

Université de Montréal

*Séquençage des peptides par spectrométrie de masse  
en tandem à ionisation par électronébulisation.*

par

Annik Bergeron

Département de chimie, Université de Montréal

Faculté des arts et sciences

Mémoire présenté à la Faculté des études supérieures  
en vue de l'obtention du grade de  
Maître ès sciences

Décembre, 2001



Q1

3

US/

2002

v.007

Université de Montréal  
Faculté des études supérieures

Ce mémoire intitulé :

*Séquençage des peptides par spectrométrie de masse  
en tandem à ionisation par électrospray.*

présenté par :  
Annik Bergeron

a été évalué par un jury composé des personnes suivantes :

K.C. Waldron	Présidente-rapporteuse
M.J. Bertrand	Directeur de recherche
A. Badia	Membre du jury

Mémoire accepté le : .....

## SOMMAIRE

Le couplage de la chromatographie liquide à haute performance (HPLC) et de la spectrométrie de masse (MS) mène à une technique analytique de couplage HPLC-MS, la technique la plus utilisée pour la séparation des peptides et protéines, permettant la détermination de la séquence. Une digestion enzymatique (protéolyse) préalable, en plus de produire de petits fragments peptidiques plus facilement analysables, apporte des informations relatives à la séquence primaire des protéines.

Avec plusieurs appareils LC-MS/MS sur le marché, la séquence des protéines peut être obtenue, après la protéolyse, avec l'aide de recherche sur bases de données. Cependant, cette technologie n'est pas capable d'identifier la séquence de protéines inconnues si elles sont absentes des bases de données. Ainsi, la détermination de la séquence, dans la présente recherche, se base sur les principes premiers de la spectrométrie de masse en tandem des protéines (fragments A, B, C, X, Y et Z), au lieu d'une base de données, afin d'élaborer une méthode performante de séquençage de peptides.

L'élaboration d'une méthode analytique, comme décrite dans ce mémoire de maîtrise, pour le séquençage des peptides et protéines à l'aide de l'appareil HPLC-MS disponible en laboratoire se fait en cinq étapes. D'abord, la protéolyse enzymatique des protéines à analyser produit un mélange contenant les fragments peptidiques. Cette protéolyse fût effectuée de deux façons : en solution et à l'aide d'un microréacteur contenant un gel d'enzyme.

Le mélange protéolytique doit alors subir une séparation chromatographique, la deuxième étape, afin de simplifier l'analyse par spectrométrie de masse. Les conditions chromatographiques sont optimisées pour éliminer l'acide trifluoroacétique (généralement utilisé dans la phase mobile), car il nuit à l'ionisation en spectrométrie de masse. Cette optimisation de la séparation s'est faite sur un mélange protéolytique complexe de la  $\beta$ -caséine, celle-ci menant a un grand nombre de fragments tryptiques, donc à plusieurs pics chromatographiques à séparer. Les conditions ainsi obtenues sont alors fiables pour des mélanges moins complexes.

En troisième lieu, l'analyse de masse suivant la séparation chromatographique permet l'identification de la masse moléculaire de tous les peptides protéolytiques. Cette information est importante pour l'obtention de la séquence en acides aminés des peptides et protéines. Par exemple, l'ion moléculaire obtenu a été utilisé comme ion parent pour l'analyse MS/MS subséquente. L'analyse LC-MS de l'angiotensine II et de la ranatachikinine digérées par la trypsine montre les deux fragments tryptiques attendus pour ces deux polypeptides.

La quatrième étape consiste en l'analyse par spectrométrie de masse en tandem des peptides protéolytiques. Cette analyse fournit la séquence en acides aminés des fragments tryptiques résultants de la digestion enzymatique de l'angiotensine II et d'un des fragments peptidiques de la ranatachikinine. La méthode élaborée a donc permis l'obtention de la séquence de 75% des peptides tryptiques générés.

Il est essentiel de connaître l'ordre dans lequel se trouvait les fragments protéolytiques dans la protéine, avant digestion, pour déterminer la séquence complète en acides aminés de la protéine. Ceci est la dernière étape. Pour y arriver, il y a deux possibilités : (i) la digestion de la protéine par une autre enzyme permettant l'analyse du chevauchement des séquences qui se répètent ou (ii) l'étude cinétique du clivage enzymatique en microréacteur (constituant un autre projet de recherche parallèle au sein du groupe). Après c'est cinq étapes, la séquence entière en acides aminés de la protéine native peut être déterminée.

L'optimisation des conditions d'analyse nécessaires pour l'analyse de la séquence d'une protéine par chromatographie liquide à haute performance et par spectrométrie de masse en tandem furent réalisées lors du présent projet de recherche et les résultats obtenus pour plusieurs échantillons polypeptidiques y sont démontrés.

Mots clés : Séquençage de protéines, peptides, protéines, chromatographie liquide, spectrométrie de masse, ionisation par électrobulbion, LC-MS/MS, protéolyse.

## ABSTRACT

Coupling of high performance liquid chromatography (HPLC) and mass spectrometry (MS) leads to the hyphenated analytical technique of HPLC-MS, the most widely used technique for peptide and protein separation, thus permitting their sequence determination. A preliminary enzymatic digestion (proteolysis), in addition to creating smaller peptide fragments that are more manageable for analysis, provides information relating to the primary sequence of proteins.

With today's LC-MS/MS instruments, the protein sequence can be obtained, after proteolysis, with the aide of database searching techniques. Unfortunately, this technology is not able to identify an unknown protein if its sequence is missing from the database. As a result, determining the actual protein sequence is based on the basic principles of protein tandem mass spectrometry ( generation of A, B, C, X, Y and Z fragments), instead of database searching, leading to a powerful analytical method for the sequencing of polypeptides.

The development of an analytical method, as is described in this master's thesis, for sequencing of peptides and proteins using the HPLC-MS apparatus available in our laboratory involves five steps. Firstly, enzymatic proteolysis of the proteins to be analyzed produced a mixture containing the peptide fragments. This proteolysis was carried out either in solution or using a microreactor containing an enzymatic-based gel.

The proteolytic mixture then had to undergo a chromatographic separation, the second step, in order to simplify the analysis by mass spectrometry. Optimization of the chromatographic conditions was carried out in order to eliminate trifluoroacetic acid (typically used in the mobile phase) because it impedes ionization in mass spectrometry. The separation optimization was performed on a complex  $\beta$ -casein proteolytic mixture, which led to many tryptic fragments, therefore many chromatographic peaks that needed to be separated. Optimized conditions were then considered to be reliable for less complex mixtures.

Thirdly, mass analysis following the chromatographic separation allowed the determination of the molecular mass of all proteolytic peptides. This information is important for determining the amino acid sequence of peptides and proteins. For example, the molecular ion was used as the parent ion in subsequent MS/MS analysis. LC-MS analysis of angiotensin II and ranatachykinine digested by trypsin showed the two expected tryptic fragments for these two polypeptides.

The fourth step consisted of analysis by tandem mass spectrometry of the proteolytic peptides. This analysis provided the amino acid sequence of the tryptic fragments of angiotensin II and one of the two ranatachykinine proteolytic fragments. Therefore, the method developed in this work allowed us to obtain the sequence of 75% of the generated tryptic peptides.

It is essential to know the order of the proteolytic fragments within the protein prior to its digestion in order to determine the full amino acid sequence of the protein. This was the final step. There are two possibilities to achieve this : (i) digestion of the protein by a second enzyme in order to overlap repeating sequences or (ii) kinetic study of enzymatic cleavage by microreactor (which is currently a parallel research project within the group). After these five steps, the whole amino acid sequence of the native protein could be determined.

Optimization of the analysis conditions needed for protein sequence analysis by high performance liquid chromatography and tandem mass spectrometry have been carried out for this research project and results are shown for several polypeptides samples.

Key words : Protein sequencing, peptides, proteins, liquid chromatography, mass spectrometry , electrospray ionization , LC-MS/MS, proteolysis.

## TABLE DES MATIÈRES

<b>Sommaire.....</b>	<b>I</b>
<b>Abstract.....</b>	<b>III</b>
<b>Liste des tableaux.....</b>	<b>IX</b>
<b>Liste des figures.....</b>	<b>XI</b>
<b>Liste des sigles et abréviations.....</b>	<b>XIV</b>
<b>Remerciements.....</b>	<b>XV</b>
<b>CHAPITRE 1 : INTRODUCTION.....</b>	<b>1</b>
1.1 L'importance du séquençage des peptides et des protéines.....	2
1.2 Objectifs de l'étude.....	4
1.3 Démarche expérimentale.....	4
<b>CHAPITRE 2 : REVUE DE LA LITTÉRATURE.....</b>	<b>8</b>
2.1 Historique du séquençage des protéines.....	9
2.2 Historique de l'analyse de protéines par LC-MS.....	12
<b>CHAPITRE 3 : THÉORIE.....</b>	<b>17</b>
3.1 Le couplage de la chromatographie liquide à la spectrométrie de masse (LC-MS).....	18
3.2 La spectrométrie de masse.....	22
3.2.1 <i>La source d'ionisation par électronébulisation</i> .....	22
3.2.1.1 La formation de gouttelettes chargées par dispersion d'un échantillon liquide.....	23
3.2.1.2 La formation d'ions à partir de gouttelettes chargées.....	24
3.2.1.3 Le transport des ions de la pression atmosphérique initiale à la chambre sous vide de l'analyseur .....	26
3.2.2 <i>L'analyseur de masse</i> .....	26
3.2.3 <i>Allure du spectre de masse</i> .....	26

## TABLE DES MATIÈRES

3.3	La spectrométrie de masse en tandem.....	27
3.3.1	<i>Principe de base de la spectrométrie de masse en tandem.....</i>	28
3.3.1.1	Sélection de l'ion parent à l'aide du premier analyseur de masse : un multisecteur de type EBE.....	29
3.3.1.2	Fragmentation de l'ion parent par dissociation induite par collisions .....	33
3.3.1.3	Analyse des ions filles par le second analyseur de masse : l'analyseur à temps d'envol.....	35
	<b>CHAPITRE 4 : PARTIE EXPÉRIMENTALE.....</b>	<b>38</b>
4.1	Instrumentation.....	39
4.1.1	<i>Chromatographie en phase liquide.....</i>	39
4.1.2	<i>Spectrométrie de masse.....</i>	40
4.2	Composés chimiques.....	41
4.3	Méthodes de digestion enzymatique.....	42
4.3.1	<i>Procédure de protéolyse en solution.....</i>	42
4.3.2	<i>Procédure de protéolyse en microréacteur.....</i>	42
	<b>CHAPITRE 5 : RÉSULTATS DE LA SÉPARATION CHROMATOGRAPHIQUE DE MÉLANGE PEPTIDIQUE.....</b>	<b>43</b>
5.1	Optimisation de la séparation chromatographique par la méthode des simplex.....	44
5.1.1	<i>Choix de la fonction de réponse.....</i>	45
5.1.2	<i>Fonctionnement de la méthode des simplex.....</i>	45
5.1.2.1	Formation du premier simplex.....	46
5.1.2.2	Identification des sommets.....	46
5.1.2.3	Les déplacements des simplex.....	47
5.1.2.4	Déplacements exceptionnels.....	48

## TABLE DES MATIÈRES

5.2.1	<i>Choix des paramètres et contraintes chromatographiques.....</i>	50
5.2.2	<i>Optimisation des conditions chromatographiques pour la <math>\beta</math>-caséine.....</i>	51
<b>CHAPITRE 6 : RÉSULTATS DES ANALYSES LC-MS.....</b>		<b>60</b>
6.1	Les analyses par LC-MS.....	61
6.2	Analyses LC-MS de différents peptides.....	62
6.2.1	<i>L'analyse LC-MS de l'angiotensine II.....</i>	62
6.2.2	<i>L'analyse LC-MS de la ranatatchykinine.....</i>	67
<b>CHAPITRE 7 : RÉSULTATS DES ANALYSES PAR SPECTROMÉTRIE DE MASSE EN TANDEM.....</b>		<b>71</b>
7.1	Analyses par spectrométrie de masse en tandem.....	72
7.1.1	<i>Choix du gaz de collisions.....</i>	72
7.1.1.1	<i>Résultats de la MS/MS pour la méthionine enképhalinamide.....</i>	73
7.1.1.2	<i>Résultats de la MS/MS pour la pentaphénylalanine.....</i>	77
7.1.1.3	<i>Résultats de la MS/MS pour la bradykinine.....</i>	80
7.1.1.4	<i>Résultats de la MS/MS pour la substance P.....</i>	83
7.2	Analyse de peptides tryptidiques.....	88
7.2.1	<i>Spectrométrie de masse en tandem de l'angiotensine II.....</i>	88
7.2.2	<i>Spectrométrie de masse en tandem de la ranatatchykinine.....</i>	91

## TABLE DES MATIÈRES

<b>CHAPITRE 8 : CONCLUSION.....</b>	<b>95</b>
8.1 Sommaire des résultats .....	96
8.2 Projet de recherche complémentaire .....	99
8.3 Recherches futures .....	100
<b>BIBLIOGRAPHIE.....</b>	<b>102</b>
<b>ANNEXE I: STRUCTURES, MASSES MOLÉCULAIRES ET ABRÉVIATIONS POUR LES VINGT ACIDES AMINÉS COMMUNS.....</b>	<b>XVI</b>
<b>ANNEXE II : STRUCTURES, MASSES ET FORMULES MOLÉCULAIRES DES PEPTIDES ANALYSÉS.....</b>	<b>XXI</b>

## LISTE DES TABLEAUX

Tableau		Page
I	Caractéristiques des différents fragments selon la nomenclature de Roepstorff.	34
II	Conditions chromatographiques utilisées pour l'analyse des peptides protéolytiques.	39
III	Conditions de spectrométrie de masse pour les analyses LC-MS.	40
IV	Conditions de spectrométrie de masse en tandem pour l'analyse de peptides protéolytiques.	41
V	Contraintes soumises au logiciel informatique de simplex pour chaque paramètre.	51
VI	Valeurs de CRF obtenues dans le cadre de l'optimisation par simplex de la séparation chromatographique des peptides protéolytiques de la $\beta$ -caséine.	54
VII	Masse moléculaire des peptides analysés.	73
VIII	Masse moléculaire des gaz de fragmentation dans la cellule à collisions.	73
IX	Rapport m/z des fragments A ,B et Y attendus de la méthionine enképhalinamide $(M+H)^+ = 573$ .	74
X	Détermination du bon gaz de collisions pour la méthionine enképhalinamide.	74
XI	Rapport m/z des fragments A ,B et Y attendus de la pentaphénylalanine $(M+H)^+ = 754$ .	78
XII	Détermination du bon gaz de collisions pour la pentaphénylalanine.	78
XIII	Rapport m/z des fragments A ,B et Y attendus de la bradykinine $(M+H)^+ = 1061$ .	81
XIV	Détermination du bon gaz de collisions pour la bradykinine.	81

**LISTE DES TABLEAUX**

<b>Tableau</b>		<b>Page</b>
XV	Rapport m/z des fragments A ,B et Y attendus de la substance P $(M+H)^+ = 1347$ .	84
XVI	Détermination du bon gaz de collisions pour la substance P.	84
XVII	Rapport m/z des fragments A ,B et Y attendus du peptide tryptidique #1 de l'angiotensine II $(M+H)^+ = 290$ .	88
XVIII	Rapport m/z des fragments A ,B et Y attendus du peptide tryptidique #2 de l'angiotensine II $(M+H)^+ = 775$ .	89
XIX	Rapport m/z des fragments A ,B et Y attendus du peptide tryptidique #1 de la ranatachykinine $(M+H)^+ = 699$ .	92
XX	Rapport m/z des fragments A ,B et Y attendus du peptide tryptidique #2 de la ranatachykinine $(M+H)^+ = 629$ .	92

## LISTE DES FIGURES

<b>Figure</b>		<b>Page</b>
1.1	Schéma de la démarche expérimentale.	5
1.2	Arbre de masse du glucagon.	7
2.1	Schéma de la méthode séquençage de Sanger.	9
2.2	Mécanisme du séquençage par la dégradation d'Edman.	10
3.1	Les quatre étapes qui constituent le fonctionnement de l'électronébulisation.	23
3.2	Schéma de l'appareil ESI-MS.	23
3.3	Visualisation de l'effet du potentiel sur l'électronébulisation.	24
3.4	Transformation de gouttelettes chargées en ions.	25
3.5	Allure d'un spectre de masse d'une molécule de masse moléculaire élevée.	27
3.6	Schéma de la spectrométrie de masse en tandem.	29
3.7	Principe de la double focalisation.	30
3.8	Schéma d'un analyseur multisecteur de type EBE.	32
3.9	Nomenclature de Roepstorff.	34
3.10	Structure des ions fragments selon Roepstorff.	35
3.11	Schéma d'un Tof.	36
3.12	Tof à accélération orthogonale.	36
4.1	Configuration du système LC-MS.	40
5.1	Premier simplex.	46
5.2	Détermination de la nouvelle analyse à effectuer pour le deuxième simplex.	47
5.3	Détermination de la nouvelle analyse à effectuer pour le troisième simplex.	47
5.4	Oscillation des sommets.	48
5.5	Contrecarrer l'oscillation.	48
5.6	Expansion et contraction d'un simplex.	49

