échanges isotopiques couplés à la spectrométrie de masse est complémentaire des méthodes physico-chimiques classiques et de la RMN. En effet, ces

techniques donnent des valeurs moyennes de distribution des molécules de protéine, alors que les échanges H/D couplés à la spectrométrie de masse permet la mise en évidence de plusieurs conformères et est représentative des différentes populations de protéines existant en solution. Il est alors possible de mettre en évidence différents intermédiaires de repliement pour une protéine (Miranker et al., 1993).

La méthode de digestion pepsique permet d'accéder à des informations plus détaillées sur la conformation des protéines, grâce à la détermination des vitesses d'échange H/D dans les différentes régions de la protéine. Ainsi des études sur le cytochrome c et sur un domaine SH3 d'une kinase ont pu mettre en évidence des zones des protéines qui s'échangent à des vitesses différentes (Dharmasiri et Smith, 1996 ; Engen et al., 1998).

La technique des échanges H/D couplés à la spectrométrie de masse peut également être utilisée pour étudier les chemins de dépliement/repliement des protéines. Par exemple, des intermédiaires de repliement ont été ainsi identifiés pour l'aldolase et la protéine CRABP-1 (Deng et Smith, 1998 ; Eyles et al., 2000).

Il est également possible d'étudier des changements conformationnels induits par la liaison de la protéine à des cofacteurs ou encore par l'introduction de mutations au sein de la séquence protéique (Halgand et al., 1999 ; Guy et al., 1996).

### **2.3.2 Etude des interactions protéine/protéine et protéine/ligand**

Au niveau des interfaces protéine/protéine, une diminution de l'échange isotopique peut être observée. Cette diminution est due à l'exclusion stérique du solvant de la zone d'interaction. Ainsi, la technique des échanges H/D couplés à la spectrométrie de masse peut être utilisée pour étudier les zones d'interactions entre les différents monomères d'un multimère (Zhang et al., 1996), pour caractériser un complexe antigène/anticorps (Yamada et al., 2002), ou encore pour étudier les interactions entre les protéines d'un complexe (Mandell et al., 1998).

De la même manière, cette méthode peut être utilisée pour caractériser les domaines de liaison protéine/ligand. La liaison du ligand entraîne une diminution de la vitesse d'échange au niveau du site de fixation, mais peut également causer des changements conformationnels dans des régions distantes du lieu de fixation du ligand (Anand et al., 2002).

### **2.3.3 Facilitation de la cristallogénèse**

Il est parfois impossible de produire des cristaux de protéine qui donnent des données de diffraction interprétables. Les régions non structurées des protéines jouent un rôle important dans ce problème car il est impossible d'obtenir des cartes de densité électronique de ces régions de par leur trop grande flexibilité. Ces régions peuvent être identifiées par échanges isotopiques couplés à la spectrométrie de masse. En effet, du fait de leur faible degré de structuration, elles ont une vitesse de deutération très élevée. Une fois identifiées, ces régions très désordonnées peuvent être supprimées de la protéine afin d'en faciliter la cristallogénèse (Pantazatos et al., 2004).

### **3 RESULTATS ET DISCUSSION**

#### ***3.1 Buts et principes de l'étude***

Nous avons donc voulu appliquer la technique des échanges H/D à la protéine PBP-2X. Nous avons suivi la procédure expérimentale classique précédemment décrite. Une carte peptidique de la protéine PBP-2X a été réalisée après digestion à la pepsine par LC-MS-MS sur un spectromètre de masse de type piège ionique, l'Esquire 3000+ (Bruker Daltonic). Puis des analyses LC-MS ont été réalisées en parallèle sur les peptides issus de la digestion de PBP-2X non deutérée et sur les peptides issus de la digestion de PBP-2X deutérée 30 secondes.

Après la réalisation de la carte peptidique, nous avons obtenu un recouvrement de séquence de 93%, et les expériences de deutération nous ont permis d'obtenir des informations sur l'état de deutération de 74% des hydrogènes amidiques.

Il n'a pas été possible de déterminer un taux de deutération pour tous les peptides que nous avons identifiés lors des expériences de LC-MS-MS. En fait, chaque peptide ne se deutère pas d'une façon totalement homogène, il en résulte la présence de plusieurs espèces deutérées différemment pour un même peptide. Le signal étant réparti sur l'ensemble de ces espèces, l'intensité des pics représentant les peptides deutérés est généralement plus faible que celle représentant les peptides non deutérés et le rapport signal/bruit peut parfois être insuffisant pour permettre la détection. De plus, pour deux peptides qui co-élueraient, on peut obtenir, lors de la deutération, une superposition des pics de masse qui ne permet plus de différencier les deux peptides.

Nous avons tout de même pu déterminer les régions de la protéine qui se deutéraient de façon importante. Cependant cette détermination est demeurée peu précise. En effet, la résolution spatiale des zones d'échange est limitée par la taille des peptides générés (entre 5 et 30 acides aminés), ainsi pour l'étude des protéines de taille importante, cette résolution nous est apparue insuffisante. De plus, le pourcentage de liaisons peptidiques pour lesquelles nous avons obtenu des informations de deutération est très moyen, et il se peut que certaines zones d'intérêt ne soient pas couvertes par ces résultats. Il nous est donc apparu indispensable de développer de nouvelles méthodes, dérivées de la méthode classique, qui pourraient permettre d'augmenter le recouvrement de séquence et la résolution spatiale des zones d'échange.

Pour cela, nous avons choisi l'utilisation de différentes protéases en sus de la pepsine.

La protéine PBP-2X sera à nouveau utilisée comme modèle. Sa structure tridimensionnelle étant résolue nous pourrions analyser nos résultats et valider la méthode.

Nous espérons par l'utilisation de différentes enzymes, obtenir plus de peptides et ainsi recouvrir la totalité de la séquence de la protéine étudiée. De plus, l'utilisation de plusieurs enzymes permettrait également d'obtenir des peptides chevauchant et ainsi de multiplier les informations sur une même zone. Cette redondance d'information améliorerait la précision de définition des zones d'intérêt.

Les conditions particulières d'expérimentation imposée par l'étape de « quenching » nous imposent de trouver des protéases capables de fonctionner à 0°C et à pH 2,5. Très peu de protéases, disponibles dans le commerce, sont actives dans ces conditions. Outre la pepsine, nous avons trouvé deux autres protéases compatibles avec les exigences du quenching : la protéase de type XIII d'*Aspergillus satoii* et la protéase de type XVIII de *Rhizopus species*.

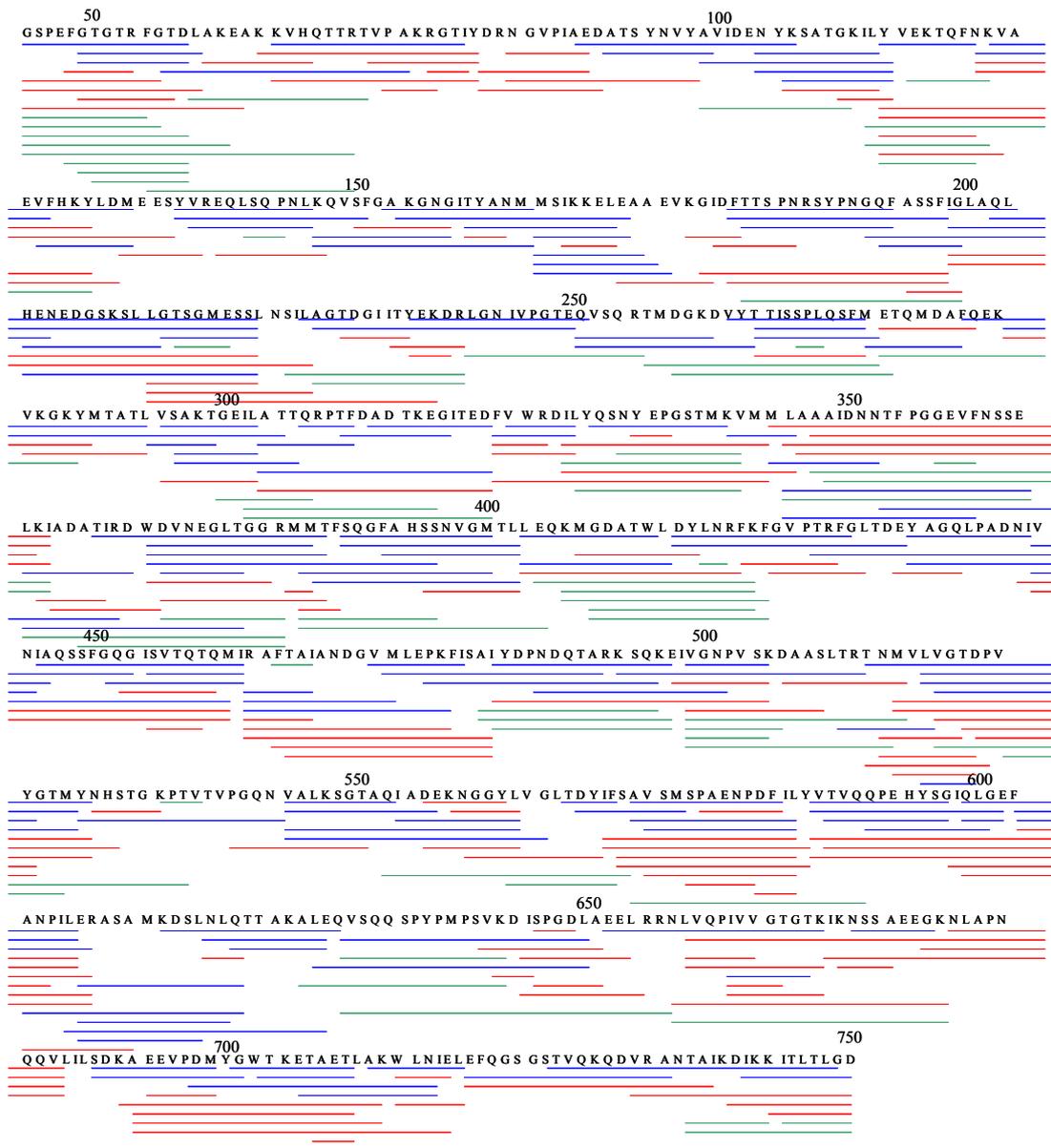
Pour chacune des enzymes, une carte peptidique de la protéine PBP-2X, puis des expériences de deutération sont réalisées. La comparaison des masses des peptides non deutérés avec celles des peptides deutérés, nous permet de déterminer un taux de deutération pour chacun des fragments protéolytiques. Le recoupement des résultats obtenus avec chacune des enzymes permet de déterminer un taux de deutération local pour les différentes régions de la protéine.

### **3.2 Recouvrement de la séquence**

Les cartes peptidiques sont réalisées pour chaque enzyme (digestion de 2 minutes) par des expériences de LC-MS-MS.

Le logiciel Data Analysis permet de générer une liste de composés, et le logiciel Biotools peut alors créer une carte peptidique théorique. Chaque spectre MS-MS est ensuite vérifié pour permettre la validation de l'identification du peptide. Seuls les peptides qui auront pu être identifiés sans ambiguïté seront conservés pour l'établissement de la carte peptidique.

Les séquences des peptides obtenus avec chaque protéase sont représentées sur la figure 34.



**Figure 34 : cartes peptidiques** : Les peptides obtenus avec chacune des trois protéases sont représentés par des traits sous la séquence de PBP-2X. Les peptides obtenus avec la pepsine, la protéase de type XIII et la protéase de type XVIII sont représentés, respectivement, en bleu, vert et rouge.

Lors de la digestion avec la pepsine, 149 peptides sont identifiés. Pour la protéase de type XIII, 69 le sont et pour la protéase de type XVIII, 144. Ces peptides représentent 93%, 40% et 84% de la séquence de PBP2X respectivement pour la pepsine, la protéase de type XIII, et la protéase de type XVIII. Finalement, aucune de ces protéases n'est réellement spécifique. Il est cependant important de noter que les profils de digestions sont dans les trois cas très

reproductibles. Le pourcentage de recouvrement est très variable d'une enzyme à l'autre, ceci peut être attribué à leur différence d'activité et dépend probablement de la nature de la protéine qu'elles digèrent.

L'utilisation d'une seule de ces enzymes ne permet pas de recouvrir entièrement la séquence, par contre si les cartes peptidiques des trois enzymes sont combinées, 99.7% de la séquence en acides aminés de PBP2X peut être recouverte.

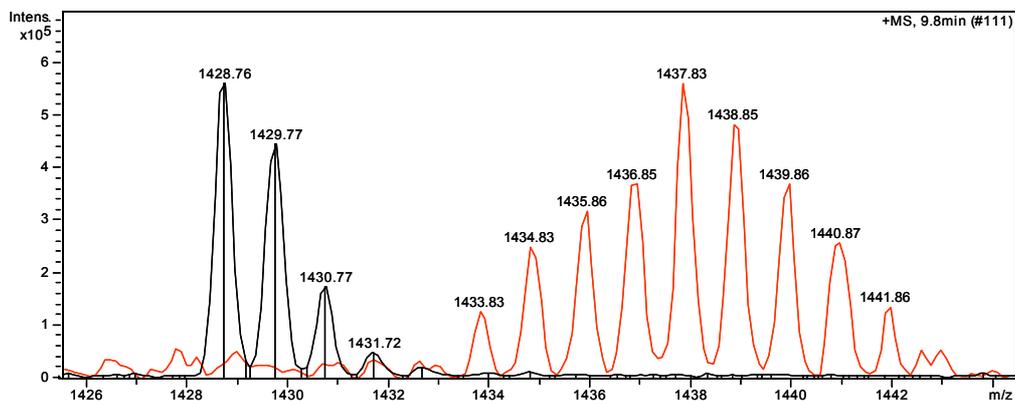
Le premier avantage dans l'utilisation combinée des trois protéases réside donc dans l'augmentation du recouvrement de la séquence en acides aminés. Par exemple, si nous considérons uniquement les résultats obtenus avec la pepsine, la région 84-92 de la protéine n'est pas couverte. Par contre la digestion à la protéase de type XVIII génère plusieurs peptides dans cette zone. De la même façon, avec la protéase de type XVIII, aucun peptide de la zone 246-267 n'est obtenu, alors que les résultats obtenus avec la protéase de type XIII et la pepsine nous permettent de recouvrir cette région. Ainsi la combinaison des résultats obtenus avec les trois protéases nous pouvons obtenir près de 100% de recouvrement de la séquence en acides aminés de PBP2X.

Un autre avantage de cette méthode est qu'elle permet l'obtention de données redondantes. Par exemple la protéase de type XIII fournit pour la protéine PBP2X des données que nous aurions pu obtenir avec l'une des deux autres protéases, mais cela nous permet de vérifier nos informations et ainsi d'obtenir des résultats plus fiables, notamment au niveau de l'identification des peptides.

### ***3.3 Deutération***

Des expériences de LC-MS sont réalisées en parallèle sur la protéine PBP2X dans un solvant deutéré et dans un solvant hydrogéné pour chacune des trois enzymes. Les profils d'élution des peptides issus de la digestion de la protéine non deutérée sont les mêmes que ceux obtenus par la digestion de la protéine deutérée.

Le nombre de deutériums incorporés est recherché pour chacun des peptides qui avaient été identifiés lors de la réalisation des cartes peptidiques. (Figure 35 et annexe 1).



**Figure 35 : exemple de deutération d'un ion** : le spectre de l'ion  $m/z$  1428.8 correspond au peptide 44-57 de la protéine PBP2X non deutérée. L'ion issu de la protéine PBP2X deutérée 30 secondes puis digérée a un rapport  $m/z$  de 1437.8 ( $m/z$  moyen). Il y a eu incorporation de 9 deutériums. Ces spectres ont été obtenus sur un appareil de type piège ionique.

La séquence des peptides deutérés obtenus avec chacune des protéases sont représentés sur la figure 36 en fonction de leur taux de deutération.

Après digestion de la protéine deutérée avec la pepsine, la protéase de type XIII et la protéase de type XVIII, l'incorporation de deutérium a pu être déterminée pour, respectivement, 84, 16, et 115 peptides, ce qui représente 72, 12.5 et 81% de recouvrement de séquence

Ainsi, pour la pepsine seule, on obtient des informations sur la deutération de 72% des hydrogènes amidiques, ce qui laisse la possibilité de manquer des zones importantes de la protéine dans les 28% restant. Cependant la combinaison des résultats obtenus avec les trois enzymes nous permet d'obtenir des informations sur près de 95% des hydrogènes amidiques de la protéine PBP2X.

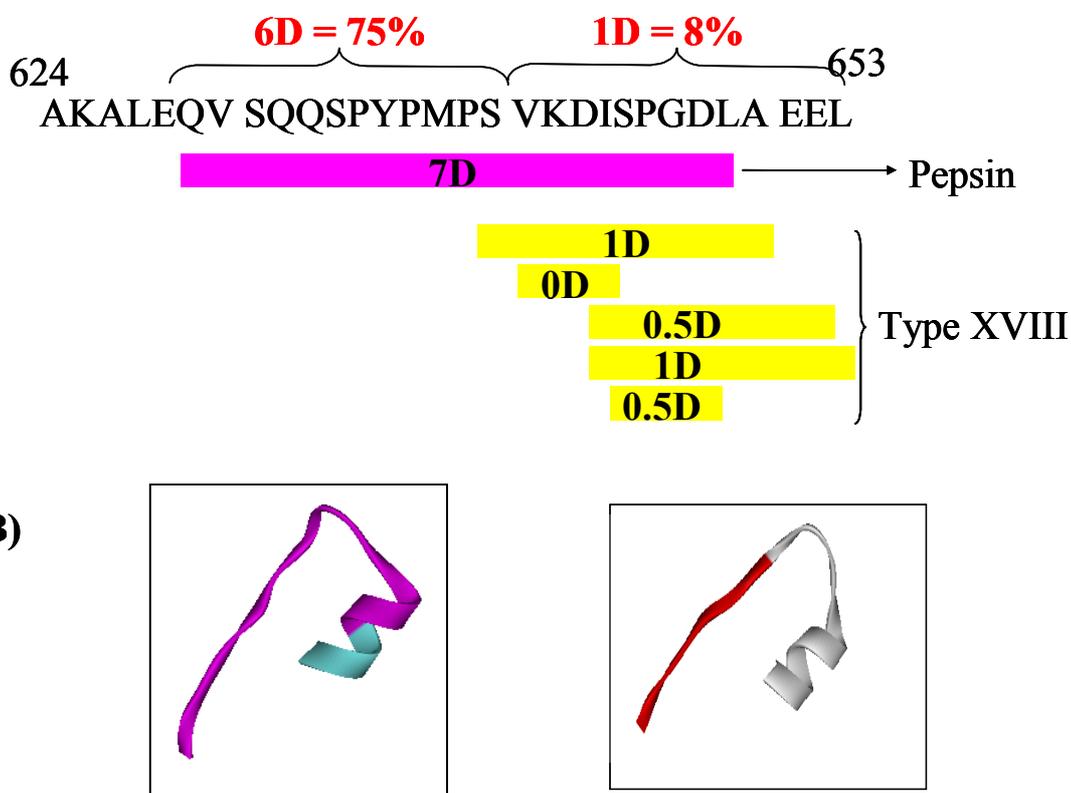
En fait, il n'est pas possible de déterminer le taux de deutération pour tous les peptides précédemment identifiés par MS-MS. Mais la combinaison des données obtenues avec chacune des trois enzymes nous permet d'augmenter significativement le taux de recouvrement de la protéine deutérée. Par exemple la digestion par la pepsine de la protéine deutérée ne nous a donné aucune information sur la région 62-102 de PBP2X, mais nous avons pu déterminer le taux de deutération de cette zone grâce aux peptides deutérés générés par la protéase de type XVIII (Figure 36). La protéase de type XVIII qui ne permet pas, au contraire d'obtenir des peptides deutérés de la région 568-573, mais l'incorporation de deutérium de cette zone peut être calculée grâce aux données obtenues avec la protéase de type XIII.



**Figure 36 : Incorporation de deutérium (après 30 secondes de deutération) pour chaque peptide obtenu avec les trois protéases.**

Les peptides obtenus avec la pepsine, la protéase de type XIII et la protéase de type XVIII sont représentées respectivement par des barres continues, des doubles soulignements et des barres hachurées. La couleur des barres indique le taux de deutération, les barres rouges, roses, jaunes et grises indiquent respectivement un taux de deutération supérieur à 50%, compris entre 50 et 30%, compris entre 30 et 10%, et inférieur à 10%.

L'utilisation combinée des trois enzymes nous permet de déterminer le taux de deutération pour des plus petites zones de la protéine grâce à la présence de peptides chevauchants. Nous avons ainsi amélioré la résolution spatiale des résultats.



**Figure 37 : Illustration de l'augmentation de la résolution spatiale par l'utilisation combinée de plusieurs protéases**

A : La séquence 624-653 est représentée, le peptide généré par la pepsine est en rose et ceux générés par la protéase de type XVIII est en jaune. Le nombre de deutériums incorporés par chaque peptide est noté sur les barres représentant le peptide.

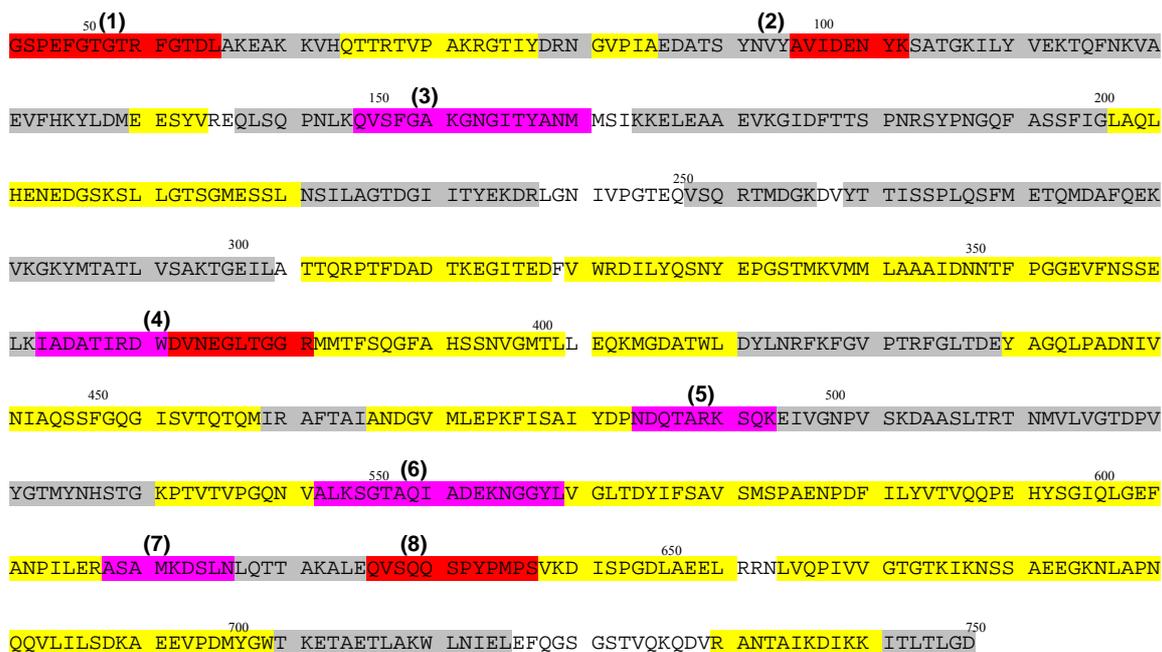
B : La structure de la région 632-653 est représentée. La première image représente les données obtenues avec la pepsine seule : 44% de deutération pour la zone 629-649 (en violet). La deuxième image représente les données obtenues avec la combinaison des résultats de la pepsine et de la protéase de type XVIII : la partie orientée vers le côté N-terminale de la zone est deutérée à 75% (représentée en rouge) et celle orientée vers le côté C-terminale à 8% (représentée en gris).

En fait la mesure de l'incorporation de deutérium d'un peptide permet seulement de calculer un taux de deutération moyen pour ce peptide. Celui-ci pouvant être long de plus de 10 acides aminés, nous ne pourrions avoir accès à l'information sur la répartition exacte des deutériums le long de la séquence. Mais les peptides chevauchants générés par les trois protéases permettront d'obtenir des données sur des plus petites parties de la protéine.

Par exemple, après digestion à la pepsine, un peptide de 2272.1 Da est identifié comme étant le fragment 629-649 de PBP2X. L'analyse de la protéine deutérée montre que ce peptide incorpore en 30 secondes 7 deutériums sur les 16 possibles. Cette région de la protéine semble donc avoir un taux de deutération plutôt important, mais il nous est impossible de définir avec précision où se trouve exactement la zone très fortement deutérée. Après digestion à la protéase de type XVIII, cinq peptides appartenant à cette région sont identifiés (les fragments 640-650; 641-644; 643-652; 643-653; 644-648). Ils incorporent respectivement 1, 0, 0.5, 1, 0.5 deutériums. La compilation des données obtenues avec la pepsine et avec la protéase de type XVIII permet de réduire la zone fortement deutérée à la région 629-640 (Figure 37). Statistiquement les deutériums se situent plus du côté N-terminal de la zone 629-649. Sur la structure tridimensionnelle cette statistique est vérifiée puisque la partie N-terminale de la zone correspond à une boucle flexible, alors que la partie C-terminale correspond à une hélice  $\alpha$  (moins sensible à la deutération).

Ainsi l'analyse combinée des résultats présentés sur la figure 36 permet d'établir le taux de deutération de chaque zone de la protéine (figure 38).

Huit zones de la protéine sont identifiées comme ayant un fort taux de deutération : les régions 44-58, 98-105, 148-163, 366-384, 487-496, 545-562, 611-619 et 629-640. Elles correspondent aux zones les plus facilement échangeables de la protéine PBP2X.



**Figure 38 : Représentation du taux de deutération sur la séquence primaire de PBP2X**

Les régions de la protéine sont surlignées en rouge, rose, jaune et gris en fonction de leur taux de deutération, respectivement, supérieur à 50%, compris entre 50 et 30%, compris entre 30 et 10%, et inférieur à 10%.

### 3.4 Validation des résultats sur la structure tridimensionnelle

Pour s'assurer de la pertinence de la nouvelle méthode développée nous avons validé nos résultats sur la structure tridimensionnelle de la protéine PBP-2X. L'utilisation combinée des trois enzymes nous a permis d'obtenir des informations sur la localisation des deutériums après 30 secondes de deutération. Lors de l'utilisation d'un temps de deutération court comme celui-ci, seules les zones les plus accessibles au solvant et les moins structurées sont deutérées.

La validité des résultats obtenus concernant la localisation des zones accessibles de la protéine PBP-2X peut être confirmée par la structure tridimensionnelle de la protéine obtenue par cristallographie d'abord à 3,5Å de résolution (Pares et al., 1996) puis à 2,4 Å de résolution sur une partie de la protéine (Gordon et al., 2000). Les huit zones ayant un taux de deutération supérieur à 30% (44-58, 98-105, 148-163, 366-384, 487-496, 545-562, 611-619 et 629-640)

sont repérées sur la structure 3D de la protéine (figure 39 respectivement zones (1), (2), (3), (4), (5), (6), (7), (8)).

Le premier acide aminé pour lequel une carte de densité électronique sur la structure cristallographique est l'arginine en position 76, ce qui indique une extrémité N-terminale très désordonnée (49-75). Ceci est en total accord avec la première zone que nous avons identifiée comme fortement deutérée (44-58) qui est donc à priori très flexible (zone (1) figure 38).

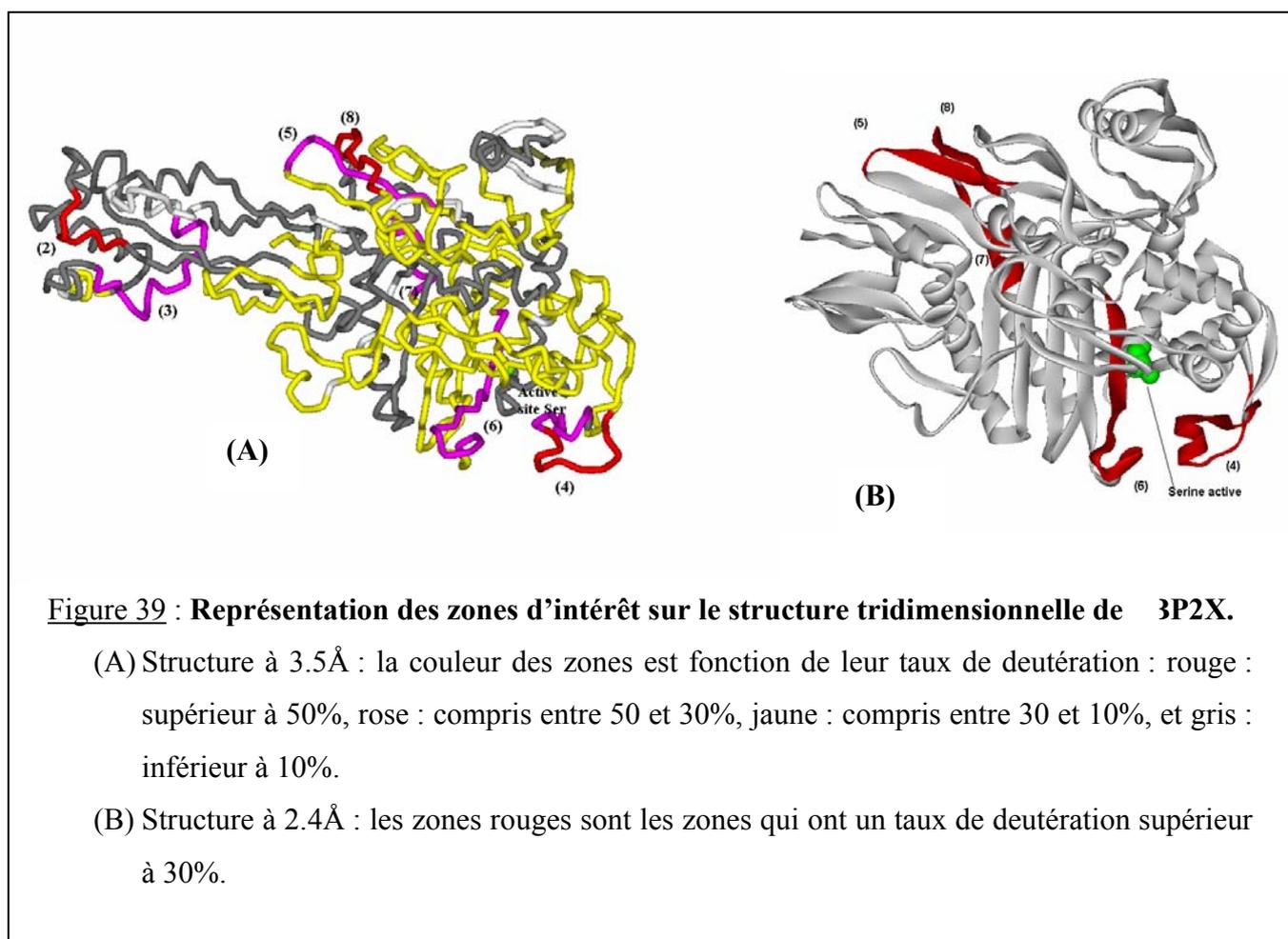
La région (2) qui possède un fort taux de deutération (98-105) correspond à une boucle très accessible au solvant à l'extrémité d'une des « pince à sucre » du domaine N-terminal.

La zone (3) identifiée (148-163) se trouve dans une région très peu structurée.

Les régions (4) et (6) sont des boucles flexibles de chaque côté de l'entrée du site actif (366-384, 545-562), et sont donc à priori très accessibles au solvant.

La région (5) constitue une petite boucle du domaine transpeptidase (487-496).

Enfin les régions (7) et (8) sont des parties du « linker » entre le domaine transpeptidase et le domaine C-terminal (611-619 et 629-640). L'absence pour la zone couvrant le « linker » de carte de densité électronique sur la structure de plus haute résolution indique que ce domaine est effectivement très flexible et/ou très peu structuré.



Toutes les zones que nous avons caractérisées comme étant fortement deutérées correspondent donc bien à des régions de la protéine peu structurées et/ou très accessibles au solvant, la méthode utilisée est donc ainsi validée. Nous avons pu obtenir des informations de structure sur la protéine PBP2X et, ce même pour des zones qui n'étaient pas observable en cristallographie.

D'un point de vue biologique, ces résultats permettent d'illustrer les nombreuses applications de la méthode des échanges H/D couplés à la spectrométrie de masse. En effet, les zones identifiées peuvent être classées en plusieurs catégories selon la raison pour laquelle nous les avons trouvées fortement deutérées.

Tout d'abord certaines, (1), (2) et (3), se trouvent dans des régions qu'il a été nécessaire de supprimer pour obtenir la structure cristallographique de plus haute résolution. Ces zones sont si désorganisées, qu'elles gênaient à la cristallogenèse et qu'il n'était pas possible d'obtenir une carte de densité électronique de ces régions. Leur identification plus précoce aurait permis de gagner du temps dans l'obtention de la structure à haute résolution, en évitant toutes les étapes de tâtonnement lors de la phase de cristallogenèse qui ont conduit à la délétion de toute la partie N-ter de la protéine.

Le deuxième type de région identifiée se trouve au niveau du site actif (zones (4) et (6)). En effet, le site actif devant par définition être très accessible, il est très rapidement deutéré.

Certaines des régions que nous avons caractérisées correspondent aux points chauds de mutations retrouvés dans les bactéries résistantes aux  $\beta$ -lactamines (zones (4), (6), (7) et (8)). Les zones sujettes aux mutations ne devant pas altérer la structure de la protéine, elles sont le plus souvent peu structurées et accessibles au solvant.

Ainsi, les zones fortement deutérées peuvent l'être pour des raisons diverses, et selon l'étude réalisée, il sera important de les différencier. Pour cela, diverses méthodes sont envisageables. Il est par exemple possible de faire de la mutagenèse dirigée dans les régions d'intérêt pour vérifier l'activité de la protéine après mutation et donc caractériser le site actif (Brier et al., 2004). Pour apprécier l'effet sur la cristallogénèse, différents essais de cristaux contenant des délétions des zones peuvent être réalisés. Enfin, il est possible d'établir plus précisément la cinétique de deutération de ces zones pour étudier leur structuration (Mazon et al., 2004).

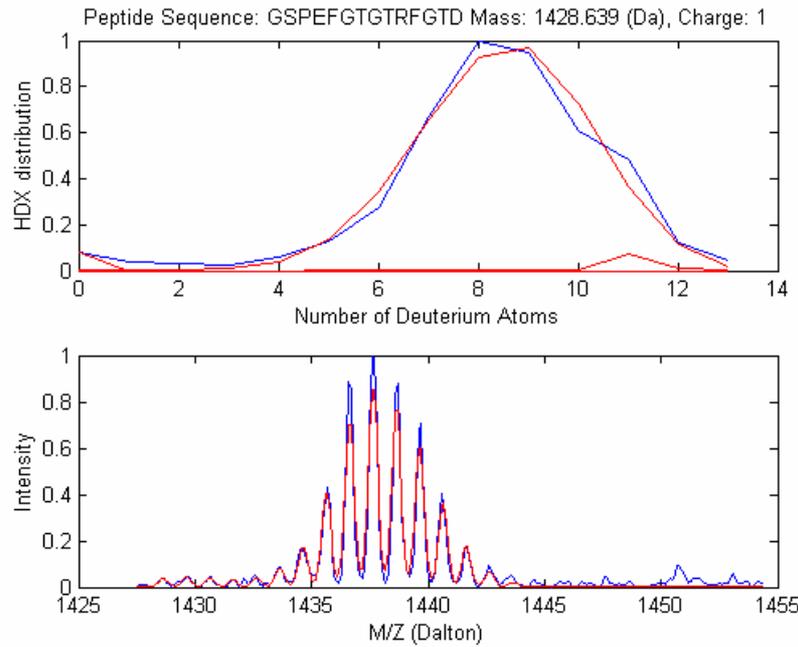
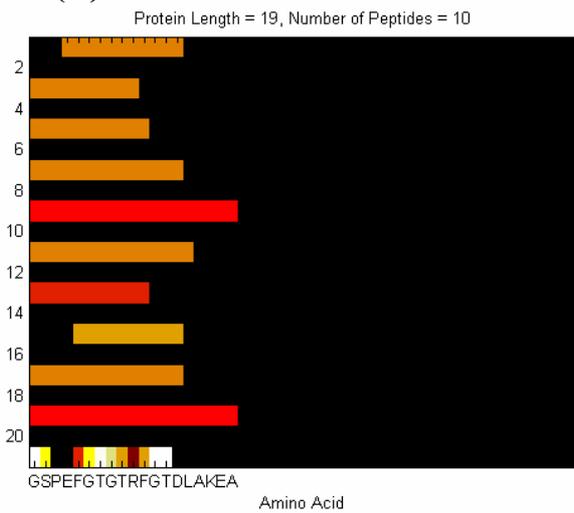
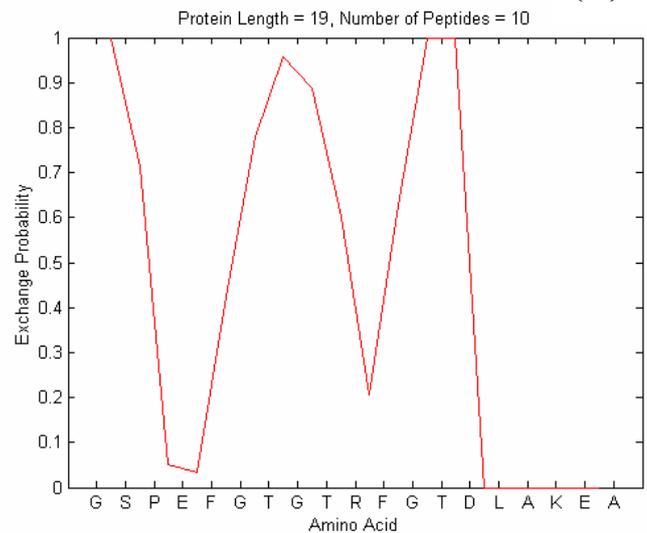
#### **4 PERSPECTIVE D'OPTIMISATION DE LA METHODE**

La compilation des données de chaque enzyme se fait manuellement. Cela représente un travail lourd et fastidieux. Nous développons actuellement un logiciel en collaboration avec Markus Muller de l'institut suisse de bioinformatique à Genève. Ce logiciel donnera la répartition statistique des deutériums sur la séquence primaire de la protéine étudiée.

Dans un premier temps, ce logiciel détermine à partir des spectres de masse de chaque peptide le nombre moyen de deutérium incorporé, en tenant compte des rapports isotopiques. Cette détermination se fait par comparaison du spectre expérimental et de spectres théoriques générés par le programme. L'identité de chaque peptide est donnée par l'expérimentateur, qui l'aura déterminée par des analyses MS-MS. Le logiciel compile les taux de deutération de chaque peptide et les associe à la séquence primaire de la protéine. Il peut de cette manière déterminer la répartition statistique des deutériums sur la séquence et la probabilité de présence d'un deutérium sur chaque liaison peptidique.

Des tests de ce logiciel sont en cours, mais il reste encore des points à optimiser. Pour l'instant, l'incorporation des spectres de masse de chaque peptide doit se faire manuellement, spectre après spectre. C'est pourquoi nous n'avons pu jusqu'à présent tester le logiciel que sur des petites parties de la séquence. Le nombre de peptide et donc de spectres de la séquence entière étant trop important pour être entré dans le logiciel manuellement.

La figure 40 montre les résultats obtenus sur l'étude de la partie 44-62 de la protéine PBP-2X\*. Ces résultats même s'ils ne concernent qu'une partie minime de la protéine indiquent le bon fonctionnement du logiciel et son fort potentiel.

**(A)****(B)****(C)**

### Figure 40 : Présentation des résultats du logiciel

*A* : Détermination de la deutération moyenne d'un peptide. Le spectre de masse du peptide est représenté en bas : en bleu, le spectre expérimental, et en rouge le spectre théorique pour le nombre de deutérium moyen considéré. Le graphique en haut de la figure A représente la distribution du nombre de deutérium pour le peptide considéré.

*B* : Représentation du taux de deutération moyen de chaque peptide. La couleur représente le taux de deutération du peptide : plus la barre est rouge et plus le peptide qu'il représente est deutéré.

*C* Répartition statistique des deutériums sur la séquence. La compilation des données de la figure B permet de déterminer la probabilité de présence d'un deutérium sur chaque liaison peptidique.

## 5 **CONCLUSION**

Afin d'améliorer une méthode très utilisée depuis près de dix ans, nous avons utilisé trois protéases différentes (pepsine, protéase de type XIII et protéase de type XVIII) dans des expériences d'échange H/D en association avec la spectrométrie de masse. Les conditions de fonctionnement de ces protéases se sont révélées compatibles avec les conditions imposées par l'étape de blocage des rééchanges. Nous avons pu par cette méthode obtenir un excellent recouvrement de séquence, ce qui permet d'éviter de manquer des zones d'intérêt. De plus, la résolution spatiale de nos résultats de deutération a été très largement améliorée grâce à l'utilisation combinée des trois enzymes. Nos résultats ont été validés sur la structure tridimensionnelle de la protéine obtenue par cristallographie. Les zones pour lesquelles nous avons observé un taux de deutération important appartiennent effectivement à des zones très accessibles et/ou peu structurées de la protéine. Cette méthode, maintenant au point, va pouvoir être appliquée en routine sur des protéines de structure inconnue, et ce quelle que soit leur taille.