**LISTE DES FIGURES**

<b>Figure</b>		<b>Page</b>
5.7	Chromatogramme du mélange de peptides protéolytiques de la $\beta$ -caséine.	52
5.8	Chromatogramme du blanc de la protéolyse avec la trypsine en solution.	53
5.9	Détermination du débit et du temps d'analyse optimum.	55
5.10	Détermination de la concentration en acide formique et la concentration initiale de l'éluant organique optimum.	56
5.11	Chromatogramme du mélange de peptides trytiques de la $\beta$ -caséine donnant une CRF=568,8.	57
5.12	Chromatogramme du mélange de peptides trytiques de la $\beta$ -caséine donnant une CRF=723,7.	57
5.13	Chromatogramme du mélange de peptides protéolytiques de la $\beta$ -caséine dans les conditions optimales.	58
6.1	Spectre de masse de la bradykinine.	62
6.2	Peptides protéolytiques obtenus pour l'angiotensine II.	63
6.3	Chromatogramme du mélange protéolytique de l'angiotensine II, obtenu par digestion de trois heures en microréacteur.	63
6.4	Chromatogramme du blanc de la digestion en microréacteur.	64
6.5	Spectres de masse de chaque pics du mélange protéolytique de l'angiotensine II.	64
6.6	Fragments protéolytiques obtenus pour la ranatachykinine.	67

**LISTE DES FIGURES**

<b>Figure</b>		<b>Page</b>
6.7	Chromatogramme du mélange protéolytique de la ranatachykinine, obtenu par digestion de trois heures en microréacteur.	68
6.7	Spectres de masse de chaque pics du mélange protéolytique de la ranatachykinine.	65
7.1	Fragmentation de la méthionine enképhalinamide.	73
7.2	Spectres de masse MS/MS de la méthionine enképhalinamide pour les trois gaz à collisions.	76
7.3	Fragmentation de la pentaphénylalanine.	77
7.4	Spectres de masse MS/MS de la pentaphénylalanine pour les trois gaz à collisions.	79
7.5	Fragmentation de la bradykinine.	80
7.6	Spectre de masse MS/MS de la bradykinine avec le méthane.	83
7.7	Fragmentation de la substance P.	83
7.8	Spectre de masse MS/MS de la substance P avec le xénon.	86
7.9	Détermination du gaz de collisions.	87
7.10	Fragmentation des peptides protéolytiques de l'angiotensine II.	89
7.11	Spectres de masse en tandem des peptides tryptidiques de l'angiotensine II.	90
7.12	Fragmentation des peptides protéolytiques de la ranatachykinine.	91
7.13	Spectres de masse en tandem des peptides tryptidiques de la ranatachykinine.	93
8.1	Allure de l'arbre de masse d'une protéine.	99

## LISTE DES SIGLES ET ABRÉVIATIONS

aa	Acide aminé.
ACN	Acétonitrile.
B	Secteur magnétique.
C <sub>18</sub>	Octadécylsilane.
Cf-FAB	Ionisation à bombardement d'atomes rapides en débit continu.
CI	Ionisation chimique.
CID	Dissociation induite par collisions.
CRF	Fonction de réponse chromatographique.
DLI	Introduction liquide directe.
E	Secteur électrique.
ESI	Ionisation par électronébulisation.
FAB	Ionisation à bombardement d'atomes rapides.
HPLC	Chromatographie liquide à haute performance.
kDa	kiloDalton.
keV	kilo électron-Volt.
LC-MS	Chromatographe en phase liquide couplé à un spectromètre de masse.
MCP	« <i>Multichannel plate</i> » ou barrette de photodiodes.
MM	Masse moléculaire.
MS	Spectromètre de masse.
MS/MS	Spectromètre de masse en tandem.
m/z	Rapport masse sur charge.
Oa-Tof	Analyseur à temps d'envol à accélération orthogonale.
RP-HPLC	Chromatographe liquide à haute performance à phase inversée.
TFA	Acide trifluoroacétique.
Tof	« <i>Time of flight</i> » ou temps d'envol.
UV	Ultraviolet.
$\lambda$	Longueur d'onde.

## REMERCIEMENTS

J'aimerais tout d'abord remercier mon directeur de recherche Michel J. Bertrand pour ses excellents conseils, ses encouragements et sa confiance en moi qu'il a eu tout au long de ma maîtrise.

Je tiens également à remercier Mike Evans et Mustapha Harraj du Centre Régional de Spectrométrie de Masse de l'Université de Montréal. Sans eux, je n'aurais sûrement pas de résultats à inclure dans ce mémoire.

Merci aussi aux techniciens de l'atelier mécanique et électronique pour leur aide si précieuse. Merci à tous les étudiants du groupe et à Anik Forest pour leur support moral et toutes les idées partagées.

Enfin, je n'aurais pu passer au travers de ces trois années de recherche sans l'appui de mon conjoint, Marc-André, de mes parents, Suzanne et Jean-Yves, et de mon frère, Mario.

## **CHAPITRE 1**

### **INTRODUCTION**

## 1.1 L'IMPORTANCE DU SÉQUENÇAGE DES PEPTIDES ET DES PROTÉINES

Avec la récente découverte de la séquence presque entière du génome humain, le séquençage des peptides et protéines est toujours d'actualité. Les protéines sont des éléments primordiaux dans les processus biologiques. Elles agissent sous différents rôles lors des réactions chimiques du corps humain ; comme enzymes, comme hormones ou comme récepteurs d'hormones. La connaissance de leur séquence permet de comprendre leurs mécanismes d'action.

Les protéines peuvent être abordées selon les quatre niveaux de structure qui les représentent. La *structure primaire* qui est l'arrangement des liens covalents, ce qui correspond à l'arrangement séquentiel des acides aminés les composant. La *structure secondaire* est la conformation que prennent les chaînes polypeptidiques, comme l'hélice- $\alpha$  ou le feuillet- $\beta$ , selon les ponts hydrogènes qui leurs sont caractéristiques. La *structure tertiaire* est l'arrangement spatiale dépendant de la *structure secondaire* pour former, par exemple, la forme globulaire typique des protéines solubles. La *structure quaternaire*, tant qu'à elle, est l'arrangement spatiale des chaînes polypeptidiques pliées des protéines multimériques.<sup>(1,2)</sup>

Il est possible de faire plusieurs types d'analyses <sup>(3)</sup> des peptides et protéines dont la cartographie peptidique et le séquençage des peptides et protéines. La cartographie peptidique est une technique comparative par laquelle l'empreinte d'une protéine sous forme de chromatogramme ou d'électrophérogramme permet de faire une recherche dans une base de donnée pour identifier la protéine. Cependant, cette méthode d'analyse est surtout utilisée pour l'identification des modifications des protéines. Les modifications de la structure primaire des protéines comme la substitution d'un acide aminé est à l'origine des maladies.<sup>(4)</sup> Par exemple, c'est la substitution d'un résidu acide glutamique par un résidu valine qui est à

l'origine de l'anémie.<sup>(5-7)</sup> Alors, l'empreinte de la protéine modifiée peut être comparée à l'empreinte de la protéine native, ce qui permet d'isoler rapidement la modification subite.

Le séquençage est une méthode permettant l'identification complète de la structure primaire des peptides et des protéines. Cette connaissance de l'arrangement séquentiel des acides aminés composants les peptides et les protéines n'est qu'une partie de la structure complète des protéines pouvant être reconnues par les récepteurs biologiques.

Depuis la détermination de la séquence en acides aminés de l'insuline par Sanger<sup>(8)</sup>, plusieurs innovations telles la dégradation d'Edman<sup>(9)</sup> et le séquenceur automatique ont amélioré le séquençage des peptides et protéines. Cependant, ces nouvelles méthodes de séquençage sont très limitées au niveau de la grosseur des peptides et protéines à analyser qui en plus d'être longue et ardue, ne doit pas dépasser 80 acides aminés.<sup>(10)</sup> L'utilisation de la digestion enzymatique ou protéolyse permet de contourner le problème du nombre d'acides aminés puisqu'elle permet de couper la protéine native en plusieurs petits peptides. Le principal problème de la digestion enzymatique est l'autoprotéolyse de l'enzyme ce qui ajoute un certain nombre de peptides dans le mélange de peptides protéolytiques, qui ne proviennent pas de la protéine native mais plutôt de l'enzyme elle-même. Afin de réduire ce problème, l'immobilisation de l'enzyme sur un support solide ou en un gel est possible<sup>(11)</sup>

De plus, l'avènement du couplage de la chromatographie liquide à la spectrométrie de masse pour l'analyse des acides aminés, des peptides et des protéines facilite et réduit considérablement le temps de l'analyse permettant le séquençage des peptides et protéines. En effet, il existe des appareils LC-MS/MS permettant d'obtenir, après digestion enzymatique, la séquence des peptides et protéines connus. Par comparaison de leur spectre de masse et de leur chromatogramme, comme empreinte digitale,

avec une base de données, il est possible d'identifier les protéines. C'est pourquoi plusieurs compagnies d'instrumentation analytique travaillent à l'élaboration de logiciels permettant la recherche systématique dans les bases de données et l'obtention de la séquence des protéines avec la fiabilité la plus élevée possible.

## ***1.2 OBJECTIF DE L'ÉTUDE***

Le présent travail de recherche vise à élaborer une méthode d'analyse LC-MS/MS permettant le séquençage de peptides inconnus et, par conséquent, introuvables dans les bases de données. Le point de départ de cette recherche est l'optimisation d'une méthode chromatographique menant à la bonne séparation des fragments protéolytiques obtenus à partir du gel d'enzyme, autre sujet de recherche dans le groupe. La méthode chromatographique doit aussi favoriser le couplage du chromatographe liquide au spectromètre de masse.

L'étude a pour deuxième objectif d'optimiser les conditions de spectrométrie de masse en tandem (MS/MS) pour toutes les grosseurs de fragments protéolytiques provenant des protéines natives. Cette étape a pour but d'établir la séquence des peptides protéolytiques. Puisqu'il n'est pas possible d'obtenir la séquence de peptides inconnus par la recherche dans les bases de données, il faut optimiser l'analyse de spectrométrie de masse en tandem afin d'obtenir le plus d'indices possibles ( fragments de séries A,B,C,X,Y et Z discutés au chapitre 5) pour l'interprétation des spectres de masse à partir des principes premiers.

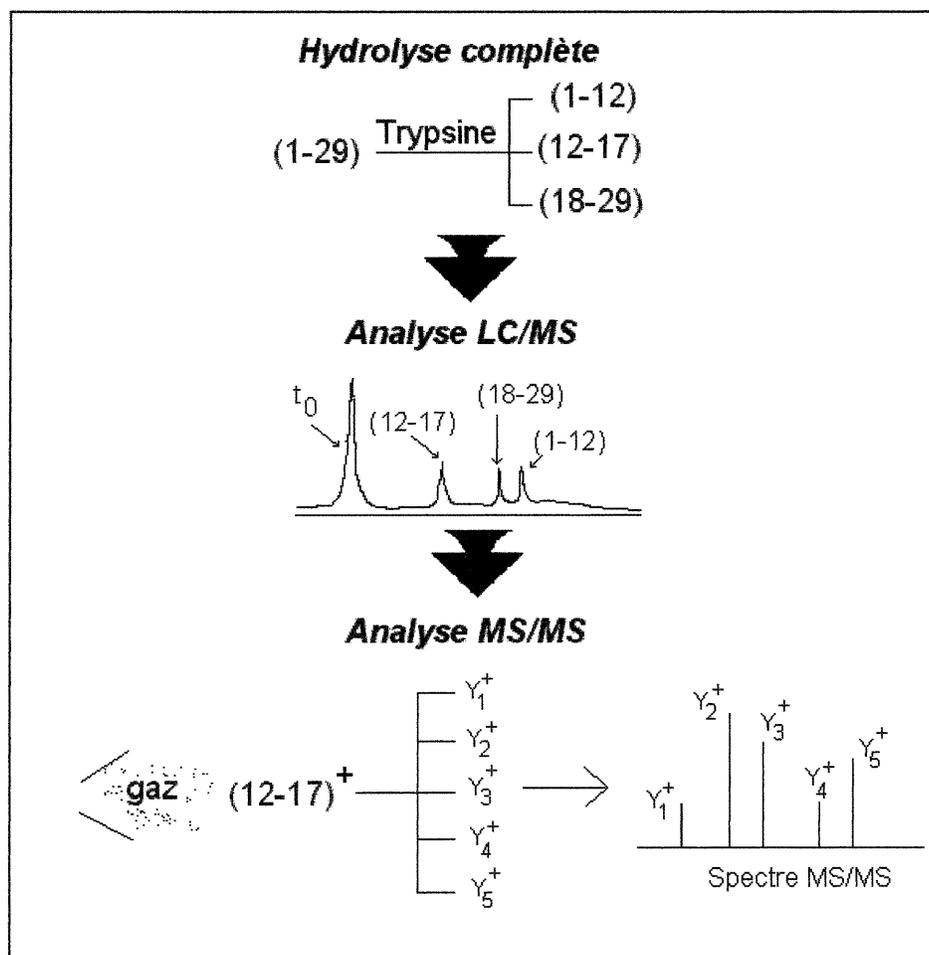
## ***1.3 DÉMARCHE EXPÉRIMENTALE***

L'approche utilisée pour permettre le séquençage des peptides et protéines est divisée en cinq parties : la protéolyse, la séparation chromatographique,

la mesure du spectre de masse de chaque pic chromatographique, la mesure des spectres de masse en tandem pour chaque fragments protéolytiques et l'étude cinétique de la réaction de protéolyse.

Le schéma de la figure 1.1 montre, en exemple, les trois étapes importantes (protéolyse, analyse LC-MS et MS/MS) qui permettent l'identification de la séquence en acides aminés des fragments protéolytiques pour le glucagon, un peptide contenant 29 acides aminés. (Exemple de la littérature)<sup>(12)</sup>.

Le clivage enzymatique, de la protéine à analyser est effectué afin d'obtenir des fragments peptidiques attendus selon les préférences de l'enzyme utilisée à couper les protéines en des liens spécifiques.



**Figure 1.1 Schéma de la démarche expérimentale.**  
 (Exemple donné pour le glucagon, un peptide à 29 acides aminés)<sup>(12)</sup>

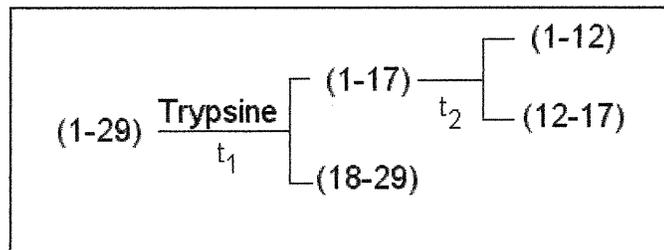
L'optimisation des conditions chromatographiques est une étape assez complexe. En utilisant un mélange de fragments protéolytiques de la  $\beta$ -caséine obtenu à l'aide de la digestion enzymatique en solution par la trypsine, il sera possible de déterminer les conditions optimales de chromatographie en phase liquide pour une bonne séparation. La  $\beta$ -caséine est une protéine de masse élevée (24KDa) contenant 210 acides aminés. Cette protéine est utilisée car elle se fragmente en 14 peptides protéolytiques en présence de la trypsine.<sup>(9)</sup>

Après l'obtention d'une bonne séparation chromatographique d'un mélange de plusieurs peptides tryptiques, il sera possible de séparer des mélanges de peptides moins nombreux. Un logiciel informatique utilisant la méthode d'optimisation par simplex facilitera la tâche pour le traitement des données.

Lorsque les conditions chromatographiques sont connues, le couplage avec le spectromètre de masse permet l'obtention du spectre de masse de chaque pic chromatographique. Ceci donnera l'information sur la masse moléculaire des peptides tryptiques séparés par la chromatographie.

Par la suite, les conditions d'analyse de spectrométrie de masse en tandem (MS/MS) seront optimisées à leur tour afin de faciliter la fragmentation de toutes les grosseurs de peptides protéolytiques susceptibles d'être obtenus. Les spectres MS/MS des ions parents correspondant à chaque peptide tryptique seront enregistrés afin de déterminer la masse des fragments obtenus par la spectrométrie de masse en tandem. Ainsi, ces quatre premières étapes permettent l'identification de la séquence en acides aminés des peptides protéolytiques de la protéine à analyser. Cependant, il demeure impossible avec cette méthode unique de déterminer la séquence dans laquelle se trouvaient les fragments peptidiques dans la protéine avant sa protéolyse complète. C'est là qu'entre en jeu le projet de recherche parallèle qui se déroule dans le groupe de recherche.

Ce projet de recherche a pour objectif l'étude cinétique de la réaction de protéolyse permettant de déterminer dans quelle ordre les fragments peptidiques sont produits. Les fragments obtenus dans le temps ( $t_i$ ) lors de la réaction de clivage enzymatique produite dans un microréacteur de gel d'enzyme, conçu en laboratoire, permettront d'élaborer un arbre généalogique de masse de la protéine à séquencer, comme celui de la figure 1.2.



**Figure 1.2** Arbre de masse du glucagon.

Seul l'optimisation de l'analyse LC-MS/MS est le sujet de ce mémoire et donc les détails de l'étude cinétique n'y seront pas discutés.

## **CHAPITRE 2**

### **REVUE DE LA LITTÉRATURE**

## 2.1 HISTORIQUE DU SÉQUENÇAGE DES PROTÉINES

Il y a 50 ans, la détermination de la séquence en acides aminés d'une protéine était perçue comme une tâche impossible. Pendant ces années, les méthodes de détermination de séquence ont été développées de façon à ce qu'aujourd'hui cette tâche soit devenue une simple routine. C'est en 1955 que Sanger <sup>(8)</sup> détermina pour la première fois la séquence complète d'acides aminés de l'insuline après dix ans de travail. L'insuline est une petite protéine comprenant 51 résidus, s'il avait choisi une plus grande protéine, il serait sûrement encore en train de chercher la séquence. La réussite de Sanger a certainement été un important accomplissement pour le séquençage des protéines. Le développement du code génétique en est une conséquence directe et ainsi, l'avancement de la biologie moléculaire. <sup>(1,13)</sup> Sa méthode pour la détermination de séquence d'acides aminés, schématisée à la figure 2.1, est celle décrite dans les prochaines lignes. <sup>(8)</sup>

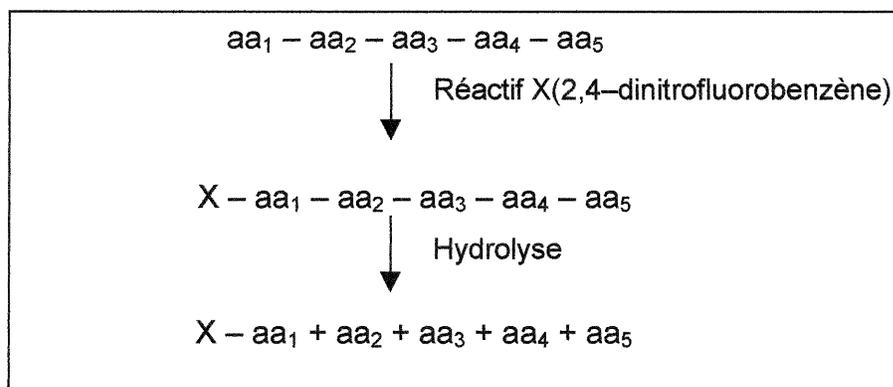
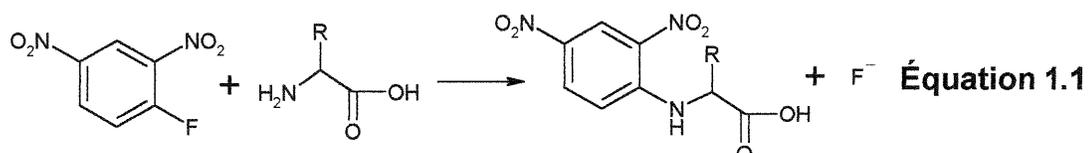


Figure 2.1 Schéma de la méthode de séquençage de Sanger. <sup>(8)</sup>

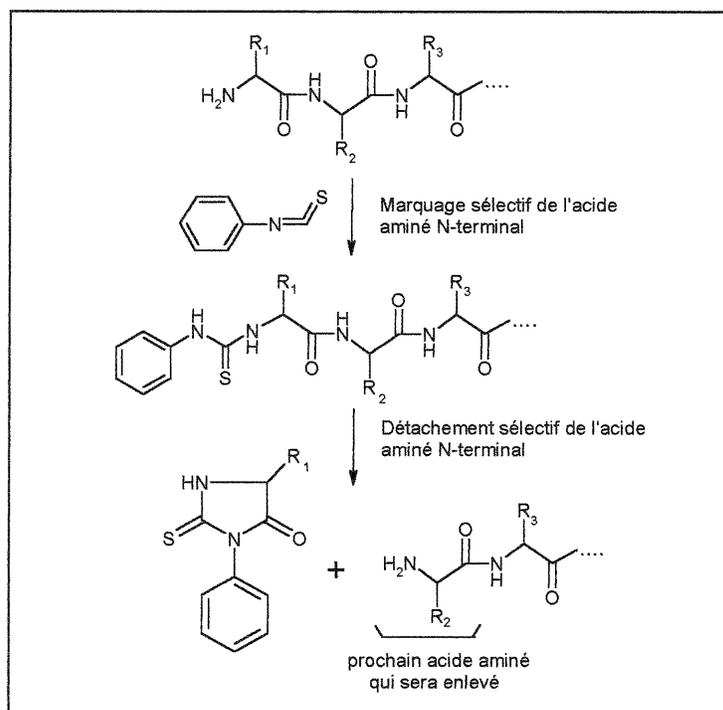
D'abord, il marquait l'acide aminé N-terminal à l'aide d'un réactif, le 2,4-dinitrofluorobenzène, qui réagit uniquement avec les groupements amines ( $\text{NH}_2$ ) libres des acides aminés (aa) et des peptides. Produisant ainsi la réaction suivante :



Cette méthode permet d'identifier qu'un seul acide aminé de la séquence, soit l'acide aminé N-terminal. Sanger a néanmoins réussi à identifier la séquence complète de l'insuline, en hydrolysant partiellement la protéine, en séparant les peptides de plus petites tailles ainsi formés et en identifiant les acides aminés (aa) N-terminaux de ces derniers.

Avec l'avancement des techniques de séquençage, la quantité requise d'analyte et le temps d'analyse ont considérablement diminué. L'introduction de deux importants développements allait propulser la biochimie vers l'avenir : l'analyseur de séquence d'acide aminé automatique et la procédure de dégradation d'Edman sur laquelle les séquenceurs modernes <sup>(3)</sup> sont basés.

Le réactif d'Edman, l'isothiocyanate de phényle ( $C_6H_5-N=C=S$ ), est aussi sélectif aux acides aminés N-terminaux mais en plus il permet de détacher les acides aminés N-terminaux un à la fois. <sup>(8)</sup>



**Figure 2.2 Mécanisme du séquençage par la dégradation d'Edman. <sup>(8)</sup>**

Le séquenceur automatique permet de faire le séquençage d'Edman d'une protéine d'environ 50 acides aminés seulement en une journée seulement (environ 30 minutes par acides aminés), ce qui est une très grande amélioration en comparaison aux dix ans de travail de Sanger.

Cependant, la technique d'Edman et le séquenceur sont limités par le nombre de résidus aminés. En effet, au-delà de 80 acides aminés, le mélange devient trop complexe pour le séquenceur, il a donc fallu penser à une préparation préalable pour permettre la séquence de grosses protéines. C'est ainsi qu'entre en jeu le clivage enzymatique ou protéolyse des protéines. Ce clivage permet de couper la protéine en petits peptides qui sont, par la suite, séquencés dans le séquenceur automatique. Les enzymes sont connues pour couper les peptides et protéines en des endroits spécifiques permettant de prédire le nombre et la taille des peptides protéolytiques obtenus pour une protéine connue. Par exemple, la trypsine est une enzyme qui coupe les peptides et protéines du côté carboxylique de tous les résidus arginine (R) et lysine (K).<sup>(14)</sup>

Le clivage enzymatique produisant plusieurs peptides protéolytiques demande une analyse chromatographique en phase liquide pour permettre la séparation des fragments peptidiques de protéines, on obtient donc la cartographie peptidiques des protéines. Par contre, la chromatographie liquide ne peut pas identifier formellement les peptides protéolytiques séparés, puisque tous les peptides peuvent absorber à la longueur d'onde spécifique aux liens peptidiques ( $\lambda=210$  nm). Bien entendu, les acides aminés comportant des groupements fonctionnels aromatiques peuvent absorber à la longueur d'onde spécifique aux cycles aromatiques ( $\lambda=280$ nm). Une spectroscopie différentielle permet de détecter la présence et le nombre de ces acides aminés aromatiques ; tryptophane, tyrosine ou phénylalanine.<sup>(15)</sup> Cependant, la chromatographie ne donne pas la masse d'un peptide tryptique et donc pas le séquence.

Cependant, en couplant la chromatographie liquide à haute performance (HPLC) à la spectrométrie de masse (MS), on obtient un instrument facilitant le séquençage des protéines à l'aide de toutes les informations fournies par la spectrométrie de masse. Jumelée à l'usage de plus d'une enzyme pour la protéolyse ou à l'étude cinétique de la réaction de protéolyse avec une enzyme, le séquençage des protéines est possible. En effet, la masse des peptides protéolytiques (les ions parents) ainsi que celle des ions fragments provenant de la fragmentation subséquente en spectrométrie en tandem (MS/MS), permettent l'identification des peptides protéolytiques par un logiciel de base de données mais ne donnent pas la séquence de la protéine de départ.

## **2.2 HISTORIQUE DE L'ANALYSE DE PROTÉINES PAR LC-MS<sup>(16)</sup>**

L'analyse des acides aminés, peptides et protéines est passée d'une analyse avec dérivatisation des acides aminés à une analyse de routine de protéines avec des masses moléculaires plus grandes que 25000 Da. Cela est dû au progrès énorme de l'interface en chromatographie liquide couplée à la spectrométrie de masse (LC-MS).

En 1984, l'analyse de peptides dérivés en utilisant une interface à courroie (*moving belt*) a d'abord été faite avec une source d'ionisation chimique (CI) opérant avec du isobutane.<sup>(17)</sup> Un an plus tard, le plus gros peptide intact analysé était le N-acétyl-N,O,S-perméthylé Met-enképhaline (MM = 713 Da), montrant les possibilités à venir de cette technique. Avec l'interface à courroie et la source d'ionisation chimique utilisant l'ammoniaque, un spectre de masse d'un peptide non-dérivé, la cyclosporine A (MM = 1201 Da), est obtenu.<sup>(18)</sup> Le couplage de cette interface et de la source d'ionisation à bombardement d'atomes rapides (FAB) offre une bonne performance pour l'analyse des peptides de plus de 1900 Da.<sup>(19,20)</sup>

Dans ces mêmes années, l'introduction liquide directe (DLI) pour l'analyse de peptides est étudiée. La détection et le séquençage de pentapeptides comme la Met- et la Leu-enképhaline donnent des informations complémentaires sur la structure en utilisant le mode d'ionisation positif et négatif.<sup>(21)</sup> Cependant, plus tard, il fût démontré que l'introduction liquide directe (DLI) n'était pas applicable au séquençage de peptides inconnus, car les fragments caractéristiques n'étaient pas toujours observés.<sup>(22)</sup>

En 1983, lors du développement de la source d'ionisation par thermonébulisation (*thermospray*), des analyses d'acides aminés et de peptides ont été effectuées.<sup>(23)</sup> Cette source d'ionisation donne des peptides mono, doublement et triplement chargés sans produire de fragmentation. Le potentiel de la thermonébulisation pour l'analyse de peptides comme la alpha-mélanocyte et le glucagon démontre que les pics d'ions multichargés dépendent directement du pH de la solution.<sup>(24)</sup> Le couplage de colonnes d'enzyme immobilisée<sup>(10)</sup> directement avec un spectromètre de masse à thermonébulisation<sup>(25)</sup> ou via un chromatographe en phase liquide pour une séparation additionnelle,<sup>(26)</sup> rend possible la détermination d'une large portion de la séquence d'un peptide à l'aide d'une base de données. Les peptides protéolytiques obtenus par la digestion enzymatique sont identifiés par la masse moléculaire de leur ions mono, doublement ou triplement chargés.<sup>(27)</sup> Des données pour de larges peptides comme le glucagon, l'insuline bovine (MM=5734 Da) et la N-acétylée eglincine C (MM=8133 Da) ont été obtenues à l'aide d'un instrument à secteur magnétique<sup>(28)</sup>, alors qu'un instrument à triple quadropole permet l'analyse entre 1000 et 4000 Da.<sup>(29)</sup>

Le développement de l'interface de la source à ionisation par bombardement d'atomes rapides en débit continu (Cf-FAB) a été préalablement utilisée pour l'analyse des peptides en 1986.<sup>(30)</sup> La Cf-FAB est alors largement appliquée à la caractérisation et l'analyse des digestions protéolytiques.<sup>(31)</sup> La Cf-FAB LC-MS et LC-MS/MS donnent des informations excellentes pour l'analyse de petits peptides, comme le Met-Leu-Phe (MM=409 Da)<sup>(32-34)</sup>

La révolution et le point tournant des applications biochimiques de la LC-MS proviennent des développements dans le domaine de l'ionisation par électronébulisation (*electrospray*). C'est en 1988, qu'il est démontré qu'il est possible d'obtenir des ions multichargés de protéines, rendant ainsi possible leur détection et la détermination de leur masse moléculaire avec un simple et peu cher quadripôle.<sup>(35,36)</sup> Depuis ce temps, le potentiel de l'électronebulisation est toujours étudié. Les deux plus importants développements dans ce domaine sont :

- l'application possible de la source d'ionisation par électronebulisation sur la majorité des spectromètres de masses de manufacturiers,
- l'introduction de la nano-électronebulisation permettant une très grande sensibilité avec de faibles quantités d'échantillon.<sup>(37)</sup>

La détermination de la masse moléculaire de la protéine est évidemment la première étape dans la caractérisation d'une protéine isolée et purifiée. Cela est rendu possible grâce à l'électronebulisation<sup>(38-42)</sup> Cependant, lorsque la masse moléculaire d'une protéine est connue, la question de la structure primaire fait surface.

Le séquençage de peptides par spectrométrie de masse est basé sur la dissociation induite par collisions de peptides protonés, générés par électronebulisation, FAB ou toutes autres techniques d'ionisation douce. Lorsque la dissociation induite par collisions (CID) à basse énergie est effectuée, les ions fragments observés proviennent des clivages de la chaîne d'acides aminés de la protéine. Ce clivage mène aux ions fragments suivants la nomenclature de Roepstroff et Fohlmann<sup>(43,44)</sup> L'approche générale fût établie pour le séquençage des peptides par FAB/MS et FAB/MS/MS. Dans la pratique, cette procédure est compliquée par deux principaux facteurs :

- l'utilisation de la source d'ionisation FAB apporte un problème majeur de signal de fond dû à la matrice utilisée.<sup>(32)</sup> L'électronébulisation peut aussi causer ce problème de signal de fond dû à la phase mobile.
- la quantité d'énergie transférée à une large protéine est insuffisante pour le clivage de la chaîne d'acides aminés de celle-ci. La façon la plus facile de contourner ce problème est l'utilisation de la digestion enzymatique qui réduit une large protéine en plusieurs petits peptides pouvant être soumis à la CID par la suite.

De nos jours, les appareils sur le marché couplés aux logiciels de recherche sur les bases de données fonctionnent de la façon suivante. En utilisant en combinaison la digestion enzymatique et la spectrométrie de masse par électronébulisation (ESI-MS) ou la LC-ESI-MS, la cartographie peptidique peut être établie, donnant une série de pics reliés aux peptides fragments provenant de la digestion enzymatique. Les peptides détectés sont par la suite analysés par CID en MS/MS. Cette méthode est un peu compliquée par la présence de modifications des acides aminés puisqu'elle ne peut pas différencier la leucine (L) de l'isoleucine (I), par exemple, alors que cela est possible par la dégradation d'Edman suivie de la chromatographie liquide. Cependant, les avantages de la spectrométrie de masse par rapport à la dégradation d'Edman pour le séquençage des peptides pèsent plus dans la balance que les légers inconvénients de cette technique. Alors, la procédure de séquençage de peptides et protéines par LC-ESI-MS/MS comprend les étapes qui suivent.<sup>(45-50)</sup>

1. Détermination de la masse moléculaire de la protéine intacte, isolée et purifiée.
2. Digestion enzymatique par la trypsine ou autre enzyme pour générer le mélange de fragments peptidiques (tryptiques).
3. Analyse directe ou LC-MS du mélange de peptides tryptiques, appelée la cartographie peptidique. Cela donne l'information sur le rapport masse sur charge ( $m/z$ ) des différents peptides tryptiques.

4. Comparer la cartographie obtenue avec les bases de données pour afin de savoir si la protéine analysée est déjà connue.
5. Production des ions fragments en MS/MS des différents peptides tryptiques pour obtenir la séquence des peptides.
6. Passer à une autre enzyme pour la protéolyse, comme la chymotrypsine, qui donnera des fragments peptidiques différents et qui permettra par le chevauchement des peptides, de connaître la séquence. Ou bien on peut effectuer une étude cinétique de la réaction de protéolyse <sup>(51-54)</sup> avec la trypsine qui permettra de connaître l'ordre dans lequel les fragments tryptiques apparaissent. C'est le sujet du projet de recherche complémentaire à cette recherche, qui se déroule simultanément dans notre groupe.