CHAPITRE 3 :  
MISE AU POINT DE LA  
TECHNIQUE DES  
RETICULATIONS CHIMIQUES  
COUPLEES A LA  
SPECTROMETRIE DE MASSE

## **INTRODUCTION**

La fonction des protéines est intimement liée à leur structure. Leur conformation donne l'agencement des sites actifs, des domaines d'interactions et des lieux de fixations des ligands entre eux. De plus, ces protéines agissent le plus souvent sous forme de complexes apparentés à des machineries moléculaires. C'est pourquoi la compréhension de la structure des protéines est une étape clé dans la connaissance de leurs fonctions cellulaires. Depuis ces dernières années, le nombre de nouvelles protéines identifiées par les projets génomiques et protéomiques est devenu très élevé. La base de données Genbank par exemple contient près d'un million de séquences, alors que la base de données Protein Databank ne contient qu'environ 25000 structures de protéine (Redfern et al., 2005). L'obtention de données structurales précises pour toutes ces protéines, et l'étude de la façon dont elles interagissent entre elles, est une tâche compliquée qui semble presque impossible à réaliser (du moins dans un laps de temps raisonnable). En fait résoudre toutes ces structures une à une n'est pas forcément nécessaire pour déterminer les relations structure/fonctions. En effet, la structure des protéines semble n'avoir changé au cours de l'évolution que par un nombre limité de voies et les motifs fonctionnels des protéines semblent avoir été largement conservés. Parfois certaines nouvelles fonctions sont apparues et un réarrangement des groupes fonctionnels a pu se produire, mais on estime qu'il existe seulement quelques milliers de familles de repliement, soit un ordre de magnitude de moins que le nombre de protéines (Redfern et al., 2005). La modélisation moléculaire joue ici un grand rôle dans la prédiction et l'affinement de la structure des protéines à partir de modèles. Mais ces données théoriques doivent impérativement être vérifiées et s'appuyer sur des données expérimentales. C'est dans ce cadre que la technique des réticulations chimiques associées à la spectrométrie de masse a émergé.

Dans cette étude, nous allons tenter d'appliquer et d'optimiser cette technique sur le cytochrome c. Nous espérons ainsi créer une méthode modèle pour l'attribution d'une famille de repliement à une protéine inconnue.

# **1 LA RETICULATION CHIMIQUE**

Le principe de la réticulation chimique est de créer un lien covalent entre deux groupements fonctionnels d'une molécule biologique par l'intermédiaire d'un agent réticulant. Les molécules biologiques concernées peuvent être de nature différente, la réticulation peut concerner des protéines, des peptides, des acides nucléiques,...

## ***1.1 Les différents agents réticulants***

Il existe plus d'une centaine d'agents réticulants décrit dans la littérature ou disponibles commercialement (Hermanson, 1996 ; Wong 1993 ; Pierce biotechnology 2004). Un agent de réticulation est caractérisé par ses groupements fonctionnels, qui vont réaliser le lien covalent, par son bras espaceur et par son organisation.

### **1.1.1 Réactivité des agents réticulants**

Parmi le grand nombre d'agents réticulant existant la plupart utilisent les mêmes principes de base de chimie organique qui peuvent être réduits à quelques réactions (Sinz et al. 2003 ; Back et al. 2003).

#### *1.1.1.1 Réticulation sur les amines réactives*

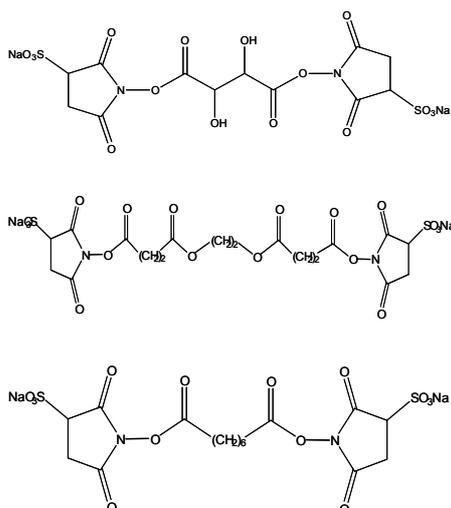
Pour qu'une réaction de réticulation sur une protéine soit optimale, plusieurs conditions doivent être remplies. Les chaînes latérales réactives doivent être accessibles aux agents réticulants, les distances entre ces sites réactifs doivent être compatibles avec la longueur des bras espaceur des agents réticulants et il est préférable de limiter le rendement de réaction. Les groupements polaires, comme les acides carboxyliques ou les amines primaires, remplissent ces critères puisqu'ils sont souvent situés à l'extérieur de la chaîne peptidique ou près des interfaces. Leur fréquence sur la chaîne peptidique est relativement élevée, ce qui permet d'avoir de nombreux points de réactions (répartis de façon assez homogène). Cependant, cette

multiplicité des sites de réaction est en contradiction avec la volonté de limiter les réticulations. De plus, ces acides aminés étant les sites de clivage des enzymes les plus couramment utilisées, on peut assister à une modification de ces sites. Un autre inconvénient est la modification de la répartition des charges sur les protéines, les lysines modifiées n'étant plus chargées. Cette différence d'environnement électronique peut entraîner la modification de la structure tridimensionnelle. C'est l'une des raisons pour lesquelles il faut limiter le nombre d'adduit par molécule.

Les agents réticulants qui ciblent les fonctions amines des protéines sont les plus couramment utilisés. Les deux réactions de modification des amines sont l'acylation et l'alkylation ; ce sont des réactions en général rapides et avec un haut rendement qui donnent des liaisons amides stables ou des amines secondaires.

- *Esters N-hydroxysuccinimide*

Les esters N-hydroxysuccinimide NHS (figure 41) sont utilisés comme réactifs pour des réactions d'acylation. La plupart des esters NHS sont insolubles dans les solvants aqueux courant en biologie. Deux protocoles sont donc utilisés, soit une préparation de forte concentration d'ester NHS dans un milieu organique est diluée dans le milieu réactif, soit un dérivé ester sulfo-NHS, hydrosoluble, est utilisé.

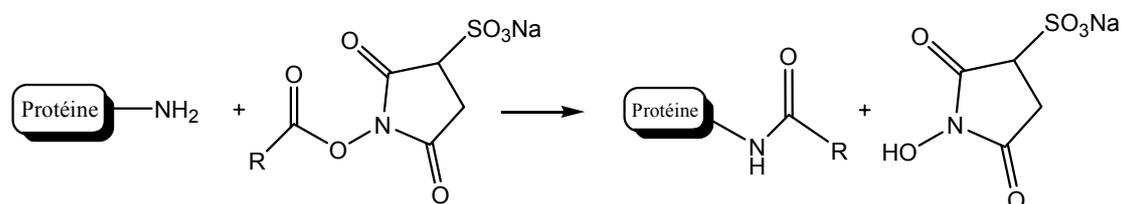


**Figure 41 : Formules développées des agents réticulants Sulfo-DST, Sulfo-EGS et BS3.**

Les esters NHS réagissent avec les nucléophiles, libérant le groupement NHS ou sulfo-NHS pour former un produit acétylé. Ils créent une liaison amide ou imide stable avec les amines primaires ou secondaires. Dans le cas de réaction sur une protéine, ils réagissent sur la fonction amine terminale et sur les amines ε des lysines (figure 42). La réactivité des lysines va alors dépendre de leur implication dans des liaisons hydrogène intramoléculaires et de leur accessibilité. En fait les informations obtenues par ces lysines vont nous donner des contraintes de distance assez larges du fait de la longueur et de la flexibilité importante des chaînes latérales des lysines.

La réactivité des histidines et des serines peut généralement être ignorée. En effet, les produits obtenus après réaction sur une histidine (N-acylimidazole) sont très instables et la réactivité des serines est très lente et ne peut pas entrer en compétition avec d'autres réactions plus rapides (celles impliquant les amines primaires par exemple) (Leavell et al., 2004). Les cystéines sont quant à elles considérées comme réactives, mais lorsqu'elles sont impliquées dans les ponts disulfures, elles ne peuvent pas réagir avec les esters NHS. Par contre, les tyrosines semblent avoir une réactivité non négligeable, voire même comparable à celles des lysines dans certaines conditions (Swaim et al., 2004 ; Leavelle et al., 2004). Ainsi un pH plutôt acide favorise les ponts entre l'acide aminé N-terminal et une tyrosine ou entre deux tyrosines, tandis qu'un pH plus alcalin favorise les ponts entre l'acide aminé N-terminal et une lysine ou entre deux lysines. De plus, quelles que soient les conditions, des ponts entre une lysine et une tyrosine peuvent être observés.

Les esters NHS ont une demi-vie de l'ordre de l'heure dans un milieu aqueux à un pH physiologique (pH 7.0 -7.5). L'hydrolyse augmente avec le pH (Hermanson, 1996).

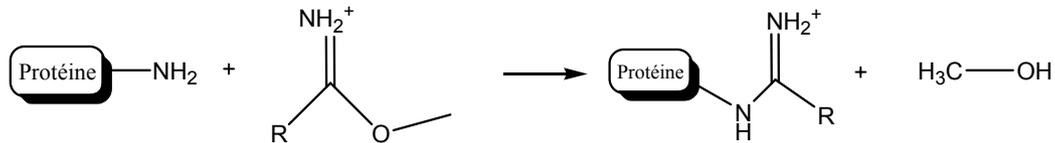


**Figure 42: Réaction des agents réticulants ester NHS sur les protéines.**

D'après Sinz et al., 2003

- *Esters imides*

Les réactifs esters imides, parmi les plus anciens pour la réticulation des protéines (Hartmann et Wold, 1966), sont les plus spécifiques pour modifier une amine primaire (figure 43).



**Figure 43 : Réaction des agents réticulants esters imides sur les protéines.**

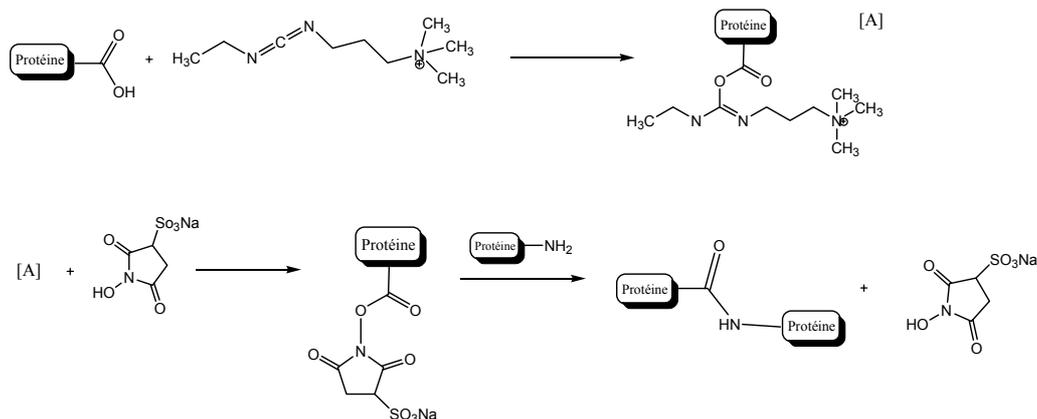
D'après Sinz, 2003

L'extrémité N-terminale de la protéine et les fonctions amines des lysines réagissent de manière optimale à un pH compris entre 7 et 10. L'un des inconvénients des esters imides est l'hydrolyse à pH élevé de la liaison amidine. Leur temps de demi vie en milieu aqueux est inférieur à 30 minutes (Hunter et al., 1962). Le produit de la réaction avec les esters imides est protoné au pH physiologique (Liu et al., 1977). Cela représente un avantage majeur des esters imides, en effet l'un des impératifs lors des réactions de cross-link ayant pour cible les amines est de maintenir la structure tridimensionnelle de la protéine étudiée tout en gardant la présence d'une charge positive sur les fonctions amines.

- *Les carbodiimides*

Les carbodiimides (figure 44) sont utilisés pour former une liaison amide entre une fonction amine de la protéine et une fonction carboxylique (Grabarek et al., 1990). Ils sont généralement utilisés en combinaison avec du sulfo-NHS qui s'hydrolyse lentement en phase aqueuse. L'avantage de l'utilisation combinée du Sulfo-NHS et des carbodiimides est de permettre la stabilisation de l'intermédiaire actif, afin qu'il puisse finalement réagir avec l'amine.

Cependant les agents réticulants de ce type peuvent former différents produits intermédiaires, qui peuvent alors fortement compliquer la détection des peptides réticulés et provoquer des modifications au niveau des sites de clivage.



**Figure 44 : Réaction de l'EDC permettant la création d'une liaison covalente entre un acide carboxylique et une fonction amine.**

D'après Sinz et al., 2003

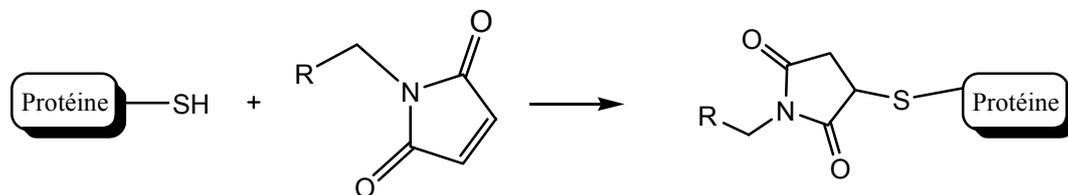
### 1.1.1.2 Réaction sur les thiols

Après les réactions sur les amines, le deuxième groupe de fonctions couramment ciblées lors des réactions de réticulation des protéines est celui des fonctions thiol des cystéines. Dans le cas de molécules hétérofonctionnelles, l'une des extrémités est souvent réactive sur les thiols. Les réactions mises en jeu sont l'alkylation et l'échange de ponts disulfures.

- *Les Maléimides*

Ces agents agissent en ajoutant un thiolate sur le groupement fonctionnel selon une addition de Michaël de type 1,4 (figure 45). Cette réaction est spécifique aux fonctions thiols lorsque le

pH de réaction est choisi dans la gamme 6.5-7.5 (Smyth et al., 1964). A pH 7, la vitesse de réaction sur les thiols est 1000 fois supérieure à la réaction sur les amines.

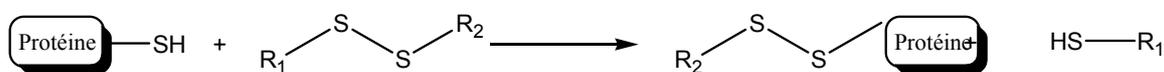


**Figure 45 : Réaction des agents réticulants maléimides sur les protéines.**

D'après Sinz et al., 2003

- *Réactifs permettant l'échange de ponts disulfures*

Les composés contenant un pont disulfure sont capables de participer à une réaction d'échange avec une fonction thiol. L'attaque de la fonction thiol sur le pont disulfure permet la rupture de la liaison S-S et la formation d'un nouveau pont (Figure 46). La réaction est réversible en utilisant des agents réducteurs des thiols, par exemple en incubant la protéine réticulée dans du DTT. Ces réactifs appartiennent à la classe des agents clivables. La réaction d'échange au niveau des sulfures peut avoir lieu dans des conditions de pH large.



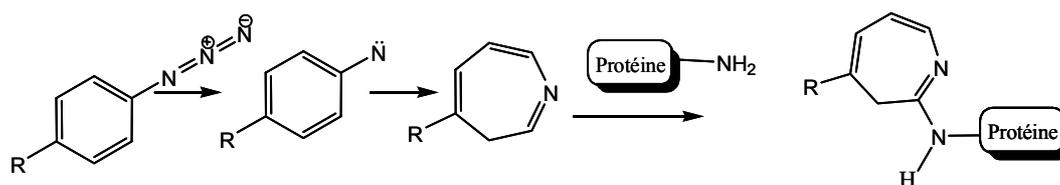
**Figure 46 : Réaction des agents réticulants permettant l'échange de ponts disulfure avec les cystéines des protéines.**

D'après Sinz et al., 2003

### 1.1.1.3 Réticulation non sélectives

Les réactions non sélectives sont réalisées en utilisant en général des agents photo-activables tel que les arylazides ou les benzophénones. Les réactions sont induites par exposition aux ultraviolets. Jusqu'à ce que les groupes fonctionnels aient été photolysés, ils sont relativement non réactifs. Ainsi les molécules de réticulation photoréactive peuvent être utilisées sous des conditions contrôlées.

Les agents photosensibles les plus utilisés sont les arylazides. Dans ce cas, l'activation par photolyse des phenyl-azides donne des nitrènes, espèces de durée de vie très courte, qui peuvent se fixer non spécifiquement sur les molécules cibles en particulier en s'additionnant sur des doubles liaisons ou en s'insérant au niveau de liaisons C-H ou N-H (Figure 47). Ce type d'agent est souvent associé à un autre agent plus spécifique dans le cadre de molécules hétérobifonctionnelles.



**Figure 47 : Réaction des agents réticulants photoactivables sur les protéines.**

D'après Sinz et al., 2003

### 1.1.2 Les bras espaceurs

Les groupements réactifs des agents réticulants sont séparés par une chaîne carbonée de longueur variable. Cette longueur pourrait donner une première approximation de la distance entre les deux groupes fonctionnels de la protéine qui ont réagi.

Les plus petits agents sont dit « de longueur nulle ». Ces composés lient de manière covalente deux fonctions en créant une liaison sans insertion d'une molécule. Parmi ces agents, la classe

la plus couramment utilisée est celle des carbodiimides. Ces composés permettent la formation d'une liaison entre un carboxylate et une amine (Hoare et al., 1996), l'EDC est le plus connu (Hermanson et al., 1996). Ce type d'agent « de longueur nulle » est préférentiellement utilisé lors de l'étude des interactions entre deux protéines, afin de ne lier que les acides aminés à l'interface.

Mais la plupart des agents réticulants utilisés introduisent des bras espaceurs, contenant des liaisons flexibles, entre les deux fonctions réticulées. Au départ, les longueurs des chaînes introduites étaient considérées comme étant celles données par les fabricants. Cependant, il est rapidement apparu que la longueur moyenne de l'agent était souvent inférieure à la longueur théorique calculée. Une étude de simulation pour déterminer la longueur de 32 agents a été réalisée (Green et al., 2001). La conformation des molécules les plus couramment utilisées pour réagir sur les fonctions amines et thiols a été étudiée dans le cas où elles étaient seules et dans le cas où elles étaient attachées à des groupements de fort encombrement stérique pour simuler l'attachement à une protéine. Les résultats obtenus pour les agents de la famille NHS-ester sont présentés dans le tableau 1.

Agent réticulant	Distance		Echelle de distances	Distance fabricant (a)
	moyenne N-N	Mode		
DSG	6.22	6.4	3.12–7.49	7.7
DSP	8.04	8.1	4.63–10.65	12
DSS	8.88	9.2	5.58–11.42	11.4
DST	4.06	3.6	2.45–5.84	6.4
EGS	9.11	9.8	3.23–14.79	16.1
BSOCOES	10.62	10.7	6.81–12.43	13

(DSG) disuccinimidyl glutarate; (DSP) dithiobis(succinimidylpropionate); (DSS) disuccinimidyl suberate; (DST) disuccinimidyl tartarate; (EGS) ethyleneglycol bis-(succinimidylsuccinate); (BSOCOES) bis(2-[succinimidooxycarbonyloxy]ethyl)sulfone. (a) voir référence Pierce Chemicals (1999).

**Tableau 1: Longueur des agents réticulants de type NHS-ester (Green et al., 2001)**

Les valeurs des fabricants considèrent l'agent réticulant comme linéaire et non flexible ce qui donne donc les distances maximales entre les fonctions amines, lors de l'exploitation des résultats, il faut tenir compte du fait que ces valeurs ne reflètent pas totalement la réalité.

Les agents réticulants possédant des bras espaceurs de longueur importante introduisent des complications dans l'analyse des résultats. En effet, plus la longueur du bras espaceur va être élevée et plus le nombre de lien créé va être important, ceci va entraîner une diminution du signal des espèces réticulées, de part la multiplicité de ces espèces. De plus, les agents possédant un fort encombrement stérique (c'est à dire un bras espaceur long) vont atteindre difficilement le cœur de la protéine, ainsi des groupements réactifs se trouvant à des distances compatibles avec celles de l'agent pourront ne pas être réticulés.

### **1.1.3 Le design des agents réticulants**

#### *1.1.3.1 Les agents homobifonctionnels*

Les agents réticulants qui possèdent deux groupements réactifs identiques sont appelés agents homobifonctionnels. Ils vont donc créer un lien covalent entre deux groupements fonctionnels identiques de la protéine. Leur nombre a considérablement augmenté durant ces vingt dernières années et une grande variété d'agents de ce type est actuellement disponible commercialement. Ces agents vont être différenciés par la longueur de leur bras espaceur et par la nature de leur groupement réactif.

#### *1.1.3.2 Les agents hétérobifonctionnels*

Les agents réticulants qui possèdent deux groupements réactifs différents sont appelés hétérobifonctionnels. Ils vont créer un lien covalent entre deux groupements réactifs de la protéine différents (par exemple une amine et un thiol). Ces agents peuvent être utilisés dans des protocoles à plusieurs étapes afin de limiter la formation d'agrégats de haut poids moléculaire. Par exemple, il est possible d'utiliser un agent hétérobifonctionnel de type NHS-ester/maléimide. Dans une première étape, la réticulation des amines réactive est réalisée par

les groupements NHS-ester. Le groupement maléimide, plus stable en solution aqueuse, n'est alors pas réactif. Après une étape de purification, la réaction sur les fonctions thiol pourra avoir lieu.

Les agents hétérobifonctionnels contenant un groupement photoréactif sont utilisés de la même manière, ces groupements étant très stables tant qu'ils n'ont pas été exposés à des rayons UV de forte intensité.

### *1.1.3.3 Les agents trifonctionnels*

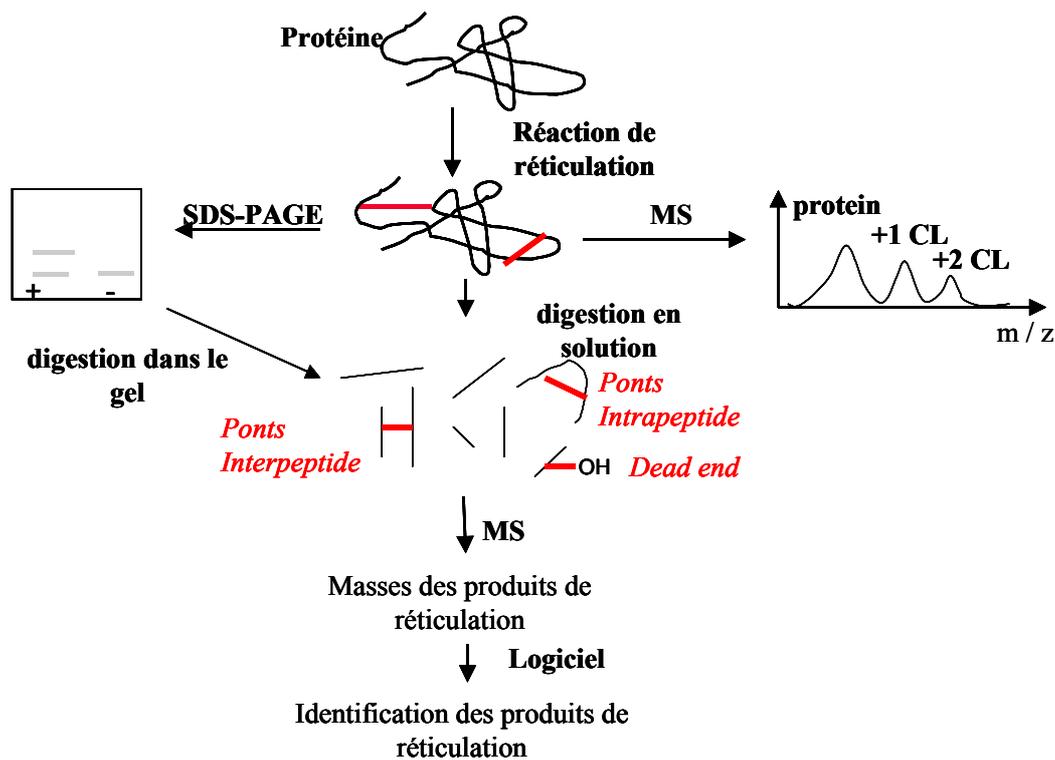
Un nouveau type d'agent a récemment émergé, il s'agit des agents trifonctionnels. Conçus comme les agents bifonctionnels, ils possèdent un troisième groupement fonctionnel. Celui-ci pourra permettre de lier spécifiquement une troisième protéine ou pourra servir à la purification du produit de réticulation, par exemple s'il possède une biotine (Trester-Zedlitz et al., 2003 ; Alley et al., 2000). Une description plus détaillée de ce type d'agent sera réalisée dans la suite de l'étude.

## ***1.2 Stratégie et applications***

### **1.2.1 La réticulation intramoléculaire**

#### *1.2.1.1 Stratégie*

L'objectif des études utilisant la réticulation intramoléculaire couplée à la spectrométrie de masse est d'obtenir des contraintes de distance entre différents résidus aminoacides d'une protéine structurée. Ces contraintes sont utilisées pour obtenir des informations structurales sur une protéine. La procédure expérimentale générale est décrite figure 48.



**Figure 48 : Stratégie générale de la méthode des réticulations intramoléculaires couplées à la spectrométrie de masse**

La réaction de réticulation est réalisée à l'aide des agents précédemment décrits. Il est important durant cette première étape de maintenir la concentration en protéine dans l'échelle du micromolaire, afin de limiter le plus possible les éventuelles réticulations intermoléculaires. La protéine doit également se trouver dans des tampons appropriés proches des conditions physiologiques, afin de conserver sa structure active. Les tampons ne devant pas interférer avec les réactions de réticulation, il est indispensable qu'ils soient non-réactifs vis à vis des agents utilisés (par exemple, il ne sera pas possible d'utiliser un tampon Tris avec les agents de type NHS-ester). Une fois la réticulation effectuée, le rendement de réaction doit être vérifié par une analyse en spectrométrie de masse. Ce rendement de réaction doit être suffisant pour la détection des produits lors des analyses futures, mais ne doit pas être trop important afin de ne pas induire des distorsions dans la structure tridimensionnelle de la protéine. Ces analyses de spectrométrie de masse et un gel SDS-Page permettent également de s'assurer qu'aucune réaction intermoléculaire n'a eu lieu. Dans le cas contraire, des séparations, par gel-filtration par exemple, peuvent être requises.

L'étape suivante consiste à digérer la protéine réticulée directement en solution ou dans le gel. Le mélange obtenu après protéolyse est plutôt complexe puisqu'il est constitué à la fois de peptides non modifiés, de peptides avec un pont intrapeptidique, de peptides avec un pont interpeptidique et de peptides modifiés avec un agent hydrolysé. Le mélange est analysé par spectrométrie de masse. La masse de tous les produits réticulés est déterminée et un logiciel permet de retrouver les acides aminés pontés. Cette identification théorique peut être confirmée par des expériences de MS-MS sur les peptides modifiés. Enfin, les contraintes de distances obtenues sont intégrées dans la détermination d'informations structurales de la protéine.

### 1.2.1.2 *Les produits obtenus*

Une des difficultés des expériences de réticulation intramoléculaire est liée à la diversité des produits de réticulation obtenus. Si la réaction de réticulation est réalisée avec un faible ratio d'agent réticulant par protéine, les agents réagissent avec un ou deux résidus d'acides aminés de la chaîne protéique (ou la fonction amine ou carboxyle terminale selon les agents). Après protéolyse, les produits obtenus peuvent être répartis en trois catégories (en plus des peptides n'ayant pas réagi) :

- Le peptide peut être modifié avec un agent réticulant hydrolysé à sa deuxième extrémité. L'agent réticulant est qualifié dans la littérature de « dead end ».
- Le peptide peut comporter deux résidus liés par un agent réticulant. La réticulation est dite intrapeptide.
- Deux peptides peuvent être liés entre eux par un agent réticulant. La réticulation est dite interpeptides.

De plus, dans le cas de réticulations multiples, des combinaisons de ces trois types de peptides peuvent être obtenues. Pour uniformiser la description des produits des expériences de réticulation une nomenclature a été proposée (Schilling et al., 2003) (figure 49).

Les peptides de type 0 ne permettent pas d'obtenir d'information concernant la distance entre deux acides aminés mais peuvent donner des informations sur la réactivité relative des différents sites de la protéine. Au contraire les peptides de type 1 et de type 2 sont le produit

d'une réaction de l'agent avec deux résidus d'acides aminés différents et peuvent donc donner des informations de distance.

Une nomenclature concernant la MS-MS des peptides réticulés a été également proposée (Schilling et al., 2003). Cette nomenclature est dérivée de la nomenclature classique (Roepstorf et Fohlman, 1984) et de celle utilisée pour les peptides cyclisés (Ngoka et Gross, 1999).

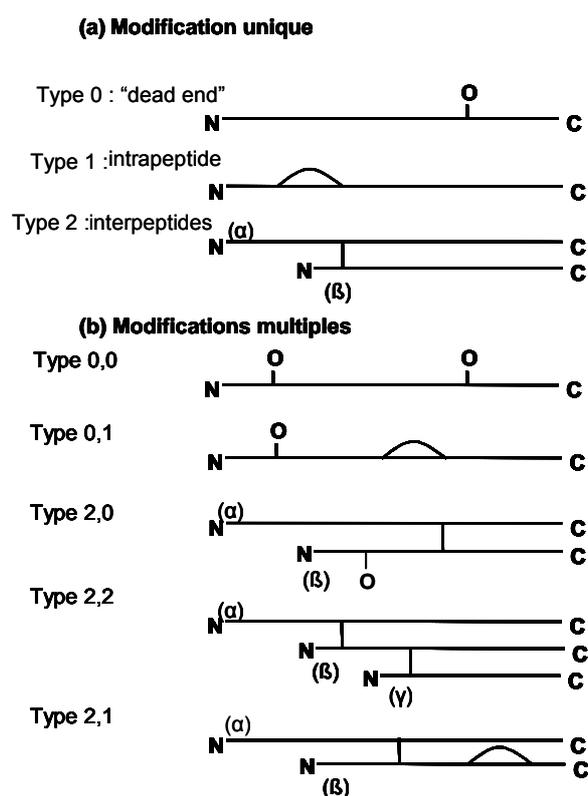


Figure 49: Nomenclature des peptides réticulés

(a) Classification des différents types de cross-link en Type 0, 1, 2.

(b) Extension de la nomenclature dans le cas de modifications multiples. Les chaînes peptidiques sont classées selon leur longueur (ou en cas de longueurs égales, selon leur masse) avec  $\alpha > \beta > \gamma$ . La position des peptides dans la séquence primaire de la protéine détermine l'ordre dans les nombres indiquant les types de réticulation.

*D'après Schilling et al., 2003.*

### 1.2.1.3 Applications

Dans la méthode des réticulations intramoléculaires couplées à la spectrométrie de masse, ce sont les contraintes de distances, finalement obtenues grâce à l'identification des produits réticulés, qui sont utilisées pour obtenir des informations structurales sur la protéine.

Il a été au départ proposé d'utiliser ces contraintes de la même façon que celles obtenues par le « nuclear Overhauser effect » (NOE) dans les expériences de RMN. Ainsi, la distance donnée par le bras espaceur de l'agent réticulant délimiterait le rayon d'une sphère imaginaire qui refléterait la distance maximale entre les deux groupements liés.

Cependant, le nombre de contraintes de distances qui peuvent être obtenues par la réticulation est bien inférieur à celui des contraintes obtenues par NOE.

En fait, pour obtenir une structure de résolution atomique d'une chaîne de N résidus, approximativement  $3N$  contraintes sont requises (Young et al., 2000). C'est pourquoi, il est pour l'instant irréaliste d'obtenir une structure de cette résolution avec des expériences de réticulations chimiques sur une protéine de structure inconnue. Par contre, le nombre de contraintes de distances nécessaires pour assigner une famille de repliement à une protéine est de l'ordre de  $N/10$ . Ainsi, le nombre de contraintes de distances obtenues par la technique des réticulations chimiques associées à la spectrométrie de masse n'est pas suffisant pour modéliser entièrement la structure d'une protéine, mais l'est pour déterminer sa famille de repliement.

La modélisation de novo de la structure des protéines basée sur les contraintes de distances obtenues par réticulation intramoléculaire n'est donc pour le moment pas possible. Il est donc nécessaire d'introduire d'autres contraintes structurales obtenues par d'autres sources expérimentales ou par des modèles informatiques.

La plupart des études réalisées par cette technique ont été jusqu'à présent des études visant à mettre au point et à optimiser la méthode. Les contraintes de distance obtenues ont alors été tout simplement validées sur des structures tridimensionnelles existantes. De telles études ont ainsi été réalisées sur le cytochrome c (Pearson et al., 2002 ; Dihazi et Sinz, 2003 ; Schilling et al., 2003), sur l'ubiquitine (Kruppa et al., 2003), sur la ribonucléase A (Pearson et al., 2002), et également sur des protéines membranaires (Schoeniger et al., 2003) ce qui laisse envisager de nouvelles applications très intéressantes (les protéines membranaires étant généralement difficilement cristallisables).

D'autres études peuvent conclure sur des informations structurales nouvelles. C'est par exemple le cas de celle réalisée sur l'erythropoïétine humaine recombinante pour laquelle des

informations structurales ont pu être obtenues sur la protéine glycosylée et sur la protéine non glycosylée (Haniu et al., 1993). Des données structurales sur la BSA ont également été obtenues grâce aux contraintes de distance déterminées par les réticulations et à la comparaison avec la structure de la protéine HSA, homologue de la BSA, pour laquelle une structure cristallographique est disponible (Huang et al., 2004).

Enfin, certaines études n'utilisent pas de structures homologues. Les contraintes de distances peuvent être traitées par bioinformatique et comparées à des bases de données structurales afin de vérifier la famille de repliement de la protéine étudiée (Young et al., 2000). Les contraintes de distances peuvent être insérées dans une structure modèle obtenue par modélisation moléculaire (Back et al., 2003). Des filtres peuvent être appliqués afin de faciliter l'attribution d'une solution bioinformatique fiable à la protéine étudiée. Ces filtres peuvent être dérivés des propriétés biochimiques des protéines (hydrophobicité, pourcentage de structure secondaire,...) ou obtenus par des alignements de structures primaires des différentes protéines d'une base de donnée.

## **1.2.2 La réticulation intermoléculaire**

### *1.2.2.1 Stratégie*

L'objectif des études utilisant la réticulation intermoléculaire couplée à la spectrométrie de masse est de cartographier et de caractériser les interfaces dans des complexes multiprotéiques. Cette cartographie peut se faire par la liaison covalente de molécules impliquées dans un complexe non-covalent. La procédure expérimentale générale est décrite figure 50. Pour la réalisation de réticulation, les agents réticulants de « longueur nulle », comme l'EDC, sont souvent utilisés avec succès dans ce type d'application, afin d'identifier les résidus en contact le plus étroit possible. Comme pour les réticulations intramoléculaires, le rendement de réaction doit être contrôlé précisément avec une attention particulière sur la formation effective du lien intermoléculaire. Les monomères et les multimères sont ensuite séparés par électrophorèse ou par gel-filtration. Une fois les multimères repérés, ils sont digérés dans le gel ou en solution. Le mélange obtenu est alors encore plus complexe que celui obtenu après réticulation intramoléculaire. En effet, il est composé de peptides non modifiés, de peptides issus de la réticulation intermoléculaire, mais aussi de peptides issus de

la réticulation intramoléculaire de chacune des protéines. En effet, même si elles sont liées covalentement entre elles, les protéines peuvent aussi incorporer des molécules d'agent réticulant dans leur structure propre. Cette multiplicité des produits de réaction complique certes l'identification de chaque produit, mais permet d'obtenir de nombreuses informations. Si les produits de réticulation intermoléculaires permettent de caractériser l'interaction entre les deux protéines, les produits de réticulation intramoléculaire vont, quant à eux, nous donner des informations sur le repliement de chacune des protéines lorsqu'elles sont en interaction. L'analyse des masses des produits de réticulation et l'identification de ces produits se fait ensuite de la même manière que pour les expériences de réticulations intramoléculaires.

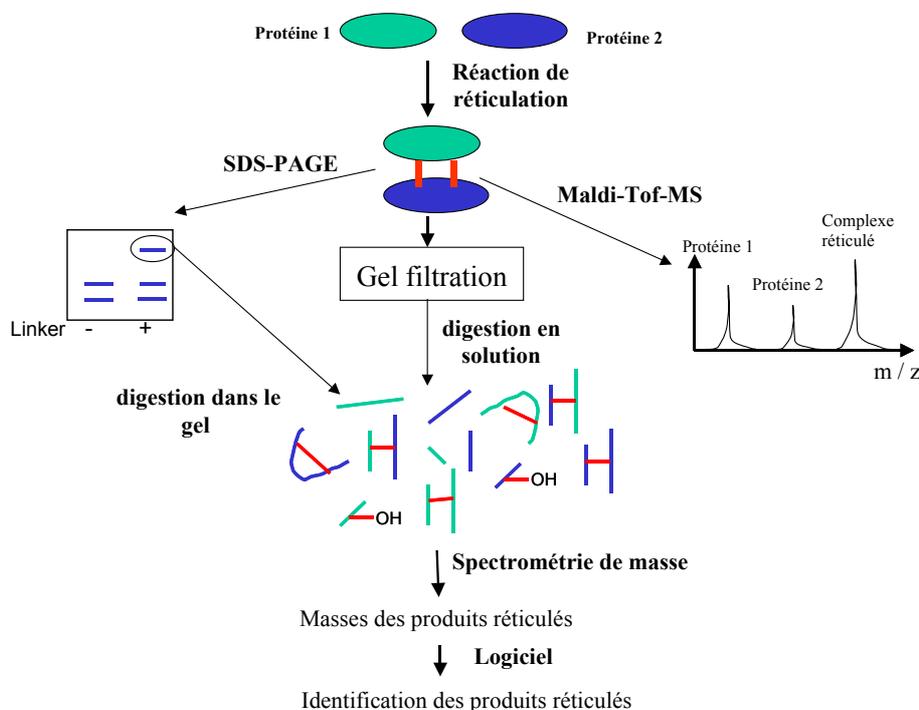


Figure 50: **Stratégie générale de la méthode des réticulations intermoléculaires couplées à la spectrométrie de masse**

### *1.2.2.2 Applications*

L'avantage de cette technique appliquée à la caractérisation des interactions est qu'elle va permettre des études conformationnelles dynamiques des complexes. En effet, la réticulation va permettre une stabilisation de certaines conformations voulues, ainsi les complexes transitoires vont pouvoir être étudiés par piégeage d'un intermédiaire conformationnel ou par inhibition d'une étape d'assemblage.

La réticulation intermoléculaire associée à la spectrométrie de masse a été appliquée à de nombreuses études très variées au niveau du contexte biologique aussi bien qu'au niveau de la nature et de la taille des complexes étudiés. Par exemple une des premières études a porté sur le complexe multiprotéique Nup85p présent dans les enveloppes nucléaires des levures. Cette étude a pu identifier les six composants de ce complexe et caractériser leur organisation spatiale (Rappsilber et al., 2000). Une autre étude a consisté à déterminer les zones d'interfaces dans l'homodimère ParR et dans le complexe glycoprotéique CD28/CD80 (Bennet et al., 2000). De la même façon des interactions ont été caractérisées pour le complexe Op18/tubuline impliqué dans la déstabilisation des microtubules (Muller et al., 2001), pour le dimère mitochondrial de levure PHB1/PHB2 (Back et al., 2002 b), entre les deux chaînes du dimère d'interleukine-6 (Taverner et al., 2002), entre la protéine gp120 de l'enveloppe du VIH et son récepteur cellulaire CD4 (Wine et al., 2002), dans le complexe hétérodimérique NC2 (Trester-Zedlitz et al., 2003), et dans le complexe protéine/peptide Calmoduline/Melittine (Schulz et al., 2004).

## ***1.3 L'analyse par spectrométrie de masse des produits réticulés***

La recherche des produits réticulés et la détermination de leur masse se fait par des analyses de spectrométrie de masse. Cet outil se révèle très efficace grâce à sa grande sensibilité qui permet d'utiliser de faible quantité de protéines (pas toujours disponibles en grande concentration) et de détecter les produits de réticulation malgré leur faible quantité comparée aux produits non modifiés.

Différents types de spectromètres de masse peuvent être utilisés dans ce type d'étude. Les spectromètres de masse comportant une source electrospray seront généralement couplés à des techniques séparatives comme la chromatographie liquide.

Quoi qu'il en soit, la précision de masse aura un rôle prépondérant dans le choix du type d'appareil. En effet, plus la précision dans la détermination de la masse du peptide réticulé sera importante, plus son identification par calcul mathématique sera aisée et fiable. Ainsi, les analyses les plus récentes utilisent des appareils de type FT-ICR soit en appliquant le protocole classique (Schultz et al., 2004), soit en utilisant une approche quelque peu différente (Kruppa et al., 2003). Dans cette dernière approche, la protéine réticulée est injectée sans digestion préalable dans l'appareil, isolée et fragmentée (grâce aux différents types de fragmentation offerts par ce genre d'appareil) afin de localiser les sites de réticulation en une seule étape.