Alors que quelques années plus tôt le séquençage des protéines était un projet s'étendant sur des années, maintenant le temps nécessaire est réduit à quelques jours. Il y a toujours des recherches en cours pour réduire le temps d'analyse et améliorer le séquençage des protéines par LC-ESI-MS/MS.

## **CHAPITRE 3**

### **THÉORIE**

### **3.1 LE COUPLAGE DE LA CHROMATOGRAPHIE LIQUIDE À LA SPECTROMÉTRIE DE MASSE (LC-MS)**

Le problème majeur du couplage d'un chromatographe en phase liquide à haute performance (HPLC) et d'un spectromètre de masse (MS) est l'énorme différence entre les grands volumes de solvant provenant du HPLC et le vide nécessaire pour le spectromètre de masse.<sup>(2,22,55)</sup> Il existe plusieurs solutions à ce problème. On peut utiliser un diviseur d'effluent par lequel l'éluant provenant de la colonne chromatographique est divisé et ainsi uniquement une faible fraction est introduite dans le spectromètre de masse. On peut aussi utiliser une colonne microbore qui ne peut permettre des débits d'éluant plus élevés que 50  $\mu\text{L}/\text{min}$  et donc est compatible avec le vide d'un spectromètre de masse. L'interface de la source d'ionisation à électronébulisation est aussi une façon de contrer ce problème puisqu'elle permet l'introduction de liquide directement dans le MS. Pour la plupart des mélanges biologiques, les analyses LC-MS peuvent être réalisées. La seule limitation qui apparaît vraiment importante, est la gamme dynamique de masse qui, jusqu'à tout récemment, ne pouvait atteindre des valeurs maximum de plus de 20000 unités de masse (20 kDa). La source d'ionisation par électronébulisation (ESI) donne des ions multichargés, ce qui diminue le ratio masse sur charge ( $m/z$ ) des ions et donc augmente considérablement la gamme dynamique de masse détectable. C'est donc la source d'ionisation favorisée pour l'analyse des protéines par LC-MS.

Pour améliorer la résolution d'un système chromatographique, il faut optimiser les différents paramètres concernant la phase stationnaire et la phase mobile. Le choix de la phase mobile en chromatographie liquide à haute performance à phase inversée (RP-HPLC) devrait être dicté, non seulement par son effet sur le mélange peptidique, mais aussi par son effet sur l'hydrophobicité de la phase stationnaire. La capacité de manipuler la résolution en changeant la composition de la phase mobile représente la caractéristique la plus avantageuse de la RP-HPLC. L'élution des peptides

ou protéines est effectuée à l'aide d'une phase mobile acide (contenant un agent de pairage d'ions) et d'un gradient croissant en composition de solvant organique.<sup>(56)</sup> Il est donc évident que le choix de la phase mobile requiert deux décisions importantes : l'agent de pairage d'ions ainsi que le solvant organique à utiliser.

Cependant, le couplage de la spectrométrie de masse à ionisation par électronébulisation à la chromatographie liquide amène des contraintes au niveau de la phase mobile. Premièrement, tous les produits à ioniser doivent être facilement volatilisables pour permettre l'ionisation par électronébulisation. Une autre contrainte, non négligeable, est qu'il faut éviter toute concentration élevée d'acide afin de ne pas perturber l'ionisation au niveau des tensions de surface du liquide. Pour ce qui est des sels, il faut éviter ceux qui sont non volatils car il y a risque d'obstruction de la source d'ionisation. De plus, même pour des sels volatils, il faut limiter leur concentration dans les éluants afin de réduire au maximum les effets sur les tensions de surface du liquide. C'est donc en tenant compte de ces contraintes, qu'il faut obtenir une bonne séparation chromatographique des peptides avant leur entrée dans le spectromètre de masse.

Puisque les peptides sont des molécules chargées à presque toutes les valeurs de pH, la présence de différents contre-ions va influencer leur comportement chromatographique. La différence de polarité des fragments peptidiques obtenus par la digestion tryptique peut être maximisée en choisissant bien l'agent de pairage d'ions. Le pouvoir de résolution des agents de pairage d'ions est affecté par leurs interactions avec les groupements ionisés d'un peptide. Les contre-ions anioniques vont interagir avec les résidus basiques protonés d'un peptide. L'utilisation d'agent de pairage d'ions cationiques est limitée par des valeurs de pH près des pKa des groupements des chaînes latérales acides ( $\text{pH} \approx 4$ ), nécessitant l'utilisation de phases mobiles tamponnées et généralement non-volatiles. Il est donc préférable d'utiliser des agents de pairage d'ions anioniques.

En général, le trifluoroacétate (TFA) est le contre-ion le plus utilisé pour la neutralisation de la charge des peptides afin de permettre le mode d'élution par partage seulement. Mais le TFA, lorsqu'utilisé en spectrométrie de masse, amenuise le signal obtenu. En effet, le TFA nuit à l'ionisation par électronébulisation dû à la combinaison des effets de pairage d'ions très fort et de la modification des tensions de surface.<sup>(57)</sup> Les acides formique et acétique peuvent aussi être utilisés.<sup>(58-60)</sup> En plus, de ne pas nuire à l'ionisation par électronébulisation, l'acide formique permet aussi la détection par UV-visible car il n'absorbe pas à la longueur d'onde d'absorption des analytes.<sup>(15)</sup>

Outre la nature de l'agent de pairage d'ions, sa concentration affecte les temps de rétention des peptides. Plus il y aura de contre-ions dans la phase mobile, plus les temps d'analyse seront élevés. Des concentrations de l'ordre du millimolaire sont normalement utilisées pour les séparations des peptides. Ces faibles concentrations sont compatibles avec la contrainte de concentration de la source d'ionisation, due aux modifications de tensions de surface qui pourraient survenir.

Le choix du solvant organique est aussi soumis à des contraintes de la part de la source d'ionisation. Il faut, donc qu'il soit volatil et peu visqueux pour permettre son ionisation. L'acétonitrile est celui le plus utilisé pour la séparation des peptides, étant donné sa transparence aux longueurs d'onde de détection pour les peptides en UV-visible. De plus, il a la plus faible viscosité ce qui est un autre avantage du point de vue de l'ionisation par électronébulisation.

Les paramètres d'opération les plus importants qui influencent la résolution des séparations de peptides sont le temps et le taux de changement de concentration en solvant organique du gradient, ainsi que le débit de l'éluant. L'effet du temps du gradient sur la séparation des peptides de la digestion enzymatique par la trypsine est très marqué. Ainsi, plus le temps du gradient

est grand, plus la résolution chromatographique est élevée car il y a une augmentation du temps de résidence des peptides dans la phase stationnaire. Par contre, il faut toujours faire un compromis entre un temps d'analyse acceptable et une bonne résolution. Il est important de limiter le temps du gradient car sinon l'analyse pourrait devenir interminable. Ainsi le temps du gradient est généralement limité à 60 minutes ce qui permet une analyse complète en environ 40 minutes.<sup>(47)</sup>

Typiquement, des gradients linéaires avec des colonnes analytiques conventionnelles sont employés de 5% d'éluant organique à entre 50 et 100% d'éluant organique. Cela se fait, en général, à des débits entre 0,5 et 2 mL/min. Le choix du gradient dépend de la sélectivité entre les solutés d'intérêt. L'effet d'une variation de débit sur les temps de rétention des peptides est faible lorsqu'on utilise un gradient comme mode d'élution. Cependant, le taux de changement de concentration de solvant organique du gradient a plus d'effet sur la résolution et le comportement chromatographique des peptides que le débit et donc un effet profond sur la séparation d'un mélange peptidique.

En utilisant un détecteur UV-visible à une longueur d'onde spécifique aux peptides,  $\lambda = 214$  nm (liens peptidiques), il est possible de détecter tous les peptides présents dans le mélange de la digestion enzymatique. Pour cette recherche, vu la présence d'acide formique absorbant au maximum d'absorption des liens peptidiques, la longueur d'onde utilisée est de 230 nm qui correspond au minimum d'absorbance de l'acide formique et à l'absorption encore suffisante des peptides.<sup>(15)</sup> Les analyses HPLC-UV ont été effectuées afin de pallier au manque de disponibilité du spectromètre de masse lors l'optimisation de la séparation chromatographique. Ainsi, bien que l'absorbance ne soit pas élevée pour les peptides, elle était suffisante pour cette étape de notre projet.

Pour ce qui est des paramètres attribués à la phase stationnaire, le choix de ceux-ci n'est pas compromis par l'utilisation de la spectrométrie de masse comme mode de détection. Alors, le type de support, les fonctions hydrophobiques et les dimensions de la colonne chromatographique sont ceux permettant le plus de types de séparations chromatographiques, soient les plus polyvalents. Pour ce qui est de la taille des particules et de leurs pores, ils ont été choisis pour des peptides de grosseur attendue après la digestion enzymatique.

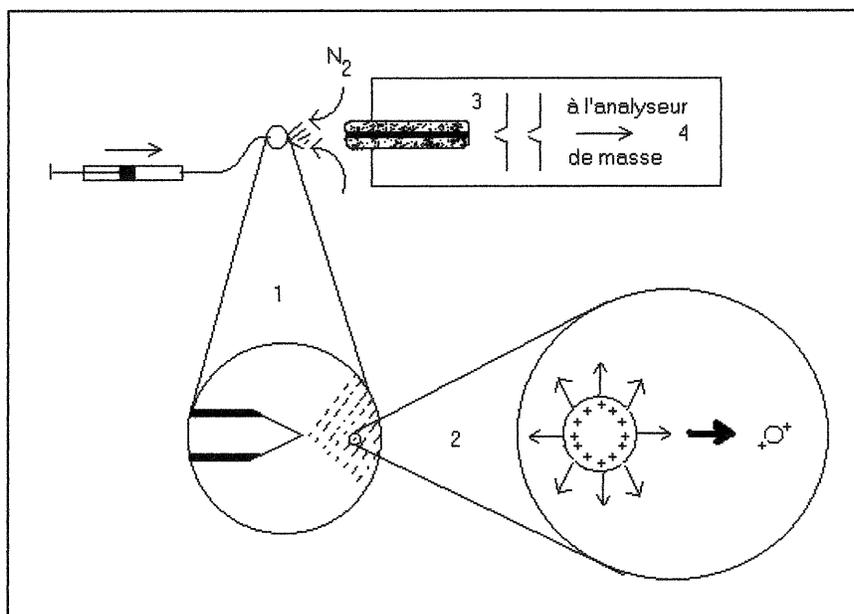
Le chromatographe liquide haute performance à phase inversée (RP-HPLC) procure uniquement la séparation des peptides fragments d'un mélange de digestion enzymatique et son couplage à la spectrométrie de masse permet d'identifier la masse de chaque fragments peptidiques élués.

### **3.2 LA SPECTROMÉTRIE DE MASSE <sup>(61)</sup>**

Un spectromètre de masse est constitué de plusieurs composantes : le système d'introduction, la source d'ionisation, l'analyseur de masse, le détecteur d'ions et le système d'enregistrement. Le système d'introduction du système utilisé lors de cette recherche est bien entendu le chromatographe liquide. La source d'ionisation est comme dit précédemment, la source d'ionisation par électronébulisation. L'analyseur de masse est un analyseur à double focalisation. La détection se fait à l'aide d'un multiplicateur d'électrons à multicanaux, alors que l'enregistrement se fait à l'aide d'un logiciel informatique.

#### **3.2.1 LA SOURCE D'IONISATION PAR ÉLECTRONÉBULISATION <sup>(62)</sup>**

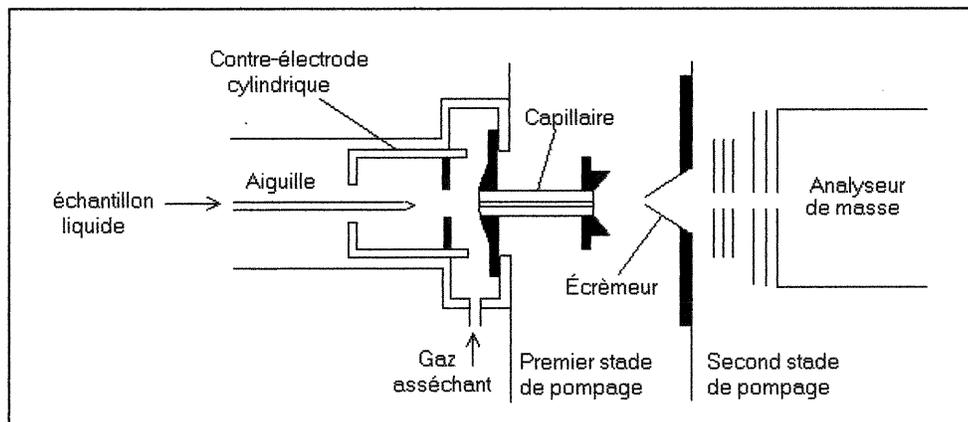
La description du fonctionnement de la spectrométrie de masse à ionisation par électronébulisation (ESI-MS) se fait par celle des quatre étapes ou procédés principaux qui la constituent et sont représentées à la figure 3.1.<sup>(2)</sup>



**Figure 3.1 Les quatre étapes qui constituent le fonctionnement de l'électronébulisation.** (1)Dispersion de gouttelettes chargées par électronébulisation.(2)Formation d'ions à partir de gouttelettes chargées.(3)Transport des ions vers le MS.(4)Analyse de masse.

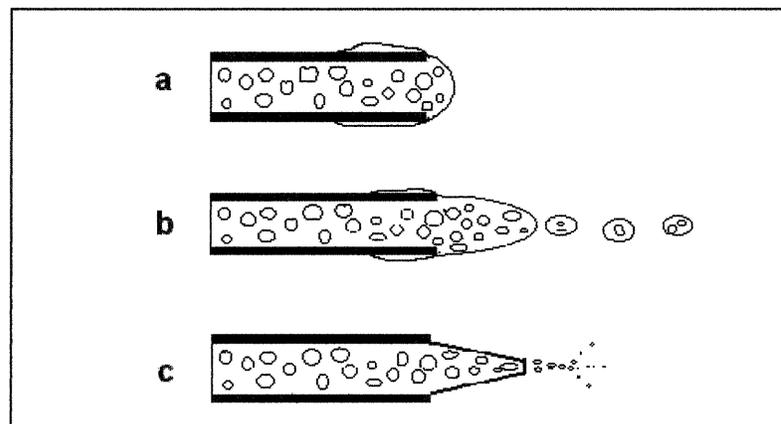
### 3.2.1.1 La formation de gouttelettes chargées par dispersion d'un échantillon liquide.

L'échantillon passe au travers un capillaire, en général une aiguille de 0.1mm de diamètre interne, à un débit variant de 1 à 100 $\mu$ L. Cette aiguille est maintenue à un haut voltage par rapport à la contre-électrode. Le champ électrique résultant à la pointe de l'aiguille permet la dispersion du liquide en fines gouttelettes sous forme conique, appelée le cône de Taylor<sup>(63)</sup>. Taylor a dit que la surface d'un liquide statique diélectrique dans un champ électrique, va prendre la forme conique.



**Figure 3.2 Schéma de l'appareil ESI-MS. <sup>(16)</sup>**

Cependant, pour obtenir un cône de Taylor, il faut bien connaître l'effet du potentiel sur les gouttelettes formées.<sup>(2)</sup> En effet, à faible potentiel ( $V$ ) appliqué entre l'aiguille et la contre-électrode, les gouttes seront trop grosses et seront recueillies par le système de pompage (figure 3.3a). À potentiel plus élevé, on obtient le « *rain mode* », qui est une fine pluie de gouttelettes toujours trop grosses pour former le cône de Taylor et aussi recueillies par le système de pompage (figure 3.3b). À potentiel critique, on obtient le cône de Taylor (figure 3.3c).



**Figure 3.3 Visualisation de l'effet du potentiel sur l'électronébulisation.**  
(a : à  $V$  trop faible, b : à  $V$  plus élevé et c : à  $V$  critique)

À une valeur critique du voltage, en général de trois à six keV entre l'aiguille et la contre-électrode, le cône de Taylor est obtenu (figure 3.3c), donnant ainsi un courant ionique stable. Il faut toujours faire attention au potentiel car en étant trop élevé, le cône sera déstabilisé par une décharge de corona.

### 3.2.1.2 La formation d'ions à partir des gouttelettes chargées.

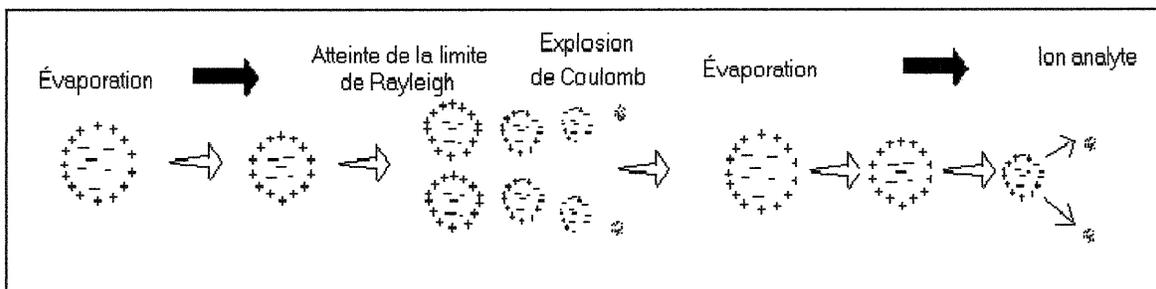
Les gouttelettes chargées sont accélérées vers le capillaire par la différence de potentiel, en passant au travers un gaz asséchant ( $N_2$  à 350°K) qui provoque la désolvatation des gouttelettes. Les gouttelettes ainsi obtenues diminuent de volume jusqu'à la limite de Rayleigh<sup>(64)</sup>, qui est celle décrite par l'équation suivante :

$$q_R = 8\pi (\gamma\epsilon_0)^{1/2} r^{3/2}$$

**Équation 3.1**

Où  $q_R$  est la charge de Rayleigh,  $r$  est le rayon de la gouttelette,  $\gamma$  est la tension de surface et  $\epsilon_0$  est la permittivité du vide.

Par la suite, lorsque les répulsions Coulombiennes deviennent de même ordre que les forces de tension superficielle, il y a des explosions de Coulomb qui mènent à la désorption en phase gazeuse. Les ions sont conduits par un haut champ à la surface des gouttes et ils sont expulsés de celles-ci avec leurs charges. En règle générale, les ions obtenus ont une charge par kDa .



**Figure 3.4 Transformation de gouttelettes chargées en ions.**

Il existe deux modèles pour expliquer le passage de l'ion en phase gazeuse : le modèle du résidu chargé et le modèle de l'évaporation ionique.

*i. Le modèle du résidu chargé.* <sup>(65)</sup>

C'est Dole qui a amené ce modèle, qui veut que la gouttelette diminue de volume jusqu'à ce qu'il ne reste plus qu'un macroion. Cependant, il s'agit d'un modèle rarement applicable mais qui ne peut être complètement rejeté.

*ii. Le modèle de l'évaporation ionique* <sup>(66)</sup>

Ce modèle, applicable dans beaucoup plus de situation, de Iribarne et Thomson veut qu'on ne soit pas limité à la formation d'un macroion par goutte. En effet, il y a formation d'agglomérats, de dimères et autres qui sont explicables par ce modèle.

### 3.2.1.3 Le transport des ions de la pression atmosphérique initiale à la chambre sous vide de l'analyseur de masse.

Les ions sont collectés dans le capillaire par l'effet de la différence de potentiel. Par la suite, le premier stade de pompage permet d'enlever les espèces neutres tandis que le deuxième système de pompage crée le vide dans l'analyseur de masse. Dans certains appareils, il peut y avoir un troisième stade de pompage compris entre les deux précédents stades discutés, qui permet d'enlever une grande quantité de gaz asséchant restant. Les vides sont alors respectivement de 1Torr,  $\leq 10^{-2}$  Torr et  $\leq 10^{-5}$  Torr.

### 3.2.2 L'ANALYSEUR DE MASSE

Il y a plusieurs choix d'analyseurs de masse, dont le quadripôle qui apporte une faible résolution et une gamme dynamique de masse limitée ainsi que le secteur magnétique qui donne une meilleure résolution et permet la détection d'une gamme dynamique de masse très grande. Les analyseurs de masse seront discutés de façon plus détaillée dans la section 3.3.

### 3.2.3 ALLURE DU SPECTRE DE MASSE

Pour des composés de haute masse moléculaire, qui ont plusieurs sites d'ionisation, l'ionisation par électronébulisation permet l'obtention d'ions multichargés en abondance. Ce qui donne un spectre de masse comme à la figure 3.5.

Alors, pour analyser ce spectre de masse il faut utiliser les deux équations suivantes afin d'obtenir la masse moléculaire de la molécule étudiée.<sup>(16)</sup>

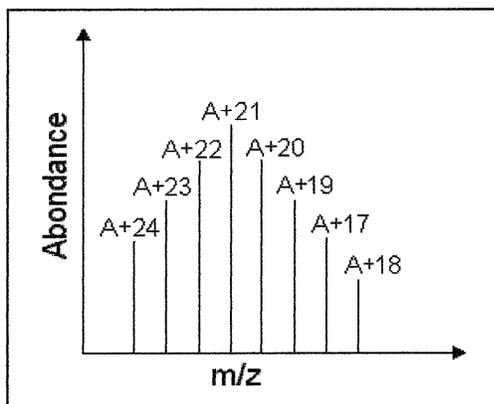
$$m_1/z_1 = (A + z_1)/z_1$$

**Équation 3.2**

$$m_2/z_2 = (A + z_1 + 1) / (z_1 + 1)$$

### Équation 3.3

Où A est la masse moléculaire de la molécule étudiée, z la charge sur un ion et m/z le rapport masse sur charge de l'ion.



**Figure 3.5** Allure d'un spectre de masse d'une molécule de masse moléculaire élevée.

L'efficacité d'ionisation par électronébulisation dépend de la limite d'une charge par kDa . En effet, si le débit de solution entrant dans l'électronebuliseur est de  $1\mu\text{L}/\text{min}$  et le courant du jet est de  $100\text{nA}$  vue la limite d'une charge par kDa, alors la limitation est atteinte pour une solution de concentration de  $0.1\text{g}/\text{L}$  d'analyte. Si ce facteur n'est pas limitant alors l'efficacité de l'ionisation par électronebulisation est dictée par l'efficacité du transport des ions car elle dépend de la fraction délivrée à l'analyseur de masse. Cependant des exceptions sont possibles, des ions doublement chargés ont déjà été obtenus pour des molécules de seulement  $1000\text{ Da}$ .

### 3.3 LA SPECTROMÉTRIE DE MASSE EN TANDEM

Le couplage d'un spectromètre de masse à un autre est une des méthodes hyphéniques importantes en chimie analytique. Le premier spectromètre de masse est généralement muni d'une source d'ionisation douce. Ainsi, il y a production, en majorité, des ions moléculaires chargés. La source d'ionisation par électronebulisation, décrite dans la section précédente, produit une ionisation ultra-douce qui ne provoque pas la décomposition des

analytes. L'ionisation par électronébulisation, en général, ne produit pas de fragmentation mais uniquement des ions moléculaires chargés. Ce type d'ionisation peut être qualifiée d'ionisation froide dans la mesure que l'énergie fournie à l'ion n'est pas assez élevée pour en permettre la fragmentation.

### 3.3.1 PRINCIPE DE BASE DE LA SPECTROMÉTRIE DE MASSE EN TANDEM (MS/MS) <sup>(67)</sup>

Le principe de base de la spectrométrie de masse en tandem est simple. Il s'agit de la mesure du rapport de masse sur charge ( $m/z$ ) des ions avant et après la réaction dans la cellule à collision du spectromètre de masse. La réaction de dissociation des ions parents, par le changement de masse, génère des ions filles ayant différentes énergies cinétiques et différents momentums que leur ion parent. La réaction de dissociation s'écrit comme suit :



Où le  $m_p^+$  réfère à l'ion parent, le  $m_f^+$  à l'ion fille et  $m_N$  sont les espèces neutres formées lors de la réaction de dissociation.

L'ion parent n'est pas nécessairement l'ion moléculaire. L'ion moléculaire est l'ion de masse équivalente à la masse moléculaire de l'analyte. L'ion moléculaire tout comme les ions fragments provenant de l'ionisation peuvent être sélectionnés comme ion parent lors d'une expérience de spectrométrie de masse en tandem (MS/MS). Cependant, puisque la source à ionisation par électronébulisation ne produit que des ions moléculaires et des ions multichargés, les ions parents ne seront pas des ions fragments dus à l'ionisation.

Même si  $m_N$  est noté comme une unique espèce neutre, dans plusieurs des cas de réaction de dissociation, il y a deux fragments neutres et plus. Alors le terme  $m_N$  est équivalent à la somme des masses de ces fragments neutres.

Puisque la spectrométrie de masse ne permet la détection que des ions chargés, alors les  $m_N$  ne seront pas détectés. Cependant, à l'aide des valeurs de  $m_p^+$  et de  $m_f^+$ , il sera possible de retrouver les valeurs de  $m_N$  associées à la réaction.

Il y a trois étapes cruciales lors d'une expérience MS/MS, démontrées à la figure suivante :

- La sélection de l'ion parent à l'aide du premier analyseur de masse.
- La fragmentation de l'ion parent dans la cellule à collision par le procédé de dissociation induite par collisions (CID).
- L'analyse des ions filles obtenus par le deuxième analyseur de masse.

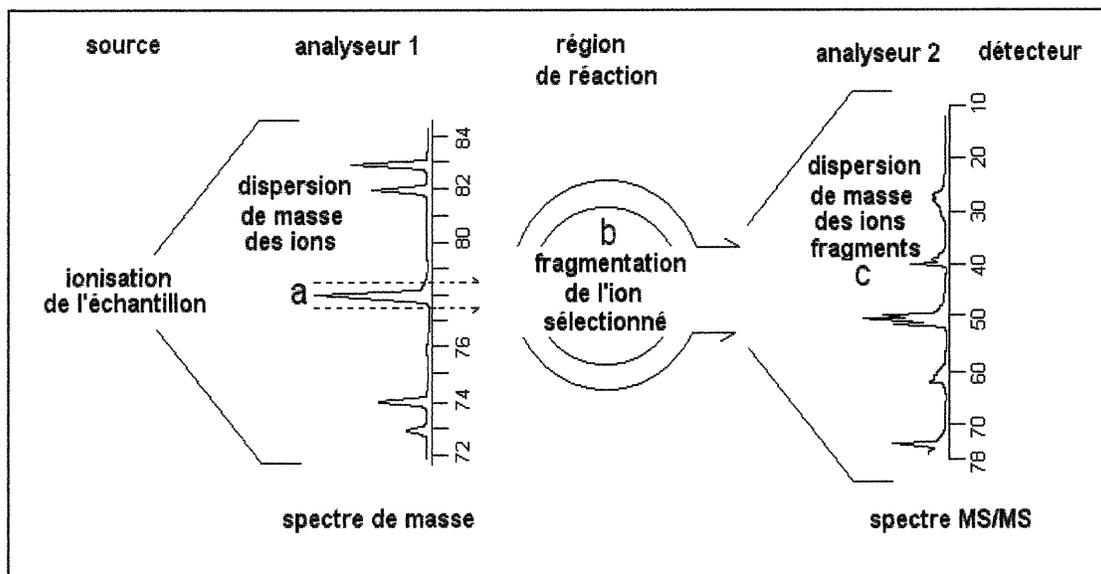


Figure 3.6 Schéma de la spectrométrie de masse en tandem.<sup>(68)</sup>

### 3.3.1.1 Sélection de l'ion parent à l'aide du premier analyseur de masse : un multisectionneur de type EBE.<sup>(16)</sup>

Le premier analyseur de masse de l'appareil utilisé est un multisectionneur de géométrie EBE : soit un secteur électrique (E) dont les deux moitiés se

retrouvent de chaque côté d'un secteur magnétique (B). Cet analyseur permet une excellente sélection de masse ce qui est bon puisque le premier analyseur a pour rôle de laisser passer uniquement l'ion parent vers la cellule à collision.

Cet analyseur est ce que l'on appelle un appareil à double focalisation dont le fonctionnement est démontré à la figure 3.7. Le fait d'augmenter la résolution en insérant un sélecteur d'énergie dans le système est appelée la double focalisation. En effet, l'utilisation d'un champ électrique, comme sélecteur d'énergie, aide à réduire l'élargissement des pics dû à la dispersion des vitesses. Le champ électrique focalise les différentes énergies dans un même plan.

L'accélération des ions hors de la source d'ionisation vers l'analyseur de masse est la première étape après leur formation. Après son passage au travers des fentes qui en réduisent l'étendue, le faisceau d'ions pénètre entre deux plaques métalliques courbes appelées un secteur électrique.

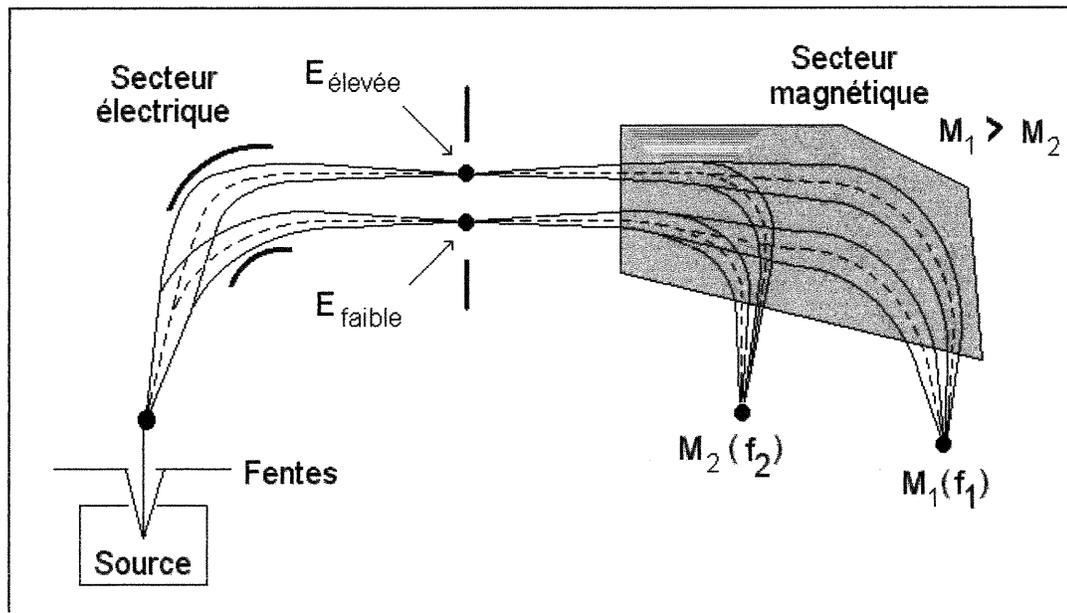


Figure 3.7 Principe de la double focalisation. <sup>(61,69)</sup>

Dans ce secteur électrique, le potentiel électrique maintenu entre les deux plaques focalise le faisceau d'ions vers le secteur magnétique. Le secteur électrique est un appareil focaliseur d'énergie. Ainsi, si  $V$  est le potentiel au travers duquel les ions sont initialement accélérés et  $E$  est le champ dans le secteur électrique, alors l'équation 3.5 est obtenue.<sup>(69)</sup>

$$R = 2V/E \quad \text{Équation 3.5}$$

Où  $R$  est le rayon de courbure de la trajectoire des ions. Les deux équations qui gouvernent le mouvement des ions à travers l'analyseur sont celles reliées à l'énergie cinétique  $E_k$  (équation 3.6) et à la force centrifuge  $F_c$  (équation 3.7).<sup>(70)</sup>

$$E_k = \frac{mv^2}{2} = zV \quad \text{Équation 3.6}$$

$$F_c = \frac{mv^2}{R} = zE \quad \text{Équation 3.7}$$

Où  $v$  est la vitesse d'un ion de masse  $m$  et de charge  $z$ . C'est une simple combinaison de ces deux équations qui donnent l'équation 3.5. Les ions accélérés à travers un potentiel ( $V$ ) et passant au travers un champ électrique uniforme ( $E$ ) ont le même rayon de courbure ( $R$ ) sans tenir compte du ratio masse sur charge. Donc, si le champ électrique ( $E$ ) est gardé constant, le secteur électrique focalise les ions selon leurs énergies cinétiques.

Le faisceau d'ions d'énergie focalisée est passé à travers un secteur magnétique où la séparation de masse est effectuée. Comme décrit ci-dessus, après l'accélération provenant de la source d'ionisation, les ions possèdent une énergie interne,  $zV$  ( $= \frac{1}{2} mv^2$ ). En passant au travers le champ magnétique, les ions sont soumis à une force centrifuge ( $mv^2/r$  où  $r$  est le rayon de courbure) due à la contrainte ( $Bzv$ ) communiquée par le champ magnétique ( $B$ ). Par ces quantités, l'équation 3.8 suivante est obtenue.

$$zV = \frac{1}{2} mv^2$$

$$Bzv = mv^2/r$$

$$m/z = \frac{B^2 r^2}{2V}$$

Équation 3.8

À des valeurs constantes de champ magnétique ( $B$ ) et à un potentiel d'accélération ( $V$ ), à chaque valeur de  $m/z$  correspondra un rayon de courbure  $r$ . Ainsi, après la séparation de masse dans le champ magnétique, les ions monochargés de masse  $m_1, m_2, m_3, \dots, m_n$  vont être amenés aux points focaux  $f_1, f_2, f_3, \dots, f_n$  correspondants aux rayons de courbure  $r_1, r_2, r_3, \dots, r_n$ . Il est aussi possible de varier le voltage d'accélération ( $V$ ), avec un champ magnétique  $B$  constant, pour amener chaque espèce  $m/z$  au même point focal, où se situe l'unique détecteur de l'appareil (équation 3.8). Puisque faire varier le rayon de courbure du secteur magnétique serait impossible du point de vue monétaire et pratique, la variation du potentiel d'accélération est le mode d'analyse utilisé. Pour chaque valeur de voltage un seul ion de rapport  $m/z$  correspondant se rendra au détecteur, tandis que les autres se fracasseront sur les parois internes du secteur magnétique. Il est alors possible de déterminer le ratio  $m/z$  d'ion, selon le potentiel d'accélération appliqué. Ainsi, un analyseur à double focalisation permet la sélection de l'ion parent pour le laisser passer dans la cellule à collisions en utilisant la variation du voltage d'accélération.