Mais les laboratoires n'étant que rarement équipés en spectromètres de masse de type FT-ICR, les précisions de masses plus faibles des autres appareils ont été compensées par la réalisation d'analyses MS-MS sur les peptides réticulés. Ces analyses permettent de valider les identifications théoriques des peptides réalisés par les logiciels.

Malgré tout, en dépit de la performance de la spectrométrie de masse comme outil d'analyse, la recherche des produits de réticulation est souvent compliquée et leur identification n'est pas toujours très précise. Ceci est dû notamment au faible rendement de la réaction qui génère une faible quantité de peptides modifiés par rapport aux peptides non modifiés. Plusieurs stratégies ont été envisagées pour pallier ce problème. Elles consistent soit à enrichir la solution en produits réticulés (par séparation des produits réticulés et non réticulés), soit à générer sur les produits modifiés une signature reconnaissable en masse pour en faciliter la détection. Ces différentes approches ont également eu pour but de discriminer les produits issus de réticulation intramoléculaire et des produits issus de réticulation intermoléculaire, ainsi que de différencier rapidement les produits modifiés par un pont interpeptidique de ceux modifiés par un pont intrapeptidique et de ceux modifiés par un agent réticulant hydrolysé.

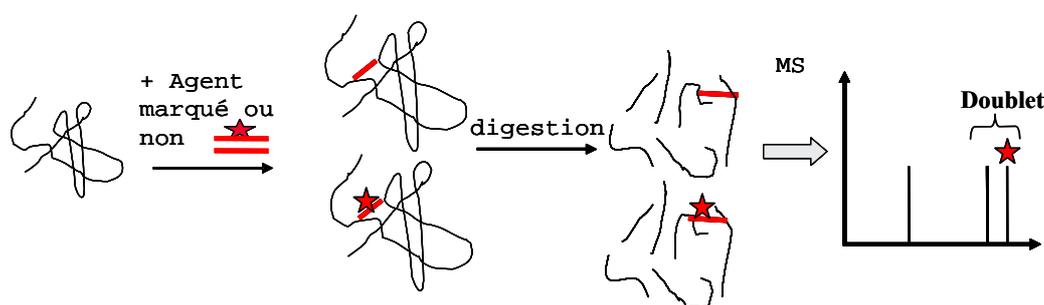
## 1.4 Amélioration de la détection des peptides réticulés

### 1.4.1 Les marquages isotopiques

#### 1.4.1.1 Marquage isotopique de l'agent réticulant

L'incorporation d'isotopes lourds à l'intérieur même de l'agent réticulant permet la détection plus aisée des produits réticulés par leur profil isotopique caractéristique. Certains pré-requis sont alors nécessaires. Seuls les agents insérant une entité chimique peuvent être utilisés, cette condition exclut donc l'ensemble des agents dit de « longueur nulle ». Un deuxième critère est le ratio entre agent marqué et l'agent non marqué. Le meilleur ratio serait de 1:1. Ce ratio permet la détection des deux signaux avec le meilleur rapport signal sur bruit. Les spectres de masses générés après digestion des protéines réticulées marquées ou non, contiendront des massifs sous forme de singulets, caractéristiques de peptides non réticulés, et des massifs sous forme de doublets, caractéristiques des espèces réticulées (figure 51).

Plusieurs études ont utilisé avec succès cette technique de marquage des agents. Par exemple, des agents BSP, BSG et BSS marqué  $d_0$ ,  $d_4$  ont permis l'obtention d'informations structurales sur la protéine Op18 et la caractérisation des interactions dans le complexe Op18/tubuline impliqué dans la déstabilisation des microtubules (Muller et al., 2001).



**Figure 51 : Stratégie analytique pour détecter des réticulations avec des agents marqués à 50% avec des atomes lourds**

Dans une approche similaire, les agents DSA et DMA sont marqués  $d_0$ ,  $d_8$  et utilisés dans le cadre de réticulations intramoléculaires sur le cytochrome c et la ribonucléase A (Pearson et al., 2002). Dans ces deux études les isotopes de l'hydrogène sont utilisés, mais il est également possible d'utiliser d'autres isotopes pour marquer les agents réticulants. Collins et ses collaborateurs ont utilisé des agents doublement marqués à l' $^{18}\text{O}$  et les ont testés sur des peptides modèles. Il semblerait que ce type de marquage ne causerait aucun effet isotopique lors des expériences de LC-MS, ce qui leur conférerait un avantage par rapport aux agents marqués avec du deutérium (Collins et al., 2003), les temps de rétention des produits marqués sont alors exactement les mêmes que ceux des produits non marqués. De plus, dans cette même étude, les agents marqués ou non ont été utilisés dans des solvants contenant de l' $\text{H}_2^{18}\text{O}$ , ainsi les peptides modifiés par un agent hydrolysé ont pu être différenciés des peptides réticulés (intra ou inter).

Ces techniques de marquage des agents réticulants possèdent également un autre avantage remarquable : les peptides ayant incorporé un seul agent sont facilement différenciés des peptides avec deux modifications. En effet ces derniers apparaîtront sur le spectre de masse sous forme de triplet d'intensité 1:2:1 (Pearson et al., 2002).

Le choix de la quantité d'isotopes lourds insérés dans l'agent déterminera la différence de masse entre les deux pics du doublet. Dans les études décrites précédemment, l'écart est de 4 ou 8 Da. Cet espacement doit permettre une séparation suffisante pour détecter les deux molécules même si elles sont fortement chargées, mais pas trop important pour permettre la sélection de tous les isomères lors d'éventuelles analyses MS-MS.

Le principal inconvénient du marquage isotopique des agents réticulant réside dans la diminution du signal des peptides modifiés.

#### *1.4.1.2 Marquage isotopique des protéines*

Le marquage des protéines après ou lors de la digestion enzymatique, se révèle être également un outil performant pour le repérage des produits réticulés dans les spectres de masse. Deux stratégies sont alors envisageables, il est possible de marquer soit l'extrémité N-terminale, soit l'extrémité C-terminale des peptides.

La première étude reportant l'utilisation de cette méthode a consisté à identifier les produits issus de la réticulation de l'ubiquitine (Chen et al., 1999). Dans un premier temps, les groupements amine primaire de la protéine réticulée sont méthylés. Puis, après la digestion

trypsique, les nouvelles extrémités N-terminale des peptides générés sont dérivées avec du 2,4-dinitrofluorobenzène marqué avec 3 deutériums et non marqué (rapport 1:1). Le profil isotopique des peptides non réticulés sera un doublet 1:1, tandis que le profil des peptides réticulés (pont interpeptidique) sera un triplet 1:2:1, puisque ces peptides réagiront deux fois. Pour le marquage des extrémités C-terminale, une propriété intéressante de la trypsine a été utilisée. Il a en effet été observé qu'après la digestion cette enzyme avait la capacité d'échanger les oxygènes de la nouvelle extrémité carboxy-terminale avec les oxygènes du solvant (Rose et al., 1988). Ainsi, la masse des peptides marqués diffère de celle des peptides non marqués de 4 Da. Cette propriété a été utilisée dans l'étude de la réticulation du complexe Rad18/Rad6 (Back et al., 2002 a). Lorsque la protéine réticulée est digérée soit dans de l'eau normale, soit dans de l'eau enrichie avec de l' $^{18}\text{O}$ . Pour les peptides non réticulés ou réticulés de façon intrapeptidique présenteront une différence de masse de 4 Da, tandis que les espèces réticulées de façon interpeptidique présenteront une différence de masse de 8 Da. Dans leur étude de la BSA, Huang et ses collaborateurs utilisent également la digestion dans de l'eau enrichie avec de l' $^{18}\text{O}$ . Cette utilisation permet, en plus de l'identification des produits de réticulation sur les spectres de masse, de faciliter l'attribution des ions fragments lors des expériences de MS-MS (Huang et al., 2004).

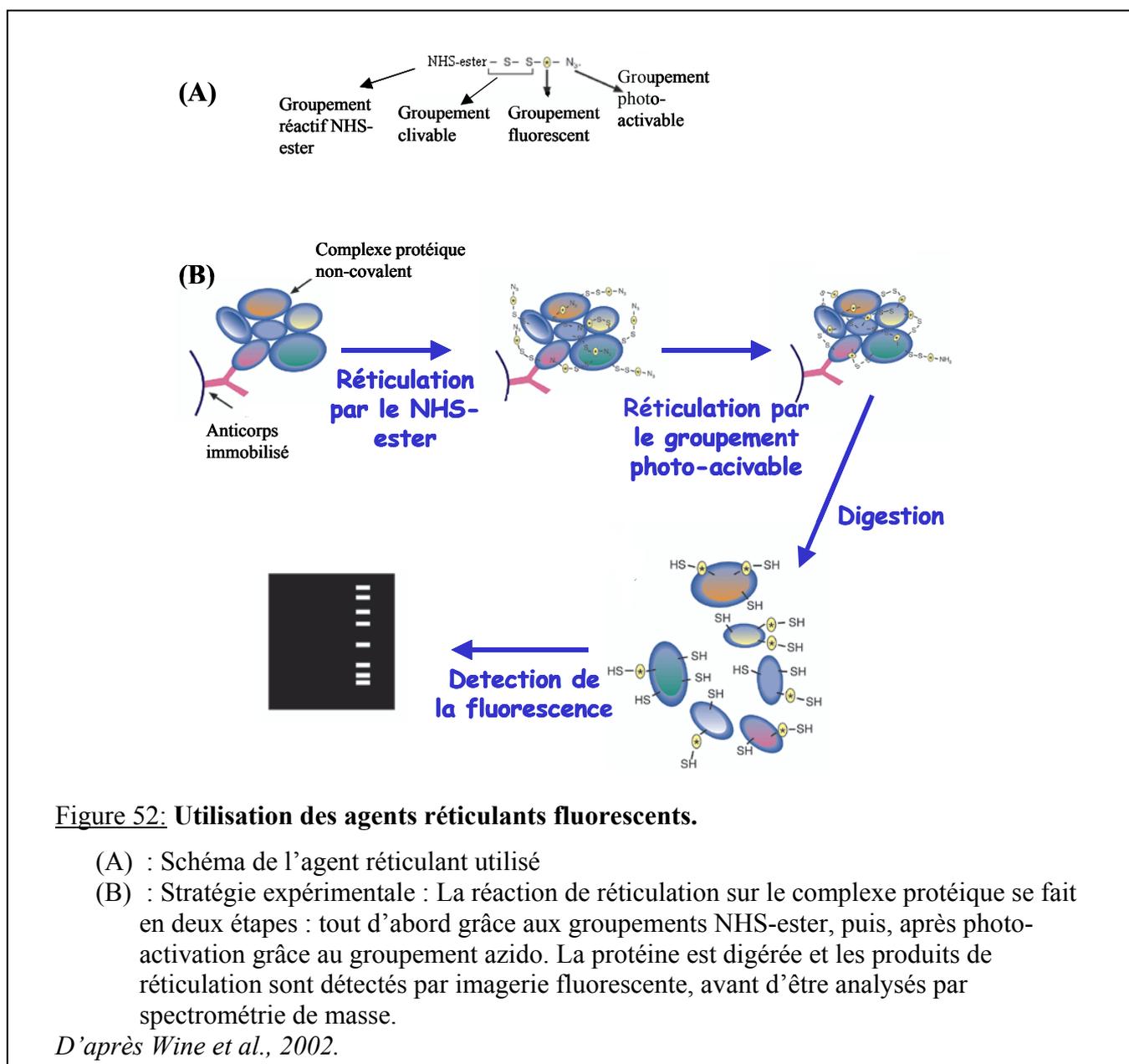
Une autre utilisation du marquage isotopique lors des réticulations chimiques a permis de différencier les réticulations intramoléculaires et les réticulations intermoléculaires dans l'études d'homodimères d'interleukine 6 (Taverner et al., 2002). L'étude a consisté à mélanger en quantité équimolaire des molécules d'interleukine 6 non marquées et marquées avec du  $^{15}\text{N}$ . Les homodimères obtenus seront alors de trois types :  $^{14}\text{N}$ ,  $^{14}\text{N}/^{15}\text{N}$  et  $^{15}\text{N}$ . Après la réaction de réticulation et la digestion, les espèces présentant un profil isotopique sous forme de triplet ou de quadriplet indiqueront une réaction de réticulation entre les deux chaînes de protéine.

### **1.4.2 La fluorescence**

Une méthode alternative aux marquages isotopiques consiste à utiliser des marquages fluorescents sur les agents réticulants. Cette technique a été utilisée par exemple dans l'étude de l'interaction entre la calmoduline et un fragment de la nebuline (Sinz et Wang, 2001). Un agent réticulant bifonctionnel liant les groupements thiol est utilisé (le DBB). Cet agent n'est pas fluorescent en solution mais le devient lorsque ses deux groupements réactifs ont agi.

Ainsi, après réticulation et digestion, les peptides modifiés sont détectés par fluorescence lors d'une HPLC avant l'analyse en spectrométrie de masse.

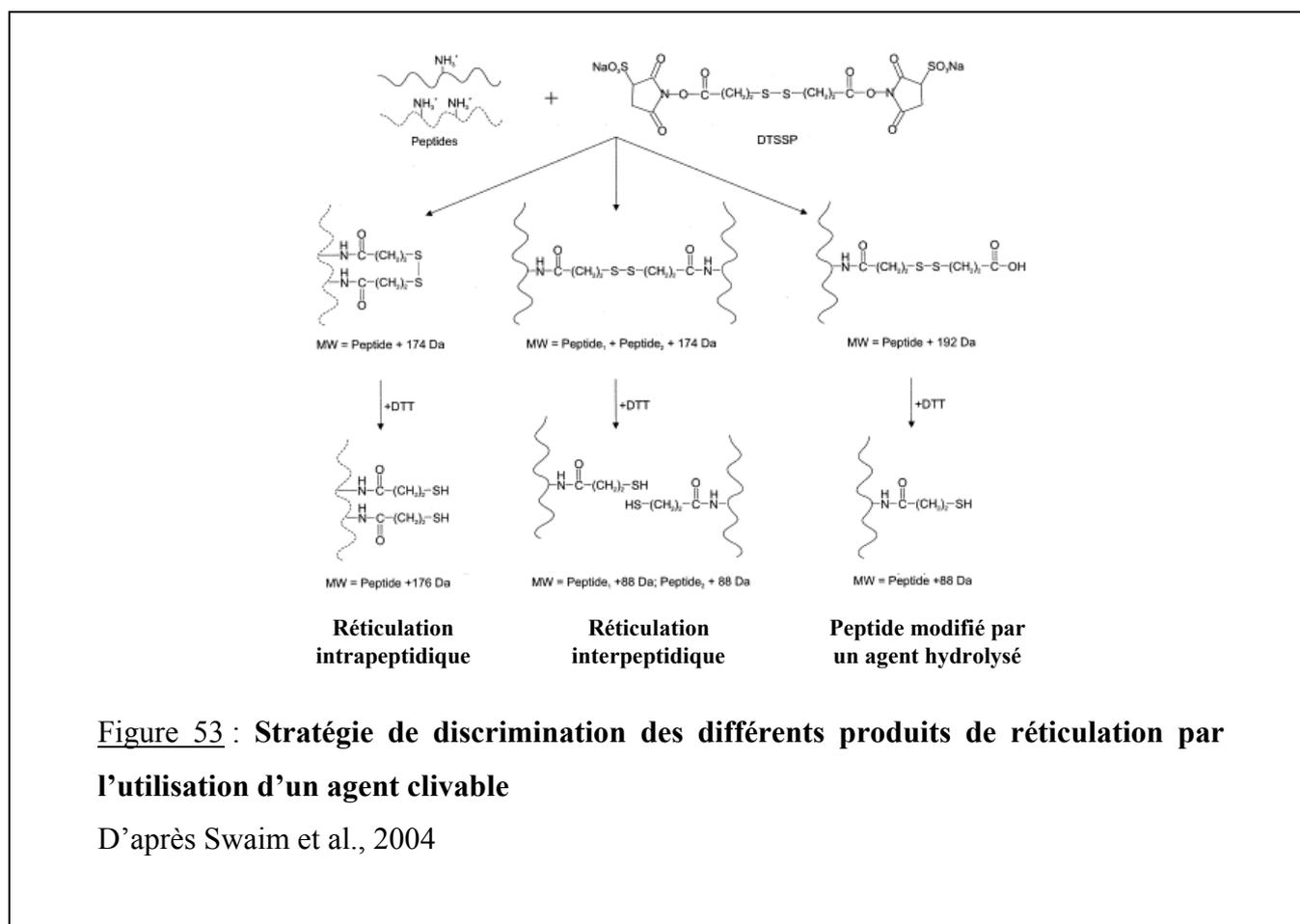
Dans une autre étude l'agent SAED est utilisé pour la caractérisation du complexe entre la protéine gp120 et son récepteur cellulaire CD4 (Wine et al., 2002). Cet agent est composé d'un site de clivage (pont disulfure), d'un groupement réactif NHS-ester d'un groupement réactif photoactivable et d'un groupement fluorescent. Une première réaction de réticulation est réalisée sur les amines primaires, puis le deuxième groupement réactif est activé. La coupure au niveau du site S-S permet d'éliminer le marquage fluorescent des protéines non réticulées. Les protéines réticulées sont détectées par imagerie fluorescente sur un gel SDS, digérées dans le gel et les peptides générés sont analysés par spectrométrie de masse MALDI (figure 52).



### 1.4.3 Les agents réticulants clivables

Le dernier agent que nous avons décrit possédait un site de clivage. Ce type d'agent réticulant peut se révéler très efficace dans les expériences de réticulation notamment pour différencier les types de produits obtenus. L'agent clivable le plus couramment utilisé est le DTSSP, qui possède un pont disulfure pouvant être réduit par du DTT par exemple. Cet agent a été utilisé dans les études de l'homodimère ParR et du complexe CD28/CD80 (Bennet et al., 2000), et dans l'étude structurale d'un complexe de levure (Back et al., 2002).

Les peptides modifiés sont présents sur les spectres de masse obtenus avant réduction et absent de ceux obtenus après réduction. Les différences de masse observées entre peptides modifiés et peptides non modifiés, avant et après réduction, permettront de discriminer les réticulation intrapeptidiques des réticulations interpeptidiques et des peptides présentant un agent hydrolysé (Swaim et al., 2004) (figure 53).



**Figure 53 : Stratégie de discrimination des différents produits de réticulation par l'utilisation d'un agent clivable**

D'après Swaim et al., 2004

#### **1.4.4 Les systèmes de purification**

Pour enrichir le spectre de masse en signal de peptides réticulé, un tag d'affinité peut être fixé sur l'agent réticulant, un agent trifonctionnel est alors obtenu. Cette stratégie a été utilisée pour étudier l'hétérodimère NC2 (Trester-Zedlitz et al., 2003). Un tag d'affinité de type biotine est greffé sur l'agent réticulant. Après l'étape de réticulation et la digestion, une purification sur colonne d'affinité (ici une colonne d'avidine) est réalisée. Seuls les peptides modifiés sont récupérés, ils peuvent donc être analysés seuls par spectrométrie de masse.

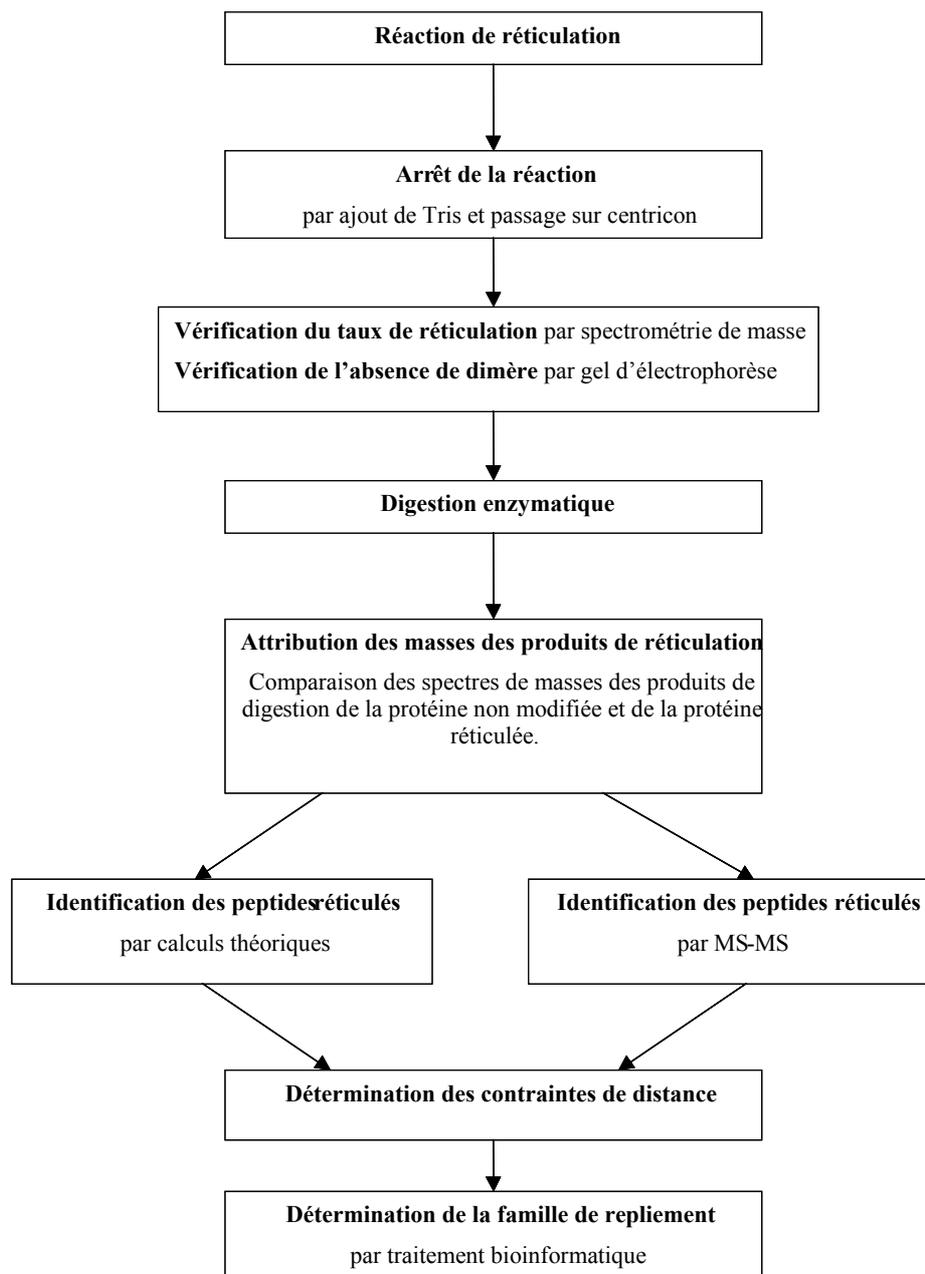
## **2 BUTS ET PRINCIPES DE L'ETUDE**

Dans cette étude, nous allons tenter d'appliquer, et d'adapter la méthode classique de réticulation intramoléculaire associée à la spectrométrie de masse sur une protéine modèle, le cytochrome c. Nous espérons ainsi obtenir des contraintes de distances entre différents acides aminés, qui nous permettront après traitement bioinformatique de déterminer la famille de repliement de la protéine. Nous espérons sur ce modèle développer une méthode applicable aux protéines dont la structure est inconnue.

Les réticulations seront effectuées à l'aide de trois agents réticulants différents : le BS3, le DST et l'EGS. Ces trois agents sont des esters sulfo-NHS. Ils diffèrent par la longueur de leur bras espaceur : 6.3Å, 11.4 Å et 16.1 Å pour le DST, le BS3 et l'EGS respectivement (distances théoriques données par les fabricants). Ces agents hydrosolubles vont donc pouvoir réagir avec les chaînes latérales des lysines. Le NH<sub>2</sub> terminal étant dans notre cas acétylé, il ne pourra pas réagir.

Le cytochrome c a été choisi comme modèle car il est disponible facilement en grande quantité et que sa structure tridimensionnelle a pu être résolue précisément par RMN et par cristallographie par diffraction des rayons X. De plus, ses lysines sont en nombre important et réparties sur l'ensemble de la séquence, ce qui présage un taux de réticulation suffisant pour nos mises aux points.

Le protocole expérimental est représenté figure 54. Après la réaction entre l'agent réticulant et la protéine, la réticulation est arrêtée par l'ajout de Tris, qui va hydrolyser les agents qui n'ont pas (ou que partiellement) réagi. Un passage sur un centricon éliminera l'excès d'agents réticulant. Puis la protéine réticulée est digérée avec de l'endoprotéinase Lys-C. Les produits de réticulation sont repérés par spectrométrie de masse et identifiés par calculs informatiques. Des expériences de MS-MS sur certains peptides réticulés permettent de valider l'identification théorique. Les contraintes de distances entre les lysines ainsi déterminées sont analysées par bioinformatique (figure 54).



**Figure 54: Principe des réticulations intramoléculaires couplées à la spectrométrie de masse pour la détermination de familles de repliement.**

### **3 RESULTATS ET DISCUSSION**

#### ***3.1 La réticulation intramoléculaire***

##### **3.1.1 Les conditions de réticulation**

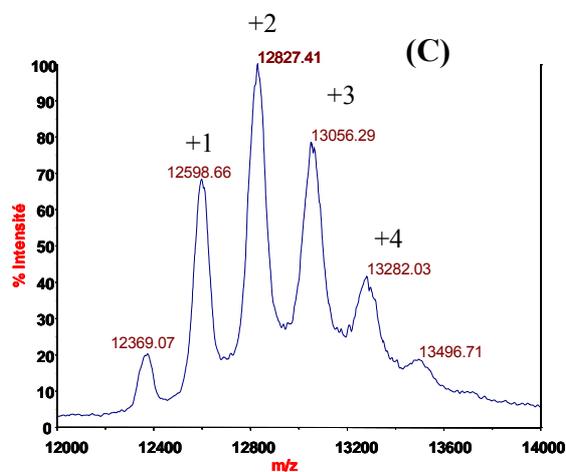
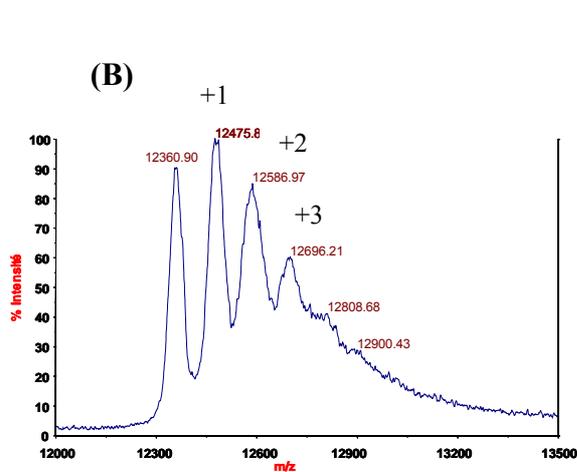
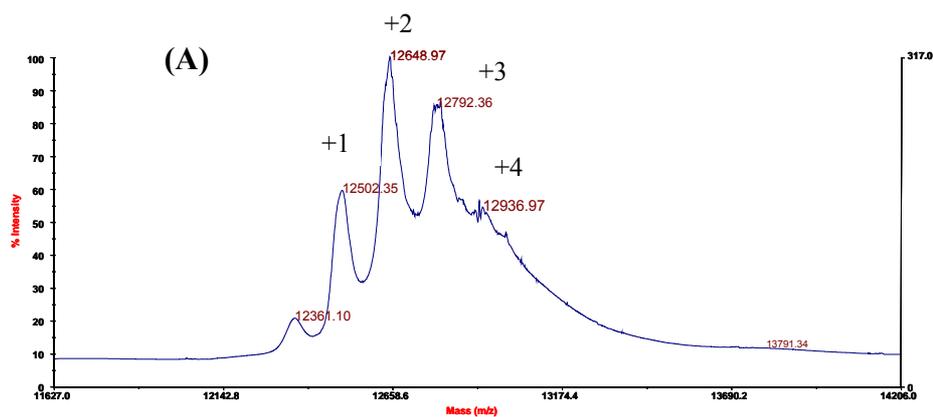
Lors de la réaction de réticulation, deux facteurs principaux peuvent être contrôlés : la concentration en protéine et le rapport entre l'agent réticulant et son substrat. La concentration en protéine doit être suffisante pour permettre les analyses ultérieures en spectrométrie de masse, mais si elle est trop importante, nous risquons de favoriser certains ponts intermoléculaires. De la même façon, le ratio agent réticulant/protéine doit être suffisant pour les analyses, mais s'il est trop important, il risque d'induire des déformations structurales importantes au niveau de la protéine, les informations de distance obtenues ne seraient alors plus significatives.

Différentes conditions de réticulations ont été testées. Des cinétiques de réaction sont réalisées pour chaque agent. Et différentes concentrations en protéine et en agent réticulant sont testées. Il est apparu que le temps de réaction n'avait qu'un effet modéré sur la réticulation. Nous avons choisi d'utiliser le cytochrome c à une concentration de 140  $\mu\text{M}$ . Les rapports de concentration sont calculés en fonction de la concentration en en agent réticulant par rapport à la concentration lysine (19 lysines par molécule de cytochrome c). Les rapports optimaux de la concentration en lysine sur la concentration en agent réticulant sont compris entre 0.5 et 0.05.

Pour chaque agent, le taux moyen de réticulation est déterminé par spectrométrie de masse MALDI-TOF (figure 55).

La protéine incorpore donc entre zéro et quatre agents avec le BS3 et l'EGS et entre zéro et trois pour le DST. Dans ces conditions, nous sommes certains de limiter les distorsions dans la structure tridimensionnelle.

Nous avons vérifié par spectrométrie de masse MALDI-TOF et par électrophorèse en conditions dénaturantes qu'aucun dimère n'avait été formé.



**Figure 55: Spectres de masse des protéines après réticulation obtenus par MALDI-TOF**

- (A) : Cytochrome C + BS3 : entre 0 et 4 adduits
- (B) : Cytochrome C + DST : entre 0 et 3 adduits
- (C) : Cytochrome C + EGS : entre 0 et 4 adduits

A ce stade de l'expérimentation, il nous est apparu indispensable de bloquer la réaction de réticulation avec certitude. Nous avons donc choisi de diluer notre échantillon dans du tampon Tris qui possède des amines réactives. Les agents réticulants fixés à une seule extrémité à la protéine et dont l'autre extrémité n'est pas hydrolysée vont réagir avec les molécules de Tris et s'hydrolyser totalement. Si l'extrémité non fixée n'est pas hydrolysée, des réactions de réticulation peuvent se produire après la digestion. Ces réticulations ne sont alors plus

significatives car elles peuvent concerner des acides aminés qui n'étaient pas spatialement proches dans la structure tridimensionnelle.

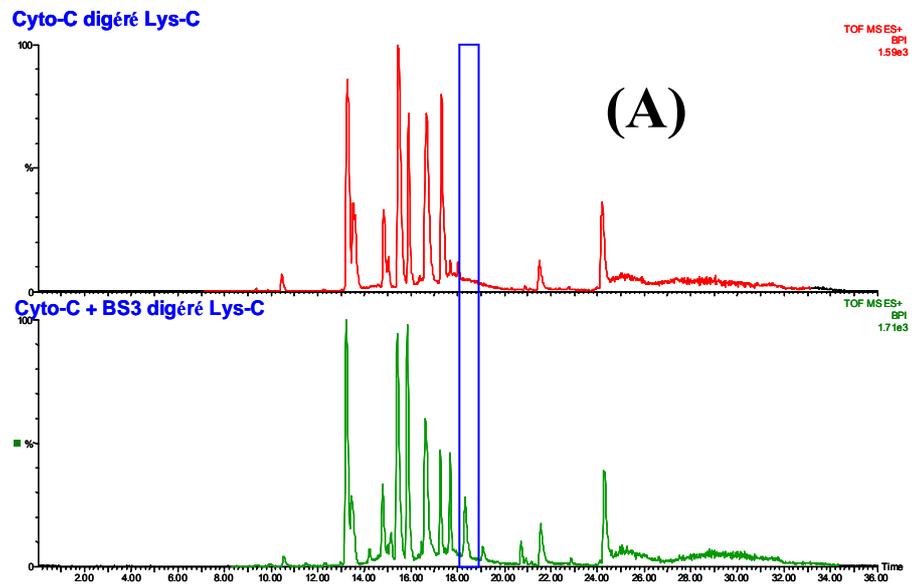
### **3.1.2 Les produits de réticulation**

Les protéines réticulées sont digérées par de l'endoprotéase Lys-C. Cette enzyme a été choisie car elle coupe les liaisons peptidiques spécifiquement après les lysines. Or lorsque la lysine est modifiée (par un adduit d'agent réticulant), elle n'est pas reconnue par l'enzyme du fait d'un changement de conformation de son carbone asymétrique et de la modification de la charge et de l'encombrement stérique. Ainsi, nous aurons un contrôle supplémentaire de nos résultats, seules les masses correspondant à des peptides comprenant une coupure théorique non réalisée (ou « missed-cleavage ») seront susceptibles d'être le résultat d'une addition d'agent réticulant.

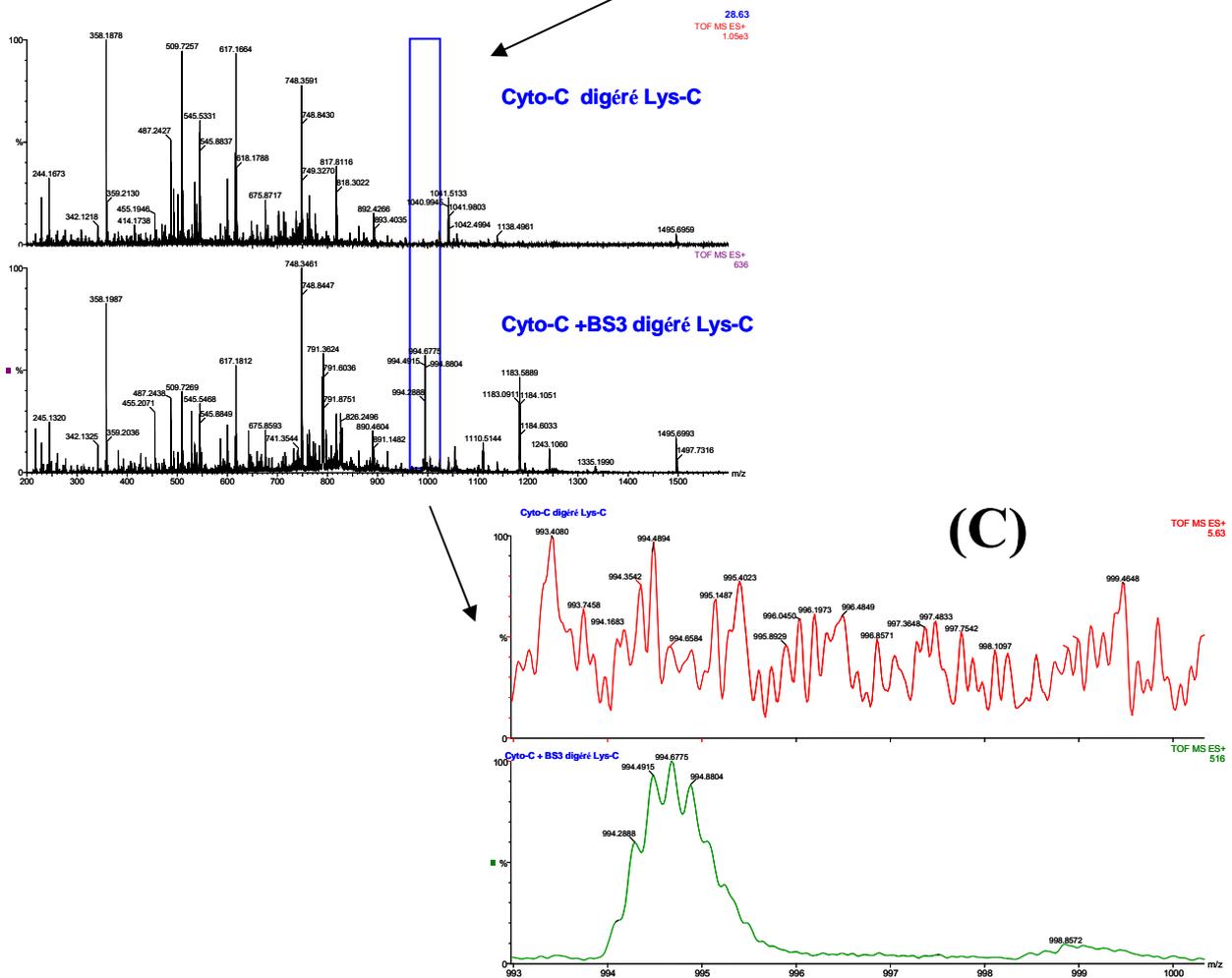
Les masses des produits issus de la réticulation sont déterminées par la comparaison des spectres de masse des produits de digestion de la protéine non modifiée et des spectres de masse des produits de digestion de la protéine réticulée avec un des trois agents. Les analyses sont réalisées de deux façons différentes : par MALDI-TOF sur le lysat non séparé (figure 57) et sur un Qq-TOF sur le lysat après séparation par chromatographie (figure 56). Ainsi, nous espérons détecter un maximum de produits avec à la fois une grande précision dans la masse et une résolution importante. La présence sur les spectres de masse des produits de la réticulation, d'ions de peptides provenant de la protéine non modifiée, nous permet de réaliser une calibration interne de nos spectres et d'attribuer ainsi une masse à chaque produit issu de la réticulation avec une grande précision.

De cette façon, nous avons pu mettre en évidence 134 produits issus de la réticulation par le BS3, 55 produits issus de celle par le DST, et 76 produits issus de celle par l'EGS.

A ce stade de l'étude, nous avons déjà pu observer une perte d'information. En effet, l'intensité du signal des peptides réticulés est généralement très inférieure à celle des peptides non modifiés. A cause de cette faible intensité certaines masses de produits de réticulation n'ont pu être attribuées avec une précision suffisante pour être exploitables dans la suite de l'étude.



(B)

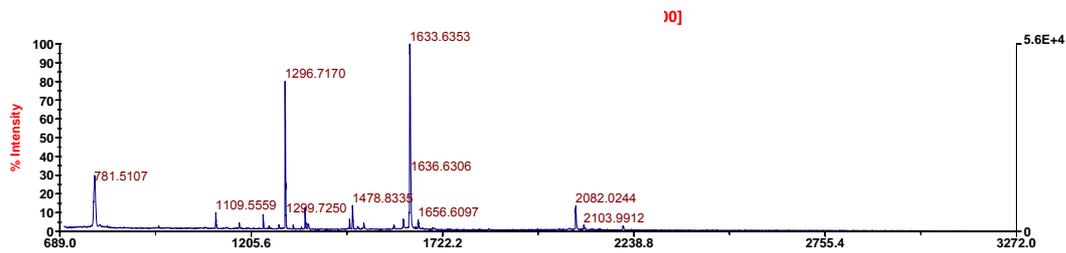


(C)

**Figure 56 : Détermination des masses des produits de réticulation par LC-MS après digestion par l'endoprotéase Lys-C.**

- (A) : chromatogrammes des lysats du cytochrome c (en haut) et du cytochrome c réticulé avec le BS3 (en bas).
- (B) : spectres de masses du temps de rétention 17.8 minutes des lysats du cytochrome c (en haut) et du cytochrome c réticulé avec le BS3 (en bas) obtenus par ESI-Qq-TOF.
- (C) : agrandissement des spectres de masse : l'ion de m/z 994.2 chargé 5 fois est présent sur le spectre du cytochrome c réticulé digéré (en bas) et pas sur celui du cytochrome c non modifié digéré (en haut)

### Cytc digéré LysC



### Cytc + BS3 digéré LysC

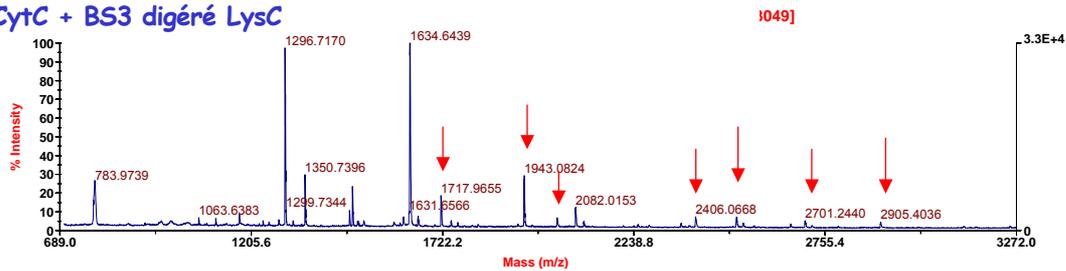


Figure 57 : Détermination des masses des produits de réticulation par MALDI-TOF  
Spectres de masse de la protéine digérée (en haut) et de la protéine réticulée avec le BS3 et digérée  
Les flèches rouges montrent des produits issus de la réticulation

## 3.2 Recherche des peptides pontés

### 3.2.1 Identification des peptides pontés par calculs mathématiques

L'ensemble des masses relevées sur les spectres comme étant des masses de produits issus de la réticulation sont analysés par un logiciel. Ce logiciel a été développé par Markus Muller (Swiss Institute of Bioinformatics) sous Matlab 6.