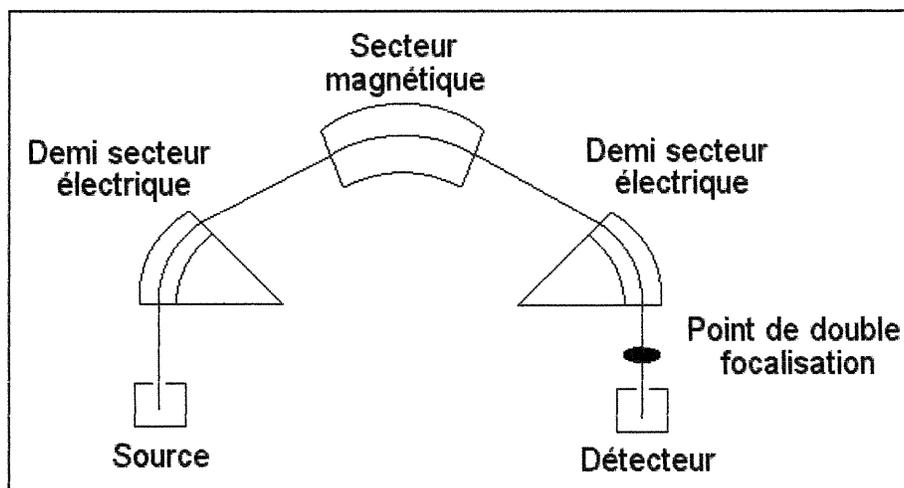


Figure 3.8 Schéma d'un analyseur multisecteur de type EBE.

Bien que dans le cas de cette recherche, l'analyseur de masse multisectoriel soit de type EBE (figure 3.8), le principe de fonctionnement est le même que celui d'un appareil de type EB, de géométrie *Nier-Johnson*, puisqu'il s'agit des deux moitiés d'un même secteur électrique disposées de part et d'autre d'un secteur magnétique.

### 3.3.1.2 Fragmentation de l'ion parent par dissociation induite par collisions.<sup>(67)</sup>

Après la sélection de la masse de l'ion parent, celui-ci passe dans une cellule à collisions où il est fragmenté par dissociation induite par collisions. Le procédé de fragmentation nommé dissociation induite par collisions (CID) consiste à introduire un gaz de collisions dans la région de réaction, soit la cellule à collisions. Ceci afin d'induire des dissociations par les collisions de l'ion parent avec le gaz neutre cible (G) comme le montre l'équation 3.9. Ce procédé est considéré comme le cœur de la spectrométrie de masse en tandem.



**Équation 3.9**

La relation 3.10 montre que la fragmentation de l'ion parent est directement reliée à la masse moléculaire du gaz de collision. Dépendant de la masse moléculaire du gaz cible ( $m_g$ ), la fragmentation sera plus ou moins forte, car l'énergie fournie ( $E_{int}$ ) à l'ion sera différente de l'énergie appliquée à la cellule à collision ( $E_{lab}$ ).

$$E_{int} = E_{lab} \left( \frac{m_g}{m_g + m_p} \right)$$

**Équation 3.10**

Ainsi, le choix du gaz de collisions est très important pour le type de fragmentation voulue. Les peptides se fragmenteront selon la nomenclature de Roepstorff (figure 3.9) dans certaines conditions de fragmentation. Il faudra donc optimiser ces conditions de fragmentation afin d'obtenir le plus d'informations possibles sur les peptides analysés.

Lorsqu'un peptide subit la CID, selon le lien covalent coupé et cependant que la charge de l'ion est retenue soit sur le N-terminal, soit sur le C-terminal, les fragments obtenus seront différents (voir tableau I). Les divers fragments (figure 3.10) obtenus par la spectrométrie de masse en tandem pour des peptides sont appelés par les lettres A, B, C, X, Y et Z. Ces fragments sont les ions filles ( $m_f$ ) de l'équation 3.9.

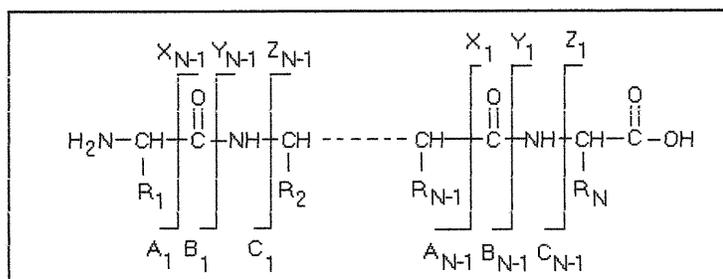


Figure 3.9 Nomenclature de Roepstorff. <sup>(43)</sup>

Tableau I : Caractéristiques des différents fragments selon la nomenclature de Roepstorff. <sup>(71)</sup>

Fragments	Liens coupé	Charge retenue sur
$A_n$	$\alpha C - CO$	N terminal
$B_n$	$CO - N$	N terminal
$C_n$	$N - \alpha C$	N terminal
$Z_n$	$\alpha C - N$	C terminal
$Y_n$	$N - CO$	C terminal
$X_n$	$CO - \alpha C$	C terminal

Les fragments de Roepstorff obtenus en majorité par la spectrométrie de masse en tandem sont les fragments A, B et Y. <sup>(12,67)</sup> En effet, il semble que ces fragments soient les plus facilement produits par la spectrométrie de masse en tandem lors de la réaction de dissociation des peptides ions parents.

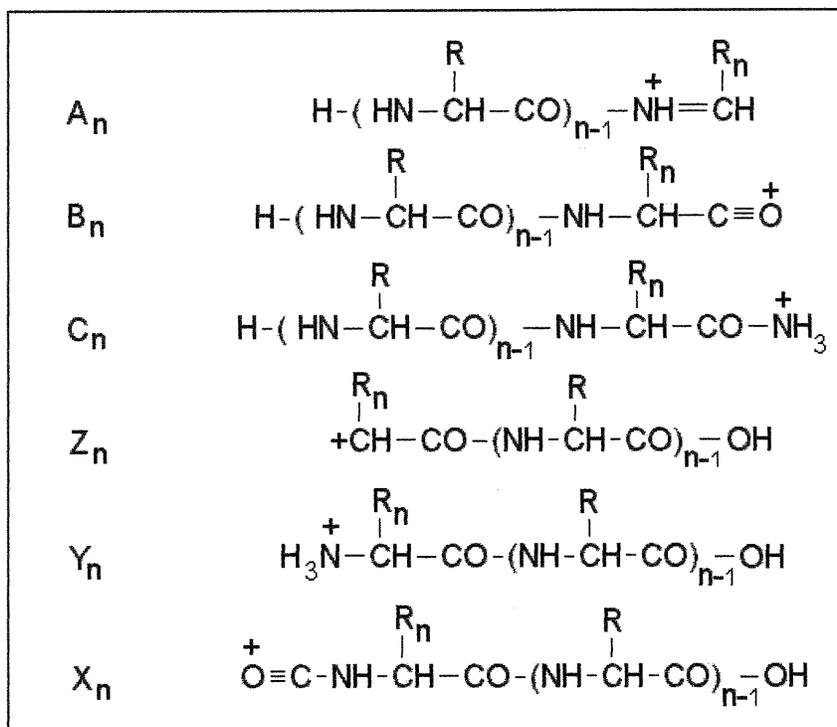


Figure 3.10 Structure des ions fragments selon Roepstorff. <sup>(12)</sup>

### 3.3.1.3 Analyse des ions filles par le second analyseur de masse : l'analyseur à temps d'envol (« *time of flight* ou *Tof* »). <sup>(72)</sup>

Alors, les ions filles tels que décrits précédemment, seront les fragments A, B et Y en majorité. Dans l'appareil de spectrométrie utilisé lors de cette recherche, le deuxième analyseur de masse est un analyseur à temps d'envol (Tof). Le principe d'opération d'un spectromètre de masse à temps d'envol (Tof) implique la mesure du temps que prend un ion pour voyager de la source d'ionisation (ou bien la cellule à collision dans ce cas-ci) jusqu'au détecteur. Ce procédé requiert la production d'un « amas » discret d'ions dans une région proche de la source d'ionisation et, ensuite, à travers une série d'événements synchronisés, l'accélération des ions vers le détecteur et la mesure de leur temps d'arrivée à ce détecteur localisé à environ 1 mètre de la source. Tous les ions reçoivent la même énergie cinétique durant leur accélération, mais puisqu'ils ont des masses différentes, ils sont séparés en des groupes selon leur vitesse (et leur masse) lorsqu'ils traversent la région sans champs entre la source d'ions et le détecteur. La valeur de  $m/z$  d'un ion

est déterminée par son temps d'arrivée au détecteur. Les ions de faible masse atteignent le détecteur avant ceux de haute masse, parce que les ions plus lourds ont une vitesse plus faible.

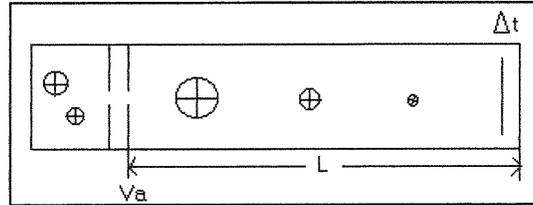


Figure 3.11 Schéma d'un ToF. <sup>(76)</sup>

En utilisant l'équation 3.6, la vitesse ( $v$ ) d'un ion peut être donnée comme suit :

$$v = \left( \frac{2zV}{m} \right)^{1/2} \quad \text{Équation 3.11}$$

L'équation 3.11, montre la relation quadratique inversée entre la vitesse ( $v$ ) et le ratio masse sur charge ( $m/z$ ). Parce qu'il n'est pas pratique de mesurer la vitesse de l'ion directement, le temps d'envol de la source au détecteur (séparé d'une distance  $L$ ) procure l'information expérimentale utile.

$$\text{Tof} = \frac{L}{v} = \left( \frac{m}{2zV_a} \right)^{1/2} \quad \text{Équation 3.11}$$

Puisque le spectromètre de masse à temps d'envol se situe après la cellule à collision dans le cas d'intérêt, l'appareil utilisé est un analyseur à accélération orthogonale (oa-ToF) qui facilite l'analyse en poussant simultanément les ions vers le détecteur, alors il n'est pas nécessaire de produire un « amas » d'ions.

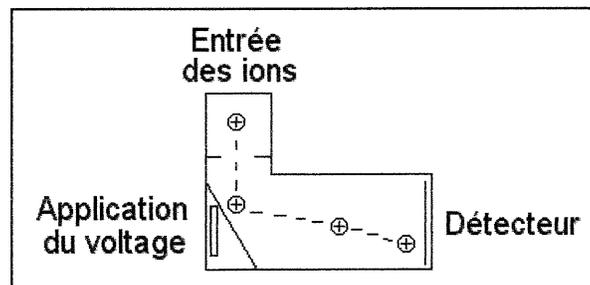


Figure 3.12 ToF à accélération orthogonale. <sup>(61)</sup>

Les ions du faisceau continu, provenant de la cellule à collision, sont pulsés à l'aide d'un voltage vers un analyseur à temps d'envol, situé à 90° du faisceau. Puisque la trajectoire dans l'analyseur est orthogonale au faisceau initial, le temps d'envol dépend uniquement de la composante orthogonale de vitesse. L'accélération orthogonale est avantageuse du point de vue de la distribution des ions de mêmes énergies. En effet, il se peut dans un analyseur à temps d'envol commun que la position ou le temps d'arrivée des ions dans l'analyseur affecte la résolution du temps d'envol. Or, puisque les seuls ions qui vont vers le détecteur sont envoyés par une application de voltage, ces problèmes ne se produisent pas.

La chromatographie liquide et la spectrométrie de masse, telles que décrites dans ce présent chapitre, seront utilisées de la façon suivante lors de ce projet de recherche. La première étape de cette étude est l'optimisation des conditions chromatographiques pour la séparation des peptides protéolytiques obtenus, après la digestion enzymatique. Ces conditions doivent aussi être compatibles avec le couplage de la chromatographie liquide à haute performance et la spectrométrie de masse. L'identification de la masse moléculaire des peptides protéolytiques sera faite grâce à l'analyse LC-MS du mélange résultant de la digestion enzymatique du peptide. La source d'ionisation par électronébulisation, étant une source d'ionisation froide, l'analyse LC-ESI-MS donnera des spectres de masse ne comportant pas de produits de fragmentation, car l'énergie fournie aux ions est trop faible. Alors, il faudra utiliser la dissociation induite par collisions en spectrométrie de masse en tandem afin d'obtenir une fragmentation des ions moléculaires. Ce sont donc les analyses MS/MS qui permettront de connaître la séquence des peptides protéolytiques obtenus. L'ordre dans lequel se trouvaient les peptides protéolytiques à l'intérieur de la protéine native, avant la digestion enzymatique demeure inconnue. C'est alors que l'étude cinétique de la réaction de digestion enzymatique entre en jeu. L'arbre de masse obtenu par cette étude donnera l'information nécessaire pour résoudre la séquence complète de la protéine native.

## **CHAPITRE 4**

### **PARTIE EXPÉRIMENTALE**

## 4.1 INSTRUMENTATION

### 4.1.1 CHROMATOGRAPHIE EN PHASE LIQUIDE

Le système de chromatographie en phase liquide est composé d'une pompe de série 410 de la compagnie Perkin Elmer. L'injecteur est de type Rheodyne 7125 avec une boucle d'injection de 20  $\mu$ L. La colonne (218TP54) de la compagnie Vydac a un diamètre interne de 0,46 cm et une longueur de 15 cm. Elle contient des particules de phase stationnaire octadécylsilane ( $C_{18}$ ) de 5  $\mu$ m de grosseur et ayant des pores de 300 Å de taille. Avant la colonne chromatographique, une pré-colonne de marque Security Guard de Phenomenex de même phase stationnaire  $C_{18}$  est utilisée. Le détecteur UV-Visible programmable, modèle 490E, à longueur d'onde variable et multiple provient de la compagnie Waters.

L'analyse chromatographique des peptides tryptiques utilisés se fait sous les conditions décrites dans le tableau II. Le logiciel de méthode de simplex est le « Multisimplex 2.0 » de la compagnie Statistical Designs.

**Tableau II : Conditions chromatographiques utilisées pour l'analyse des peptides protéolytiques .**

---

Éluants :     A : 0,17% acide formique : 5% acétonitrile : eau (v/v/v)

                  B : 0,06% acide formique : 5% eau : acétonitrile (v/v/v)

Gradient :    7,6% à 65% d'éluant B en 46,9 minutes

                  Mode linéaire

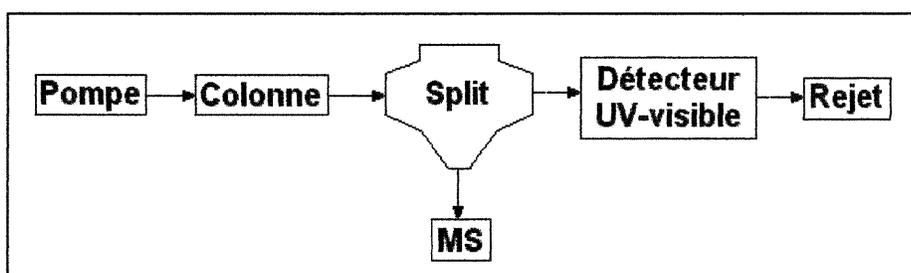
Débit :        0,7 mL/min

Longueur d'onde : 230 nm

---

#### 4.1.2 SPECTROMÉTRIE DE MASSE

Le couplage du chromatographe en phase liquide et du spectromètre de masse se fait à l'aide d'un diviseur d'effluent de type P-450 de la compagnie Upchurch Scientific. La longueur de tuyau, d'un diamètre interne (D.I.) de 0,25mm, reliant le diviseur d'effluent à la source d'ionisation du spectromètre de masse est de 13 cm, ce qui correspond à la limite minimum physique de la distance entre le diviseur d'effluent et la source d'ionisation. Ceci afin de limiter l'élargissement des pics due à la longueur du trajet à parcourir.<sup>(55)</sup>



**Figure 4.1 Configuration du système LC-MS.**

Le spectromètre de masse AutoSpec-Tof de la compagnie Micromass comprend une source d'ionisation par électrospray (ESI), un premier analyseur de masse multisectionnel de type EBE qui est séparé du deuxième analyseur de masse à temps d'envol (Tof) par une cellule à collision. Le détecteur est un multiplicateur d'électrons à multicanaux (MCP).