Une liste de peptides théoriquement obtenus après digestion à la Lys-C est d'abord générée, avec le nombre de « missed-cleavage » désiré. Puis, ces peptides sont virtuellement réticulés,

afin de créer une liste de masse de tous les produits de réticulation pouvant exister. Tous les types de réticulation sont pris en compte : les type 0 (un peptide modifié par un agent hydrolysé une extrémité), les type 1 (un peptide avec un pont intrapeptidique), les type 2 (deux peptides réticulés) et les réticulations combinant plusieurs type de modifications.

La liste des masses théoriques est ensuite comparée à la liste des masses expérimentales relevées sur les spectres.

Cette comparaison peut se faire avec des tolérances diverses, nous avons choisi une tolérance en masse de 150 ppm. La grande majorité des masses obtenues le sont en fait avec une bien meilleure précision, mais de cette façon, nous voulons nous affranchir des éventuelles erreurs effectuées lors de l'analyse de certains spectres (notamment pour les ions d'états de charges élevés).

Pour le BS3, 74 des 134 masses que nous avons trouvées peuvent être attribuées mathématiquement à des peptides réticulés. Pour le DST, 30 des 55 produits sont identifiés et 36 des 76 produits pour l'EGS.

Seules environ 50% des masses relevées sur les spectres de masse ont pu être retrouvées par le logiciel. Les autres masses pourraient provenir de la réticulation de pollutions au niveau de l'agent réticulant, au niveau de la protéine ou encore au niveau des tampons (Swaim et al., 2004).

Pour certaines masses, il est possible d'obtenir plusieurs solutions mathématiques différentes (13 masses pour le BS3, 9 masses pour le DST, et 13 pour l'EGS). Par conséquent, l'analyse mathématique seule ne suffit pas, c'est pourquoi nous avons tenté de réaliser des expériences de MS-MS sur les peptides d'intérêt.

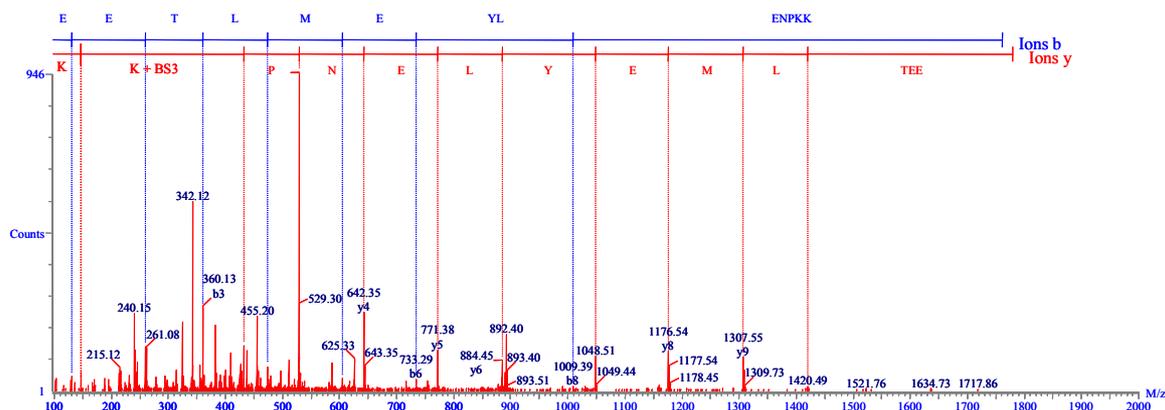
### **3.2.2 Identification des peptides pontés par MS-MS**

Des expériences de MS-MS ont été réalisées sur des appareils de type MALDI-TOF-TOF et par LC-MS-MS sur un Qq-TOF après création d'une liste contenant les masses des produits de réticulation qui seront spécifiquement fragmentés.

La fragmentation des différents produits de réticulation n'a pas toujours permis d'obtenir des spectres de masses interprétables. En effet, comme nous l'avons vu précédemment, l'intensité du signal de ces produits est généralement très faible. Ainsi, l'intensité des précurseurs est parfois insuffisante pour l'obtention de spectres MS-MS interprétables. De plus, la

fragmentation des produits contenant l'hème conduit à la création de l'ion caractéristique de l'hème à 616 m/z qui a tendance à supprimer le signal des autres ions fils.

Malgré tout, nous avons pu obtenir des spectres de fragmentation pour différents produits de réticulation, pour des peptides modifiés par un agent réticulant hydrolysé, pour des peptides réticulés deux à deux et pour des produits possédant un pont intrapeptidique.



**Figure 58 : Spectre de masse MS-MS de l'ion 890.4 obtenu par ESI-Qq-TOF**

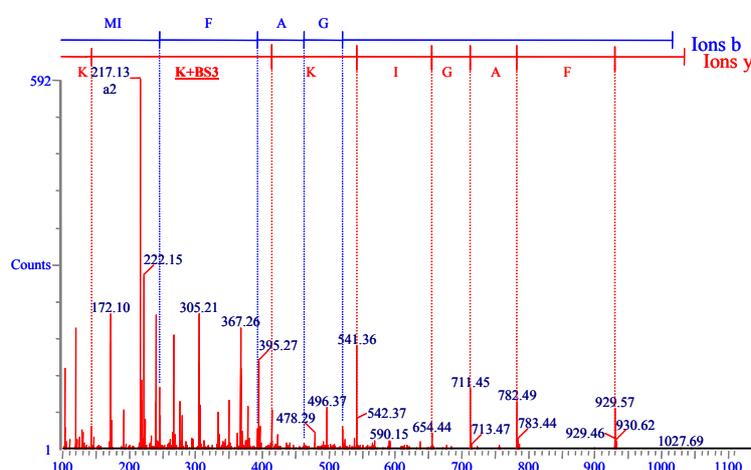
Les fragments observés permettent de valider l'identification théorique du peptide comme étant le peptide 61-73 modifié par une molécule de BS3 hydrolysée.

La séquence du peptide est : EETLM EYLENPKK.

La figure 58 montre le spectre de masse de la fragmentation de l'ion doublement chargé de m/z 890.43. Cet ion appartient donc à un produit de masse 1778.86 Da que l'analyse mathématique a identifié comme étant le peptide 61-73 modifié par une molécule de BS3 hydrolysée à une de ses extrémités. Sur le spectre de masse il est possible de reconstituer la séquence de ce peptide d'après les ions y et les ions b. Nous pouvons remarquer que la lysine modifiée par le BS3 est repérable sur le spectre. Cette liaison ne s'est donc pas dissociée lors de la fragmentation, et l'acide aminé modifié a le même profil de fragmentation qu'un acide aminé normal. Ceci laisse présager que ces expériences de MS-MS peuvent permettre de localiser précisément le lieu de fixation de l'agent réticulant.

La figure 59 représente le spectre de fragmentation de l'ion doublement chargé de m/z 587.36. Cet ion est issu d'un produit de masse 1172.70 Da identifié mathématiquement comme étant le peptide 80-88 réticulé de façon intrapeptidique par le BS3 entre les lysines 87 et 86. Sur le

spectre, les ions y et les ions b permettent de reconstituer la séquence du côté N-terminal du peptide. Les informations concernant l'extrémité C-terminale du peptide sont moins nombreuses, probablement à cause du pont intrapeptidique. Il est cependant possible d'identifier certains fragments qui comprendraient une lysine modifiée par le BS3. Ainsi l'ion de m/z 147 correspondrait à l'ion y1 contenant la lysine terminale du peptide. L'ion de m/z 413.3 correspondrait à l'ion y2 contenant la lysine terminale et la lysine 87 et la masse d'un adduit de BS3 (138 Da). Enfin, l'ion de m/z 541.4 correspondrait au fragment y3 comprenant les trois lysines (88, 87 et 86) et une molécule de BS3, liée à une ou deux lysines. A priori, la fragmentation permet de casser un pont intrapeptidique, créant un adduit sur une seule des lysines concernées.



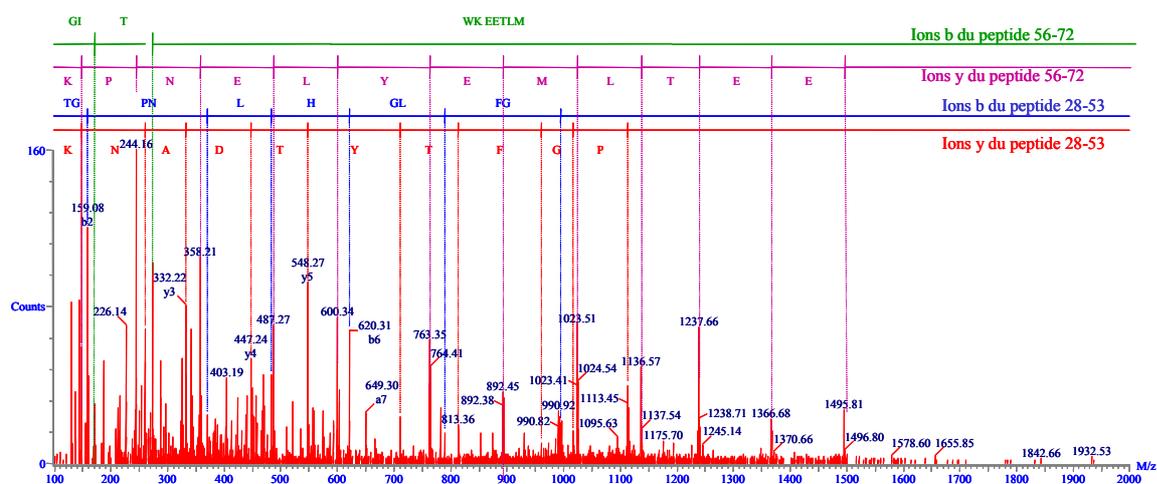
**Figure 59: Spectre de la fragmentation de l'ion 587.36 obtenu par ESI-Qq-TOF**

Les fragments observés permettent de valider l'identification théorique du peptide comme étant le peptide 80-88 avec un pont de BS3 entre les lysines 86 et 87.

La séquence du peptide est : MIFAGIKKK.

La fragmentation de l'ion 994.11 chargé cinq fois donne le spectre de masse présenté figure 60. Cet ion appartient à un produit de masse 4965.55 Da identifié par le logiciel comme étant le peptide 28-53 lié au peptide 56-72 par une molécule de BS3. Les lysines concernées par le pont seraient alors les lysines 39 et 60. L'interprétation d'un tel spectre est plus délicate que dans les exemples précédents. Sur ce spectre il est possible d'identifier des fragments

correspondant à la partie C-terminale de chacun des deux peptides. Quelques ions correspondant à leur partie N-terminale sont également présents. Cependant nous ne pouvons observer d'ions appartenant à leur partie centrale. Le produit a donc été formellement identifié comme étant le produit de réticulation du peptide 28-53 et du peptide 56-72, mais aucune information précise n'a pu être obtenue au niveau du pont. Il est impossible de déterminer si le pont a été rompu lors de la fragmentation.



**Figure 60: Spectre de masse de la fragmentation de l'ion 994.1 obtenu par ESI-Qq-TOF**

Les fragments observés permettent de valider l'identification théorique du peptide comme étant le peptide 56-72 lié par une molécule de BS3 au peptide 28-53 entre les lysines 39 et 60. La séquence du peptide 56-72 est GITWKEETLM EYLENPK, et la séquence du peptide 28-53 est TGPLHLGLFGRKTGQAPGFYTDANK.

### 3.2.3 Détermination des contraintes de distance

Grâce aux expériences de MS-MS, nous avons pu confirmer l'identification de certains produits de réticulation. Pour le BS3, dix produits ont été identifiés : quatre d'entre eux sont des peptides modifiés par un agent réticulant hydrolysé (type 0), cinq sont des peptides avec un pont interne (type 1) et un est le produit de la réticulation de deux peptides (type 2).

De la même façon pour le DST six produits sont identifiés : quatre de type 0 et deux de type 1.

Et pour l'EGS nous avons confirmé six produits : deux type 0 et quatre type 1.

A ces peptides formellement identifiés, nous avons ajouté certains des résultats obtenus uniquement de façon mathématique. Nous n'avons en fait conservé que les produits qui donnaient une unique solution mathématique et pour lesquels l'attribution des lysines touchées a pu se faire sans ambiguïté. La totalité des produits de réticulation identifiés sont répertoriés tableau 2.

La première information que nous pouvons tirer de ces résultats concerne l'accessibilité des lysines. Lors de la réticulation avec le BS3, toutes les lysines sauf les lysines 55 et 79 ont été touchées par l'agent réticulant. Comme tous les produits de réticulation n'ont pas été identifiés nous ne pouvons conclure quant à l'accessibilité des lysines 55 et 79, par contre nous pouvons affirmer que toutes les autres lysines sont bien accessibles au solvant.

Les réticulations de type 1 et de type 2 vont nous permettre d'obtenir des informations de distance. Pour chaque pont, une distance est associée en fonction de l'agent réticulant utilisé. Pour chaque agent la longueur maximale du pont créée sera la longueur de l'agent donnée par le fabricant. Mais il faut également tenir compte de la flexibilité des chaînes latérales des lysines (et des tyrosines). La distance entre le carbone  $\alpha$  et l'azote de la chaîne latérale d'une lysine est approximativement de 6 Å (la distance est la même pour les chaînes latérales de tyrosines). Ainsi, les agents pourront lier des groupements réactifs distants au maximum de la longueur théorique de l'agent plus 12 Å. Les valeurs de distance utilisées sont celles données par le constructeur plus ou moins l'erreur due à la flexibilité des chaînes latérales des lysines, soit 18.4 Å pour le Sulfo-DST, 25 Å pour le BS3 et 28.1 Å pour le Sulfo-EGS. L'utilisation combinée des trois agents aurait pu nous permettre d'obtenir des informations plus fines sur les distances interlysines. Cependant, tous les produits de réticulation n'étant pas identifiés, il n'est pas possible de savoir si l'absence d'un produit de réticulation est réellement due à son absence physique, ou si elle est due à une impossibilité d'identification. Ainsi, dans notre cas, l'utilisation combinée des trois agents ne nous donne une information très précise que lorsque la réticulation est présente avec chacun d'entre eux, c'est le cas par exemple des ponts 85-88 et 25-27. La distance maximale entre les deux lysines sera donnée par le plus petit agent (ici 18.4 Å).

		Masse théorique du produit	type de réticulation	peptide(s) concerné(s)	lysine(s) modifiée(s)
BS3	Identification par MS-MS	1173.67	type 1	80-88	86-87
		1718.02	type 0	26-39	27
		1745.05	type 1	87-99	87-88
		1779.86	type 0	61-73	72
		1814.06	type 1	23-38	25-27
		2365.18	type 0	56-73	72
		2405.19	type 0	9-22	13
		2700.41	type 1	6-22	7-8
		2856.50	type 1	8-22	8-13
		4965.42	type 2	56-72 et 28-53	60-39
	Identification mathématique	918.55	type 0	8-13	8
		962.58	type 0	73-79	73
		1063.59	type 0	80-87	86
		1085.65	type 1	6-13	7-8
		1589.90	type 0	26-38	27
		1673.91	type 2	1-7 et 8-13	5-8
		1851.16	type 0	74-88	87
		1868.98	type 0	40-55	53
		1942.04	type 1	23-29	25-27
		2031.93	type 0	14-25	22
		2093.08	type 2	6-8 et 61-73	7-72
		2237.10	type 0	56-72	60
		2259.10	type 0	6-22	7-8ou13
		2261.33	type 2	100-104 et 26-39	100-27
		2421.10	type 1	61-79	72-73
		2668.40	type 2	80-87 et 61-73	86-72
		2904.40	type 0	29-53	39
		3049.30	type 2	26-38 et 89-100	27-99
		3128.70	type 1	28-55	39-53
		DST	Identification par MS-MS	1693.93	type 0
1720.97	type 1			87-99	87-88
1918.06	type 1			23-39	25-27
2068.10	type 0			23-39	25-27
2381.09	type 0			9-22	13
2880.58	type 0		28-53	39	
Identification mathématique	1610.87		type 0	89-100	99
	1649.93		type 2	1-7 et 8-14	5-8
	1693.93		type 0	26-39	27
	1755.83		type 0	61-73	72
	1826.02	type 0	74-88	87-86	
EGS	Identification par MS-MS	2341.26	type 0	56-73	72
		2883.31	type 0	8-25	8-13
		3277.87	type 0	26-53	27-39
		1261.62	type 1	80-88	86-87
		1805.98	type 0	26-39	27
	Identification mathématique	1833.03	type 1	87-99	87-88
		2030.13	type 1	23-39	25-27
		2493.15	type 0	9-22	13
		2603.25	type 1	8-22	8-13
		1151.55	type 0	80-87	86
		1173.53	type 1	6-13	7-8
		1818.00	type 2	6-8 et 14-25	7-22
		2119.06	type 2	6-8et 26-39	7-27
		2375.17	type 2	61-73 et 23-27	25-72
		3216.75	type 1	28-55	39-53

**Tableau 2 : Identification des produits de réticulation du cytochrome c par le BS3, le Sulfo-DST et le Sulfo-EGS.**

### 3.3 Exploitation des données par bioinformatique

Avec l'aide des trois agents de réticulation, nous avons pu identifier six ponts confirmés par MS-MS et dix autres par calcul mathématique (tableau 3). Sur toutes les contraintes de distance obtenues, seules certaines d'entre elles vont être réellement exploitables en terme de structure tridimensionnelle. En effet, des ponts entre des acides aminés très proches sur la structure primaire de la protéine, n'apporteront aucune information de distance entre les différentes structures secondaires de la protéine.

	BS3		EGS		DST	
	Lysines réticulées	Distances dans la structure cristallographique	Lysines réticulées	Distances dans la structure cristallographique	Lysines réticulées	Distances dans la structure cristallographique
ponts confirmés par MS-MS	K87-K88	14.19	K86-K87	3.44	K87-K88	14.19
	K25-K27	5.26	K87-K88	14.19	K25-K27	5.26
	K7-K8	10.8	K25-K27	5.26		
	K8-K13	11.7	K8-K13	11.7		
	K86-K87	3.44				
	K60-K39	8.3				
ponts mathématiquement possibles	K72-K73	12.9	K7-K8	10.8	K5-K8	9.36
	K27-K99	24.9	K39-K53	15.63		
	K39-K53	15.63	K7-K22	20.1		
	K5-K8	9.36	K7-K27	13.92		
	K7-K72	25.3	K25-K72	21.4		
	K27-K100	18.5				
	K86-K72	15				

**Tableau 3 : Récapitulatif des lysines réticulées.**

Toutes ces distances sont compatibles avec celles données par la structure cristallographique du cytochrome c. Les distances obtenues avec le Sulfo-DST sont toutes inférieures à 18 Å, celles obtenues avec le BS3 sont inférieures à 25 Å et celles obtenues avec le Sulfo-EGS sont inférieures à 28.1 Å (tableau 3). Ces distances sont celles données par la cristallographie entre les deux amines réactives, les distances données par la structure RMN sont un peu différentes. Dans les structures les chaînes latérales des lysines sont figées, ce qui n'est pas toujours le cas en solution.

Dans l'étude du cytochrome c par réticulation par le DSA et par le DMA qui avait été réalisée en 2002 par Pearson et ses collaborateurs, six ponts avaient été identifiés (Pearson et al.,

2002). Parmi ces six ponts, nous en avons retrouvé cinq. L'étude réalisée par l'équipe de Dihazi en 2003 par réticulation du cytochrome c avec le BS3, le Sulfo-EGS et le Sulfo-DST a permis l'identification de quatre ponts (Dihazi et al., 2003). Nous avons également identifié trois de ces ponts. L'étude réalisée par Schilling et ses collaborateurs en 2003 portait sur la réticulation du cytochrome c par le BS3 et a permis la détermination de quatre contraintes de distance. Nous avons pu également par notre méthode retrouver ces quatre contraintes de distance. Notre méthode s'est donc avérée performante puisque nous avons réussi à identifier de six à seize ponts.

Ces seize distances ont été transmises au laboratoire de bioinformatique et RMN structurale de l'Institut de Biologie et Chimie des Protéines de Lyon. Ces données expérimentales leur serviront à mettre au point et à affiner leur méthode de détermination des familles de repliement de protéines inconnues.

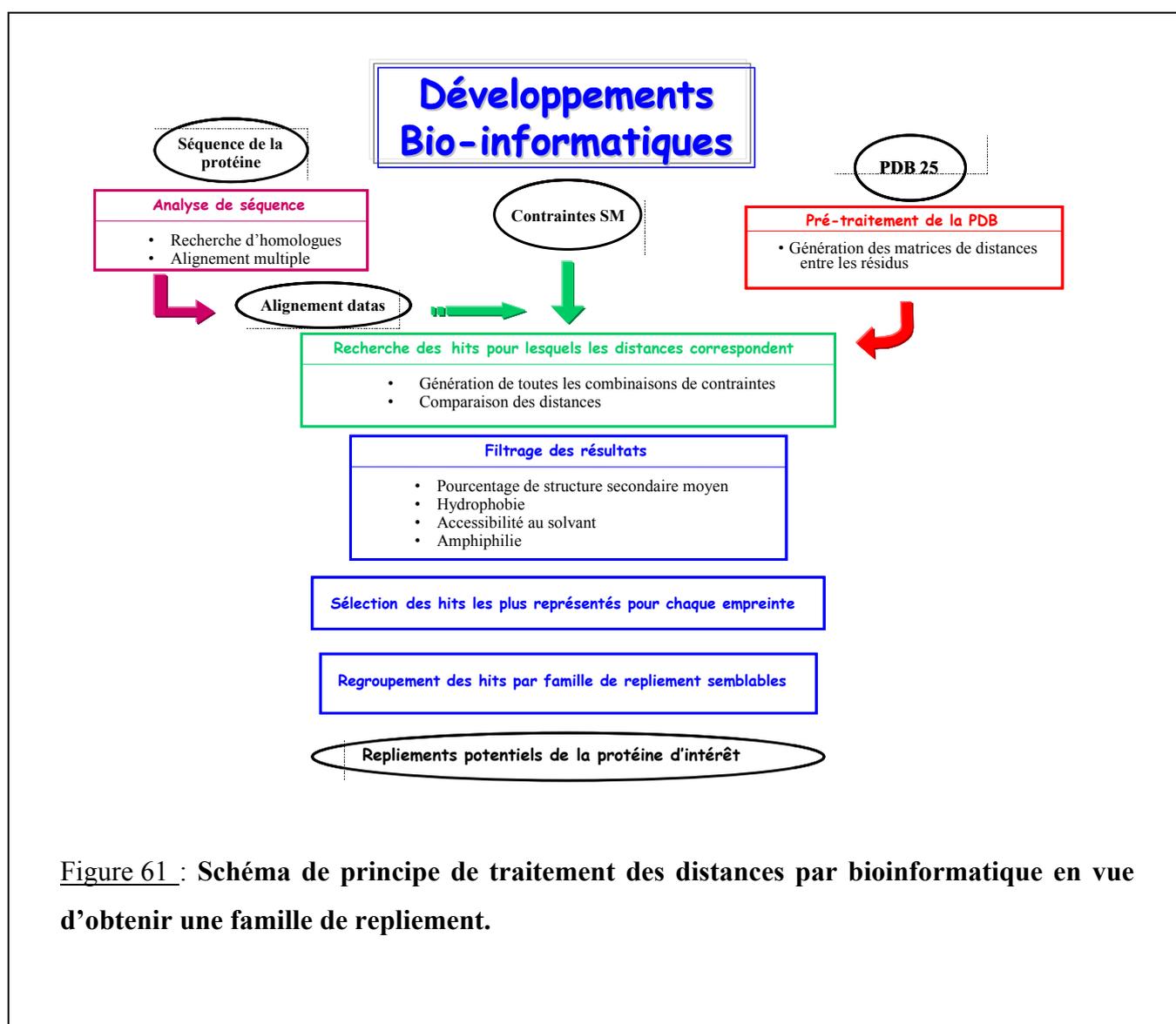


Figure 61 : Schéma de principe de traitement des distances par bioinformatique en vue d'obtenir une famille de repliement.

Les étapes du traitement bioinformatique sont représentées figure 61. Les contraintes de distances obtenues par réticulation chimique associée à la spectrométrie de masse et les analyses de séquences primaires (recherche de protéine homologue, prédiction de structure secondaire) sont compilées et appliquées sur une base de donnée structurale type PDB modifiée. Cette base de données contient des informations de distances entre les différents résidus amino-acides de toutes les familles de repliement répertoriées dans la banque PDB. Les résultats positifs sont ensuite triés grâce à des filtres biochimique et structuraux expérimentaux et modélisés (hydrophobicité, pourcentage de structure secondaire, accessibilité au solvant, flexibilité...). Ce tri permet ensuite de générer des modèles tridimensionnels et d'attribuer une famille de repliement à la protéine étudiée.

Avec nos résultats expérimentaux, nous n'avons pu obtenir clairement la famille de repliement du cytochrome c. Cependant en partant de toutes les familles possibles issues de la PDB, nous sommes arrivés à seulement quelques familles de repliement potentielles dont celle du cytochrome c. Ainsi nous sommes sur la bonne voie. Il nous manque toutefois encore des informations de distances pour affiner nos résultats.

## 4 PERSPECTIVES D'AMELIORATION DE LA METHODE

### 4.1 *L'augmentation du nombre de contraintes de distances*

#### 4.1.1 Les agents réticulants immobilisés

Nous avons vu qu'il nous manquait des informations de distance pour affiner nos résultats. Ce manque est dû en grande partie à l'impossibilité d'identifier tous les produits de réticulation. Le signal des peptides réticulé étant très faible comparé à celui des peptides non réticulés, leur détection est difficile et leur fragmentation pas toujours efficace. Différentes méthodes ont été développées pour pallier ce problème. Nous avons choisi d'explorer plus particulièrement celle qui consistera à séparer physiquement les peptides pontés des peptides non modifiés.

Le principe de la technique que nous tentons de développer est d'utiliser des agents réticulants fixés qui permettraient de ne retenir que les peptides d'intérêt contenant des modifications. L'agent utilisé est alors un agent trifonctionnel avec deux fonctions réactives, un espaceur, et un système de purification sur le troisième bras. Ce système de purification est alors séparé de la partie réactive par un site de clivage et un bras d'éloignement (figure 62).

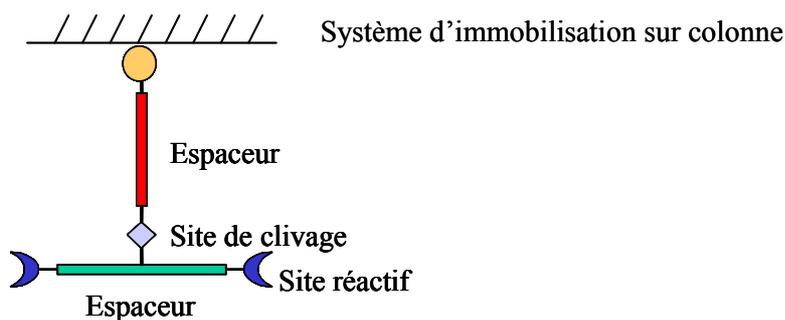
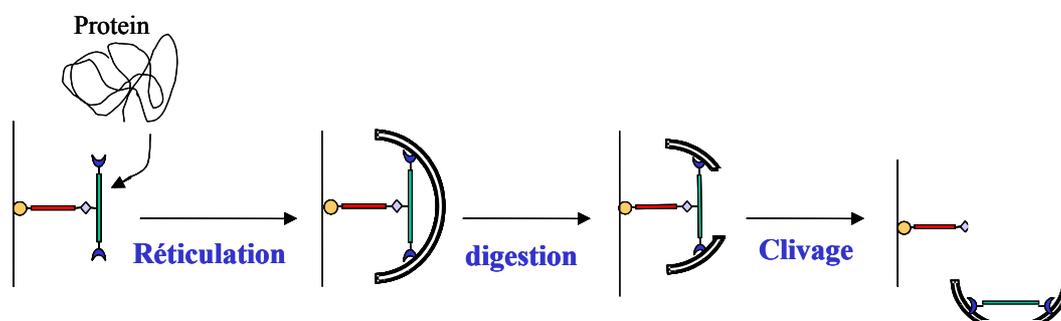


Figure 62: Schéma d'un agent réticulant immobilisé.

Le principe de fonctionnement de ce type d'agent est représenté figure 63. L'agent est immobilisé. La réaction de réticulation est effectuée, les protéines qui n'ont pas réagi sont éluées. L'étape de digestion a lieu, les peptides non fixés à un agent réticulant sont élués. Le clivage spécifique de l'agent réticulé permet ensuite de récupérer les peptides d'intérêt.



**Figure 63 : Principe d'utilisation des agents réticulants immobilisés.**

Différents essais de synthèse et d'utilisation de ce type d'agent sont en cours au Laboratoire des Aminoacides Peptides et Protéines de Montpellier. Pour l'instant, le fort encombrement stérique au voisinage de l'agent serait critique pour la fixation des protéines à réticuler. Différents essais tendent à augmenter la distance du bras d'éloignement entre le site de fixation de l'agent et les fonctions réactives.

#### **4.1.2 Utilisation d'autres agents et d'autres enzymes**

Pour obtenir de nouvelles contraintes de distances, il pourrait être intéressant d'utiliser des agents réticulants qui ne réagissent pas que sur les lysines. L'EDC, par exemple, qui crée une liaison covalente entre un acide carboxylique et une fonction amine pourrait être un candidat très intéressant. Sa capacité à cibler des acides aminés acides peut être notamment très utile pour des protéines qui possèdent peu de lysines. De plus, la longueur « nulle » de son bras espaceur permet d'envisager des applications sur des protéines seules, où il pourra atteindre des acides aminés du cœur de la protéine, mais aussi dans des complexes protéiques, en vue d'une caractérisation du site d'interaction.

Nous avons également envisagé d'utiliser d'autres protéases pour obtenir des peptides d'intérêt plus petits. Leur plus faible poids moléculaire permettrait une analyse en spectrométrie de masse plus aisée et plus précise et les informations pour un même pont seraient multipliées.

## ***4.2 L'amélioration de la précision des contraintes de distances***

Nous l'avons vu, les contraintes de distances que nous pouvons déterminer par cette méthode sont assez larges. Ceci est dû à la fois à la flexibilité des chaînes latérales des lysines ciblées et à la flexibilité des agents réticulants que nous utilisons.

Nous avons songé utiliser d'autres agents réticulants moins flexibles. C'est le cas des agents photo-activables par exemple. Cependant, la non-spécificité de ces agents, si elle nous permettrait de cibler des acides aminés moins flexibles, rendrait l'interprétation des résultats très complexe et la localisation exacte de la réticulation serait difficile à déterminer.

Nous essayons donc plutôt, en collaboration avec les organiciens du Laboratoire des Aminoacides Peptides et Protéines de Montpellier, de synthétiser de nouveaux agents qui cibleraient toujours les lysines mais qui auraient un bras espaceur rigide. Nous pensons notamment utiliser des chaînes de type poly-proline. Ainsi, les distances obtenues avec ce type d'agent seraient plus précises puisqu'elles ne dépendraient plus que de la flexibilité des lysines.

## ***4.3 L'amélioration du logiciel de traitement bioinformatique.***

Pour faciliter l'analyse bioinformatique des résultats de réticulation, il peut être intéressant de créer de nouveaux filtres. En collaboration avec l'IBCP, nous tentant actuellement d'incorporer des données issues d'expériences de d'échange H/D associé à la spectrométrie de masse. En effet, les informations données par cette technique peuvent indiquer les parties les plus accessibles et les moins structurées de la protéine.

Plutôt que d'intégrer ses résultats dans les filtres, nous pensons nous en servir pour discriminer les solutions finales obtenues. Ainsi les résultats de deutération peuvent être essayés sur les différents modèles tridimensionnels proposés.

## 5 CONCLUSION

La méthode des réticulations intramoléculaires associées à la spectrométrie de masse, s'avère donc être une technique très prometteuse pour la détermination de la famille de repliement d'une protéine. Les résultats obtenus sur le cytochrome c nous ont permis d'obtenir de 6 à 16 contraintes de distance. Ces contraintes de distances ont pu être utilisées pour mettre au point une méthode de détermination bioinformatique de la famille de repliement d'une protéine de structure inconnue à partir d'expériences de réticulations intramoléculaires associées à la spectrométrie de masse. Certains progrès restent encore à faire au niveau de la réaction en elle-même, de la détection des produits par spectrométrie de masse, des logiciels informatiques permettant l'identification des produits de réticulation et des programmes bioinformatiques qui permettent la détermination de la famille de repliement. Cette étude nécessite donc l'interaction de différentes équipes : massistes, informaticiens et organiciens.

La modélisation informatique des données structurale est un excellent outil pour éviter l'étude longue et fastidieuse de la structure des protéines par les méthodes classiques. Cependant, il faut toujours garder à l'esprit que les résultats obtenus sont purement théoriques et qu'un retour à l'expérience est absolument indispensable. C'est dans ce cadre que la technique des réticulations chimique associées à la spectrométrie de masse pourrait jouer un rôle essentiel.

Cette méthode, même si elle peut encore être améliorée, peut déjà être appliquée à des problématiques biologiques dans le but d'obtenir des informations structurales. En association avec la méthode des échanges H/D couplés à la spectrométrie de masse, elle peut s'avérer être un outil très intéressant pour obtenir des informations structurales sur des protéines.

Nous allons donc utiliser ces deux méthodes pour l'étude structurale de deux protéines bactériennes : PcrV et PcrG et du complexe qu'elles forment.

## CHAPITRE 4 :

ETUDES STRUCTURALES DES  
PROTEINES PCR V ET PCR G ET  
CARACTERISATION DU  
COMPLEXE PCR V-PCR G

## **INTRODUCTION**

*Pseudomonas aeruginosa* est un pathogène opportuniste qui utilise le système de sécrétion de type III (TTSS) pour injecter des toxines dans le cytoplasme des cellules eucaryotes. PcrV, une composante de ce système, est un élément essentiel pour la cytotoxicité de la bactérie et les anticorps dirigés contre cette protéine augmentent la survie des souris infectées (Sawa et al., 1999 ; Frank et al., 2002 ; Shime et al., 2001). *In vitro*, cette protéine forme un complexe avec la protéine PcrG qui est impliqué dans la régulation du TTSS (Nanao et al., 2003). Les seules informations structurales disponibles sur ces protéines et sur le complexe qu'elles forment sont dérivées des données obtenues sur leurs protéines homologues chez *Yersinia* : LcrV et LcrG. Une meilleure connaissance de la structure de PcrV et de PcrG et une caractérisation du complexe permettraient de mieux appréhender leur rôle au sein du TTSS.

Nous allons appliquer les méthodes précédemment décrites, les échanges H/D et la réticulation chimique associés à la spectrométrie de masse, aux deux protéines et au complexe.

Les échanges H/D et la réticulation de chacune des protéines seules, nous donneront des informations sur leur structure propre. Les données obtenues sur PcrV permettront d'affiner la structure modélisée par bioinformatique.

La comparaison des taux de deutération des protéines seules et des protéines impliquées dans le complexe permettra la caractérisation de l'interaction.

# **1 LE CONTEXTE BIOLOGIQUE DE L'ETUDE**

## **1.1 La bactérie *Pseudomonas aeruginosa***

*Pseudomonas aeruginosa* est bacille à Gram négatif. C'est une bactérie ubiquitaire qui nécessite un environnement humide. Pathogène opportuniste, il est peu virulent chez l'individu normal mais très pathogène chez le patient immunodéprimé (sidéens, patients sous chimiothérapie, grand brûlés, leucémiques...) (Lyczak et al., 2000). Chez la personne normale, *P. aeruginosa* ne se retrouve guère que dans des otites dont la cause est souvent la contamination massive par l'eau (piscine, jacuzzi...). C'est chez le patient immunodéprimé que *P. aeruginosa* provoque les infections les plus graves. 10 à 20 % des infections nosocomiales lui sont attribuées. Chez le patient mucoviscidose, *P. aeruginosa* est la cause principale des infections pulmonaires chroniques. Il colonise les poumons des individus atteints de mucoviscidose sous forme de biofilm (Lyczak et al., 2002).

*P. aeruginosa* possède une résistance naturelle à de nombreux antibiotiques. Sa sensibilité aux antibiotiques réputés actifs est modérée. Il acquiert d'autre part de multiples mécanismes de résistances vis à vis des molécules actives (Lyczak et al., 2000, 2002).

Cette bactérie possède un large arsenal de facteurs de virulence : adhésines, toxines, pili et système de sécrétion. La pathogénicité de *P. aeruginosa* est conférée en grande partie par le système de sécrétion de type III (TTSS). Ce système permet en effet injection directe des toxines dans la cellule hôte infectée (Hueck 1998).

## **1.2 Le système de sécrétion de type III**

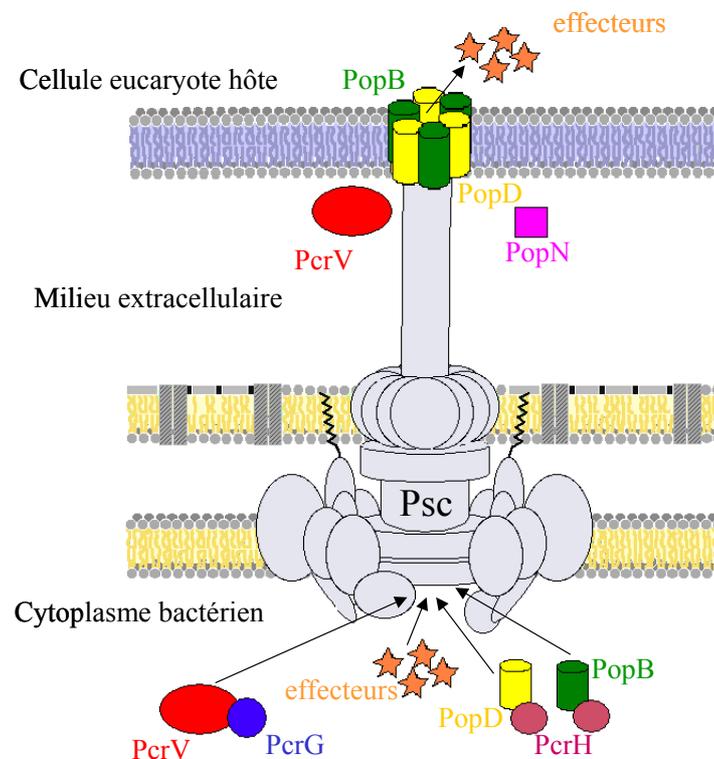
Le système de sécrétion de type III (TTSS) permet le passage des toxines de la bactérie jusque dans le cytoplasme de la cellule eucaryote infectée. La traversée des deux membranes de la bactérie et de la membrane de la cellule hôte se fait en une seule étape énergétique.

Il est important de bien différencier la sécrétion, c'est à dire le passage de l'enveloppe bactérienne, de la translocation, qui est le passage de la membrane eucaryote.

Le TTSS est présent dans de nombreuses bactéries Gram négatives qu'elles soient pathogènes de végétaux ou d'animaux. Il est très conservé d'une espèce à l'autre même si les toxines transloquées sont différentes (Galan et Collmer, 1999).

Chez *Pseudomonas aeruginosa*, le TTSS est codé sur un seul locus du chromosome par cinq opérons corégulés par un activateur de transcription central : ExsA (Dacheux et al., 2001 ; Frank, 1997).

Plus d'une trentaine de protéines interviennent dans ce système. Elles peuvent être classées en trois groupes selon leur fonction. Les protéines injectées sont les effecteurs, elles sont associées à des protéines chaperonnes. Les protéines du secreton vont permettre le passage à travers l'enveloppe bactérienne. Les protéines du translocon assurent la translocation des effecteurs dans le cytoplasme de la cellule hôte (figure 64) (Hueck, 1998).



**Figure 64: Schéma de fonctionnement du système de sécrétion de type III de *Pseudomonas aeruginosa*.**

Les protéines sont classées en trois groupes : les effecteurs, le secreton (constitués des protéines Psc en gris) et le translocon (PcrV, PcrG, PopB, PopD, PcrH, PopN).

### 1.2.1 Les molécules effectrices

Quatre toxines cytosoliques ont été identifiées chez *Pseudomonas aeruginosa* comme effecteurs du TTSS : ExoS, ExoT, ExoY et ExoU.

ExoS et ExoT possèdent une activité ADP-ribosyltransférase (Frank, 1997 ; Liu et al., 1997) et peuvent activer les GTPases contre les protéines liant le GTP, comme Rho et Ras (Goehring et al., 1999 ; Krall et al., 2000). Elles vont ainsi affecter les voies de transduction du signal de la cellule hôte.

Exo Y possède une activité adénylate cyclase (Yahr et al., 1998) et ExoU est une cytotoxine qui entraîne la mort cellulaire des cellules épithéliales et des macrophages (Finck-Barbancon et al., 1997).

Ces toxines vont provoquer des perturbations dans le cytosquelette d'actine, initier les inflammations et l'apoptose, ce qui pourra conduire à la mort cellulaire.

Le système de sécrétion de type III requiert des petites protéines chaperonnes (Wattiau et al., 1994). Ces protéines, souvent codées à proximité du gène codant pour leur cible, sont spécifiques d'une seule protéine cible. Elles vont permettre la stabilisation des protéines sécrétées avant la phase de sécrétion et ainsi éviter toutes les interactions prématurées (Woestyn et al., 1996). Elles ont également un rôle dans la sécrétion et pourraient servir de guides (Cheng et al., 1997). Enfin, elles pourraient intervenir dans la hiérarchisation de la sécrétion lorsque plusieurs protéines sont sécrétées simultanément (Boyd et al., 2000 ; Cornelis, 2000).

Aucun signal d'adressage n'a pu être observé sur la partie N-terminale des différentes toxines. L'adressage pourrait être assuré par la partie 5' de l'ARNm des protéines, même si ces parties sont peu similaires d'une toxine à l'autre (Anderson et Schneewind, 1997 ; Sory et al., 1995 ; Ghosh, 2004).

### 1.2.2 Le secreton

Le secreton est constitué d'une vingtaine de protéines, appelée Psc chez *Pseudomonas aeruginosa* (Cornelis, 2000). La plupart de ces protéines sont très conservées dans tous les TTSS quelle que soit l'espèce concernée. Ces protéines ont également de grandes homologies de séquence avec les protéines qui forment les flagelles bactériens (Ghosh, 2004). Soit les

gènes codant pour le TTSS sont issus de la duplication de certains gènes des flagelles (Nguyen et al., 2000 ; Saier et al., 2004), soit les deux systèmes dérivent d'un ancêtre génétique commun (Gophna et al., 2003)

Le sécrétion est composé de protéines ancrées dans les membranes bactériennes et de protéines formant une aiguille. Ce système, qui pourrait être comparé à une seringue, permettra l'injection des protéines à sécréter du cytoplasme bactérien jusqu'à l'espace extracellulaire.

### 1.2.3 Le translocon

Les protéines regroupées sous le terme translocon vont participer à la translocation, c'est à dire au passage des toxines du milieu extracellulaire au cytoplasme de la cellule hôte. Chez *Pseudomonas aeruginosa*, les protéines impliquées dans cette étape sont PopB, PopD, PopN, PcrG, PcrV, PcrH. Elles sont toutes, sauf PopN, codées par un opéron : l'opéron *pcrGVH-popBD* (Yahr et al., 1997).

#### 1.2.3.1 L'étape de translocation

Les mutants pour les gènes *popB*, *popD* et *pcrV* sont incapables d'infecter les cellules eucaryotes (Sawa et al., 1999). Ces trois protéines sont donc indispensables dans la translocation des toxines. Les protéines PopB et PopD possèdent des domaines transmembranaires. De plus, *in vitro*, le complexe recombinant PopB/PopD est capable de lier des membranes artificielles et d'y induire des perturbations (Schoehn et al., 2003). Ces deux protéines seraient impliquées dans la formation d'un pore au sein de la membrane eucaryote.

La protéine PcrH est une protéine chaperonne de PopB et PopD. Elle permet de masquer leurs régions transmembranaires pour éviter une interaction prématurée ou une agrégation. Elle permet également de les maintenir dans une conformation inactive vis à vis de la liaison aux membranes tant que leur activation n'est pas requise (Schoehn et al., 2003).

La protéine PcrV est également indispensable à la formation du pore, mais elle n'interviendrait pas directement dans le cœur de ce pore. Elle est capable d'interagir avec la protéine PopD, mais pas avec la protéine PopB. Elle est requise pour l'ancrage de PopD sur les membranes (Goure et al., 2004).

### 1.2.3.2 La régulation du TTSS

Les protéines PcrG et PopN, si elles ne sont pas requises pour la translocation, sont impliquées dans la régulation négative du système (Sundin et al., 2004). *In vitro*, les mutants pour les gènes *pcrV*, *pcrG* et *popN* présentent une expression de la toxine ExoS bien supérieure à celle retrouvée dans les souches sauvages. De plus, cette surexpression est observée chez les mutants *popN*, que le signal d'induction soit présent ou non. En fait, l'étape de transduction est induite *in vivo* par le contact cellulaire entre bactérie et cellule eucaryote, et *in vitro* par une diminution de la concentration en  $Ca^{2+}$ . La façon dont le système de sécrétion de type III est activée reste à ce jour inconnue, cependant elle pourrait mettre en jeu un senseur bactérien. PopN pourrait être ce senseur impliqué dans la reconnaissance du signal d'induction et transmettre ensuite un signal qui activerait l'expression des composants du TTSS (Sundin et al., 2004).

Les trois protéines PcrV, PcrG et PopN interviennent également dans la polarisation du mécanisme de translocation. En effet, lorsque des cellules sont infectées par une souche sauvage de *P. aeruginosa*, aucune toxine n'est détectée dans le milieu de culture. Inversement, après infection par des souches mutées pour les gènes *pcrV*, *pcrG* ou *popN*, des toxines sont retrouvées dans le milieu (Sundin et al., 2004). Ainsi, PopN et PcrG préviendrait la sécrétion des toxines dans le milieu extracellulaire, l'une depuis l'intérieur de la bactérie et l'autre depuis l'extérieur, et PcrV aurait un rôle dans le maintien des effecteurs dans la zone de contact bactérie/cellule hôte.

PcrG et PcrV peuvent également interagir dans le cytoplasme bactérien. Ce complexe a un rôle dans la régulation du TTSS (Nanao et al., 2003).

## 1.3 Les protéines PcrV et PcrG

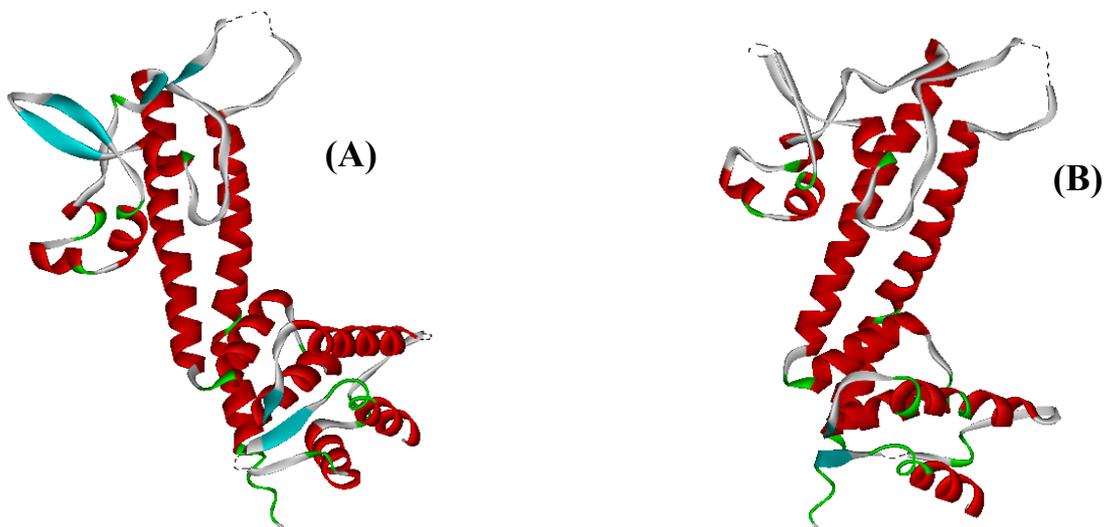
### 1.3.1 La protéine PcrV

Comme nous l'avons déjà vu, PcrV est une composante essentielle du système de sécrétion de type III. Les anticorps directement dirigés contre PcrV protègent les cellules en culture et les animaux modèles des infections par *P. aeruginosa*. L'immunisation passive ou active contre PcrV augmente la survie des animaux modèles infectés (Sawa et al., 1999 ; Frank et al., 2002 ;

Shime et al., 2001), notamment en désactivant l'effet anti-phagocytaire de PcrV vis à vis des macrophages.

PcrV est présente à la fois dans le cytoplasme des bactéries et dans le milieu extracellulaire. Elle agit aussi bien dans la translocation, en participant à la formation du pore, que dans la régulation du TTSS en interaction avec PcrG.

La structure de PcrV n'est pour l'instant pas disponible. Mais certaines informations peuvent être obtenues par comparaison avec la structure de son homologue chez *Yersinia* : LcrV. LcrV et PcrV possèdent en effet 42 % d'homologie de séquence. La structure de LcrV a été obtenue par cristallographie (Derewenda et al., 2004). Cette protéine possède une structure centrale en « coiled-coil » et deux domaines globulaires de part et d'autre de ces hélices (figure 65 A). Les domaines en « coiled-coil » sont retrouvés dans un grand nombre des protéines du système de sécrétion de type III. Cette super structure est un ensemble de plusieurs hélices, avec un profil répété tous les sept résidus, le premier et le quatrième résidus sont hydrophobes. Le tour d'hélice contient donc 3.5 résidus, à la différence des hélices  $\alpha$  classiques qui en contiennent 3.6. Dans la structure de LcrV, le coiled-coil concerne les hélices  $\alpha 4$  et  $\alpha 12$  et forme un « zipper » hydrophobe, faisant de ces zones des régions propices à des interactions hydrophobes.



**Figure 65 : Structures des protéines LcrV et PcrV.**

(A) : Structure de LcrV obtenue par cristallographie (Derewenda et al., 2004)

(B) : Structure de PcrV obtenue par modélisation par Manuelle Quinaud, LCM, IBS.

Une modélisation de la structure de PcrV a pu être réalisée par Manuelle Quinaud au Laboratoire de Cristallographie Macromoléculaire de l'IBS à partir de la structure de LcrV et avec l'aide des alignements de séquences, des prédictions de ponts salins et de liaisons hydrogènes, et des logiciels de prédiction de structures secondaires et d'étude du potentiel de surface. La protéine PcrV présenterait, comme son homologue LcrV, une partie en « coiled-coil » et deux parties globulaires (figure 65 B).

### 1.3.2 Le complexe PcrV-PcrG

PcrV et PcrG forment *in vitro* un complexe de stœchiométrie 1:1. Elles s'associent avec une cinétique rapide et une affinité de l'ordre du nanomolaire (Nanao et al., 2003).

Les protéines homologues de *Yersinia*, LcrV et LcrG, forment également un complexe. Sous les conditions induisant la sécrétion (diminution du calcium *in vitro* ou contact cellulaire *in vivo*), la protéine LcrV est surexprimée, elle lierait LcrG et bloquerait ainsi l'action répressive de cette dernière sur la sécrétion des molécules effectrice (Nilles et al., 1997 et 1998). Cependant un tel mode d'action peut difficilement être envisagé chez *P. aeruginosa*, puisque les mutants pour le gène *pcrV* sont capables de sécrétion sous conditions d'induction comme sous conditions de non-induction (Sawa et al., 1999). Ainsi même si les TTSS de *Yersinia* et de *P. aeruginosa* sont très proches, leur système de régulation semble différer. Cependant, à ce jour, la fonction exacte et le mode d'action du complexe PcrG-PcrV restent obscurs.

Une meilleure connaissance de la structure des deux protéines et une caractérisation de leur interaction permettrait de mieux appréhender leur mécanisme et leur rôle dans le système de sécrétion de type III. Dans cette étude, nous allons donc tenter d'obtenir des données structurales sur les protéines et de localiser les zones d'interaction du complexe. Pour cela, nous utiliserons les deux méthodes développées dans les chapitres précédents. Des expériences d'échange H/D seront réalisées sur chacune des protéines. Des réticulations chimiques intramoléculaires seront effectuées sur PcrV. Et enfin, des expériences d'échange H/D sur le complexe permettront de mieux caractériser l'interaction.

## **2 RESULTATS ET DISCUSSION**

### **2.1 *Etudes structurales des protéines PcrV et PcrG seules***

#### **2.1.1 La protéine PcrV**

##### *2.1.1.1 Deutération de la protéine PcrV*

La protéine PcrV est digérée par la pepsine ou par la protéase de type XIII ou par la protéase de type XVIII. Les cartes peptidiques sont réalisées pour chaque enzyme par des expériences de LC-MS-MS comme décrit dans le chapitre 2.

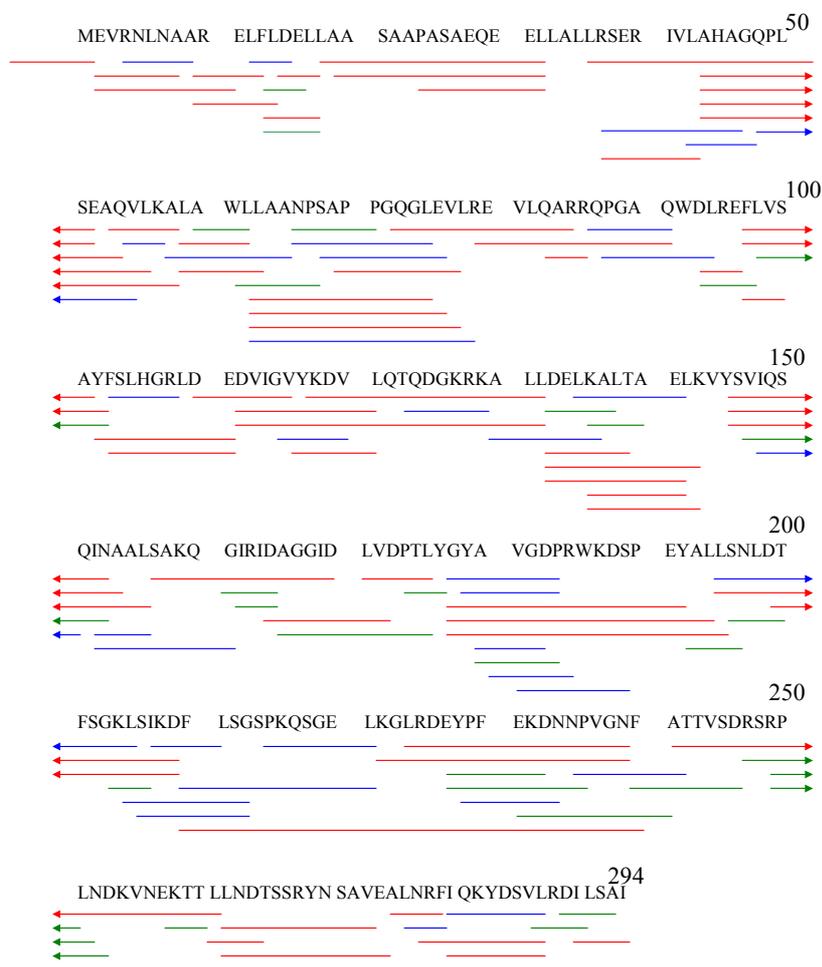
Après digestion à la pepsine 66 peptides sont identifiés, 29 peptides le sont après digestion avec la protéase de type XIII, et 35 peptides sont obtenus par la digestion avec la protéase de type XVIII. Ces peptides permettent un recouvrement de la séquence de PcrV de 98 %, 40 % et 59 %, respectivement pour la pepsine, la protéase de type XIII, et la protéase de type XVIII. L'utilisation combinée des trois enzymes nous permet donc au final de recouvrir 98% de la séquence primaire de PcrV (figure 66).

En fait, les analyses ont été réalisées sur des protéines qui possédaient un his-tag. C'est pourquoi, sur toutes les figures représentant les séquences de PcrV (et de PcrG), certains peptides sont symbolisés comme commençant avant l'acide aminé N-terminal des protéines, ils représentent les peptides comprenant une partie de l'his-tag et une partie de la protéine.

Des expériences de LC-MS sont réalisées en parallèle sur la protéine PcrV dans un tampon deutéré et dans un tampon hydrogéné pour chacune des trois enzymes. Le temps de deutération est de 30 secondes. Les profils d'élution des peptides issus de la digestion de la protéine non deutérée sont les mêmes que ceux obtenus par la digestion de la protéine deutérée.

Le nombre de deutériums incorporés est recherché pour chacun des peptides qui avaient été identifiés lors de la réalisation des cartes peptidiques. Nous obtenons des informations de deutération pour 15 des peptides générés par la protéase de type XIII (soit 30 % des hydrogènes amidiques), pour 33 peptides générés par la protéase de type XVIII (soit 53 % des

hydrogènes amidiques) et pour 56 peptides générés par la pepsine (83 % des hydrogènes amidiques).

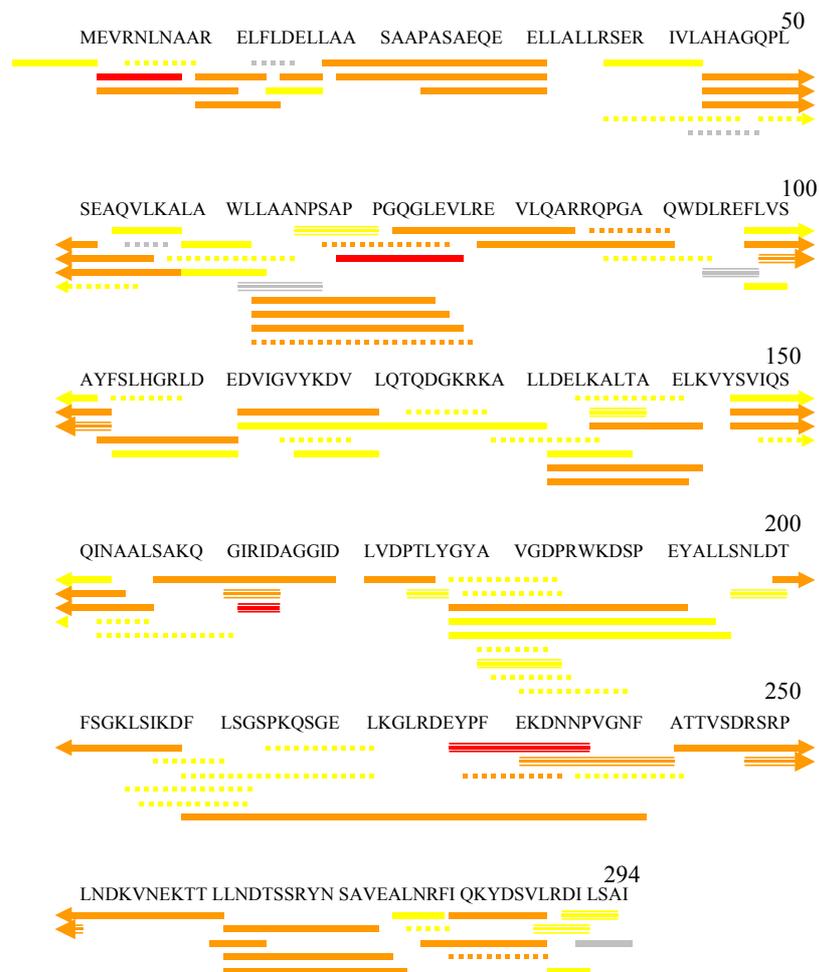


**Figure 66 : Recouvrement de la séquence de PcrV.**

Les peptides obtenus avec chacune des trois protéases sont représentés par des traits sous la séquence de PcrV. Les peptides obtenus avec la pepsine, la protéase de type XIII et la protéase de type XVIII sont représentés, respectivement, en rouge, vert et bleu.

L'utilisation combinée des trois enzymes nous permet d'obtenir des informations sur la deutération de 96 % des hydrogènes amidiques de la protéine PcrV. Une fois encore l'utilisation d'enzymes multiples nous permet d'augmenter le pourcentage de la séquence pour lequel il est possible d'avoir des informations de deutération (figure 67 et annexe 2).

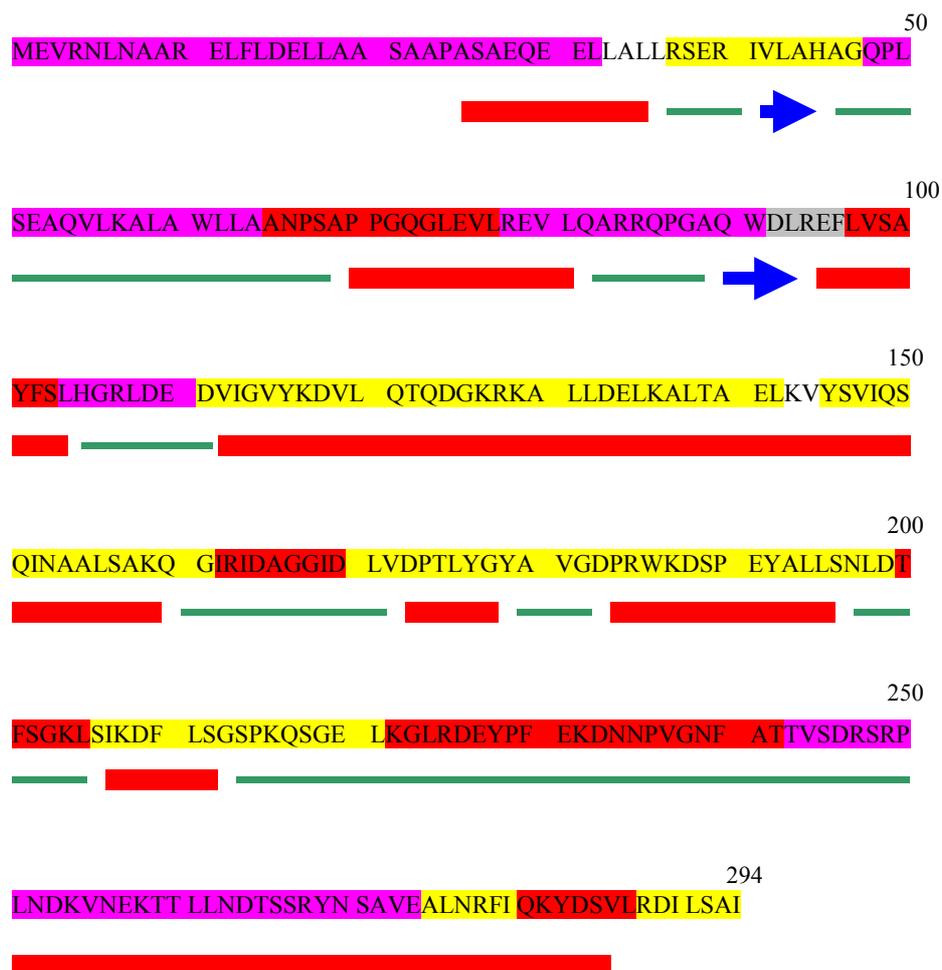
Le recouplement des informations obtenues par chacune des enzymes, et notamment le recouplement des taux de deutération des peptides chevauchant, nous permet de déterminer des taux d'échanges locaux pour différentes régions de la protéine (figure 68).



**Figure 67 : Incorporation de deutérium (après 30 secondes de deutération) pour chaque peptide obtenu avec les trois protéases.**

Les peptides obtenus avec la pepsine, la protéase de type XIII et la protéase de type XVIII sont représentées respectivement par des barres continues, des triples soulignements et des barres pointillées. La couleur des barres indique le taux de deutération, les barres rouges, orange, jaunes et gris indiquent respectivement un taux de deutération supérieur à 50%, compris entre 50 et 30%, compris entre 30 et 10% et inférieur à 10%.

Le taux d'échange global de la protéine apparaît comme plutôt élevé. En fait, d'après le modèle de structure et la structure de la protéine homologue LcrV, PcrV serait accessible au solvant. Aucune partie de la protéine n'est réellement enfouie, ainsi, les différences de deutération dans certaines zones de la protéine vont plutôt refléter l'implication des acides aminés de ces zones dans les structures secondaires, le paramètre accessibilité sera ici moins important dans les différences.



**Figure 68 :** Représentation du taux de deutération sur la séquence primaire de PcrV et structures secondaires présumées.

Les régions de la protéine sont surlignées en rouge, rose, jaune et gris en fonction de leur taux de deutération, respectivement, supérieur à 50%, compris entre 50 et 30%, compris entre 30 et 10%, et inférieur à 10%.

Structures secondaires présumées : les hélices  $\alpha$  sont représentées par des rectangles rouges, les feuillets  $\beta$  par des flèches bleues et les boucles par des traits verts.

Le recouplement des informations de deutération dans les peptides chevauchant va nous permettre de déterminer des taux de deutération locaux pour différentes zones. Ces zones sont comparées aux structures secondaires présumées issues de la structure modélisées de PcrV (figure 68).

La délimitation des différentes zones pour lesquelles nous avons déterminé un taux de deutération est seulement due aux sites de coupures des différentes enzymes. Ainsi, son caractère imprévisible ne va en aucun cas refléter des délimitations de structures secondaires. Une zone, pour laquelle nous avons déterminé une incorporation de deutérium global, peut en fait contenir plusieurs éléments de structure secondaire entiers ou une partie d'entre eux seulement. De plus, la modélisation de la structure, si elle permet d'obtenir des informations globales, ne délimite pas non plus très précisément les différents éléments de structure secondaire. Ainsi, il va être impossible de réaliser une corrélation exacte entre éléments de structures secondaires et taux de deutération des différentes zones.

D'après les résultats de deutération obtenus, nous allons tout de même pouvoir faire une corrélation globale avec la structure modèle.

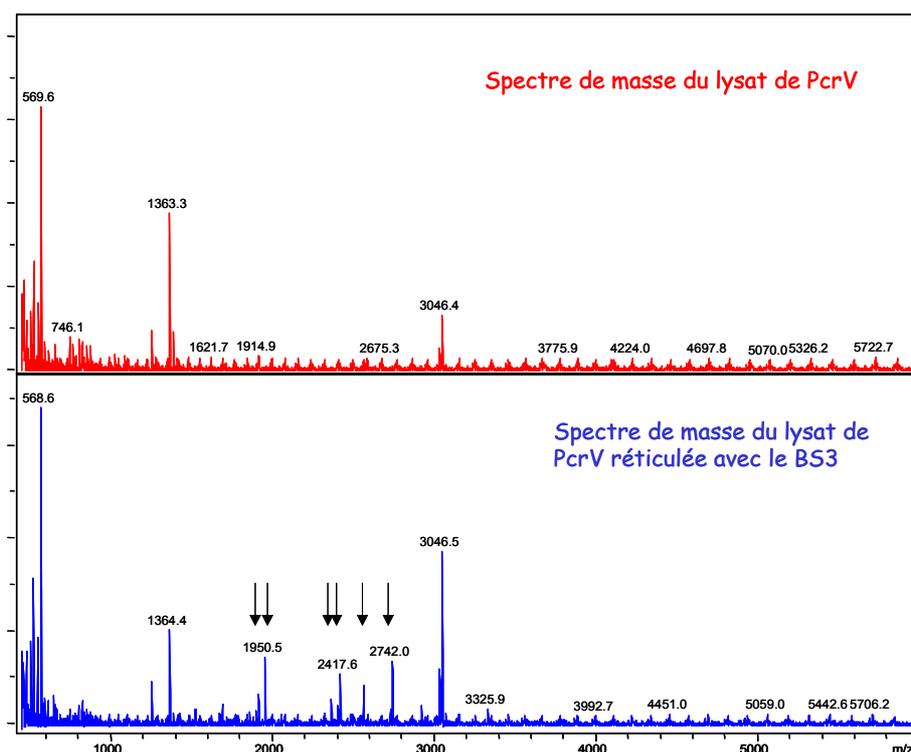
Tout d'abord, l'agencement des différentes parties de PcrV semble bien être celui attendu. Les zones correspondant aux deux hélices impliquées dans le « coiled-coil » (110-150 et 250-285) incorporent relativement peu de deutériums, indiquant bien la présence de structures secondaires dans ces zones. En dehors de ces deux structures, nous pouvons observer une alternance de zones peu deutérées et de zone très deutérées. Ceci est en adéquation avec la présence de deux petites parties globulaires de part et d'autre des structures en « coiled-coil ». Dans ces zones globulaires, des régions structurées (en hélice  $\alpha$  ou en feuillet  $\beta$ ) alterneraient avec des zones moins structurées. Même si la délimitation exacte de ces zones n'est pas possible, aucune contradiction entre taux de deutération et structure secondaire présumée n'est observée. Par exemple, aucune des hélices présumées ne se trouve en totalité dans une région de la protéine identifiée comme ayant un fort taux de deutération. De la même façon, aucune des boucles non structurées ne se trouve totalement dans une zone qui ne s'échange pas ou très peu.

Ainsi, même si nous ne pouvons pas valider totalement le modèle de la structure tridimensionnelle de PcrV, nous pouvons affirmer qu'il représente bien globalement la réalité. Pour confirmer encore un peu plus ce modèle, nous avons réalisé des expériences de réticulations intramoléculaires en association avec la spectrométrie de masse.

### 2.1.1.2 Réticulation chimique intramoléculaire de la protéine PcrV

La protéine PcrV possède seize lysines réparties sur l'ensemble de sa séquence primaire. Elle fait donc un bon candidat à la réticulation intramoléculaire par les agents réticulants de type NHS-ester.

Des réactions de réticulation sont réalisées avec le BS3, avec le Sulfo-DST et avec le Sulfo-EGS pendant 15 minutes. Les rapports de concentration entre lysine de PcrV et agent réticulant sont de 1.6 pour le BS3, 1 pour le Sulfo-EGS et 0.1 pour le Sulfo-DST. Les protéines réticulées sont ensuite digérées par l'endoprotéase Lys C pendant 18 heures. Les produits de réticulation sont identifiés par spectrométrie de masse MALDI-TOF, ou par LC-MS sur un Q-TOF. Ces produits sont identifiés grâce à leur masse par calcul informatique et les réticulations peuvent ainsi être localisées.



**Figure 69:** Spectres de masse MALDI-TOF des lysats de PcrV non réticulée et réticulée avec du BS3.

Les flèches indiquent des produits de réticulation.

La figure 69 montre la comparaison des spectres de masse MALDI du lysat de PcrV non modifiée et du lysat de PcrV réticulée par le BS3. Six ions sont présents sur le spectre des produits réticulés et pas sur celui des produits non modifiés. Ils permettent de repérer les produits de réticulation. L'analyse des mêmes échantillons par un spectromètre de masse de type Q-TOF, nous permet de trouver d'autres produits de réticulation. Au total pour le BS3, 19 produits de réticulation sont repérés.

De la même façon, 8 produits de réticulation par le Sulfo-EGS et 11 produits de réticulation par le Sulfo-DST sont repérés.

L'ensemble des masses relevées sur les spectres comme étant des masses de produits issus de la réticulation par l'un des agents sont analysés par un logiciel dont le fonctionnement est décrit dans le chapitre 3.

De cette manière, nous avons pu identifier 7 produits issus de la réticulation avec le BS3, 2 issus de celle avec le Sulfo-EGS et un issu de celle avec le Sulfo-DST.

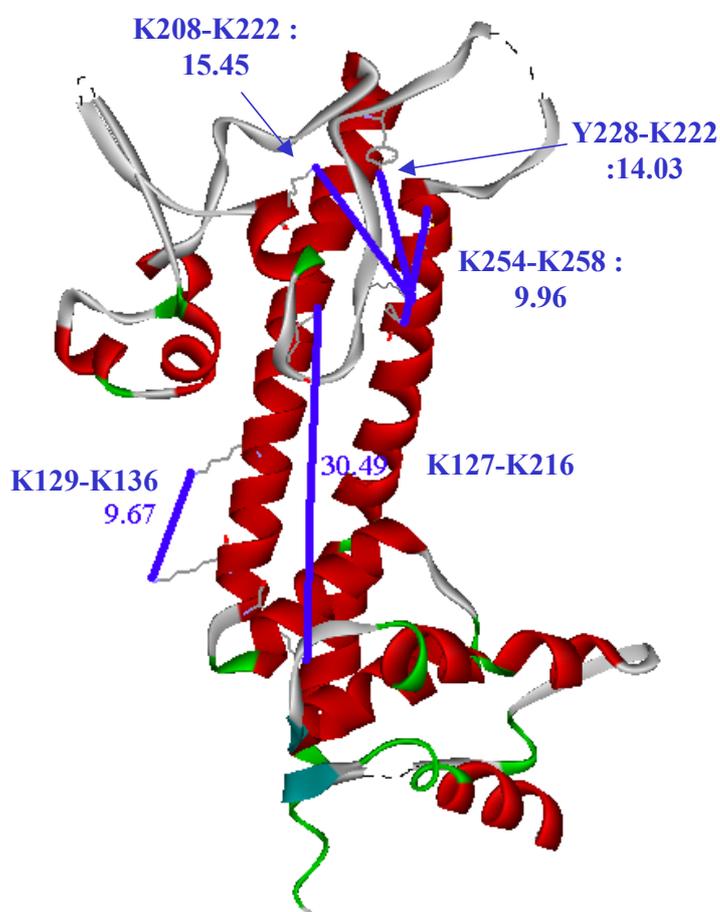
L'ensemble de ces peptides est répertorié sur le tableau 4.

	Masse du produit	type de réticulation	peptide(s) concerné(s)	lysine(s) modifiée(s)	Distance N-N sur la structure
BS3	3324.73	type 2	205-216 et 217-232	K208 et K222	15.4
	2416.34	type 2	205-216 et 217-225	K208 et K222	15.4
	2917.52	type 2	119-129 et 209-222	K127 et K216	30.5
	1443.78	type 0	119-129	K127	
	2052.02	type 0	217-232	K222	
	2416.34	type 0	205-225	K208 ou K216 ou K222	
	2034.017	type 1	217-232	K222 et Y228	14
Sulfo-EGS	2038.143	type 1	128-143	K129 et K136	9.7
	1531.759	type 0	119-129	K127	
Sulfo-DST	4089.9	type 1	233-268	K254 et K258	10

**Tableau 4 : Identification des produits de réticulation de PcrV par le BS3, le Sulfo-DST et le Sulfo-EGS.**

Tous les types de réticulation (0, 1 et 2) touchent des lysines communes (K127, K208, K222 et K116), ce qui nous assure de la validité des résultats obtenus sur ces lysines, même s'ils ne sont issus que de solutions mathématiques. Seules les informations sur les réticulations de type 1 et de type 2 pourront être utilisées pour affiner la structure tridimensionnelle modèle de PcrV. Nous pouvons remarquer qu'une tyrosine est touchée par la réticulation. Même si dans nos conditions elles sont moins réactives que les lysines, elles peuvent tout de même réagir avec les agents réticulants de type NHS-ester (Swaim et al., 2004 ; Leavelle et al., 2004). Les

ponts 208-222, 127-216, 222-228, 129-136 et 254-258 sont reportés sur la structure modèle (figure 70). Les distances indiquées sont les longueurs des ponts qui relieraient les azotes des chaînes latérales des acides aminés concernés.



**Figure 70: Représentation des ponts générés par la réticulation chimique sur la structure modèle de PcrV.**

Les distances indiquées sont les distances entre les groupements réactifs des acides aminés.

Le pont 208-222 relie des lysines qui seraient à une distance de 15.45 Å et le pont 222-228 relie des acides aminés distants de 14.03 Å. La longueur du BS3 donnée par le fabricant est de 13 Å. Si on considère la flexibilité des chaînes latérales des lysines (mobiles dans un rayon de 6 Å) et des tyrosines (mobiles dans un rayon de 6 Å), le BS3 peut théoriquement lier des lysines distantes de moins de 25 Å. La structure modèle est tout à fait compatible avec les résultats expérimentaux obtenus pour ces deux ponts.

De la même manière, le pont 129-136 réalisé par le Sulfo-EGS fait 9.67 Å, distance compatible avec l'agent utilisé qui peut théoriquement lier des lysines distantes de moins de 28.1 Å. Le pont 254-258 réalisé par le Sulfo DST fait 9.95 Å, ce qui correspond bien à la longueur maximale de l'agent (6.4 Å) plus ou moins l'apport de flexibilité des deux chaînes latérales.

Par contre, le pont entre les lysines 127 et 216 fait 30.49 Å. Or ce pont, réalisé par le BS3, devrait mesurer moins de 25 Å. La distance entre ces deux lysine a donc probablement été surévaluée lors de la réalisation de la structure modèle. Nous pouvons supposer que les deux parties globulaires de la protéine sont en fait plus proches l'une de l'autre, « refermées » sur les hélices en coil-coiled.

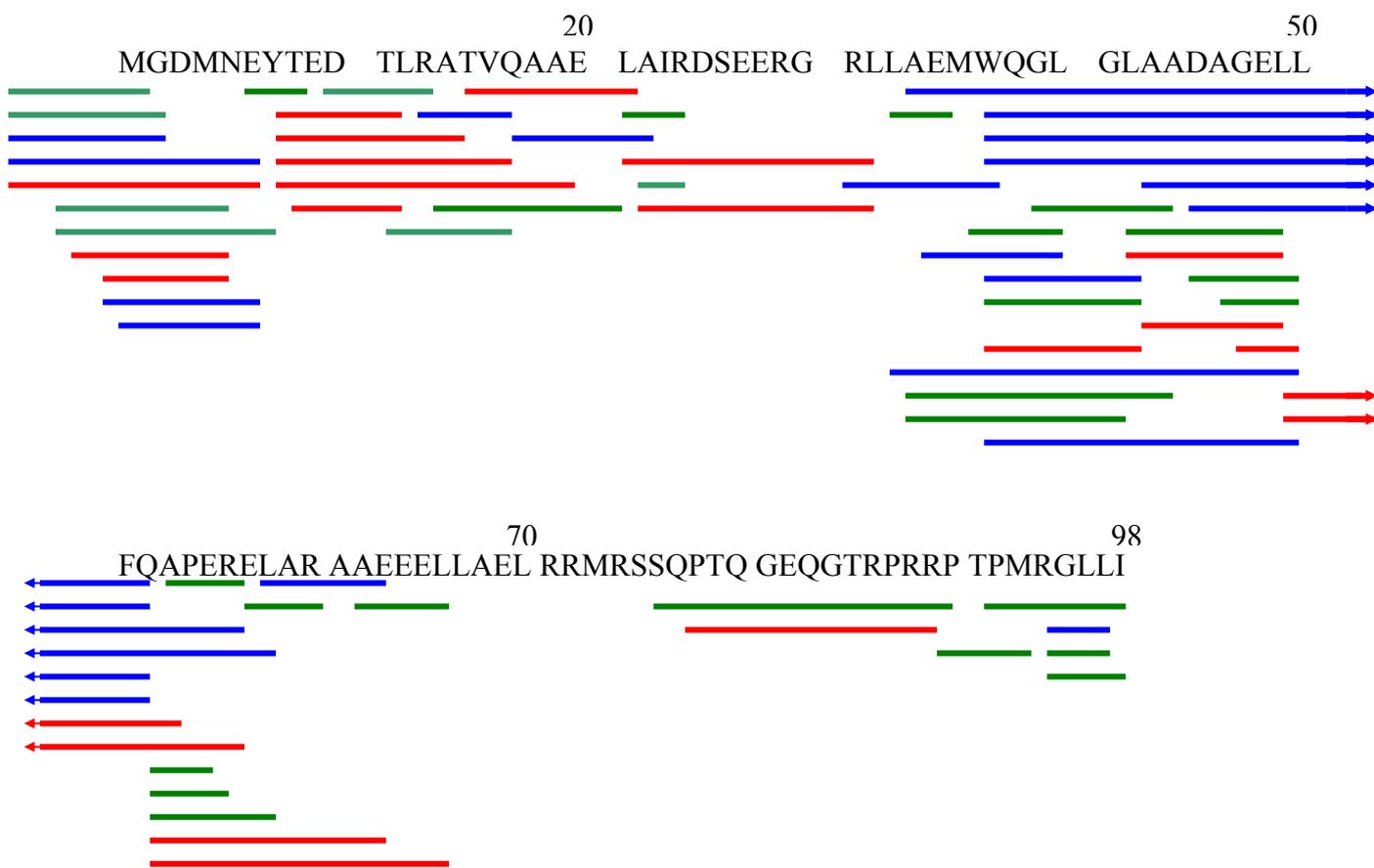
Ces résultats, après avoir été vérifiés par MS-MS, seront transmis aux informaticiens qui modéliseront à nouveau la structure de PcrV en prenant en compte les nouvelles contraintes de distance définies dans cette étude.

### **2.1.2 Deutération de la protéine PcrG**

La protéine PcrG est à son tour digérée par la pepsine ou par la protéase de type XIII ou par la protéase de type XVIII. Des cartes peptidiques sont réalisées pour chaque enzyme par des expériences de LC-MS-MS comme décrit dans le chapitre 2.

La digestion avec la pepsine, nous permet d'identifier 21 peptides qui recouvrent 80 % de la séquence primaire de PcrG. Après digestion avec la protéase de type XIII, nous pouvons identifier 29 peptides qui recouvrent 81 % de la séquence. Et 19 peptides (soit 54 % de la séquence) sont identifiés après digestion avec la protéase de type XVIII. Le recoupement des résultats obtenus avec les trois enzymes permet un recouvrement de 91 % de la séquence en acides aminés de PcrG (figure 71).

Nous pouvons déjà remarquer qu'il nous manque une partie relativement importante de la séquence. En fait certaines études ont montré que cette protéine était très peu structurée et que ses digestions généraient des fragments très petits, ne comprenant parfois qu'un ou deux acides aminés. Cela explique donc le fait qu'aucun peptide faisant partie de la zone 67-75 n'a pu être identifié.