**Tableau III : Conditions du spectromètre de masse pour les analyses LC-MS.**

---

Résolution du spectromètre de masse : 1500  
 Intervalle de masse : 100 à 1200 m/z  
 Vitesse de balayage : 37 décades/secondes

Voltage d'accélération : 3500 V  
 Voltage au capillaire : 7000V

Éluant : Éluants A et B utilisés pour la chromatographie.  
 Débit entrant dans la source : 80  $\mu$ L/min

---

L'analyse MS/MS s'effectue sous différentes conditions, décrites dans les tableau IV, pour chaque gamme de masse des fragments protéolytiques (ions parents) à analyser.

**Tableau IV : Conditions de spectrométrie de masse en tandem pour l'analyse de peptides protéolytiques.**

---

<u>Masse des peptides</u>	<u>Gaz</u>
< 1300 Da	CH <sub>4</sub> ou Ar
> 1300 Da	Xe

Pression de gaz : Jusqu'à ce qu'il ne demeure plus que 10% du signal de l'ion parent.

Voltage appliqué à la cellule à collision ( $E_{lab}$ ) : 150 V

---

## 4.2 COMPOSÉS CHIMIQUES

Les éluants chromatographiques en HPLC ont été faits à partir de l'acétonitrile de grade HPLC provenant de la compagnie A&C, d'acide formique 99% et d'acide trifluoracétique de la compagnie Sigma. L'eau a été purifiée par un système Milli-Q (Millipore). Tous les éluants ont subi un dégazage de 20 minutes à l'hélium de qualité « zero gas ». Les éluants sont filtrés à leur entrée dans le système chromatographique par des filtres à solvant de deux microns de la compagnie Waters.

Les standards des peptides et protéines utilisés ( la  $\beta$ -caséine lyophilisée pure à plus de 90% , la [ $\beta$ -Asp<sup>1</sup>]-angiotensine II pure à plus de 97% ainsi que la bradykinine, la méthionine enképhalinamide , la pentaphénylalanine, la ranatachykinine et de la substance P pures à 98%) proviennent tous de la compagnie Sigma. Il en est de même pour l'enzyme utilisée pour la digestion

enzymatique, la trypsine. La trypsine traitée par la L-1-Tosylamide-2-phényléthylchlorométhyle cétone a une activité enzymatique de 13.4 unités/mg de protéine.

### **4.3 MÉTHODES DE DIGESTION ENZYMATIQUE**

#### **4.3.1 PROCÉDURE DE PROTÉOLYSE EN SOLUTION <sup>(73)</sup>**

Afin de préparer une solution de 2,5 mg/mL de protéine digérée, il faut peser 2,5 mg de la protéine à digérer dans un tube Eppendorf de 1,5mL. Solubiliser celle-ci dans 250  $\mu$ L d'une solution d'urée 8 M dans un tampon de borate 0,4 M à pH 8. Ajouter 50  $\mu$ L d'une solution de dithiothréitol 45 mM et par la suite incuber, dans une étuve à 50°C, pour 15 minutes.

Lorsque le tube Eppendorf est revenu à la température de la pièce, ajouter 50  $\mu$ L d'une solution d'iodoacétamide 100 mM. (Cette solution est sensible à la lumière, elle devra donc être entreposée au réfrigérateur dans un contenant opaque.) Laisser réagir à la température de la pièce pour 15 minutes. Ajouter 600  $\mu$ L d'eau et 0,1 mg d'enzyme dissoute dans un volume de 50  $\mu$ L d'eau, ce qui représente un ratio poids : poids de 1:25 , enzyme pour protéine. Ensuite incuber à 37°C pour 24 heures. La solution peut être congelée afin d'arrêter la réaction.

#### **4.3.2 PROCÉDURE DE PROTÉOLYSE EN MICRORÉACTEUR <sup>(74)</sup>**

Le microréacteur est une colonne remplie d'un gel d'enzyme créé en laboratoire. Ce microréacteur permet, selon le temps de séjour du peptide dans celui-ci, de faire une protéolyse complète ou partielle du peptide à analyser. La procédure était toujours en développement dans notre groupe de recherche au moment d'écrire cet ouvrage.

## **CHAPITRE 5**

### **RÉSULTATS DE LA SÉPARATION**

#### **CHROMATOGRAPHIQUE DE MÉLANGE PEPTIDIQUE**

## **5.1 OPTIMISATION DE LA SÉPARATION CHROMATOGRAPHIQUE PAR LA MÉTHODE DES SIMPLEX**

La séparation chromatographique du mélange de peptides obtenu après la digestion enzymatique par la trypsine est la première étape à optimiser. La séparation par chromatographie liquide à phase inversée des peptides est un sujet élaboré dans plusieurs documents.<sup>(11,15,56,58,75)</sup> Cependant, les conditions utilisées comportent en général un éluant contenant de l'acide trifluoroacétique (TFA), qui lorsqu'utilisé en spectrométrie de masse amenuise le signal obtenu comme discuté au chapitre 3. Alors, pour éviter ce problème, le TFA est remplacé par l'acide formique.<sup>(58,59)</sup> Afin d'optimiser les conditions chromatographiques avec des éluants contenant de l'acide formique au lieu de TFA, il faut faire varier plusieurs paramètres.

L'optimisation d'une séparation ne dépendant que d'un seul paramètre est facile. Cependant, en chromatographie en phase liquide, l'optimisation d'une séparation dépend de plusieurs paramètres ; comme le débit, les concentrations en acide des éluants, la composition initiale du gradient et le temps de durée du gradient. Il est bien évident que la variation de chacun des paramètres tout en gardant tous les autres paramètres constants est une méthode d'optimisation à proscrire dans un cas comme celui-ci où le nombre de paramètres à varier est élevé. En effet, cette expérience se révélerait longue et ardue. L'utilisation de la méthode d'optimisation par simplex facilite les choses, car elle permet de trouver les conditions spécifiques optimales d'une manière simple, directe et efficace en manipulant plusieurs variables à la fois.<sup>(76,77)</sup> Cette méthode réduit de façon considérable le nombre d'analyses chromatographiques à effectuer. Le logiciel informatique « Multisimplex 2.0 » a été utilisé afin d'aider au traitement des données qui étaient très nombreuses.

### 5.1.1 CHOIX DE LA FONCTION DE RÉPONSE

Le première étape de la méthode d'optimisation par simplex est la détermination d'une fonction de réponse chromatographique (CRF) qui permettra de quantifier la qualité de la séparation chromatographique. Il existe plusieurs fonctions de réponse qui peuvent être utilisées pour l'optimisation d'une séparation chromatographique.<sup>(76)</sup> Cependant, la fonction de réponse choisie devra tenir compte du plus de paramètres possibles. Ainsi, celle choisie ne tient pas seulement compte de la résolution entre chaque paire de pic ( $R_{s_i}$ ) mais aussi du temps de rétention du dernier pic ( $T_L$ ), du temps d'analyse acceptable ( $T_M$ ) et du nombre de pics que l'on devrait obtenir(L).<sup>(78-81)</sup>

$$CRF = \sum R_{s_i} + L^x - a |T_M - T_L| - b(T_1 - T_0) \quad \text{Équation 5.1}$$

Où  $T_1$  est le temps de rétention du premier pic,  $T_0$  est le temps mort et les facteurs  $a$ ,  $b$  et  $x$  sont arbitraires . Dans ce cas-ci, la valeur des facteurs  $a$ ,  $b$  et  $x$  a été fixée à deux afin d'amplifier l'effet de variation des termes. Ainsi, une faible variation dans les termes, auxquels ils sont attribués, sera perçue.

### 5.1.2 FONCTIONNEMENT DE LA MÉTHODE DES SIMPLEX<sup>(78,79,82)</sup>

L'optimisation par simplex est effectuée avec l'aide du logiciel informatique, afin de faciliter la compilation des valeurs numériques qui deviennent d'autant plus nombreuses qu'il y a de pics chromatographiques. Pour débiter une expérience d'optimisation par simplex il faut d'abord déterminer les paramètres qui affectent la séparation chromatographique. Il faut porter une attention particulière à choisir des paramètres indépendants les uns des autres. Par exemple, la composition initiale en éluant organique donne aussi la composition initiale en éluant aqueux puisqu'il n'y a que deux éluants. Alors, il faut en choisir qu'une seule comme paramètre car il est évidemment

impossible de varier les deux simultanément puisqu'elles sont dépendantes l'une de l'autre.

### 5.1.2.1 Formation du premier simplex.

Afin de faciliter l'explication de la méthode des simplex, le système témoin se limitera à deux paramètres variables indépendants, A et B. Pour former le premier simplex, il faut effectuer  $k+1$  analyses, où  $k$  est le nombre de paramètres, donc trois analyses. Les conditions pour ces analyses sont aléatoires mais doivent être comprises dans les frontières délimitées pour chaque paramètre (voir 5.2.1). Les valeurs attribuées aux paramètres doivent être différentes pour chacune des analyses. Pour chaque analyse effectuée, on calcule la valeur de CRF.

### 5.1.2.2 Identification des sommets.

Pour le système témoin, le simplex sera donc une forme géométrique triangulaire dont les trois sommets représentent chaque analyse effectuée. Ces sommets seront identifiés selon la valeur de CRF obtenue pour chacun d'eux (figure 5.1). Le sommet ayant la valeur de CRF la plus élevée sera le sommet M (meilleur), celui ayant la valeur la plus faible sera le P (pire) alors que celui entre les deux sera le sommet I (intermédiaire).

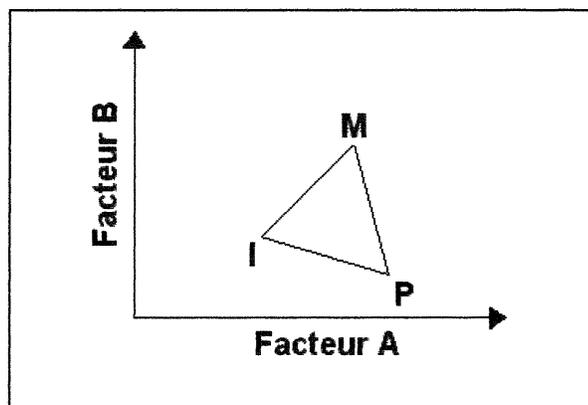


Figure 5.1 Premier simplex.

### 5.1.2.3 Les déplacements des simplex.

Pour obtenir le nouveau simplex, il faut réaliser une autre analyse ( $R_1$ ) selon ce que dicte le premier simplex de la figure 5.1. En effet, en rejetant le pire sommet (P) et en effectuant sa réflexion (trait pointillé) par le milieu de l'arête du triangle joignant les autres sommets, les paramètres A et B à ajuster pour la prochaine analyse sont connus (point  $R_1$ ).

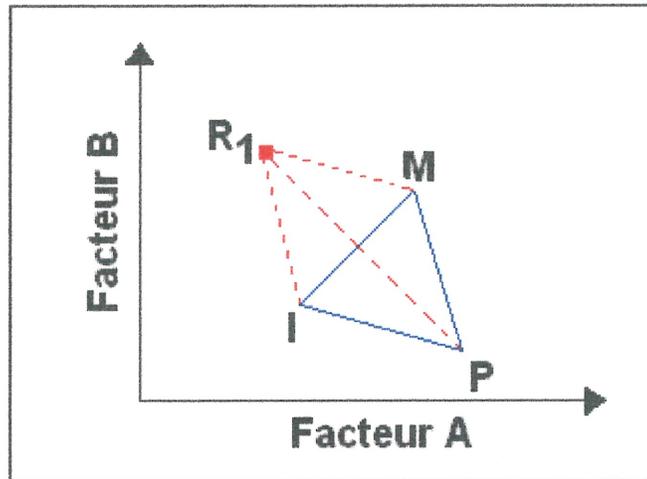


Figure 5.2 Détermination de la nouvelle analyse à effectuer pour le deuxième simplex.

La valeur de CRF est évaluée pour la nouvelle analyse ( $R_1$ ) et les trois sommets du deuxième simplex sont identifiés (triangle rouge) selon les appellations M, I et P.

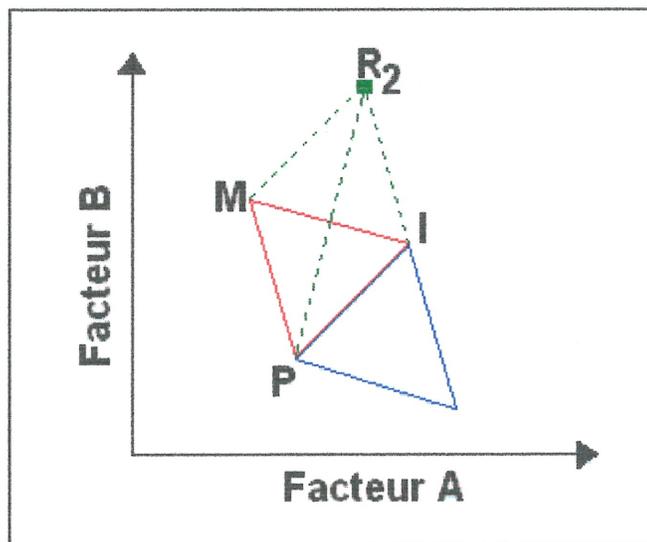


Figure 5.3 Détermination de la nouvelle analyse à effectuer pour le troisième simplex.

Alors, les paramètres A et B de l'analyse suivante ( $R_2$ ) peuvent être déterminés de la même façon que la précédente et ainsi de suite.

#### 5.1.2.4 Déplacements exceptionnels.

Il y a certaines situations, où les déplacements ne peuvent pas être faits de la manière décrite précédemment. En effet, lorsque le nouveau sommet obtenu représente le pire sommet du nouveau simplexe, il y a oscillation et aucune convergence vers le sommet n'est possible (fig.5.4).

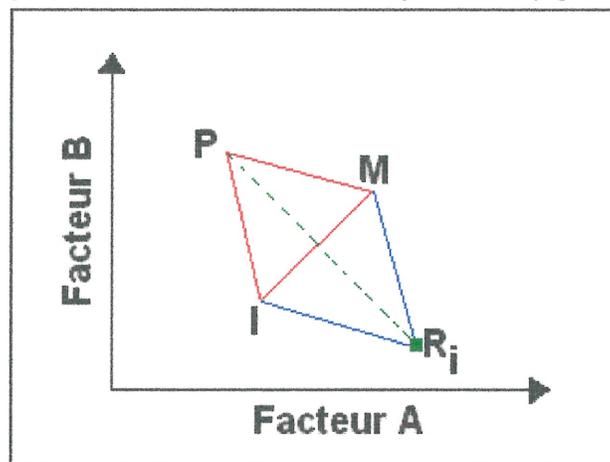


Figure 5.4 Oscillation des sommets.

Afin d'éliminer ce problème, s'il se présente, il faut plutôt rejeter le sommet qui précède le pire sommet et en faire la réflexion (trait pointillé). Ainsi, le sommet rejeté serait plutôt le sommet intermédiaire I.

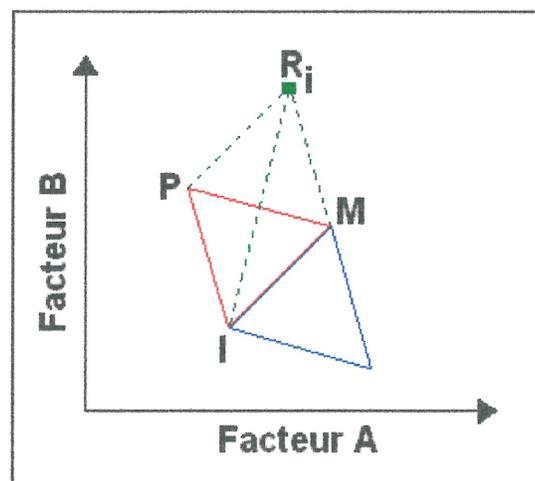


Figure 5.5 Contrecarrer l'oscillation.

Lorsqu'un sommet réfléchi dépasse les limites assignées à un facteur, on ne fait aucune analyse et on assigne à ce sommet une valeur non désirable de CRF et on poursuit selon les règles préalablement établies. Si un même sommet demeure après  $n+1$  simplex, il est probable que ce sommet soit la valeur optimale du système, ou une valeur qui en est très proche. Cependant, il est préférable de réévaluer la réponse à ce sommet par un autre analyse afin de s'assurer qu'il n'a pas été engendré par des erreurs de manipulations. Si ce sommet demeure, alors on peut dire que les conditions de chromatographie optimales ont été trouvées et donc faire toutes les analyses subséquentes dans ces conditions.

Afin d'accélérer la convergence vers la valeur optimale, on peut faire une expansion. C'est-à-dire que lorsqu'un sommet réfléchi représente le meilleur sommet du nouveau simplex, une expansion vers un sommet (E) situé sur l'axe à une distance de 1,5 fois plus grande du sommet rejeté que le premier sommet réfléchi est effectuée. De la même façon, si le sommet réfléchi est le pire sommet du nouveau simplex, on peut réaliser une contraction vers un sommet (C) situé sur l'axe de réflexion mais à une distance 25% moindre du sommet rejeté que le premier sommet réfléchi. Ainsi, le simplex se dirige moins rapidement vers la mauvaise direction. Un schéma représentant ces deux conditions exceptionnelles se retrouve à la figure 5.6 .