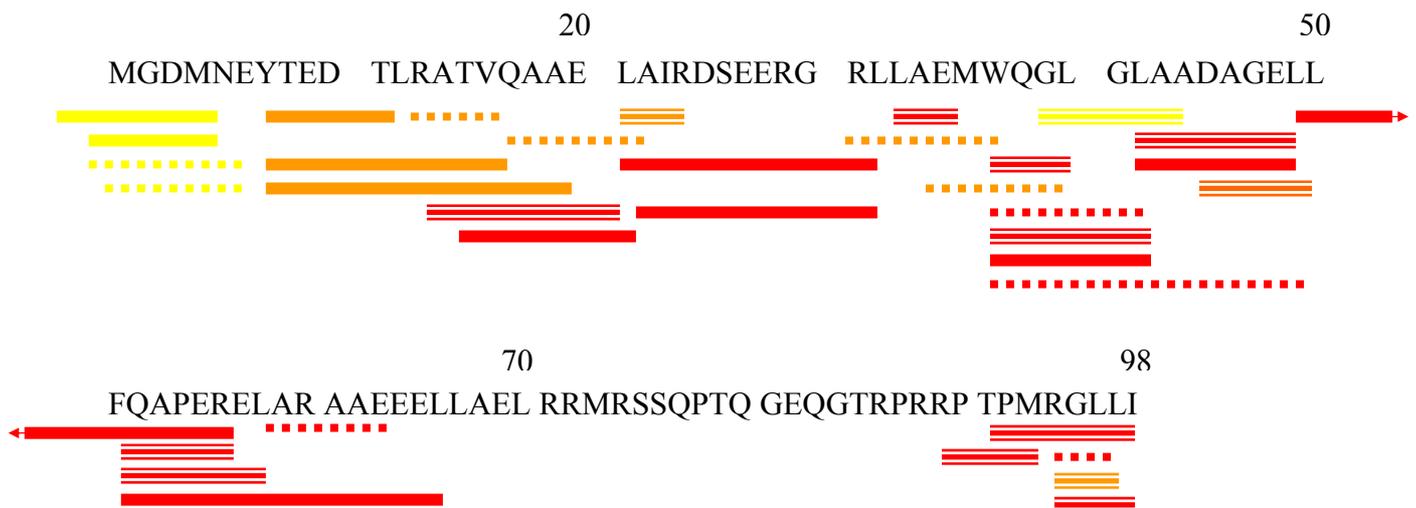


**Figure 71 : Recouvrement de la séquence de PcrG.**

Les peptides obtenus avec chacune des trois protéases sont représentés par des traits sous la séquence de PcrG. Les peptides obtenus avec la pepsine, la protéase de type XIII et la protéase de type XVIII sont représentés, respectivement, en rouge, vert et bleu.

Les expériences de deutération sont réalisées avec un temps d'échange de 30 secondes. Les protéines deutérées ou non, sont ensuite digérées avec l'une des trois enzymes et les peptides sont analysés par LC-MS.

Le nombre de deutériums incorporés est recherché pour chacun des peptides qui avaient été identifiés lors de la réalisation des cartes peptidiques. Nous obtenons des informations de deutération pour 13 des peptides générés par la protéase de type XIII (soit 45 % des hydrogènes amidiques), pour 10 peptides générés par la protéase de type XVIII (soit 46 % des hydrogènes amidiques) et pour 13 peptides générés par la pepsine (65 % des hydrogènes amidiques) (figure 72 et annexe 3).



**Figure 72 : Incorporation de deutérium (après 30 secondes de deutération) pour chaque peptide obtenu avec les trois protéases.**

Les peptides obtenus avec la pepsine, la protéase de type XIII et la protéase de type XVIII sont représentées respectivement par des barres continues, des triples soulignements et des barres pointillées. La couleur des barres indique le taux de deutération, les barres rouges, orange et jaunes indiquent respectivement un taux de deutération supérieur à 50%, compris entre 50 et 30% et compris entre 30 et 10 %.

Les recouvrements obtenus avec chacune des enzymes sont plutôt moyens, mais une fois encore le recoupement des résultats va nous permettre d'augmenter le recouvrement. Ainsi, avec les trois enzymes, nous obtenons des données de deutération sur 77 % des hydrogènes amidiques. Après recoupement des différents taux de deutération des peptides, nous pouvons déterminer des taux d'échange pour différentes zones de la protéine (figure 73).

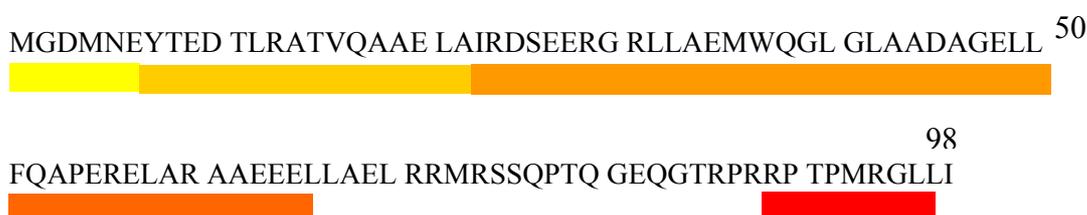
La deutération sur l'ensemble de la protéine est élevée. Ceci indique que la protéine est très accessible au solvant et possède peu de structures secondaires. Ces résultats sont en adéquation avec ceux obtenus par dichroïsme circulaire (Nanao et al., 2003). Ces expériences indiquent que l'hélicité de PcrV est de 6 %, ce qui dénote d'une protéine très peu repliée. Il est donc normal que le taux d'échange soit élevé.

Nous pouvons remarquer que les zones les moins deutérées de la protéine se situent en partie N-terminale, alors que la partie C-terminale est très deutérée.

Les prédictions de structures secondaires données par informatique, suggèrent que la partie située après le résidu 74 ne présente pas de structuration stable. Le manque d'information de deutération sur la zone 67-88 et la deutération de plus de 70 % des hydrogènes amidiques de la zone 89-98 en 30 secondes d'échange confirment bien que la partie C-terminale de PcrG ne comporte pas de structures secondaires stables.

La partie N-terminale pourrait comporter quelques éléments de structure secondaire.

Certaines études indiquent que la protéine PcrG serait stabilisée par PcrV lors de la formation du complexe (Nanao et al., 2003). C'est ce que nous allons vérifier en caractérisant le complexe PcrG-PcrV par échange H/D.



**Figure 73: Représentation du taux de deutération sur la séquence primaire de PcrG.**

Les couleurs représentent le taux de deutération de la zone :

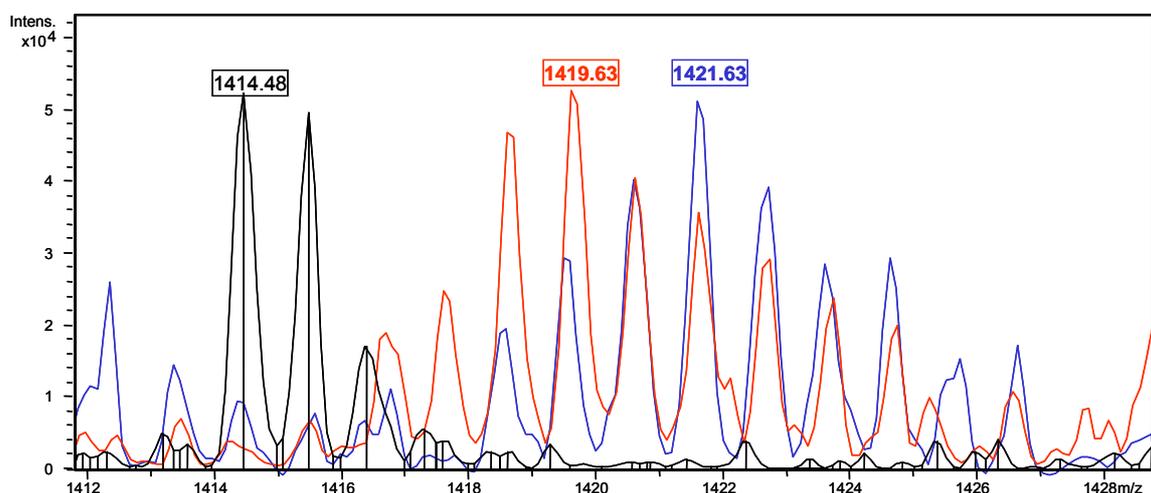
- (  ) : inférieur à 30 %
- (  ) : entre 30 et 50 %
- (  ) : entre 50 et 60 %
- (  ) : entre 60 et 70 %
- (  ) : supérieur à 70 %

## ***2.2 Caractérisation du complexe PcrV-PcrG par échange H/D***

Le complexe PcrV-PcrG est digéré soit par la pepsine, soit par la protéase de type XVIII, soit par la protéase de type XIII. Les peptides générés sont identifiés comme appartenant à l'une ou l'autre des protéines grâce aux cartes peptidiques réalisées lors de l'étude des protéines seules.

Les expériences de deutération sont réalisées sur le complexe avec un temps d'échange de 30 secondes et une digestion de 2 minutes avec chacune des trois enzymes. L'incorporation de deutérium de chaque peptide est déterminée et comparée avec celle établie lors de l'étude des protéines non complexées.

La figure 74 montre l'exemple d'un peptide de 1413.5 Da identifié comme étant le peptide L21-L32 de PcrG. Ce peptide incorpore en moyenne 7 deutériums lorsque la protéine est incubée 30 seconde dans du tampon deutéré. Lorsque le complexe PcrG-PcrV est incubé le même temps dans le même tampon, ce peptide n'incorpore plus que 5 deutériums en moyenne. Deux des liaisons amidiques de ce peptide ne sont donc plus deutérées quand PcrG forme le complexe avec PcrV. Les deux acides aminés correspondant participent donc, directement ou indirectement, à la formation ou à la stabilisation du complexe.



**Figure 74 : exemple de deutération d'un peptide de PcrG impliquée dans le complexe ou non.**

L'ion de m/z 1414.48 correspond au peptide 21-32 de la protéine PcrG. Le spectre noir est celui du peptide de la protéine non deutérée, le bleu est celui du peptide de PcrG deutérée 30 secondes et le rouge est celui du peptide de PcrG lorsqu'elle est en interaction avec PcrV.

Lorsque la protéine est impliquée dans le complexe elle incorpore 2 deutériums de moins dans la zone 21-32 que lorsqu'elle est seule. Les spectres sont obtenus sur un appareil de type piège ionique.

Le nombre de deutérium incorporés a pu être déterminé pour 85 peptides de PcrV (toutes protéases confondues) et pour 33 peptides de PcrG (toutes protéases confondues). Des

différences de deutération protéine seule/protéine complexée sont observées pour chacune des deux protéines (annexes 2 et 3).

Masse du peptide	Identification du peptide	Différence de deutération protéine seule - protéine dans le complexe
801.4	14-21	1.0
730.4	15-21	1.2
601.3	17-22	2.0
1413.8	21-32	1.8
1300.7	22-32	1.3
1045.5	31-38	0.7

**Tableau 5 : Différence de deutération entre peptides issus de PcrG seule et peptides issus de PcrG impliquée dans le complexe PcrG-PcrV.**

Pour la protéine PcrG, une différence significative de deutération entre protéine seule et protéine impliquée dans le complexe est observée pour 6 des peptides (tableau 5). Pour chacun de ces peptides, l'incorporation de deutérium est plus faible lorsqu'ils sont issus de la protéine complexée que lorsqu'ils sont issus de la protéine non-complexée.

Le recoupement des données de ces différents peptides nous permet d'identifier une zone pour laquelle le taux d'échange varie entre protéine seule et protéine complexée. Par exemple, les peptides 14-21 et 15-21 incorporent respectivement 1 et 1.2 deutériums en moyenne. Or le peptide 17-22 en incorpore 2 en moyenne. Statistiquement les deutériums incorporés ne se trouvent pas dans la zone 14-17.

La région que nous avons identifiée comme ayant un taux de deutération différent lorsque PcrG est impliqué dans le complexe et lorsqu'elle ne l'est pas, est la zone 17-38. Cette zone intervient donc dans la formation du complexe PcrG-PcrV. Cette diminution du taux d'incorporation indique que soit cette zone est la zone d'interaction, elle est alors moins accessible au solvant, soit cette zone est structurée lors de la formation du complexe.

Des études réalisées sur le complexe LcrG-LcrV ont permis la mise en évidence d'une zone de LcrG impliquée dans le complexe (Matson et Nilles, 2002). La plus petite zone de LcrG encore capable d'interagir avec LcrV a été identifiée comme étant la zone 7-40. Or les prédictions informatiques de structures secondaires indiquent que cette zone forme un

domaine structuré en « coiled-coil ». L'interaction LcrV-LcrG requière donc la partie N-terminale de LcrG est impliquerait un domaine en « coiled-coil ».

L'alignement des séquences de PcrG et LcrG est présenté figure 75. Le domaine de PcrG que nous avons identifié comme impliqué dans l'interaction PcrG-PcrV correspond tout à fait à celui de LcrV identifié comme impliqué dans l'interaction LcrG-LcrV. Ainsi, l'interaction PcrV-PcrG, comme celle LcrG-LcrV, impliquerait un motif en « coiled-coil » situé en partie N-terminale de PcrG.

LcrG	7	DEYDKTLKQA ELAIADSDHR AKLLQEMCAD IGLT	66
PcrG	10	DTLRATVQAA ELAIRDSEER GRLLAE MWQ	69
		PEAVMKIFAGRSAAEIKPAERELLDE	
		GLGFAADAGEL LFQAPERELARAAEEELLAE	
LcrG		IKRQRERQPQ HPYDGKRPRK PTMMRGQII	95
PcrG		LRRMRSSOPT QGEQGTRPRR PTPMRGLLI	98

**Figure 75 : Alignement de séquence de LcrG et PcrG.**

La zone surlignée en vert sur la séquence de LcrG correspond à la plus petite région identifiée comme interagissant avec LcrV (Matson et Nilles, 2002).

La zone surlignée en rouge sur la séquence de PcrG correspond à la séquence pour laquelle le taux d'échange H/D est différent entre PcrG seule et PcrG en interaction avec PcrV.

Les acides aminés marqués d'un astérisque sur la séquence de LcrG correspondent aux acides aminés impliqués dans l'interaction avec LcrV définis par Matson et Nilles (2002).

Pour la protéine PcrV, une différence significative de deutération entre protéine seule et protéine impliquée dans le complexe est observée pour 13 des peptides. Pour chacun de ces peptides, l'incorporation de deutérium est plus faible lorsqu'ils sont issus de la protéine complexée que lorsqu'ils sont issus de la protéine non-complexée. Il n'y a donc pas de déstructuration de la protéine lors de la complexation. Tous les peptides identifiés appartiennent à des zones soit situées au niveau de la zone de contact, soit qui subissent un changement de conformation lors de la formation du complexe (tableau 6).

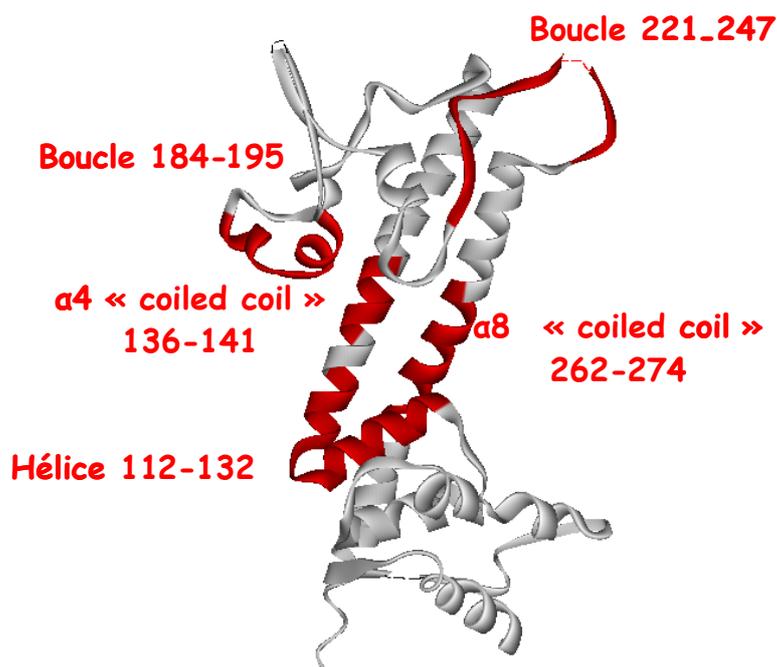
Masse du peptide	Identification du peptide	Différence de deutération protéine seule - protéine dans le complexe
2358.3	112-132	3
679.4	115-120	1
735.4	116-121	1
988.5	133-141	1.3
745.5	136-142	1.3
1901.8	177-192	1
2086	177-194	1.5
2199	177-195	2
3468.7	210-240	1.5
2344.3	241-261	1.7
473.3	261-264	0.8
1454.7	262-274	1.7
1638.8	262-276	1.5

**Tableau 6 : Différence de deutération entre peptides issus de PcrG seule et peptides issus de PcrG impliquée dans le complexe PcrG-PcrV.**

Le recouplement de ces données de deutération et de celles des peptides chevauchants pour lesquels aucun changement n'est observé, nous permet d'identifier cinq zones pour lesquelles le taux d'échange diffère lorsque la protéine est impliquée dans le complexe. Ces zones sont les régions 112-132, 136-141, 184-195, 221-247 et 262-274. Elles sont représentées sur la structure modélisée de PcrV dans la figure 76.

Les zones 136-141 et 262-274 correspondent aux deux hélices impliquées dans le domaine en « coiled-coil ». Les zones 184-195 et 221-247 sont deux boucles situées dans la séquence primaire entre les deux hélices et localisées sur la structure dans le même domaine globulaire. La région 112-132 contient une partie de l'hélice du domaine globulaire opposé, situé à proximité du domaine « coiled-coil » et une partie de l'hélice  $\alpha 4$  du domaine « coiled-coil ». Toutes ces zones sont donc impliquées dans l'interaction PcrV-PcrG.

L'étude structurale de LcrV suggère que la zone d'interaction avec LcrG est située au niveau d'une des hélice du coiled-coil (l'hélice  $\alpha 4$ ) (Derewenda et al., 2004). Nos données expérimentales montrent que l'interaction PcrG-PcrV pourrait bien avoir lieu dans le domaine en « coiled-coil », mais impliquerait les deux hélices.



**Figure 76 :** Représentation des zones pour lesquelles le taux de deutération entre PcrV seule et PcrV complexée avec PcrG est différent.

Ces zones sont potentiellement impliquées dans la formation du complexe.

Les trois autres zones pourraient subir un changement de conformation. Leur disposition dans la structure tridimensionnelle suggère un repliement des deux parties globulaires vers les hélices du coiled-coil. Il pourrait se former une sorte de poche qui stabiliserait PcrG.

Les données obtenues au niveau des deux protéines indiquent donc que l'interaction PcrV-PcrG a lieu au niveau des domaines coiled-coil de chaque protéine.

Lors de l'étude du complexe LcrG-LcrV, des expériences de mutations ponctuelles des acides aminés hydrophobes du domaine coiled-coil potentiel de LcrV suggéraient que l'interaction avait lieu au sein d'une interface hydrophobe (Matson et Nilles., 2002). Cependant, il semblerait que les interactions hydrophobes ne soient pas le seul type d'interaction impliqué dans la formation du complexe. En effet des études réalisées en 2003 ont utilisé la technique de spectrométrie de masse non-covalent pour caractériser le complexe PcrG-PcrV (Nanao et al., 2003). Cette technique ne permet pas la mise en évidence des interactions hydrophobes

qui ne peuvent subsister en phase gaz. Or, le complexe a tout de même pu être observé par cette technique, ce qui suggère que d'autres types d'interaction sont mis en jeu. Les expériences d'échange H/D réalisées dans cette étude vont également dans ce sens. En effet, les hydrogènes amidiques des acides aminés hydrophobes ne s'échangent que très lentement, une deutération de 30 secondes ne devrait pas suffire à les échanger. Or, nous observons un changement dans l'incorporation de deutérium au niveau des coiled-coil lorsque les protéines sont en interaction. Cela supposerait que les acides aminés impliqués ne sont pas qu'hydrophobes.

### **3 CONCLUSIONS ET PERSPECTIVES**

Nous avons pu, en utilisant la réticulation chimique et les échanges H/D, conforter la structure tridimensionnelle modélisée de PcrV. Nous pouvons toutefois supposer que la distance entre les domaines globulaires et le domaine « coiled-coil » a été légèrement surévaluée sur cette structure modélisée. Les données structurales obtenues par échange H/D associé à la spectrométrie de masse indique que la protéine PcrG est très peu structurée. Elle semble toutefois plus structurée du côté N-terminal que du côté C-terminal.

L'interaction PcrG-PcrV implique les zones 17-38 de PcrG et les zones 112-132, 136-141, 184-195, 221-247 et 262-274 de PcrV. Cette interaction pourrait avoir lieu entre les domaines « coiled-coil » de chacune des protéines. La protéine PcrG pourrait être stabilisée par PcrV qui formerait alors une poche. Les interactions peuvent être de plusieurs types, notamment de type hydrophobe.

Pour confirmer ces résultats, il est envisageable de tester la stabilité du complexe avec différentes constructions de PcrV et de PcrG. Une protéine recombinante de PcrG ne contenant que les 40 premiers acides aminés de la protéine sauvage permettrait de confirmer la zone d'interaction sur PcrG. Des protéines recombinante de PcrV sur lesquelles la zone 184-195 ou la zone 221-247 seraient supprimées confirmeraient l'importance des ces deux zones dans la stabilisation du complexe.

La spectrométrie de masse associée aux réticulations chimiques et aux échanges H/D, nous a permis dans cette étude d'obtenir des informations structurales essentielles sur des protéines et un complexe qui n'avaient jamais pu être cristallisés. Ces deux méthodes sont donc une alternative très intéressante aux études structurales par cristallographie lorsque celles-ci ne sont pas possibles.

# CONCLUSION GENERALE

Dans cette étude, nous avons développé deux méthodes innovantes utilisant la spectrométrie de masse pour l'étude structurale des protéines, et nous les avons appliquées à une problématique biologique avec succès.

L'utilisation combinée de trois protéases (la pepsine, la protéase de type XIII et la protéase de type XVIII) pour les échanges H/D en association avec la spectrométrie de masse a permis l'obtention de résultats complets et précis. Nous avons augmenté le recouvrement de la séquence pour laquelle des informations de deutération sont obtenues et amélioré la résolution spatiale dans la localisation des zones d'intérêt.

La méthode associant les réticulations chimiques à la spectrométrie de masse est encore en cours d'optimisation mais elle permet déjà d'obtenir des résultats fort prometteurs. Les contraintes de distances obtenues par cette technique ont été utilisées par les bio-informaticiens pour mettre au point une méthode de détermination de la famille de repliement d'une protéine de structure inconnue.

Une fois les méthodes mises au point, nous avons pu les appliquer sur un système biologique : les protéines PcrV et PcrG du système de sécrétion de type III de *Pseudomonas aeruginosa*. Nos résultats expérimentaux ont conforté la structure modélisée de PcrV. Les contraintes de distances établies indiqueraient que les parties globulaires sont probablement plus proches que prévu du domaine « coiled-coil ». De même, la protéine PcrG apparaît globalement peu structurée, surtout dans sa partie C-terminale. Nous avons également caractérisé l'interaction PcrG-PcrV. Cette interaction se fait probablement par les domaines « coiled-coil » de chacune des deux protéines. La formation du complexe induit un changement de la conformation de PcrV qui pourrait avoir pour conséquence la stabilisation de PcrG.

La spectrométrie de masse est un outil très performant dans l'analyse de la structure des protéines et dans la caractérisation de leurs interactions. Elle consomme peu d'échantillon et permet de travailler dans des conditions proches des conditions physiologiques. Elle permet de réaliser des analyses avec une grande sensibilité et est relativement rapide à mettre en œuvre.

Les méthodes développées dans cette étude vont pouvoir être appliquées à d'autres protéines et systèmes biologiques. Elles vont notamment être complémentaires des techniques d'analyse structurale classiques que sont la RMN et la cristallographie. Même si la résolution des données structurales obtenues n'est pas comparable, nos méthodes permettront d'obtenir des informations préliminaires sur la structure de protéines pour lesquelles la cristallogénèse est difficile, comme les protéines membranaires ou les protéines qui possèdent des parties très flexibles. Ainsi, la spectrométrie de masse, déjà largement utilisée dans les approches

protéomiques, peut aussi légitimement être considérée comme un outil à part entière en biologie structurale.

# MATERIELS ET METHODES

# **1 MATÉRIELS**

## **1.1 La protéine PBP-2X**

La protéine utilisée dans l'étude est PBP-2X\*. Il s'agit de la protéine PBP-2X à laquelle les acides aminés 1 à 48, correspondants à la région hydrophobe d'ancrage de la protéine dans la membrane bactérienne, ont été retirés pour permettre sa solubilisation dans l'eau.

La protéine est produite et purifiée par le Laboratoire de Cristallographie Macromoléculaire (LCM) et le Laboratoire d'Ingénierie des Macromolécules (LIM) de l'IBS. Elle est exprimée comme une protéine de fusion avec la glutathion transférase dans une souche d'*Escherichia coli* et purifiée sur une colonne d'affinité.

Elle nous est fournie à une concentration de 2,12mg/mL dans du tampon 50mM Tris pH=8, 100mM NaCl, 1mM EDTA.

## **1.2 Le cytochrome c**

La solution de cytochrome c de départ est du cytochrome c de cœur de cheval (Sigma) à 1.7 mg/mL dans un tampon phosphate 20mM, NaCl 150mM, pH7.5.

## **1.3 Les protéines PcrV et PcrG**

L'expression et la purification des deux protéines et du complexe ont été réalisées au Laboratoire de Cristallographie Macromoléculaire de l'IBS.

L'expression de PcrV a été réalisée dans une souche d'*E.coli* BL21(DE3), contenant le plasmide pET15b qui code pour la protéine, dans du milieu terrific broth (Sigma) après induction par 1mM d'IPTG 5h à 30°C. Les cellules sont ensuite centrifugées, ressuspendues dans du tampon de lyse (Tris 25mM pH 8, 200mM NaCl, 10mM benzamidine) et soniquées 20 minutes. Le surnageant est ensuite récupéré par centrifugation de 30 minutes à 18000 rpm et injecté sur une colonne de nickel sepharose (Qiagen). La colonne est lavée avec 25 mM Tris pH 8, 200mM NaCl et 20mM imidazole et la protéine est éluée par un gradient

d'imidazole. La protéine est concentrée sur centricon 3000 (Amicon Bioseparation) et dialysée contre du tampon 25 mM Tris pH 8, 200mM NaCl. Elle est finalement obtenue à une concentration finale de 21  $\mu$ M.

La protéine PcrG est produite et purifiée de façon similaire. Elle est obtenue à une concentration finale de 21  $\mu$ M.

Le complexe PcrV/PcrG est obtenu après mélange des deux protéines à un rapport molaire 1:1.5 PcrV/PcrG 15 minutes à 4°C. Le mélange est ensuite purifié sur une colonne de « gel filtration » Superdex. La présence du complexe dans les fractions récoltées est vérifiée par gel d'électrophorèse en conditions dénaturantes. Le complexe est concentré sur centricon 3000 (Amicon Bioseparation) et finalement obtenu à une concentration de 21  $\mu$ M dans du tampon 25 mM Tris pH 8, 200mM NaCl.

## **2 METHODES**

### **2.1 *L'étude de la protéine PBP-2X***

#### **2.1.1 Digestion de la protéine**

La protéine PBP-2X\* (à 2,12mg/mL) est tout d'abord dialysée contre du tampon phosphate 5mM pH=6,8 sur une membrane de 10000 MWCO grâce à une cassette de dialyse (Pierce). La protéine est digérée 2 min à 0°C par l'une des trois proteases. La pepsine (Sigma Aldrich), la protéase de type XIII d'*Aspergillus saitoi* (Sigma Aldrich) et la protéase de type XVIII de *Rhizopus species* (Sigma Aldrich) sont utilisés à des rapports massiques de 1, 10.5 and 17 respectivement. La détermination du rapport optimal enzyme/protéine a été déterminé par l'observation par électrophorèse en conditions dénaturantes, de la disparition de la protéine en fonction de la concentration d'enzyme. Les solutions de protéases sont préparées dans 0.11 M H<sub>3</sub>PO<sub>4</sub>, pH 1.6.

#### **2.1.2 Réalisation des cartes peptidiques par LC/MS/MS**

La protéine PBP-2X\* est digérée comme décrit plus haut par chacune des trois enzymes. Les peptides sont séparés par HPLC sur une colonne de phase inverse de type C<sub>18</sub> (1 mm \* 100 mm, Interchrom). Le tampon A est une solution de 0.1% TFA (Sigma Aldrich), et la phase B est un mélange 90/10/0.1 acétonitrile (SDS)/eau/TFA. L'échantillon est d'abord dessalé sur la colonne avec 2% de B à 50 µL/min. Les peptides sont ensuite élués par un gradient de 2 à 65 % de B en 45 minutes, suivi d'une augmentation jusqu'à 100% de B en 15 minutes et d'un palier de 10 minutes à 100% de B, toujours à 50 µL/min. La moitié du débit de sortie est amené jusqu'au spectromètre de masse et les peptides sont analysés en ligne par MS-MS.

### **2.1.3 Deutération des peptides et analyses LC/MS**

4  $\mu$ L de PBP-2X\* à 2.1 mg/mL sont dilués 20 fois dans un tampon phosphate à 5mM (pD 6.8) constitué de  $\text{Na}_2\text{HPO}_4$  et  $\text{NaH}_2\text{PO}_4$  (Sigma Aldrich) dissous dans du  $\text{D}_2\text{O}$  (Sigma Aldrich). Les échanges isotopiques sont réalisés pendant 30 secondes à température ambiante. Les ré-échanges sont stoppés par l'addition de 9 $\mu$ L de la solution de protéase. La digestion est réalisée pendant 2 minutes dans la glace. L'échantillon est ensuite préconcentré et dessalé sur une colonne peptide MacroTrap (Michrom Bioresources) avec 2% de tampon B à 300  $\mu$ L/min, avant d'être élué sur une colonne C18 (1 mm \* 100 mm, Interchrom). Le gradient utilisé est le même que celui écrit dans l'étape des cartes peptidiques. Pour éviter le ré-échange, tampons, vannes et colonnes sont conservés dans la glace. La totalité du débit est ensuite amenée jusqu'au spectromètre de masse. Les mêmes expériences sont réalisées avec un tampon non-deuté, ainsi les masses des peptides non deutérés sont comparées à celles des peptides deutérés pour un même temps de rétention.

### **2.1.4 ESI-MS and ESI-MS-MS**

Les analyses de spectrométrie de masse sont réalisées sur un appareil de type piège ionique ESQUIRE 3000+ (Bruker Daltonics) équipé d'une source électrospray. La tension du capillaire est fixée à 4kV et celle des skimmers à 500V. La pression du gaz nébulisateur est de 10psi et le gaz de séchage a un débit de 8 L/min avec une température de 250°C.

Lors des analyses MS-MS, les trois ions les plus intenses des spectres MS sont fragmentés et ensuite exclus pendant 1 minute. La fenêtre d'isolation des ions est de 4 m/z et l'amplitude de fragmentation est fixée à 2 V. Les spectres de masse sont acquis sur une gamme de m/z de 50 à 2000. Les spectres sont ensuite traités avec l'aide des logiciels Data Analysis 3.0 and Biotools 2.1.

## **2.2 Les réticulations chimiques sur le cytochrome c**

### **2.2.1 L'étape de réticulation**

Trois agents réticulants sont utilisés le BS3 (pierce), le Sulfo-EGS (Interchim) et le Sulfo-DST (Interchim). L'agent réticulant est directement dissous dans une solution de cytochrome c à 1.7g/L dans du tampon phosphate 20mM pH 7.5, NaCl 150mM. La concentration finale d'agent est telle que le rapport concentration d'agent / concentration de lysine soit compris entre 0.05 et 0.5. La réticulation s'effectue pendant 15 minutes à température ambiante. Cette réaction est arrêtée par l'ajout de Tris 3M et par centrifugation sur centricon 3000 (Amicon Bioseparation) à 10000 tours minutes. Deux rinçages sur le centricon avec du tampon phosphate 20mM pH 7.5, NaCl 150mM sont ensuite réalisés.

### **2.2.2 contrôle de la réticulation**

Le contrôle du rendement de réticulation et de la présence de dimère est effectué par gel SDS-PAGE à 15 % d'acrylamide. Un contrôle est également effectué par MALDI-TOF après dépôt de 2.5 pmol de protéine sur une matrice d'acide sinapinique (Fluka). L'appareil utilisé est un MALDI-TOF Voyager Elite XL (Applied Biosystems). L'analyse s'effectue en mode linéaire, extraction retardée (délai d'extraction de 80nsec). La tension d'extraction est de 20000V, le « grid voltage » de 96% et le « guide wire » de 0.15. L'acquisition se fait pour des m/z de 5000 à 30000.

### **2.2.3 Digestion peptidique**

100 µL de la solution de protéine (soit 170µg de cytochrome c) sont ajoutés à 5µg d'endoprotéinase Lys-C (Roche). La digestion a lieu pendant 18 heures à température ambiante.

## 2.2.4 Mesures de la masse des produits de réticulation

### 2.2.4.1 Mesures par MALDI-TOF

Les analyses des mélanges peptidiques réticulés ou non sont fait après dépôt de 0.5pmol du mélange sur une matrice acide  $\alpha$ -cyano-4-hydroxycinnamique (Sigma). L'appareil utilisé est un MALDI-TOF Voyager Elite XL (Applied Biosystems). L'analyse s'effectue en mode réflectron, extraction retardée (délai d'extraction de 80nsec). Le voltage d'extraction est de 20000V, le grid voltage de 62% et le guide wire de 0.005. L'acquisition se fait pour des m/z de 700 à 5000.

Les spectres sont calibrés en interne d'après les masses calculées des peptides issus de la digestion théorique par l'endoprotéinase Lys-C du cytochrome c non modifié.

Les spectres sont analysés à l'aide du logiciel Data Explorer TM.

### 2.2.4.2 Mesures par LC-MS et LC-MS-MS

Les mesures sont effectuées sur un spectromètre de masse Q-TOF Micro (Waters) couplé un système de chromatographie liquide CapLC pump micromass (Waters).

Pour les expériences de MS et de MS-MS, la séparation chromatographique est la même. La colonne utilisée est une colonne de type C18 (0.32 \* 100mm , Waters) et la cartouche de dessalage est une peptide micro-trap de type C18 (Interchim). Les solvants utilisés sont le solvant A (0.2 % d'acide formique), le solvant B (90 % de méthanol, 0.2% d'acide formique) et le solvant C (0.2 % d'acide formique).

5  $\mu$ L d'échantillon (soit 60 pmol) sont injecté sur la précolonne et dessalé pendant 5 minutes à 20  $\mu$ L/min avec du solvant C. L'élution est effectuée par à un gradient de débit 5  $\mu$ L/min. Dans un premier temps le pourcentage de solvant B passe de 2 % à 40 % en 8 minutes, puis de 40 % à 100% en 10 minutes. Un pallier est réalisé pendant 7 minutes à 100% de B, puis le pourcentage de B passe à 2 % en 6 minutes.

L'acquisition par spectrométrie de masse se fait en ligne avec une tension de capillaire de 3000 V, une tension de cône de 42 V, et une tension d'extraction de cône de 1.2 V. La source est à une température de 80°C, le débit du gaz de nébulisation (N<sub>2</sub>) est de 60L/h. La désolvatation se fait par un gaz (N<sub>2</sub>) à 450L/h et 150°C. Pour les analyses en MS simple, la

gamme de  $m/z$  étudiée est de 200-1600  $m/z$ , avec un temps de balayage de une seconde. Pour les expériences de MS-MS, les MS sont réalisées pour une gamme de  $m/z$  de 500-1500 avec un temps de balayage de une seconde. Les trois précurseurs les plus intenses de notre liste d'inclusion sont fragmentés avec des rampes d'énergie de fragmentation dépendant de l'état de charge du précurseur. Les spectres MS2 sont réalisés pour une gamme de  $m/z$  de 100-2000. Les spectres sont ensuite analysés avec l'aide du logiciel MassLynx. Ils sont calibrés en interne grâce aux masses théoriques des peptides issus de la digestion du cytochrome c par la Lys-C.

#### 2.2.4.3 Mesures par MALDI-TOF-TOF

Les expériences de fragmentation par MALDI-TOF-TOF sont réalisées sur des appareils de démonstration de type Autoflex chez le constructeur Bruker Daltonics.

La matrice utilisée pour les dépôts est l'acide  $\alpha$ -cyano-4-hydroxycinnamique.

Les analyses s'effectuent en mode reflectron, extraction retardée avec une fréquence laser de 50Hz.

En mode MS le potentiel d'accélération est de 19kV et l'acquisition se fait pour une gamme de  $m/z$  de 700 à 6000.

En mode MS-MS, Le potentiel d'accélération initiale est de 6kV et la ré-accélération au niveau du LIFT est de 19kV. L'acquisition se fait pour une gamme de  $m/z$  de 0 à 5000.

Les spectres sont traités par le logiciel FlexAnalysis et interprétés grâce au logiciel Biotools 2.2.

## 2.3 L'étude de PcrV, PcrG et du complexe PcrV/PcrG

### 2.3.1 Cartes peptidiques des protéines

Pour chacune des protéines, la digestion s'effectue avec l'une des protéases à des rapports de concentration de 1, 10 et 17 pour la pepsine, la protéase de type XIII et la protéase de type XVIII respectivement. Les protéases sont diluées dans du tampon 0.14M  $H_3PO_4$ . La digestion s'effectue 2 minutes à 0°C et à pH 2.5.

Les peptides sont séparés par HPLC sur une colonne de phase inverse de type C<sub>18</sub> (1 mm \* 100 mm, Interchrom). 150 pmol de protéine digérée sont injectées. Le gradient HPLC utilisé est le même que pour les expériences réalisées sur PBP-2X. La moitié du débit de sortie est amené jusqu'au spectromètre de masse et les peptides sont analysés en ligne par MS-MS sur un appareil de type piège ionique ESQUIRE 3000+ (Bruker Daltonics) équipé d'une source électrospray. Les paramètres d'acquisition sont les mêmes que pour les cartes peptidiques de PBP-2X.

### **2.3.2 La deutération des protéines et du complexe**

60 pmol de protéine (ou de complexe) sont dilués dans dix volumes de tampon deutéré (tampon Tris 25mM, NaCl 200mM, pD 8). La deutération s'effectue pendant des temps variables à température ambiante ou dans la glace. Elle est arrêtée par l'ajout des protéases. La digestion s'effectue comme décrit plus haut. L'échantillon est ensuite préconcentré et dessalé sur une colonne peptide MacroTrap (Michrom Bioresources) avec 2% de tampon B à 300 µL/min, avant d'être élué sur une colonne C18 (1 mm \* 100 mm, Interchrom). Le gradient utilisé est le même que celui écrit dans l'étape des cartes peptidiques. Pour éviter les ré-échanges tampons, vannes et colonnes sont conservés dans la glace. La totalité du débit est ensuite amenée jusqu'au spectromètre de masse. Les mêmes expériences sont réalisées avec un tampon non-deuté, ainsi les masses des peptides non deutérés sont comparées à celles des peptides deutérés pour un même temps de rétention.

Les analyses de spectrométrie de masse s'effectuent sur l'ESQUIRE 3000+, avec les mêmes paramètres que précédemment.

### **2.3.3 La réticulation chimique de PcrV**

La protéine est d'abord dialysée contre du tampon phosphate 20mM pH 7.5, NaCl 150mM. Elle est obtenue à une concentration de 0.8g/L. L'agent réticulant est ajouté 200µL de PcrV à 24µM : 1µL d'une solution de BS3 26.5g/L, 1µL de Sulfo-EGS à 47.5 g/L et 10 µL de Sulfo-DST 39.5 g/L. Les solutions d'agents réticulants sont réalisées dans du tampon phosphate 20mM pH 7.5, NaCl 150mM.

La réticulation s'effectue pendant 15 minutes à température ambiante et est arrêtée par l'ajout de Tris 3M et par centrifugation sur centricon 3000 (Amicon Bioseparation) à 10000 tours minutes. Deux rinçages sur le centricon avec du tampon phosphate 20mM pH 7.5, NaCl 150mM sont ensuite réalisés.

200  $\mu$ L de la solution de protéine (soit 160 $\mu$ g) sont ajoutés à 5 $\mu$ g d'endoprotéinase Lys-C (Roche). La digestion a lieu pendant 18 heures à température ambiante.

## **2.3.4 Détermination de la masse des produits de réticulation**

### *2.3.4.1 Mesures par LC-MS*

Les mesures sont effectuées sur un spectromètre de masse Q-TOF Micro (Waters) couplé un système de chromatographie liquide CapLC pump micromass (Waters).

Les gradients et les paramètres d'acquisition des spectres sont les mêmes que ceux utilisés pour l'étude de la réticulation du cytochrome c

Les spectres sont ensuite analysés avec l'aide du logiciel MassLynx. Ils sont calibrés en interne grâce aux masses théoriques des peptides issus de la digestion de PcrV par la Lys-C.

### *2.3.4.2 Mesures par MALDI-TOF*

Les mesures sont effectuées sur un Autoflex (Bruker Daltonics).

Les échantillons sont dessalés sur ziptip et déposés (2.5 pmol déposées) sur de l'acide  $\alpha$ -cyano-4-hydroxycinnamique.

Les analyses s'effectuent en mode reflectron et en extraction retardée. Le potentiel d'accélération est de 20kV et l'acquisition se fait pour une gamme de m/z de 500 à 6000.

Les spectres sont calibrés en interne grâce aux masses théoriques des peptides issus de la digestion de PcrV par la LysC et analysés sur le logiciel Data Analysis 3.0.

REFERENCES  
BIBLIOGRAPHIQUES

- Alley S.C., Ishmael F.T., Jones A.D., and Benkovic S.J. (2000) « Mapping protein-protein interactions in the bacteriophage T4 DNA polymerase holoenzyme using a novel trifunctional photo-cross-linking and affinity agent. » J. Am. Chem. Soc. ; **122** : 6126-6127.
- Anand, G. S., Hughes, C. A., *et al.* (2002). "Amide H/2H exchange reveals communication between the cAMP and catalytic subunit-binding sites in the R(I)alpha subunit of protein kinase A." J Mol Biol **323**(2): 377-86.
- Anderson DM, Schneewind O. (1997) « A mRNA signal for the type III secretion of Yop proteins by *Yersinia enterocolitica*. » Science. ; **278**(5340):1140-3.
- Asahi Y, Takeuchi Y, Ubukata K. (1999) « Diversity of substitutions within or adjacent to conserved amino acid motifs of penicillin-binding protein 2X in cephalosporin-resistant *Streptococcus pneumoniae* isolates. » Antimicrob Agents Chemother. ; **43**(5):1252-5.
- Back JW, De Jong L, Muijsers AO, de Koster CG. (2003) "Chemical Cross-linking and Mass Spectrometry for Protein Structural Modeling." J Mol Biol ; **331**: 303-313.
- Back JW, Notenboom V, de Koning LJ, Muijsers AO, Sixma TK, de Koster CG, de Jong L. (2002a) « Identification of cross-linked peptides for protein interaction studies using mass spectrometry and 18O labeling. » Anal Chem. ; **74**(17):4417-22.
- Back JW, Sanz MA, De Jong L, De Koning LJ, Nijtmans LG, De Koster CG, Grivell LA, Van Der Spek H, Muijsers AO. (2002b) « A structure for the yeast prohibitin complex: Structure prediction and evidence from chemical crosslinking and mass spectrometry. » Protein Sci. ; **11**(10):2471-8.
- Bai, Y., Milne, J. S., *et al.* (1993). "Primary structure effects on peptide group hydrogen exchange." Proteins **17**(1): 75-86.
- Beavis RC, Chait BT. (1989). "Cinnamic acid derivated as matrices for ultraviolet laser desorption mass spectrometry of proteins." Rapid Commun Mass Spectrom , **3** : 432-435
- Beavis RC, Chait BT. (1992). "Alpha-cyano-4-hydrocinnamic acid as a matrix for matrix-assisted laser desorption mass spectrometry." Org. Mass Spectrom., **27** : 156-158
- Bennett KL, Kussmann M, Björk P, Godzwon M., Mikkelsen M, Sorensen P, Roepstorff P. (2000) "Chemical cross-linking with thiol-cleavable reagents combined with differential mass spectrometric peptide mapping – A novel approach to assess intermolecular protein contact." Protein Sci. ; **9**: 1503-1518.
- Bonner RF. (1977) "Derivation of the field equations and stability parameters for three operating modes of the three-dimensional quadripole" Int J Mass Spectrom Ion Physics ; **23** : 249-257.

- Botta G.A., Park J.T. (1981) « Evidence for involvement of penicillin-binding protein 3 in murein synthesis during septation but not during cell elongation. » J. Bacteriol. ; **145**(1) : 333-340.
- Boyd AP, Lambermont I, Cornelis GR. (2000) « Competition between the Yops of *Yersinia enterocolitica* for delivery into eukaryotic cells: role of the SycE chaperone binding domain of YopE. » J Bacteriol. ; **182**(17):4811-21.
- Bramhill D. (1997) « Bacterial cell division. » Annu Rev Cell Dev Biol. ;**13**:395-424.
- Brier S, Lemaire D, Debonis S, Forest E, Kozielski F. (2004) « Identification of the protein binding region of S-trityl-L-cysteine, a new potent inhibitor of the mitotic kinesin Eg5. » Biochemistry. ;**43**(41):13072-82.
- Busch K. (1995). "Desorption Ionization Mass Spectrometry", J Mass Spectrom , **30** : 233-240
- Canet, D., Sunde, M., *et al.* (1999). "Mechanistic studies of the folding of human lysozyme and the origin of amyloidogenic behavior in its disease-related variants." Biochemistry **38**(20): 6419-27.
- Chen X, Chen YH, Anderson VE. (1999) « Protein cross-links: universal isolation and characterization by isotopic derivatization and electrospray ionization mass spectrometry. » Anal Biochem. ; **273**(2):192-203.
- Cheng LW, Anderson DM, Schneewind O. (1997) « Two independent type III secretion mechanisms for YopE in *Yersinia enterocolitica*. » Mol Microbiol. ; **24**(4):757-65.
- Chernushevich IV, Loboda AV, Thomson BA. (2001) « An introduction to quadrupole-time-of-flight mass spectrometry. » J Mass Spectrom.; **36**(8):849-65.
- Chesnel L, Pernot L, Lemaire D, Champelovier D, Croize J, Dideberg O, Vernet T, Zapun A. (2003) « The structural modifications induced by the M339F substitution in PBP2x from *Streptococcus pneumoniae* further decreases the susceptibility to beta-lactams of resistant strains. » J Biol Chem. ;**278**(45):44448-56.
- Collins C, Schilling B, Young M, Dollinger G, Guy RK. (2003) "Isotopically labelled crosslinking reagents: resolution of mass degeneracy in the identification of crosslinked peptides." Bioorg Med Chem Lett. ; **13**: 4023-4026.
- Cornelis GR. (2000) « Molecular and cell biology aspects of plague. » Proc Natl Acad Sci USA ; **97**(16):8778-83.
- Cotter RJ. (1992). "Time of flight mass spectrometry for the structural analysis of biological molecules." Anal Chem. **64** : 1027A-1039A

- Dacheux D, Attree I, Toussaint B. (2001) « Expression of ExsA in trans confers type III secretion system-dependent cytotoxicity on noncytotoxic *Pseudomonas aeruginosa* cystic fibrosis isolates. » Infect Immun. ; **69**(1):538-42.
- Dawson, P. H. (1986). "Quadrupole mass analysers : Performance, design and some recent applications." Mass Spectrom. Rev. **5**(1): 1-37.
- De Hoffmann, E., Charette, J., *et al.* (1994). "Spectrométrie de masse." (ed Masson).
- Deng, Y. et Smith, D. L. (1998) « Identification of unfolding domains in large proteins by their unfolding rates. » Biochemistry. ; **37**(18):6256-62.
- Deng, Y., Zhang, Z., *et al.* (1999). "Comparison of continuous and pulsed labeling amide hydrogen exchange/mass spectrometry for studies of protein dynamics." J Am Soc Mass Spectrom **10**(8): 675-84.
- Derewenda U, Mateja A, Devedjiev Y, Routzahn KM, Evdokimov AG, Derewenda ZS, Waugh DS. (2004) « The structure of *Yersinia pestis* V-antigen, an essential virulence factor and mediator of immunity against plague. » Structure (Camb). ; **12**(2):301-6.
- Dessen A, Mouz N, Gordon E, Hopkins J, Dideberg O. (2001) « Crystal structure of PBP2x from a highly penicillin-resistant *Streptococcus pneumoniae* clinical isolate: a mosaic framework containing 83 mutations. » J Biol Chem. ; **276**(48):45106-12.
- Dharmasiri, K. et Smith, D. L. (1996). "Mass spectrometric determination of isotopic exchange rates of amide hydrogens located on the surfaces of proteins." Anal Chem **68**(14): 2340-4.
- Di Guilmi A.M., Mouz N., Petillot Y., Forest E., Dideberg O., Vernet T. (2000) « Deacylation kinetics analysis of *Streptococcus pneumoniae* penicillin-binding protein 2x mutants resistant to beta-lactam antibiotics using electrospray ionization- mass spectrometry. » Anal. Biochem. ; **284**(2) : 240-246.
- Di Guilmi AM, Dessen A, Dideberg O, Vernet T. (2003) « Functional characterization of penicillin-binding protein 1b from *Streptococcus pneumoniae*. » J Bacteriol. ; **185**(5):1650-8.
- Di Guilmi AM, Mouz N, Andrieu JP, Hoskins J, Jaskunas SR, Gagnon J, Dideberg O, Vernet T. (1998) « Identification, purification, and characterization of transpeptidase and glycosyltransferase domains of *Streptococcus pneumoniae* penicillin-binding protein 1a. » J Bacteriol. ; **180**(21):5652-9.
- Di Guilmi AM, Mouz N, Martin L, Hoskins J, Jaskunas SR, Dideberg O, Vernet T. (1999) « Glycosyltransferase domain of penicillin-binding protein 2a from *Streptococcus pneumoniae* is membrane associated. » J Bacteriol. ; **181**(9):2773-81.

- Dideberg O, Charlier P, Wery JP, Dehottay P, Dusart J, Ercicum T, Frere JM, Ghuysen JM. (1987) « The crystal structure of the beta-lactamase of *Streptomyces albus* G at 0.3 nm resolution. » Biochem J. ;**245**(3):911-3.
- Dihazi GH, Sinz A. (2003) « Mapping low-resolution three-dimensional protein structures using chemical cross-linking and Fourier transform ion-cyclotron resonance mass spectrometry. » Rapid Commun Mass Spectrom. ; **17**(17):2005-14.
- Dole, M., Mack, L. L., *et al.* (1968). "Molecular beams of macroions." J Chem Phys **49**: 2240.
- Dowson CG, Coffey TJ, Spratt BG. (1994) « Origin and molecular epidemiology of penicillin-binding-protein-mediated resistance to beta-lactam antibiotics. » Trends Microbiol. ; **2**(10):361-6.
- Dowson CG, Hutchison A, Spratt BG. (1989) « Nucleotide sequence of the penicillin-binding protein 2B gene of *Streptococcus pneumoniae* strain R6. » Nucleic Acids Res. ; **17**(18):7518.
- Engen JR, Smithgall TE, Gmeiner WH, Smith DL. (1998) « Identification and localization of slow, natural, cooperative unfolding in the hematopoietic cell kinase SH3 domain by amide hydrogen exchange and mass spectrometry. » Biochemistry.; **36**(47):14384-91.
- Engen, J. R. et Smith, D. L. (2000). « Investigating the higher order structure of proteins. » Methods in Molecular Biology : Mass Spectrometry of Proteins and Peptides. J. R. Chapman. Totowa, NJ, Humana Press Inc. **146**: 95-112.
- Englander , S. W. et Kallenbach, N. R. (1984). "Hydrogen exchange and structural dynamics of proteins and nucleic acids." Quarterly Review of Biophysics **16**(4): 521-655.
- Eyles, S. J. et Gierasch, L. M. (2000). "Multiple roles of prolyl residues in structure and folding." J Mol Biol **301**(3): 737-47.
- Fenn JB, Mann M, Meng CK, Wong SF, Whitehouse CM. (1989) "Electrospray ionization for mass spectrometry of large biomolecules." Science. ; **246**(4926):64-71.
- Finck-Barbancon V, Goranson J, Zhu L, Sawa T, Wiener-Kronish JP, Fleiszig SM, Wu C, Mende-Mueller L, Frank DW. (1997) « ExoU expression by *Pseudomonas aeruginosa* correlates with acute cytotoxicity and epithelial injury. » Mol Microbiol. ; **25**(3):547-57.
- Frank DW, Vallis A, Wiener-Kronish JP, Roy-Burman A, Spack EG, Mullaney BP, Megdoud M, Marks JD, Fritz R, Sawa T. (2002) « Generation and characterization of a protective monoclonal antibody to *Pseudomonas aeruginosa* PcrV. » J Infect Dis. ; **186**(1):64-73.
- Frank DW. (1997) « The exoenzyme S regulon of *Pseudomonas aeruginosa*. » Mol Microbiol. ; **26**(4):621-9.

- Frere JM, Ghuysen JM, Perkins HR. (1975) « Interaction between the exocellular DD-carboxypeptidase-transpeptidase from *Streptomyces* R61, substrate and beta-lactam antibiotics. A choice of models. » Eur J Biochem. ; **57**(2):353-9.
- Frere JM, Joris B, Dideberg O, Charlier P, Ghuysen JM. (1988) « Penicillin-recognizing enzymes. » Biochem Soc Trans. ; **16**(6):934-8.
- Galan JE, Collmer A. (1999) « Type III secretion machines: bacterial devices for protein delivery into host cells. » Science. ; **284**(5418):1322-8.
- Garcia-Bustos J, Tomasz A. (1990) « A biological price of antibiotic resistance: major changes in the peptidoglycan structure of penicillin-resistant pneumococci. » Proc Natl Acad Sci U S A. ; **87**(14):5415-9.
- Garcia-Bustos JF, Chait BT, Tomasz A. (1988) « Altered peptidoglycan structure in a pneumococcal transformant resistant to penicillin. » J Bacteriol. ; **170**(5):2143-7.
- Gevaert K. et Vanderkerckhove J. (2000) “Protein identification methods in proteomics.” Electrophoresis. ; **21**(6):1145-54.
- Ghosh P. (2004) « Process of protein transport by the type III secretion system. » Microbiol Mol Biol Rev. ; **68**(4):771-95.
- Ghuysen JM. (1968) « Use of bacteriolytic enzymes in determination of wall structure and their role in cell metabolism. » Bacteriol Rev. ; **32**(4 Pt 2):425-64.
- Ghuysen JM. (1994) « Molecular structures of penicillin-binding proteins and beta-lactamases. » Trends Microbiol. ; **2**(10):372-80.
- Goehring UM, Schmidt G, Pederson KJ, Aktories K, Barbieri JT. (1999) « The N-terminal domain of *Pseudomonas aeruginosa* exoenzyme S is a GTPase-activating protein for Rho GTPases. » J Biol Chem. ; **274**(51):36369-72.
- Goffin C, Ghuysen JM. (1998) « Multimodular penicillin-binding proteins: an enigmatic family of orthologs and paralogs. » Microbiol Mol Biol Rev. ; **62**(4):1079-93.
- Gophna U, Ron EZ, Graur D. (2003) « Bacterial type III secretion systems are ancient and evolved by multiple horizontal-transfer events. » Gene. ; **312**:151-63.
- Gordon E. J., Mouz N., Duee E., Dideberg O. (2000) « The Crystal Structure of the Penicillin Binding Protein 2X from *Streptococcus Pneumoniae* and its Acyl-Enzyme Form: Implication in Drug Resistance. » J. Mol. Biol. ; **299** : 501.
- Goure J, Pastor A, Faudry E, Chabert J, Dessen A, Attree I. (2004) « The V antigen of *Pseudomonas aeruginosa* is required for assembly of the functional PopB/PopD translocation pore in host cell membranes. » Infect Immun. ; **72**(8):4741-50.
- Grabarek Z, Gergely J. (1990) “Zero-length cross-linking procedure with the use of active ester.” Anal Biochem. ; **36**: 131-135.

- Grebe T, Hakenbeck R. (1996) « Penicillin-binding proteins 2b and 2x of *Streptococcus pneumoniae* are primary resistance determinants for different classes of beta-lactam antibiotics. » Antimicrob Agents Chemother. ; **40**(4):829-34.
- Green NS, Reisler E, Houk KN. (2001) "Quantitative evaluation of the lengths of homobifunctional protein cross-linking reagents used as molecular rulers." Protein Sci. ; **10**: 1293-1304.
- Guy, P., Jaquinod M., *et al.* (1996). "New conformational properties induced by the replacement of tyr-64 in *Desulfovibrio vulgaris* Hildenborough ferricytochrome c553 using isotopic exchanges monitored by mass spectrometry." FEBS Lett **395**(1) : 53-57.
- Hakenbeck R, Coyette J. (1998) « Resistant penicillin-binding proteins. » Cell Mol Life Sci. ;**54**(4):332-40.
- Hakenbeck R, Grebe T, Zahner D, Stock JB. (1999) « beta-lactam resistance in *Streptococcus pneumoniae*: penicillin-binding proteins and non-penicillin-binding proteins. » Mol Microbiol ; **33**(4):673-8.
- Hakenbeck R, Kohiyama M. (1982) « Purification of penicillin-binding protein 3 from *Streptococcus pneumoniae*. » Eur J Biochem. ; **127**(2):231-6.
- Halgand, F., Dumas, R., *et al.* (1999). " Characterization of the conformational changes of acetohydroxy acid isomerase induced by the binding of Mg<sup>2+</sup> ions, NADPH, and a competitive inhibitor." Biochemistry **38**(19) : 6025-34.
- Haniu M, Narhi LO, Arakawa T, Elliott S, Rohde MF. (1993) « Recombinant human erythropoietin (rHuEPO): cross-linking with disuccinimidyl esters and identification of the interfacing domains in EPO. » Protein Sci. ; **2**(9):1441-51.
- Hartmann FC., Wold F. (1966) "Bifunctional reagents. Cross-linking of pancreatic ribonuclease with a diimido ester" J. Am. Chem. Soc.; **84** : 3491.
- Hermanson GT. (1996) Bioconjugate Techniques, Academic Presse: San Diego, CA.
- Hilser VJ, Freire E. (1996) « Structure-based calculation of the equilibrium folding pathway of proteins. Correlation with hydrogen exchange protection factors. » J Mol Biol. ;**262**(5):756-72.
- Hoare DG, Koshland DE. (1996) "A procedure for selective modification of carboxyl groups in proteins." J Am Chem Soc.; **88**: 2057.
- Holtje JV. (1998) « Growth of the stress-bearing and shape-maintaining murein sacculus of *Escherichia coli*. » Microbiol Mol Biol Rev. ; **62**(1):181-203.
- Huang BX, Kim HY, Dass C. (2004) « Probing three-dimensional structure of bovine serum albumin by chemical cross-linking and mass spectrometry. » J Am Soc Mass Spectrom. ; **15**(8):1237-47.

- Hueck CJ. (1998) « Type III protein secretion systems in bacterial pathogens of animals and plants. » Microbiol Mol Biol Rev. ; **62**(2):379-433.
- Hunter MJ, Ludwig ML. (1962) “Reactions of imidoesters with proteins and related small molecules.” J Am Chem Soc. ; **84**: 3491.
- Iribarne, J. V. et Thomson, B. A. (1976). "On the evaporation of small ions from charged droplets." J Chem Phys **64**: 2287.
- Jamin M, Damblon C, Millier S, Hakenbeck R, Frere JM. (1993) « Penicillin-binding protein 2x of *Streptococcus pneumoniae*: enzymic activities and interactions with beta-lactams. » Biochem J.; **292**:735-41.
- Jelsch C, Mourey L, Masson JM, Samama JP. (1993) « Crystal structure of *Escherichia coli* TEM1 beta-lactamase at 1.8 Å resolution. » Proteins.; **16**(4):364-83.
- Joris B, Ghuysen JM, Dive G, Renard A, Dideberg O, Charlier P, Frere JM, Kelly JA, Boyington JC, Moews PC, et al. (1988) « The active-site-serine penicillin-recognizing enzymes as members of the *Streptomyces* R61 DD-peptidase family. » Biochem J. ; **250**(2):313-24.
- Karas MB, Ingehod A, Nordhoff E, Stahl B, Strupat K, Hillenkamp F. (1990). “Principles and applications of matrix-assisted laser desorption/ionization mass spectrometry.” Anal Chim Acta , **241** : 175-185
- Kebarle, P. et Tang, L. (1993). "From ions in solution to ions in the gas phase, the mechanism of electrospray mass spectrometry." Anal. Chem. **65**(22): 972A-986A.
- Kell C.L., Sharma U.K., Dowson C.G., Town C., Balganesch T.S., Spratt B.G. ; (1993) « Deletion analysis of the essentiality of penicillin-binding proteins 1A, 2B and 2X of *Streptococcus pneumoniae*. » ; FEMS Microbiol. Lett. ; **106**(2) : 171-175.
- Kelly JA. and Kuzin AP. (1995) « The refined crystallographic structure of a DD-peptidase penicillin-target enzyme at 1.6 Å resolution. » J Mol Biol.; **254**(2):223-36.
- Kim KS, Woodward C. (1993) « Protein internal flexibility and global stability: effect of urea on hydrogen exchange rates of bovine pancreatic trypsin inhibitor. » Biochemistry. ; **32**(37):9609-13.
- Klein D.L. (1999) “ Pneumococcal disease and the role of conjugate vaccines. “ Microb. Drug Resist. **5** : 147-157.
- Krall R, Schmidt G, Aktories K, Barbieri JT. (2000) « *Pseudomonas aeruginosa* ExoT is a Rho GTPase-activating protein. » Infect Immun. ; **68**(10):6066-8.
- Kruppa GH, Schoeniger J, Young MM. (2003) « A top down approach to protein structural studies using chemical cross-linking and Fourier transform mass spectrometry. » Rapid Commun Mass Spectrom. ; **17**(2):155-62.

- Laible G, Hakenbeck R, Sicard MA, Joris B, Ghuysen JM. (1989) « Nucleotide sequences of the pbpX genes encoding the penicillin-binding proteins 2x from *Streptococcus pneumoniae* R6 and a cefotaxime-resistant mutant, C506. » Mol Microbiol. ; **3**(10):1337-48.
- Laible G., Spratt B.G., Hakenbeck R. (1991) « Interspecies recombinational events during the evolution of altered PBP 2x genes in penicillin-resistant clinical isolates of *Streptococcus pneumoniae*. » Mol. Microbiol. ; **5**(8) : 1993-2002.
- Laiko VV, Baldwin MA, Burlingame AL. (2000). “Atmospheric pressure matrix-assisted laser desorption/ionization mass spectrometry.” Anal Chem. **72**(4):652-7.
- Laiko VV, Taranenko NI, Berkout VD, Musselman BD, Doroshenko VM. (2002). “Atmospheric pressure laser desorption/ionization on porous silicon.” Rapid Commun Mass Spectrom. **16**(18):1737-42.
- Laurans G., Murbach V., Cattier B., Chomarant M., Cottin J., Demachy M., Dupond M., Fauchere J., Fosse T., Laaberki M., Lemozy J., Maugein J., Pechinot A., Ploy M., Romaszko J., Roussel-Delvallez M., Vaucel J., Vernet-Garnier V., Weber M., and Chardon H. (2001) « Surveillance de la sensibilité aux antibiotiques des pneumocoques isolés en situation pathogène en France en 1999. » Bulletin épidémiologique hebdomadaire. **33**.
- Leavell MD, Novak P, Behrens CR, Schoeniger JS, Kruppa GH. (2004) « Strategy for selective chemical cross-linking of tyrosine and lysine residues. » J Am Soc Mass Spectrom. ; **15**(11):1604-11.
- Liu S, Yahr TL, Frank DW, Barbieri JT. (1997) « Biochemical relationships between the 53-kilodalton (Exo53) and 49-kilodalton (ExoS) forms of exoenzyme S of *Pseudomonas aeruginosa*. » J Bacteriol. ; **179**(5):1609-13.
- Liu SC, Fairbanks G, Palek J. (1977) “Spontaneous, reversible protein cross-linking in human erythrocyte membranes – temperature and pH dependence.” Biochemistry.; **16**: 4066.
- Loo, J. A., Udseth, H. R., *et al.* (1988). "Solvent effect on the charge distribution observed with electrospray ionization mass spectrometry of large molecules." Biomed. Environ. Mass Spectrom. **17**: 411-414.
- Lu WP, Kincaid E, Sun Y, Bauer MD. (2001) « Kinetics of beta-lactam interactions with penicillin-susceptible and -resistant penicillin-binding protein 2x proteins from *Streptococcus pneumoniae*. Involvement of acylation and deacylation in beta-lactam resistance. » J Biol Chem. ; **276**(34):31494-501.
- Lyczak JB, Cannon CL, Pier GB. (2000) « Establishment of *Pseudomonas aeruginosa* infection: lessons from a versatile opportunist. » Microbes Infect. ; **2**(9):1051-60.

- Lyczak JB, Cannon CL, Pier GB. (2002) « Lung infections associated with cystic fibrosis. » Clin Microbiol Rev. ; **15**(2):194-222.
- Mandell JG, Falick AM, Komives EA. (1998) « Identification of protein-protein interfaces by decreased amide proton solvent accessibility. » Proc Natl Acad Sci U S A. ;**95**(25):14705-10.
- March, R. E. (2000). "Quadrupole ion trap mass spectrometer." Encyclopedia of Analytical Chemistry. Chichester, John Wiley & Sons Ltd: 24.
- Massidda O, Anderluzzi D, Friedli L, Feger G. (1998) « Unconventional organization of the division and cell wall gene cluster of *Streptococcus pneumoniae*. » Microbiology. ;**144**:3069-78.
- Massova I, Mobashery S. (1998) « Kinship and diversification of bacterial penicillin-binding proteins and beta-lactamases. » Antimicrob Agents Chemother. ; **42**(1):1-17.
- Matagne A, Dubus A, Galleni M, Frere JM. (1999) « The beta-lactamase cycle: a tale of selective pressure and bacterial ingenuity. » Nat Prod Rep. ; **16**(1):1-19.
- Matson JS, Nilles ML. (2002) « Interaction of the *Yersinia pestis* type III regulatory proteins LcrG and LcrV occurs at a hydrophobic interface. » BMC Microbiol. ; **2**(1):16.
- Mazon H, Marcillat O, Forest E, Vial C. (2004) « Hydrogen/deuterium exchange studies of native rabbit MM-CK dynamics. » Protein Sci. ;**13**(2):476-86.
- McLuckey, S. A., Van Berkel, G. J., *et al.* (1994). "Ion trap mass spectrometry. Using high-pressure ionization." Anal Chem **66**(14): 737A-743A.
- Medzihradzky KF, Campbell JM, Baldwin MA, Falick AM, Juhasz P, Vestal ML, Burlingame AL. (2000). "The Characteristics of Peptide Collision-Induced Dissociation Using a High-Performance MALDI-TOF/TOF Tandem Mass Spectrometer." Anal. chem. ; **72**(3):552-8
- Miller, D. W. et Dill, K. A. (1995). "A statistical mechanical model for hydrogen exchange in globular proteins." Protein Sci **4**(9): 1860-73.
- Miranker A, Robinson CV, Radford SE, Aplin RT, Dobson CM. (1993) « Detection of transient protein folding populations by mass spectrometry. » Science. ;**262**(5135):896-900.
- Morlot C, Zapun A, Dideberg O, Vernet T. (2003) « Growth and division of *Streptococcus pneumoniae*: localization of the high molecular weight penicillin-binding proteins during the cell cycle. » Mol Microbiol. ;**50**(3):845-55.
- Mouz N, Di Guilmi AM, Gordon E, Hakenbeck R, Dideberg O, Vernet T. (1999) « Mutations in the active site of penicillin-binding protein PBP2x from *Streptococcus pneumoniae*. Role in the specificity for beta-lactam antibiotics. » J Biol Chem. ;**274**(27):19175-80.

- Mouz N., Gordon E., Di Guilmi A.M., Petit I., Petillot Y., Dupont Y., Hakenbeck R., Vernet T., Dideberg O. (1998) « Identification of a structural determinant for resistance to beta-lactam antibiotics in Gram-positive bacteria. » Proc. Natl. Acad. Sci. U S A. ; **95**(23) : 13403-13406.
- Müller DR, Schindler P, Towbin H, Wirth U, Voshol H, Hoving S, Steinmetz MO. (2001) “Isotope-tagged cross-linking reagents. A new tool in mass spectrometric protein interaction analysis.” Anal. Chem. ; **73**: 1927-1934.
- Nanao M, Ricard-Blum S, Di Guilmi AM, Lemaire D, Lascoux D, Chabert J, Attree I, Dessen A. (2003) « Type III secretion proteins PcrV and PcrG from *Pseudomonas aeruginosa* form a 1:1 complex through high affinity interactions. » BMC Microbiol. ; **3**(1):21.
- Nazabal A., Laguerre M., Schmitter J.M., Vaillier J., Chaignepain S. and Velours J. (2003) « Hydrogen/deuterium exchange on yeast ATPase supramolecular protein complex analyzed at high sensitivity by MALDI mass spectrometry. » J Am Soc Mass Spectrom ; **14** : 471–481.
- Ngoka LC, Gross ML. (1999) « A nomenclature system for labeling cyclic peptide fragments. » J Am Soc Mass Spectrom. ; **10**(4):360-3.
- Nguyen L, Paulsen IT, Tchieu J, Hueck CJ, Saier MH Jr. (2000) « Phylogenetic analyses of the constituents of Type III protein secretion systems. » J Mol Microbiol Biotechnol. ; **2**(2):125-44.
- Nilles ML, Fields KA, Straley SC. (1998) « The V antigen of *Yersinia pestis* regulates Yop vectorial targeting as well as Yop secretion through effects on YopB and LcrG. » J Bacteriol. ; **180**(13):3410-20.
- Nilles ML, Williams AW, Skrzypek E, Straley SC. (1997) « *Yersinia pestis* LcrV forms a stable complex with LcrG and may have a secretion-related regulatory role in the low-Ca<sup>2+</sup> response. » J Bacteriol. ; **179**(4):1307-16.
- Pantazatos D, Kim JS, Klock HE, Stevens RC, Wilson IA, Lesley SA, Woods VL Jr. (2004) « Rapid refinement of crystallographic protein construct definition employing enhanced hydrogen/deuterium exchange MS. » Proc Natl Acad Sci U S A. ; **101**(3):751-6.
- Parès S., Mouz N., Petillot Y., Hakenbeck R., Dideberg O. ; (1996) ; « X-ray structure of *Streptococcus pneumoniae* pbp2x, a primary penicillin target enzyme. » ; Nat. Struct. Biol. ; **3** : 284-289.
- Pearson KM, Pannell LK, Fales HM. (2002) “Intramolecular cross-linking experiments on cytochrome c and ribonuclease A using an isotope multiplet method.” Rapid Commun Mass Spectrom. ; **16**: 149.
- Pierce Biotechnology (2004) Applications Handbook and Catalog. Pierce: Rockford, IL.

- R. Kaufmann, D. Kirsch and B. Spengler (1994) « Sequencing of peptides in a time-of-flight mass spectrometer: evaluation of postsorce decay following matrix-assisted laser desorption ionisation (MALDI) » Int J Mass Spectrom Ion Proc ; **131** : 355-384.
- Rappsilber J, Siniosoglou S, Hurt EC, Mann M. (2000) « A generic strategy to analyze the spatial organization of multi-protein complexes by cross-linking and mass spectrometry. » Anal Chem. ; **72**(2):267-75.
- Redfern O, Grant A, Maibaum M, Orengo C. (2005) « Survey of current protein family databases and their application in comparative, structural and functional genomics. » J Chromatogr B Analyt Technol Biomed Life Sci. ; **815**(1-2):97-107.
- Resing KA, Hoofnagle AN, Ahn NG. (1999) « Modeling deuterium exchange behavior of ERK2 using pepsin mapping to probe secondary structure. » J Am Soc Mass Spectrom. ;**10**(8):685-702.
- Roder H, Wagner G, Wuthrich K. (1985) « Individual amide proton exchange rates in thermally unfolded basic pancreatic trypsin inhibitor. » Biochemistry. ;**24**(25):7407-11.
- Roepstorff P., Fohlman J. ; (1984) « Proposal for a common nomenclature for sequence ions in mass spectra of peptides. » ; Biomed. Mass Spectrom. ; **11**(11) : 601.
- Roepstorff P., Fohlman J. ; (1984) « Proposal for a common nomenclature for sequence ions in mass spectra of peptides. » ; Biomed. Mass Spectrom. ; **11**(11) : 601.
- Rosa, J. J. et Richards, F. M. (1979). "An experimental procedure for increasing the structural resolution of chemical hydrogen-exchange measurements on proteins: application to ribonuclease S peptide." J Mol Biol **133**(3): 399-416.
- Rose K, Savoy LA, Simona MG, Offord RE, Wingfield P. (1988) « C-terminal peptide identification by fast atom bombardment mass spectrometry. » Biochem J. ; **250**(1):253-9.
- Saier MH Jr. (2004) « Evolution of bacterial type III protein secretion systems. » Trends Microbiol. ; **12**(3):113-5.
- Sawa T, Yahr TL, Ohara M, Kurahashi K, Gropper MA, Wiener-Kronish JP, Frank DW. (1999) « Active and passive immunization with the Pseudomonas V antigen protects against type III intoxication and lung injury. » Nat Med.; **5**(4):392-8.
- Schilling B, Row RH, Gibson BW, Guo X, Young MM. (2003) «MS2Assign, automated assignment and nomenclature of tandem mass spectra of chemically cross-linked peptides.» J Am Soc Mass Spectrom ; **14**: 834-850.
- Schleifer KH, Kandler O. (1972) « Peptidoglycan types of bacterial cell walls and their taxonomic implications. » Bacteriol Rev. ; **36**(4):407-77.

- Schoehn G, Di Guilmi AM, Lemaire D, Attree I, Weissenhorn W, Dessen A. (2003) « Oligomerization of type III secretion proteins PopB and PopD precedes pore formation in *Pseudomonas*. » EMBO J. ; **22**(19):4957-67.
- Schoeniger S.J., Jacobsen R.B., Ayson M., Lane P., Novak P., Kruppa G.H., and Young M.M. (2003) « Using proteolysis and LC/MS to identify chemical crosslinks in the G-protein-coupled receptor rhodopsin. » Proceedings of the 51st ASMS Conference on Mass Spectrometry and Allied Topics, Montreal, Quebec, Canada.
- Schulz DM, Ihling C, Clore GM, Sinz A. (2004) « Mapping the topology and determination of a low-resolution three-dimensional structure of the calmodulin-melittin complex by chemical cross-linking and high-resolution FTICRMS: direct demonstration of multiple binding modes. » Biochemistry. ; **43**(16):4703-15.
- Shime N, Sawa T, Fujimoto J, Faure K, Allmond LR, Karaca T, Swanson BL, Spack EG, Wiener-Kronish JP. (2001) « Therapeutic administration of anti-PcrV F(ab')(2) in sepsis associated with *Pseudomonas aeruginosa* » J Immunol. ; **167**(10):5880-6.
- Sinz A, Wang K. (2001) « Mapping protein interfaces with a fluorogenic cross-linker and mass spectrometry: application to nebulin-calmodulin complexes. » Biochemistry. ; **40**(26):7903-13.
- Sinz A. (2003) “Chemical cross-linking and mass spectrometry for mapping three-dimensional structures of protein and complexes.” J Mass Spectrom ; **38** : 1225-1237.
- Smyth JR, Blumenfeld OO, Konigsberg W. (1964) “Reactions of N-ethylmaleimide with peptides and amino acids.” Biochem J. ; **91**: 589.
- Sory MP, Boland A, Lambermont I, Cornelis GR. (1995) « Identification of the YopE and YopH domains required for secretion and internalization into the cytosol of macrophages, using the *cyaA* gene fusion approach. » Proc Natl Acad Sci U S A. ; **92**(26):11998-2002.
- Spratt B.G., Pardee A.B. (1975) « Penicillin-binding proteins and cell shape in *E. coli*. » Nature. ; **254**(5500) : 516-517.
- Strife R.J., Kelley P.E., Weber-Grabau M. (1988) “Tandem mass spectrometry of prostaglandins : a comparison of an ion trap and a reversed geometry sector instrument” ; Rapid Commun Mass Spectrom. ; **2** : 105-109.
- Suckau D, Resemann A, Schuerenberg M, Hufnagel P, Franzen J, Holle A. (2003) “A novel MALDI LIFT-TOF/TOF mass spectrometer for proteomics.” Anal Bioanal Chem. ; **376**(7):952-65.
- Sundin C, Thelaus J, Broms JE, Forsberg A. (2004) « Polarisation of type III translocation by *Pseudomonas aeruginosa* requires PcrG, PcrV and PopN. » Microb Pathog. ; **37**(6):313-22.

- Swaim CL, Smith JB, Smith DL. (2004) « Unexpected products from the reaction of the synthetic cross-linker 3,3'-dithiobis(sulfosuccinimidyl propionate), DTSSP with peptides. » J Am Soc Mass Spectrom. ; **15**(5):736-49.
- Tabet JC. et Rebuffat S. (2003) “ Nobel Prize 2002 for chemistry: mass spectrometry and nuclear magnetic resonance.” Med. Sci. ; **19**(8-9):865-72.
- Taverner T, Hall NE, O'Hair RA, Simpson RJ. (2002) « Characterization of an antagonist interleukin-6 dimer by stable isotope labeling, cross-linking, and mass spectrometry. » J Biol Chem. ; **277**(48):46487-92.
- Tipper DJ, Strominger JL. (1965) « Mechanism of action of penicillins: a proposal based on their structural similarity to acyl-D-alanyl-D-alanine. » Proc Natl Acad Sci U S A. ; **54**(4) : 1133-41.
- Trester-Zedlitz M, Kamada K, Burley SK, Fenyő D, Chait BT, Muir TW. (2003) “A modular approach for exploring protein interactions.” J Am Chem Soc.; **125**: 2416.
- Verentchikov AN, Ens W, Standing KG. (1994) “Reflecting time-of-flight mass spectrometer with an electrospray ion source and orthogonal extraction.” Anal Chem. ; **66**(1):126-33.
- Vestal ML, Juhasz P, Martin SA. (1995). "Delayed extraction matrix-assisted laser desorption time-of-flight mass spectrometry." Rapid Commun Mass Spectrom ; **9** : 1044-1050
- Wattiau P, Bernier B, Deslee P, Michiels T, Cornelis GR. (1994) « Individual chaperones required for Yop secretion by Yersinia. » Proc Natl Acad Sci U S A. ; **91**(22):10493-7.
- Wei J, Buriak JM, Siuzdak G. (1999). “Desorption-ionization mass spectrometry on porous silicon.” Nature. ; **399**(6733):243-6.
- Wine RN, Dial JM, Tomer KB, Borchers CH. (2002) « Identification of components of protein complexes using a fluorescent photo-cross-linker and mass spectrometry. » Anal Chem.; **74**(9):1939-45.
- Woestyn S, Sory MP, Boland A, Lequenne O, Cornelis GR. (1996) « The cytosolic SycE and SycH chaperones of Yersinia protect the region of YopE and YopH involved in translocation across eukaryotic cell membranes. » Mol Microbiol. ; **20**(6):1261-71.
- Wong S. F., Meng C. K., and Fenn J. B. (1988) “Multiple Charging in Electrospray Ionization of Poly(ethylene glycols)” J. Phys. Chem. ; **92**, 546-550.
- Wong SS. (1993) Chemistry of Protein Conjugation and Crosslinking, CRC Press: Boca Raton, FL.
- Woodward C, Simon I, Tuchsén E. (1982) « Hydrogen exchange and the dynamic structure of proteins. » Mol Cell Biochem.; **48**(3):135-60.

- Yahr T.L., Mende-Mueller L.M., Friese M.B. et Frank D.W. (1997). « Identification of type III secreted products of the *Pseudomonas aeruginosa* exoenzyme S regulon. » J Bacteriol ; **179** :7165–7168
- Yahr TL, Vallis AJ, Hancock MK, Barbieri JT, Frank DW. (1998) « ExoY, an adenylate cyclase secreted by the *Pseudomonas aeruginosa* type III system. » Proc Natl Acad Sci U S A ; **95**(23):13899-904.
- Yamada N, Suzuki E, Hirayama K. (2002) « Identification of the interface of a large protein-protein complex using H/D exchange and Fourier transform ion cyclotron resonance mass spectrometry. » Rapid Commun Mass Spectrom. ;**16**(4):293-9.
- Young MM, Tang N, Hempel JC, Oshiro CM, Taylor EW, Kuntz ID, Gibson BW, Dollinger G. (2000) “High Throughput Fold Identification by Using Experimental Constraints Derivated from Intramolecular Cross-Links and Mass Spectrometry.” Proc Natl Acad Sci U S A ; **97**: 5802-5806.
- Yousif SY, Broome-Smith JK, Spratt BG. (1985) « Lysis of *Escherichia coli* by beta-lactam antibiotics: deletion analysis of the role of penicillin-binding proteins 1A and 1B. » J Gen Microbiol. ; **131**:2839-45.
- Zhang Z, Post CB, Smith DL. (1996) « Amide hydrogen exchange determined by mass spectrometry: application to rabbit muscle aldolase. » Biochemistry. ;**35**(3):779-91.
- Zhang, Z. et Smith, D. L. (1993). "Determination of amide hydrogen exchange by mass spectrometry: a new tool for protein structure elucidation." Protein Sci **2**(4): 522-31.

# ANNEXES

**ANNEXE I : Incorporation de deutérium dans les peptides issus de la digestion de PBP-2X**

Les valeurs indiquées sont des moyennes de plusieurs analyses.

Les peptides générés par la pepsine sont en style normal, ceux générés par la protéase de type XVIII sont en gras, et ceux générés par la protéase de type XIII sont en italique.

masse calculée	m/z mesuré	séquence début	séquence fin	nombre de deutériums possible	nombre de deutérium incorporé	taux deutération
<b>436.2</b>	<b>437.15</b>	<b>341</b>	<b>343</b>	<b>2</b>	<b>0.5</b>	<b>25.0</b>
<b>445.2</b>	<b>446.15</b>	<b>33</b>	<b>36</b>	<b>3</b>	<b>0</b>	<b>0.0</b>
<i>450.2</i>	<i>451.1</i>	<i>313</i>	<i>316</i>	<i>3</i>	<i>0</i>	<i>0.0</i>
473.3	474.16	183	187	4	0	0.0
<b>473.3</b>	<b>474.2</b>	<b>598</b>	<b>601</b>	<b>3</b>	<b>0</b>	<b>0.0</b>
<b>483.3</b>	<b>484.25</b>	<b>617</b>	<b>621</b>	<b>4</b>	<b>0</b>	<b>0.0</b>
<b>486.2</b>	<b>487.2</b>	<b>575</b>	<b>578</b>	<b>3</b>	<b>0.5</b>	<b>16.7</b>
<b>487.2</b>	<b>488.25</b>	<b>601</b>	<b>605</b>	<b>4</b>	<b>0.5</b>	<b>12.5</b>
<b>492.2</b>	<b>493.2</b>	<b>478</b>	<b>481</b>	<b>3</b>	<b>0</b>	<b>0.0</b>
<b>505.2</b>	<b>506.3</b>	<b>627</b>	<b>631</b>	<b>4</b>	<b>1</b>	<b>25.0</b>
<b>506.3</b>	<b>507.2</b>	<b>539</b>	<b>542</b>	<b>3</b>	<b>0</b>	<b>0.0</b>
<b>521.2</b>	<b>522.2</b>	<b>289</b>	<b>292</b>	<b>3</b>	<b>0</b>	<b>0.0</b>
<i>523.2</i>	<i>524.1</i>	<i>173</i>	<i>177</i>	<i>4</i>	<i>1</i>	<i>25.0</i>
<b>523.25</b>	<b>524.2</b>	<b>151</b>	<b>155</b>	<b>4</b>	<b>0</b>	<b>0.0</b>
<b>528.2</b>	<b>529.15</b>	<b>342</b>	<b>345</b>	<b>3</b>	<b>0</b>	<b>0.0</b>
<b>533.3</b>	<b>534.2</b>	<b>663</b>	<b>667</b>	<b>4</b>	<b>0</b>	<b>0.0</b>
<b>546.3</b>	<b>547.1</b>	<b>411</b>	<b>415</b>	<b>4</b>	<b>1</b>	<b>25.0</b>
<b>553.2</b>	<b>554.2</b>	<b>193</b>	<b>196</b>	<b>3</b>	<b>0</b>	<b>0.0</b>
<b>569.3</b>	<b>571.1</b>	<b>29</b>	<b>33</b>	<b>4</b>	<b>0</b>	<b>0.0</b>
<b>578.3</b>	<b>579.3</b>	<b>133</b>	<b>137</b>	<b>4</b>	<b>0</b>	<b>0.0</b>
<b>580.3</b>	<b>581.2</b>	<b>115</b>	<b>119</b>	<b>4</b>	<b>1</b>	<b>25.0</b>
<b>592.3</b>	<b>593.2</b>	<b>66</b>	<b>70</b>	<b>4</b>	<b>0</b>	<b>0.0</b>
<b>593.3</b>	<b>594.25</b>	<b>477</b>	<b>481</b>	<b>4</b>	<b>0</b>	<b>0.0</b>
602.3	603.3	473	478	5	0.5	10.0
<b>606.3</b>	<b>607.3</b>	<b>279</b>	<b>282</b>	<b>3</b>	<b>0</b>	<b>0.0</b>
<b>610.2</b>	<b>611.2</b>	<b>4</b>	<b>9</b>	<b>5</b>	<b>0</b>	<b>0.0</b>
<b>613.2</b>	<b>614.2</b>	<b>533</b>	<b>538</b>	<b>5</b>	<b>0</b>	<b>0.0</b>
613.3	614.3	155	160	5	1	20.0
617.3	618.2	250	255	5	0	0.0
<b>618.3</b>	<b>619.2</b>	<b>187</b>	<b>192</b>	<b>5</b>	<b>0.5</b>	<b>10.0</b>
620.3	621.2	297	301	4	0	0.0
<b>629.4</b>	<b>630.3</b>	<b>123</b>	<b>127</b>	<b>4</b>	<b>0</b>	<b>0.0</b>
<b>631.3</b>	<b>632.2</b>	<b>471</b>	<b>476</b>	<b>5</b>	<b>0</b>	<b>0.0</b>
<b>642.3</b>	<b>643.3</b>	<b>486</b>	<b>491</b>	<b>5</b>	<b>0.5</b>	<b>10.0</b>
<b>650.3</b>	<b>651.3</b>	<b>514</b>	<b>519</b>	<b>5</b>	<b>1</b>	<b>20.0</b>
<b>671.3</b>	<b>672.2</b>	<b>601</b>	<b>607</b>	<b>6</b>	<b>2</b>	<b>33.3</b>
673.4	674.3	560	565	5	0	0.0
677.3	678.3	419	424	5	0	0.0

<b>681.2</b>	<b>682.2</b>	<b>388</b>	<b>393</b>	<b>5</b>	<b>1</b>	<b>20.0</b>
691.3	692.2	78	83	5	0.5	10.0
692.3	693.1	628	634	6	0	0.0
<b>698.3</b>	<b>699.2</b>	<b>40</b>	<b>46</b>	<b>6</b>	<b>1</b>	<b>16.7</b>
699.4	700.1	456	462	6	0	0.0
<b>718.3</b>	<b>719.2</b>	<b>651</b>	<b>656</b>	<b>5</b>	<b>0</b>	<b>0.0</b>
<b>721.4</b>	<b>722.3</b>	<b>240</b>	<b>245</b>	<b>5</b>	<b>0</b>	<b>0.0</b>
<b>730.4</b>	<b>732.2</b>	<b>322</b>	<b>328</b>	<b>6</b>	<b>2</b>	<b>33.3</b>
731.4	732.3	701	707	6	0	0.0
<b>732.4</b>	<b>733.3</b>	<b>379</b>	<b>385</b>	<b>6</b>	<b>0</b>	<b>0.0</b>
<b>733.3</b>	<b>734.3</b>	<b>472</b>	<b>478</b>	<b>6</b>	<b>0</b>	<b>0.0</b>
<b>736.3</b>	<b>737.3</b>	<b>107</b>	<b>114</b>	<b>7</b>	<b>1</b>	<b>14.3</b>
748.4	749.3	262	267	5	1	20.0
749.3	750.3	475	481	6	0	0.0
<b>756.25</b>	<b>757.3</b>	<b>89</b>	<b>94</b>	<b>5</b>	<b>1</b>	<b>20.0</b>
<b>766.5</b>	<b>767.2</b>	<b>614</b>	<b>620</b>	<b>6</b>	<b>0</b>	<b>0.0</b>
770.4	771.1	522	527	5	1.5	30.0
<b>780.3</b>	<b>781.3</b>	<b>136</b>	<b>142</b>	<b>6</b>	<b>0.5</b>	<b>8.3</b>
<b>781.4</b>	<b>782.4</b>	<b>512</b>	<b>518</b>	<b>6</b>	<b>0</b>	<b>0.0</b>
<b>786.4</b>	<b>787.3</b>	<b>670</b>	<b>675</b>	<b>5</b>	<b>0</b>	<b>0.0</b>
<b>788.4</b>	<b>789.4</b>	<b>534</b>	<b>540</b>	<b>6</b>	<b>0</b>	<b>0.0</b>
800.45	801.25	280	285	5	2	40.0
802.4	803.3	560	566	6	1	16.7
<b>807.4</b>	<b>809.2</b>	<b>352</b>	<b>359</b>	<b>7</b>	<b>2</b>	<b>28.6</b>
<b>815.2</b>	<b>817.2</b>	<b>128</b>	<b>135</b>	<b>7</b>	<b>0</b>	<b>0.0</b>
817.4	818.3	252	259	7	0	0.0
<b>832.4</b>	<b>835.2</b>	<b>409</b>	<b>416</b>	<b>7</b>	<b>1</b>	<b>14.3</b>
833.4	834.4	468	474	6	0	0.0
<b>847.6</b>	<b>848.2</b>	<b>471</b>	<b>478</b>	<b>7</b>	<b>0.5</b>	<b>7.1</b>
<b>859.5</b>	<b>860.3</b>	<b>700</b>	<b>707</b>	<b>7</b>	<b>0</b>	<b>0.0</b>
867.3	868.2	172	180	8	1	12.5
<b>870.5</b>	<b>871.4</b>	<b>617</b>	<b>625</b>	<b>8</b>	<b>0</b>	<b>0.0</b>
873.5	874.3	501	509	8	3	37.5
<b>894.4</b>	<b>895.5</b>	<b>512</b>	<b>519</b>	<b>7</b>	<b>2</b>	<b>28.6</b>
<b>909.5</b>	<b>910.1</b>	<b>16</b>	<b>23</b>	<b>7</b>	<b>0</b>	<b>0.0</b>
910.4	911.3	6	14	8	3	37.5
<b>913.4</b>	<b>914.3</b>	<b>70</b>	<b>76</b>	<b>6</b>	<b>1</b>	<b>16.7</b>
913.4	914.2	70	76	6	2	33.3
916.5	917.3	251	259	8	0.5	6.3
<b>916.5</b>	<b>917.3</b>	<b>251</b>	<b>259</b>	<b>8</b>	<b>1</b>	<b>12.5</b>
920.4	921.3	260	267	7	0	0.0
<b>923.4</b>	<b>925.2</b>	<b>191</b>	<b>197</b>	<b>6</b>	<b>1</b>	<b>16.7</b>
927.5	928.3	455	463	8	2	25.0
931.4	932.2	300	308	8	1	12.5
<b>941.5</b>	<b>942.3</b>	<b>454</b>	<b>462</b>	<b>8</b>	<b>0</b>	<b>0.0</b>
<b>942.4</b>	<b>943.4</b>	<b>5</b>	<b>13</b>	<b>8</b>	<b>1</b>	<b>12.5</b>
947.4	948.2	390	398	8	2	25.0
947.5	948.4	279	285	6	2	33.3
<b>947.5</b>	<b>948.5</b>	<b>279</b>	<b>285</b>	<b>6</b>	<b>1</b>	<b>16.7</b>
<b>947.5</b>	<b>948.5</b>	<b>554</b>	<b>562</b>	<b>8</b>	<b>0</b>	<b>0.0</b>
<b>950.5</b>	<b>951.35</b>	<b>376</b>	<b>383</b>	<b>7</b>	<b>1</b>	<b>14.3</b>
<b>958.5</b>	<b>959.4</b>	<b>560</b>	<b>567</b>	<b>7</b>	<b>1</b>	<b>14.3</b>
<b>968.4</b>	<b>969.2</b>	<b>231</b>	<b>238</b>	<b>7</b>	<b>1</b>	<b>14.3</b>
<b>969.5</b>	<b>970.3</b>	<b>38</b>	<b>46</b>	<b>8</b>	<b>1</b>	<b>12.5</b>

<b>979.6</b>	<b>980.4</b>	<b>61</b>	<b>69</b>	<b>8</b>	<b>0.5</b>	<b>6.3</b>
<b>980.4</b>	<b>981.2</b>	<b>171</b>	<b>180</b>	<b>9</b>	<b>1</b>	<b>11.1</b>
982.4	983.2	266	274	8	1	12.5
<b>985.6</b>	<b>986.2</b>	<b>667</b>	<b>674</b>	<b>7</b>	<b>0</b>	<b>0.0</b>
985.6	986.4	668	675	7	1	14.3
988.5	989.3	572	586	14	1	7.1
992.5	993.2	71	78	7	0	0.0
<b>995.5</b>	<b>996.4</b>	<b>636</b>	<b>644</b>	<b>8</b>	<b>1</b>	<b>12.5</b>
1007.5	1008.25	1	10	9	6	66.7
1016.5	1017.3	332	341	9	6	66.7
1034.5	1035.2	104	113	9	1.5	16.7
1038.5	1039.3	475	484	9	1	11.1
1039.5	1040.3	519	527	8	0	0.0
<b>1040.6</b>	<b>1042.4</b>	<b>561</b>	<b>570</b>	<b>9</b>	<b>1</b>	<b>11.1</b>
<b>1043.5</b>	<b>1044.3</b>	<b>456</b>	<b>466</b>	<b>10</b>	<b>1</b>	<b>10.0</b>
<b>1044.5</b>	<b>1045.25</b>	<b>600</b>	<b>609</b>	<b>9</b>	<b>0.5</b>	<b>5.6</b>
<b>1045.5</b>	<b>1046.5</b>	<b>470</b>	<b>479</b>	<b>9</b>	<b>1</b>	<b>11.1</b>
<b>1046.45</b>	<b>1047.15</b>	<b>331</b>	<b>340</b>	<b>9</b>	<b>1.5</b>	<b>16.7</b>
<b>1055.6</b>	<b>1056.3</b>	<b>96</b>	<b>104</b>	<b>8</b>	<b>0</b>	<b>0.0</b>
1057.5	1058.3	5	14	9	7	77.8
1057.5	1058.25	5	14	9	2	22.2
<b>1059.5</b>	<b>1060.3</b>	<b>323</b>	<b>331</b>	<b>8</b>	<b>4</b>	<b>50.0</b>
<b>1070.4</b>	<b>1071.2</b>	<b>364</b>	<b>372</b>	<b>8</b>	<b>1</b>	<b>12.5</b>
<b>1070.6</b>	<b>1071.4</b>	<b>497</b>	<b>507</b>	<b>10</b>	<b>0.5</b>	<b>5.0</b>
<b>1078.5</b>	<b>1079.2</b>	<b>463</b>	<b>472</b>	<b>9</b>	<b>1</b>	<b>11.1</b>
1079.5	1080.3	220	229	9	1	11.1
<b>1100.5</b>	<b>1101.2</b>	<b>597</b>	<b>607</b>	<b>10</b>	<b>1</b>	<b>10.0</b>
<b>1104.6</b>	<b>1105.4</b>	<b>28</b>	<b>37</b>	<b>9</b>	<b>1.5</b>	<b>16.7</b>
<b>1110.6</b>	<b>1111.25</b>	<b>241</b>	<b>250</b>	<b>9</b>	<b>0.5</b>	<b>5.6</b>
1114.6	1115.5	321	330	9	4	44.4
1132.5	1133.4	523	532	9	1	11.1
1148.5	1149.4	207	216	9	1	11.1
1148.5	1149.3	408	418	10	1	10.0
<b>1154.5</b>	<b>1155.3</b>	<b>1</b>	<b>11</b>	<b>10</b>	<b>7</b>	<b>70.0</b>
1154.5	1155.25	1	11	10	7	70.0
<b>1157.5</b>	<b>1158.3</b>	<b>600</b>	<b>610</b>	<b>10</b>	<b>1</b>	<b>10.0</b>
1162.5	1163.3	349	359	10	2	20.0
1177.5	1178.3	360	369	9	0.5	5.6
1186.5	1187.3	4	14	10	7	70.0
1192.5	1193.25	268	278	10	2	20.0
1210.6	1211.3	220	230	10	1	10.0
1219.6	1220.3	399	410	11	1	9.1
1227.6	1228.3	161	171	10	0	0.0
1227.6	1228.5	548	558	10	1	10.0
1250.6	1251.4	473	484	11	3.5	31.8
1256.6	1257.4	60	70	10	2	20.0
1273.6	1274.3	390	401	11	4	36.4
1275.5	1276.4	286	296	10	3	30.0
1290.6	1291.6	360	370	10	2	20.0
1297.6	1298.3	657	667	10	1	10.0
1301.7	1302.4	501	513	12	4	33.3
<b>1309.7</b>	<b>1310.3</b>	<b>633</b>	<b>644</b>	<b>11</b>	<b>2</b>	<b>18.2</b>
1333.6	667.7	331	342	11	8	72.7
1333.6	667.8	566	577	11	5	45.5

1339.6	1340.4	647	658	11	3	27.3
1341.7	1342.4	218	229	11	1	9.1
<b>1342.7</b>	<b>1343.5</b>	<b>555</b>	<b>566</b>	<b>11</b>	<b>1</b>	<b>9.1</b>
1350.6	1351.3	528	540	12	2	16.7
1367.6	1368.3	345	357	12	3	25.0
1377.7	1378.4	419	431	12	2	16.7
1406.6	1407.3	543	554	11	2	18.2
<b>1407.7</b>	<b>1408.4</b>	<b>171</b>	<b>184</b>	<b>13</b>	<b>0</b>	<b>0.0</b>
<b>1418.6</b>	<b>1419.3</b>	<b>530</b>	<b>542</b>	<b>12</b>	<b>2</b>	<b>16.7</b>
1418.7	1419.4	530	542	12	2	16.7
1427.6	1428.5	1	14	13	9	69.2
1427.6	1428.3	1	14	13	9	69.2
<b>1540.7</b>	<b>1541.4</b>	<b>1</b>	<b>15</b>	<b>14</b>	<b>10</b>	<b>71.4</b>
1565.7	783.8	331	344	13	8	61.5
<b>1576.7</b>	<b>1577.3</b>	<b>528</b>	<b>542</b>	<b>14</b>	<b>2</b>	<b>14.3</b>
1576.75	1577.4	528	542	14	2	14.3
1581.7	1582.3	345	359	14	3	21.4
1592.8	797.2	77	89	12	1	8.3
<b>1611.7</b>	<b>1612.4</b>	<b>40</b>	<b>54</b>	<b>14</b>	<b>0</b>	<b>0.0</b>
1614.7	808.6	137	150	13	1	7.7
<b>1620.8</b>	<b>811.3</b>	<b>472</b>	<b>487</b>	<b>15</b>	<b>0</b>	<b>0.0</b>
1634.8	818.2	301	316	15	2	13.3
1642.8	822.3	237	250	13	0.5	3.8
1698.8	850.2	485	500	15	3	20.0
1758.9	880.4	371	384	13	0.5	3.8
1781.8	891.75	402	418	16	2.5	15.6
1783.9	892.8	55	70	15	5	33.3
1783.9	892.8	443	458	15	5	33.3
1784.8	893.25	104	120	16	5	31.3
1805.9	903.8	258	273	15	0	0.0
1813.8	907.7	343	359	16	4	25.0
<b>1922.9</b>	<b>962.4</b>	<b>422</b>	<b>439</b>	<b>17</b>	<b>2</b>	<b>11.8</b>
1934	967.7	501	519	18	7	38.9
<b>1939.9</b>	<b>970.7</b>	<b>1</b>	<b>19</b>	<b>18</b>	<b>10</b>	<b>55.6</b>
1946.9	974.3	543	559	16	6	37.5
<b>1955.1</b>	<b>978.3</b>	<b>21</b>	<b>37</b>	<b>16</b>	<b>4</b>	<b>25.0</b>
<b>1974.9</b>	<b>988.25</b>	<b>283</b>	<b>299</b>	<b>16</b>	<b>4</b>	<b>25.0</b>
1992	997.3	419	436	17	1	5.9
<b>1992</b>	<b>996.8</b>	<b>419</b>	<b>436</b>	<b>17</b>	<b>6</b>	<b>35.3</b>
<b>1999.9</b>	<b>1000.75</b>	<b>526</b>	<b>543</b>	<b>17</b>	<b>0.5</b>	<b>2.9</b>
<b>2006.9</b>	<b>1004.2</b>	<b>137</b>	<b>154</b>	<b>17</b>	<b>1</b>	<b>5.9</b>
<b>2039</b>	<b>1020.3</b>	<b>626</b>	<b>644</b>	<b>18</b>	<b>4</b>	<b>22.2</b>
2051.9	1026.75	301	320	19	4	21.1
<b>2068.9</b>	<b>1035.3</b>	<b>650</b>	<b>667</b>	<b>17</b>	<b>0</b>	<b>0.0</b>
<b>2069.2</b>	<b>1035.4</b>	<b>689</b>	<b>707</b>	<b>18</b>	<b>2</b>	<b>11.1</b>
<b>2083.2</b>	<b>1042.3</b>	<b>20</b>	<b>37</b>	<b>17</b>	<b>0</b>	<b>0.0</b>
2090.1	1046.35	501	521	20	8	40.0
2094.9	1048.3	260	278	18	6	33.3
<b>2094.9</b>	<b>1048.3</b>	<b>260</b>	<b>278</b>	<b>18</b>	<b>4</b>	<b>22.2</b>
2108	1054.8	399	418	19	2	10.5
<b>2108</b>	<b>1055.3</b>	<b>399</b>	<b>418</b>	<b>19</b>	<b>0</b>	<b>0.0</b>
<b>2109</b>	<b>1005.5</b>	<b>303</b>	<b>322</b>	<b>19</b>	<b>1</b>	<b>5.3</b>
<b>2129</b>	<b>1065.5</b>	<b>70</b>	<b>86</b>	<b>16</b>	<b>0</b>	<b>0.0</b>
<b>2180</b>	<b>1090.9</b>	<b>302</b>	<b>322</b>	<b>20</b>	<b>4</b>	<b>20.0</b>

2205	1103.25	279	296	17	4	23.5
<b>2263.2</b>	<b>1132.3</b>	<b>419</b>	<b>439</b>	<b>20</b>	<b>1.5</b>	<b>7.5</b>
2272.1	1136.8	586	606	20	7	35.0
2274.1	1137.9	371	389	18	1	5.6
<b>2292</b>	<b>1146.7</b>	<b>134</b>	<b>154</b>	<b>20</b>	<b>0</b>	<b>0.0</b>
<b>2357</b>	<b>1179.5</b>	<b>70</b>	<b>88</b>	<b>18</b>	<b>0</b>	<b>0.0</b>
<b>2368.3</b>	<b>1184.8</b>	<b>613</b>	<b>635</b>	<b>22</b>	<b>6</b>	<b>27.3</b>
2389.1	1195.7	158	180	22	5	22.7
<b>2556.2</b>	<b>1278.8</b>	<b>171</b>	<b>195</b>	<b>24</b>	<b>2</b>	<b>8.3</b>
<b>2563.2</b>	<b>1282.4</b>	<b>279</b>	<b>299</b>	<b>20</b>	<b>3</b>	<b>15.0</b>
<b>2586.3</b>	<b>1293.8</b>	<b>440</b>	<b>462</b>	<b>22</b>	<b>5</b>	<b>22.7</b>
<b>2672.3</b>	<b>1337.3</b>	<b>155</b>	<b>180</b>	<b>25</b>	<b>1</b>	<b>4.0</b>
2673.3	1336.9	155	180	25	5	20.0
<b>2923.4</b>	<b>1462.4</b>	<b>650</b>	<b>674</b>	<b>24</b>	<b>2</b>	<b>8.3</b>
<b>3099.5</b>	<b>1550.3</b>	<b>155</b>	<b>184</b>	<b>29</b>	<b>1</b>	<b>3.4</b>

## ANNEXE 2 : Incorporation de deutérium dans les peptides issus de la digestion de PcrV

Les valeurs indiquées sont des moyennes de plusieurs analyses.

Les peptides générés par la pepsine sont en style normal, ceux générés par la protéase de type XVIII sont en gras, et ceux générés par la protéase de type XIII sont en italique.

masse calculée	m/z mesuré	séquence début	séquence fin	nombre de deutériums possibles	nombre de deutérium incorporé	taux de deutération	nombre de deutérium incorporé dans le complexe	différence de deutération protéine seule / protéine dans le complexe
464.2	465.2	97	100	3	0.5	16.7	0.3	0.2
473.2	474.3	261	264	3	1.4	45.0	1.2	0.2
<b>474.3</b>	<b>475.3</b>	<b>153</b>	<b>157</b>	<b>4</b>	<b>0.2</b>	<b>4.2</b>		
<b>486.3</b>	<b>487.4</b>	<b>54</b>	<b>57</b>	<b>3</b>	<b>0</b>	<b>0.0</b>		
492.3	493.3	174	177	3	0.4	13.3		
501.5	502.3	59	62	3	0.5	16.7	0.6	-0.1
502.3	503.3	136	140	4	1	25.0		
<b>506.3</b>	<b>507.3</b>	<b>12</b>	<b>15</b>	<b>3</b>	<b>0</b>	<b>0.0</b>		
515.3	516.3	162	165	3	1.75	58.3		
515.3	516.4	288	291	3	0.5	16.7	0.5	0
515.3	516.3	290	294	4	0.25	6.3	0.25	0
517.3	518.3	289	293	4	0.8	20.8	1	-0.2
<b>523.3</b>	<b>524.3</b>	<b>179</b>	<b>183</b>	<b>4</b>	<b>0.8</b>	<b>20.8</b>	<b>0.8</b>	<b>0</b>
535.3	536.3	97	101	4	1	25.0	1	0
548.2	549.3	196	200	4	0	0.0		
551.3	552.4	98	102	4	2.8	71.9		
572.3	573.5	161	165	4	1.3	33.3		
<b>573.3</b>	<b>574.4</b>	<b>147</b>	<b>151</b>	<b>4</b>	<b>0.3</b>	<b>8.3</b>		
581.3	582.3	66	71	3	0.8	26.7		
<b>595.3</b>	<b>596.4</b>	<b>43</b>	<b>48</b>	<b>5</b>	<b>0.2</b>	<b>3.3</b>		
601.3	602.4	14	18	4	1.25	31.3	1	0.3
<b>613.3</b>	<b>614.5</b>	<b>180</b>	<b>185</b>	<b>4</b>	<b>0.8</b>	<b>20.8</b>		
614.4	615.4	59	63	4	1	25.0	1	0.0
619.3	620.5	275	279	4	1	25.0	1	0
620.3	620.4	179	184	5	1	20.0	1	0
627.4	628.5	247	251	3	1.4	47.0	1.4	0
628.4	629.5	53	58	5	1.5	30.0	1.3	0.2
628.4	629.1	287	291	4	0.8	20.0	1	-0.2
634.3	635	9	13	4	2	50.0	1.9	0.1
<b>634.4</b>	<b>635.4</b>	<b>207</b>	<b>211</b>	<b>4</b>	<b>1</b>	<b>25.0</b>	<b>1</b>	<b>0</b>
<b>655.4</b>	<b>656.5</b>	<b>86</b>	<b>91</b>	<b>4</b>	<b>1.5</b>	<b>37.5</b>		
656.4	657.5	171	176	4	1.4	35.0		
<b>657.4</b>	<b>658.5</b>	<b>4</b>	<b>9</b>	<b>5</b>	<b>1</b>	<b>20.0</b>	<b>1</b>	<b>0.0</b>
<b>661.4</b>	<b>662.4</b>	<b>276</b>	<b>280</b>	<b>4</b>	<b>1</b>	<b>25.0</b>		
<b>677.3</b>	<b>678.4</b>	<b>178</b>	<b>184</b>	<b>6</b>	<b>0.8</b>	<b>14.3</b>	<b>1</b>	<b>-0.2</b>
678.3	679.4	93	97	4	0	0.0		

<b>679.4</b>	<b>680.5</b>	<b>115</b>	<b>120</b>	<b>5</b>	<b>1.4</b>	<b>28.0</b>	<b>0.5</b>	<b>1</b>
684.4	685.4	62	68	5	0.2	4.0		
698.4	699.5	97	102	5	2	40.0	2	0
735.4	736.5	116	121	5	1.3	26.0	0.3	1
<b>744.4</b>	<b>745.4</b>	135	141	6	0.8	13.9		
744.4	745.5	136	142	6	3	50.0	1.7	1.3
747.4	748.5	9	14	5	2	40.0	2	0.0
748.4	749.5	13	18	5	0.5	10.0	0.5	0.0
<b>757.4</b>	<b>758.4</b>	<b>215</b>	<b>221</b>	<b>6</b>	<b>2</b>	<b>33.3</b>		
788.4	789.5	133	139	6	1.7	29.0		
<b>800.4</b>	<b>801.4</b>	<b>130</b>	<b>136</b>	<b>6</b>	<b>1</b>	<b>16.7</b>		
814.5	815.5	2	8	6	2.5	41.6	2.8	-0.3
<b>818.4</b>	<b>819.4</b>	<b>235</b>	<b>242</b>	<b>6</b>	<b>1.2</b>	<b>19.4</b>		
<b>827.3</b>	<b>828.3</b>	<b>58</b>	<b>65</b>	<b>7</b>	<b>1</b>	<b>14.3</b>		
<b>828.4</b>	<b>829.3</b>	<b>103</b>	<b>109</b>	<b>6</b>	<b>1</b>	<b>16.7</b>	<b>1</b>	<b>0</b>
<b>831.4</b>	<b>832.3</b>	<b>123</b>	<b>129</b>	<b>6</b>	<b>1.2</b>	<b>19.4</b>		
<b>840.3</b>	<b>841.3</b>	<b>177</b>	<b>184</b>	<b>7</b>	<b>1</b>	<b>14.3</b>	<b>1</b>	<b>0</b>
<b>854.4</b>	<b>855.4</b>	<b>68</b>	<b>76</b>	<b>6</b>	<b>2</b>	<b>33.3</b>		
<b>858.4</b>	<b>859.4</b>	<b>153</b>	<b>161</b>	<b>8</b>	<b>1.2</b>	<b>14.6</b>		
<b>865.4</b>	<b>866.4</b>	<b>206</b>	<b>213</b>	<b>7</b>	<b>1</b>	<b>14.3</b>	<b>1</b>	<b>0</b>
<b>870.4</b>	<b>871.4</b>	<b>48</b>	<b>55</b>	<b>6</b>	<b>1.5</b>	<b>25.0</b>		
871.5	872.6	37	43	6	1	16.7	1	0.0
908.5	909.5	44	52	7	2.2	31.3		
908.5	909.5	70	78	7	4	56.3	4	0
<b>911.4</b>	<b>912.4</b>	<b>228</b>	<b>234</b>	<b>5</b>	<b>1.3</b>	<b>26.7</b>		
<b>913.4</b>	<b>914.3</b>	<b>87</b>	<b>94</b>	<b>6</b>	<b>1</b>	<b>16.7</b>		
<b>964.5</b>	<b>965.4</b>	<b>280</b>	<b>287</b>	<b>7</b>	<b>2</b>	<b>28.6</b>	<b>2</b>	<b>0</b>
964.5	965.6	280	287	7	2.75	39.3	2.5	0.25
972.4	973.4	24	32	8	3	37.5	2.8	0.2
<b>978.5</b>	<b>979.4</b>	<b>205</b>	<b>213</b>	<b>8</b>	<b>1</b>	<b>12.5</b>	<b>1</b>	<b>0</b>
988.5	989.6	133	141	8	3	37.5	1.7	1.3
1050.6	1051.6	145	153	8	1.5	18.8	1.5	0
<b>1058.5</b>	<b>1059.4</b>	<b>181</b>	<b>189</b>	<b>7</b>	<b>2</b>	<b>28.6</b>		
1072.5	1073.6	103	111	8	2	25.0	2.2	-0.2
1101.6	1102.6	133	142	9	4.5	50.0		
1119.6	1120.7	112	121	9	3.75	41.7		
1121.6	1122.7	145	154	9	3.5	38.9	3.5	0
1170.6	1171.6	2	11	9	4.75	52.8	4.8	0.0
1191.6	1192.7	63	75	9	4.3	47.9	4.5	-0.2
1203.5	1204.5	231	241	9	1.8	20.0		
<b>1207.7</b>	<b>1208.4</b>	<b>37</b>	<b>47</b>	<b>10</b>	<b>1.5</b>	<b>16.7</b>		
1209.6	599.75	199	209	10	4.25	42.5	4.1	0.15
1235.6	1236.7	102	111	9	3.75	41.7	3.75	0
<b>1248.6</b>	<b>1249.8</b>	<b>210</b>	<b>221</b>	<b>10</b>	<b>1.2</b>	<b>11.7</b>	<b>1.2</b>	<b>0</b>
1251.5	1252.3	227	236	8	6	75.0		
1305.7	1306.7	145	156	11	4.75	43.2	4.5	0.25
1319.7	1320.8	44	56	11	4.3	39.3	4.1	0.2
1325.6	1326.7	262	273	11	4.5	40.9		
1343.6	1344.6	19	32	12	5.5	46.2	5.7	-0.2
1381.7	1382.8	277	287	10	4.25	42.5	4	0.25
1399.8	701	157	170	13	5	38.5	5	0
1419.7	1420.7	63	77	14	6.3	44.6	6.5	-0.2
1454.7	1455.7	262	274	12	5.7	47.9	4	1.7
1456.7	1457.7	18	32	13	6.3	48.2	6.2	0.1

1508	755	79	91	11	4.5	41.7	4.5	0
1518.8	760.4	44	58	13	5.5	42.9	5.5	0.0
1532.8	1533.9	63	78	12	5.6	46.7	5.9	-0.3
1565.9	784.1	72	85	13	6	46.2	5.8	0.2
1638.8	1639.8	262	276	14	7	50.0	5.5	1.5
<b>1688.9</b>	<b>845.8</b>	<b>63</b>	<b>79</b>	<b>13</b>	<b>1.8</b>	<b>14.1</b>		
1709.9	856	83	96	12	6.5	53.8	6.3	0.2
1901.8	951.9	177	192	13	6	46.7	5	1
2086	1044	177	194	15	4.5	29.4	3	1.5
2180	1091	-18	1	19	4.5	23.7	4.33	0.2
2199	1100.5	177	195	16	3.5	22.2	1.5	2
2344.3	1173.1	241	261	19	5.7	30.0	4	1.7
2358.3	1180.2	112	132	20	5.4	27.0	2.5	3
3468.7	1171.7	210	240	27	13	48.3	11.5	1.5

### **ANNEXE 3 : Incorporation de deutérium dans les peptides issus de la digestion de PcrG**

Les valeurs indiquées sont des moyennes de plusieurs analyses.

Les peptides générés par la pepsine sont en italique, ceux générés par la protéase de type XVIII sont en gras, et ceux générés par la protéase de type XIII sont en style normal.

masse calculée	m/z mesuré	séquence début	séquence fin	nombre de deutériums possibles	nombre de deutérium incorporé en 30sec	pourcentage de deutération	nombre de deutérium incorporé dans le complexe	différence deutération protéine seule / protéine dans le complexe
<b>457.3</b>	<b>458.3</b>	<b>94</b>	<b>97</b>	<b>3</b>	<b>1.5</b>	<b>50.0</b>	<b>1.5</b>	<b>0.0</b>
457.3	458.2	94	97	3	1	33.3	1.0	0.0
462.2	463.3	33	36	3	2	66.7		
471.3	472.4	21	24	3	1	33.3		
500.3	501.4	39	44	5	1	20.0	1.0	0.0
502.3	503.3	37	40	3	1.5	50.0	1.3	0.2
570.4	571.5	94	98	4	2.5	62.5	2.5	0.0
<b>573.3</b>	<b>574.4</b>	<b>13</b>	<b>17</b>	<b>4</b>	<b>1.5</b>	<b>37.5</b>	<b>1.5</b>	<b>0.0</b>
600.3	601.4	89	93	2	1.5	75.0	1.7	-0.2
<b>601.3</b>	<b>602.4</b>	<b>17</b>	<b>22</b>	<b>5</b>	<b>2</b>	<b>40.0</b>	<b>0.0</b>	<b>2.0</b>
616.3	617.3	45	50	5	2	40.0	2.0	0.0
<b>629.3</b>	<b>630.5</b>	<b>58</b>	<b>63</b>	<b>5</b>	<b>3</b>	<b>60.0</b>	<b>3.0</b>	<b>0.0</b>
<b>672.4</b>	<b>673.5</b>	<b>37</b>	<b>42</b>	<b>5</b>	<b>3</b>	<b>60.0</b>	<b>2.7</b>	<b>0.3</b>
672.4	673.5	37	42	5	3	60.0	3.0	0.0
672.4	673.5	37	42	5	3	60.0	2.7	0.3
<b>695.3</b>	<b>696.4</b>	<b>1</b>	<b>6</b>	<b>5</b>	<b>1</b>	<b>20.0</b>	<b>0.7</b>	<b>0.3</b>
703.3	703.5	0	5	5	1	20.0	1.0	0.0
730.4	731.5	15	21	6	4.5	75.0	3.3	1.2
740.3	741.5	7	12	5	2	40.0	2.0	0.0
746.4	747.4	51	56	4	2.5	62.5	2.7	-0.2
758.4	759.4	42	49	7	4	57.1	4.0	0.0
758.4	759.5	42	49	7	3.3	47.1	3.5	-0.2
801.4	802.4	14	21	7	6	85.7	5.0	1.0
<b>832.5</b>	<b>833.4</b>	<b>0</b>	<b>6</b>	<b>6</b>	<b>1</b>	<b>16.7</b>	<b>1.0</b>	<b>0.0</b>
<b>833.4</b>	<b>834.4</b>	<b>34</b>	<b>40</b>	<b>6</b>	<b>2.5</b>	<b>41.7</b>		
899.5	900.5	91	98	6	3.6	60.0	3.5	0.1
988.5	989.4	51	58	6	4	66.7	4.3	-0.3
1003.4	1004.5	-3	5	8	1.5	18.8	1.5	0.0
<b>1045.5</b>	<b>1046.5</b>	<b>31</b>	<b>38</b>	<b>7</b>	<b>2.5</b>	<b>35.7</b>	<b>1.8</b>	<b>0.7</b>
1295.6	1296.7	7	17	10	4.8	48.0	5.0	-0.2
1300.7	1301.7	22	32	10	6	60.0	4.7	1.3
<b>1412.7</b>	<b>1413.6</b>	<b>37</b>	<b>50</b>	<b>13</b>	<b>9</b>	<b>69.2</b>	<b>9.2</b>	<b>-0.2</b>
1413.8	1414.8	21	32	11	7	63.6	5.2	1.8
1566.7	784.5	7	20	13	6.3	48.5	6.0	0.3
1858	930	51	66	14	9	64.3	9.0	0.0
1971	986.5	50	66	15	10	66.7	10.0	0.0

## RESUME

Les protéines sont impliquées dans de nombreux processus biologique et peuvent, dans le domaine médical, être aussi bien des cibles que des agents thérapeutiques. Il est nécessaire pour comprendre en détail leur fonction et leur mode d'action afin d'obtenir des informations sur leur structure et sur leurs interactions éventuelles avec leurs partenaires. Le travail réalisé durant cette thèse a consisté à développer deux méthodes innovantes utilisant la spectrométrie de masse pour étudier la structure des protéines et à appliquer ces méthodes à une problématique biologique.

Nous avons optimisé une méthode associant les échanges H/D et la spectrométrie de masse sur une protéine modèle, la protéine PBP-2X. L'utilisation combinée de trois protéases nous a permis d'obtenir un meilleur recouvrement de séquence de la protéine étudiée et une plus grande résolution spatiale dans la localisation des zones d'intérêt.

Une méthode associant la réticulation chimique et la spectrométrie de masse a été mise au point sur une protéine modèle : le cytochrome c. Les contraintes de distances ainsi obtenues vont intervenir dans une démarche bioinformatique visant à déterminer la famille de repliement d'une protéine de structure inconnue.

Enfin, ces deux méthodes ont été appliquées avec succès sur des protéines du système de sécrétion de type III de *Pseudomonas aeruginosa* : PcrV et PcrG. Nos résultats expérimentaux sur PcrV (accessibilité, structures secondaires et contraintes de distances) corréleront à la structure modélisée de PcrV. Nous avons également déterminé que la protéine PcrG est globalement peu structurée. L'interaction PcrV-PcrG a été caractérisée, elle met en jeu les domaines « coiled-coil » de chacune des deux protéines. La formation du complexe induit un changement de la conformation de PcrV qui pourrait avoir pour conséquence la stabilisation de PcrG.

## ABSTRACT

Proteins are involved in many biological processes. They might be targets for medical treatments as well as therapeutic agents. A detailed knowledge of protein structure and a characterization of protein complexes are important to understand protein functions in a cell. In this study, we developed two new methods, which use mass spectrometry, to elucidate protein structure. These methods were then applied with success on a biological study.

We improved a method that combines H/D exchange experiments with mass spectrometry on a model protein: PBP-2X. We show that the combination of three proteases increases the sequence coverage of the protein and the spatial resolution in the determination of interest areas.

We developed a method, which associates intramolecular cross-linking and mass spectrometry on a model protein: the cytochrome c. Distance constraints determined by this way will be included in a bioinformatic project, that could give the folding family of a protein of which the tri-dimensional structure is unknown.

We applied these two methods on proteins which are involved in type III protein secretion system from *Pseudomonas aeruginosa*: PcrV and PcrG. Experimental data (accessibility, secondary structures and distance constraints) are in agreement with the structural predictions on PcrV. PcrG is mainly unstructured. PcrG and PcrV are in interaction through their coiled-coil domains. Complexation between the two proteins induces conformational changes on PcrV, which could stabilize PcrG.

Mots clefs : Spectrométrie de masse - échange H/D - réticulation chimique - caractérisation structurale - PcrG - PcrV.

Keywords : Mass spectrometry - H/D exchange- chemical cross-linking - structural characterization - PcrG - PcrV.