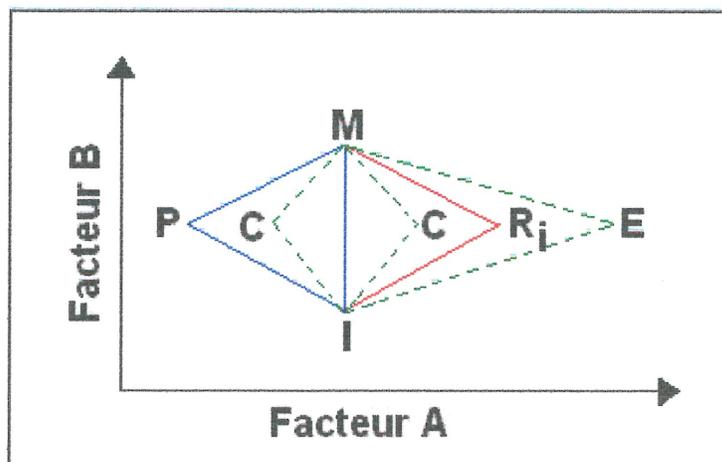


Figure 5.6 Expansion et contraction d'un simplex.

Étant donné la complexité qu'engendre l'utilisation d'un système d'optimisation par simplex à quatre différents paramètres ( $k=4$ ), au lieu de deux comme dans le système témoin présenté pour l'explication du fonctionnement des simplex, il devient évident que l'informatique est d'un grand secours dans le cas d'intérêt. Il ne reste plus qu'à effectuer les cinq analyses du simplex de départ ( $k+1$ ) et qu'à suivre ce que le logiciel recommande pour les autres analyses.

### 5.2.1 CHOIX DES PARAMÈTRES ET DES CONTRAINTES CHROMATOGRAPHIQUES

Lors de l'optimisation de la séparation chromatographique des fragments tryptiques de la  $\beta$ -caséine, les paramètres variables indépendants choisis sont les suivants : le débit, la concentration en acide formique dans l'éluant aqueux, la composition initiale en éluant organique et le temps d'analyse. Puisqu'il y a quatre paramètres ( $k=4$ ) pour cette recherche, on doit donc faire cinq analyses ( $k+1$ ) afin d'obtenir le premier simplex. Le simplex à quatre paramètres mène à une forme géométrique polyédrique à cinq sommets qui explique le choix de l'utilisation d'un logiciel informatique pour simplifier le travail. Les sommets seront identifiés de S1 à S5 selon la grandeur de la valeur de CRF. Les contraintes pour chacun des paramètres (voir tableau V) doivent aussi être déterminées afin de ne pas avoir à effectuer des analyses qui causeraient des difficultés pour la colonne chromatographique.

En effet, le débit dans la colonne chromatographique utilisée ne doit pas excéder la valeur de 1,5 mL/min sous peine d'affecter la qualité de la phase stationnaire. Pour ce qui est de la concentration en acide formique maximum elle dépend surtout du pH obtenu avec celle-ci. Le pH ne doit pas être plus acide que deux pour la préservation de la phase stationnaire et pour limiter l'acidité qui affecte l'efficacité d'ionisation dans le spectromètre de masse.

**Tableau V : Contraintes soumises au logiciel informatique de simplex pour chaque paramètres**

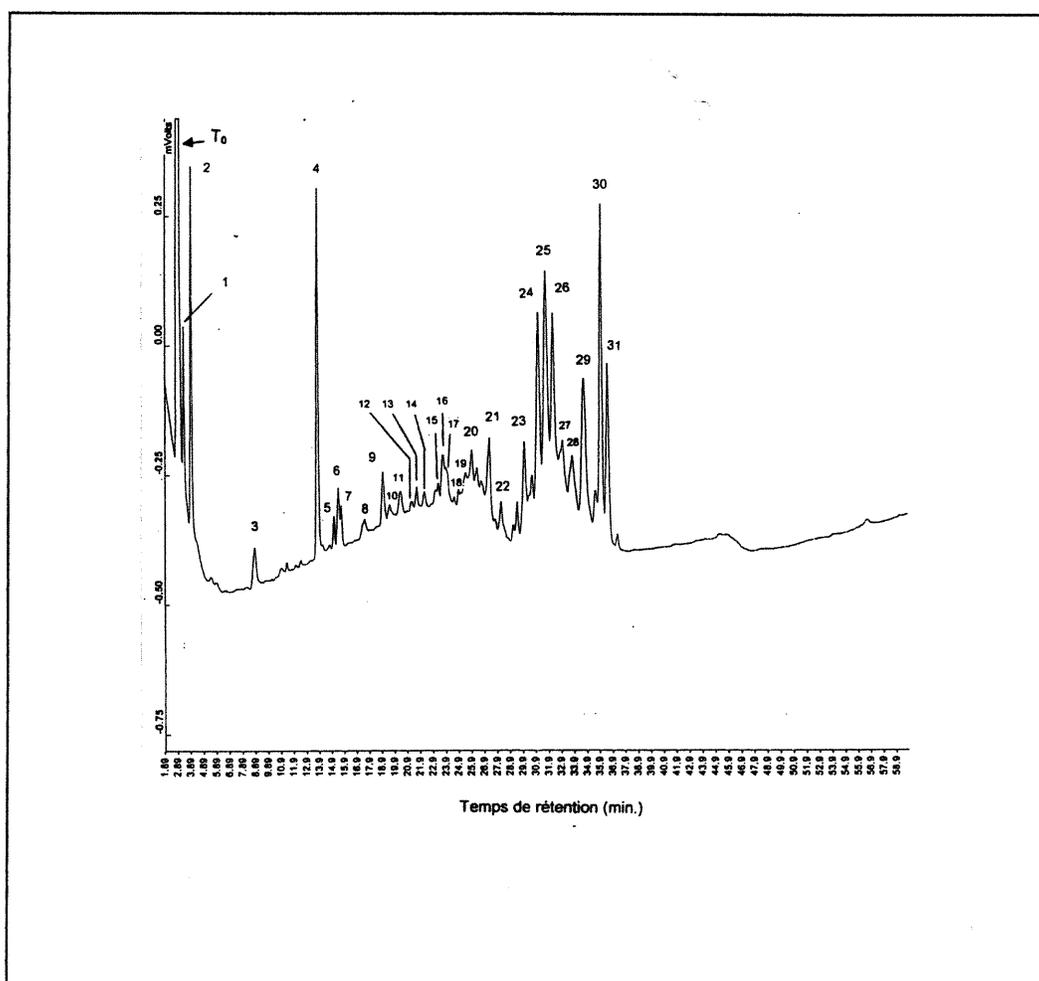
Paramètres	Valeur minimum	Valeur maximum
Débit (mL/min)	0,5	1,5
Temps d'analyse (min)	0	60
Composition initiale en éluant organique (%)	0	50
Concentration en acide formique dans l'éluant aqueux (% <sup>V</sup> /V)	0	0,3

### 5.2.2 OPTIMISATION DES CONDITIONS CHROMATOGRAPHIQUES POUR LA $\beta$ -CASÉINE

Afin de s'assurer du bon fonctionnement de cette méthode des simplex, la protéine choisie pour la tester est la  $\beta$ -caséine. En effet, la  $\beta$ -caséine étant une protéine de masse élevée (24kDa), elle contient plusieurs acides aminés, ce qui permet d'obtenir théoriquement 14 fragments protéolytiques par la digestion enzymatique.<sup>(83)</sup> Par la suite, les conditions chromatographiques pourront être utilisées pour d'autres peptides de plus petite taille. En effet, les conditions d'une bonne séparation chromatographique pour un mélange ayant plusieurs peptides tryptiques, donc plusieurs pics, permettra la séparation d'un mélange d'un plus petit nombre de fragments protéolytiques.

La solution de protéolyse utilisée lors de ces expériences d'optimisation des conditions chromatographiques est celle provenant d'une digestion enzymatique en solution et non de la protéolyse en microréacteur. Cependant, les conditions chromatographiques ne sont pas affectées par le fait que ce soit une digestion en solution ou une digestion en microréacteur. Pour débiter l'optimisation, la séparation chromatographique des peptides protéolytiques a été faite dans les conditions d'éluants contenant de l'acide trifluoroacétique (TFA), afin de voir ce que serait une bonne séparation en utilisant les conditions optimales publiées.<sup>(47)</sup> Le chromatogramme obtenu

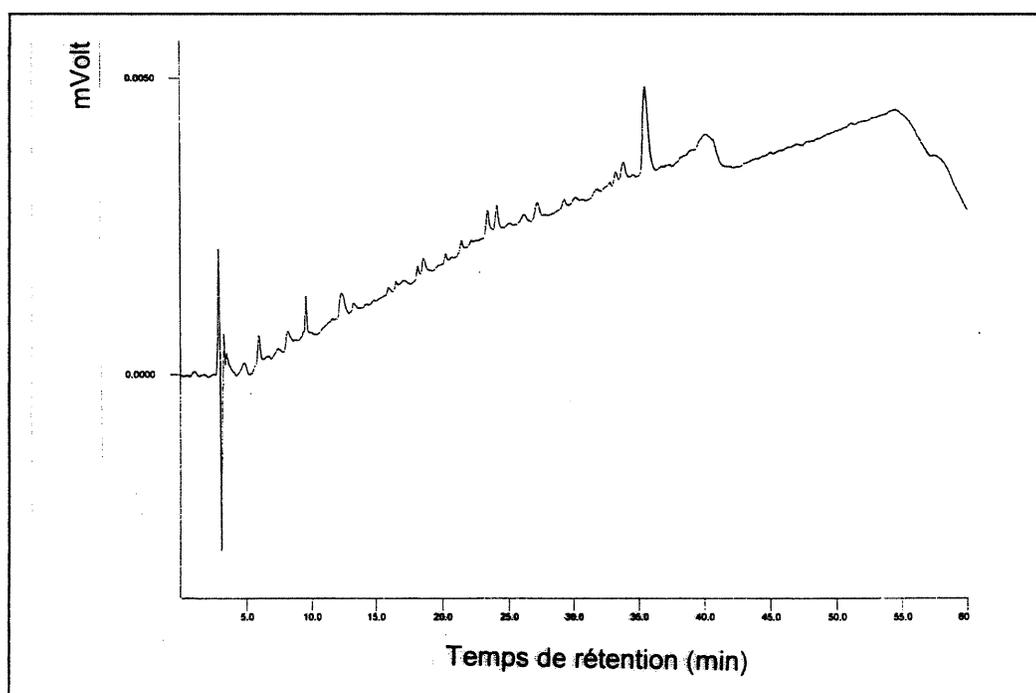
(figure 5.7) permet ainsi de déterminer le nombre de pics théoriques (L) de l'équation 5.1 précédente. Ce nombre L a donc été fixé à 31 pics puisqu'il semble que même s'il ne devrait y avoir que 14 fragments protéolytiques, il y a aussi des isoformes de la  $\beta$ -caséine, des peptides provenant de l'autoprotéolyse de la trypsine en solution (figure 5.8) ainsi que des impuretés, qui doivent être bien résolus des analytes d'intérêt.<sup>(83)</sup> Plus il y a de pics chromatographiques à séparer, plus les conditions chromatographiques seront optimales pour un grand nombre de peptides différents.



**Figure 5.7 Chromatogramme du mélange de peptides protéolytiques de la  $\beta$ -caséine.**

(Conditions: Éluant A : 0,1% TFA /H<sub>2</sub>O , Éluant B : 0,1% TFA/ACN, gradient de 5% à 70% B en 60 minutes, débit : 1mL/min)

Les pics # 1,2,4,5,6,10,19 et 30 correspondent à des fragments tryptiques de la  $\beta$ -caséine selon une analyse LC-MS subséquente. Alors, les autres pics correspondent aux isoformes de la  $\beta$ -caséine (>90% pure), aux peptides conséquents de l'autoprotéolyse de la trypsine en solution ainsi qu'à des peptides provenant de clivages incomplets ou non spécifiques de la  $\beta$ -caséine par la trypsine.<sup>(3)</sup> Par analyse LC-MS, on obtient la masse moléculaire des produits correspondant à chaque pics chromatographiques. Afin de déterminer quels pics est dû aux fragments tryptiques du peptides d'intérêt, il faut tenir compte du fait que ce projet a été fait pour être complémentaire à l'étude cinétique de la réaction de protéolyse. Alors, lorsque l'étude cinétique est faite, les masses des fragments tryptiques provenant de la protéolyse du peptide natif sont connus et par conséquent les pics leur sont facilement attribués.



**Figure 5.8 Chromatogramme du blanc de la protéolyse avec la trypsine en solution.**

(Conditions: Éluant A : 0,17% Acide formique/5%ACN/H<sub>2</sub>O ,  
Éluant B : 0,06% Acide formique /5%H<sub>2</sub>O /ACN,  
gradient de 7,6% à 65% B en 46,9 minutes, débit : 0,7mL/min)

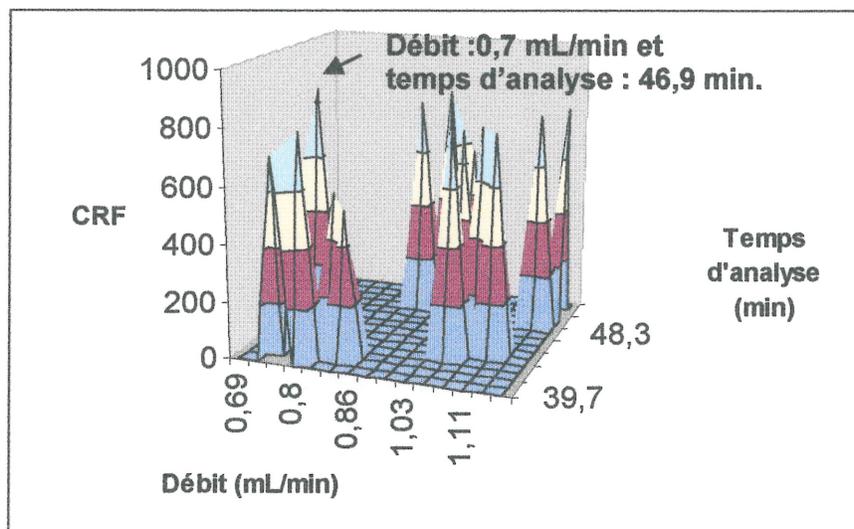
La présence des peptides conséquents de l'autoprotéolyse de la trypsine en solution est prouvée par le chromatogramme du blanc de la digestion en

solution (figure 5.8). En effet, bien qu'il ne devrait pas y avoir de pics dans ce chromatogramme, on remarque très bien la présence de plusieurs pics qui seront interférents dans les chromatogrammes de la digestion en solution de la  $\beta$ -caséine. Il peut aussi y avoir des réactifs en excès dans la solution qui peuvent mener à des pics chromatographiques obtenus pour le blanc.

Les résultats de CRF obtenus pour chaque analyse demandée par le logiciel se retrouvent dans le tableau VI avec les cinq résultats obtenus pour le premier simplex. Selon les résultats obtenus, les conditions chromatographiques optimales sont celles de l'essai 8 du tableau. Les graphiques en trois dimensions (figure 5.9 et 5.10) montrent la variation de la valeur de CRF en fonction de la variation de chaque paramètres dans les frontières données comme contraintes.

**Tableau VI : Valeurs de CRF obtenues dans le cadre de l'optimisation par simplex de la séparation chromatographique des peptides protéolytiques de la  $\beta$ -caséine.**

Essai	%Acide formique/H <sub>2</sub> O (%)	Concentration initiale en éluant organique (%)	Temps d'analyse (min)	Débit (mL/min)	CRF
1	0,3	5	60	1	569,3
2	0,25	8	45	1,1	785,0
3	0,2	7	50	1,2	783,9
4	0,15	6	55	0,9	802,6
5	0,1	9	40	0,8	793,8
6	0,05	10	35	1	194,2
7	0,24	6,25	53,8	1	625,9
<b>8</b>	<b>0,17</b>	<b>7,6</b>	<b>46,9</b>	<b>0,7</b>	<b>850,1</b>
9	0,10	9,0	39,7	0,75	722,0
10	0,13	8,3	43,2	0,81	563,5
11	0,15	8,0	44,7	1,11	777,2
12	0,19	7,2	49,1	1,14	779,0
13	0,20	7,0	49,8	0,86	784,2
14	0,16	7,8	45,8	0,69	451,4
15	0,18	7,3	48,3	1,03	723,7
16	0,14	8,1	44,3	1,05	841,7
17	0,15	6	55	0,9	841,7
18	0,14	8,2	43,8	0,9	841,6
19	0,1	9	40	0,8	568,8
20	0,12	8,7	41,6	0,83	521,0
21	0,25	8	45	1,1	669,3

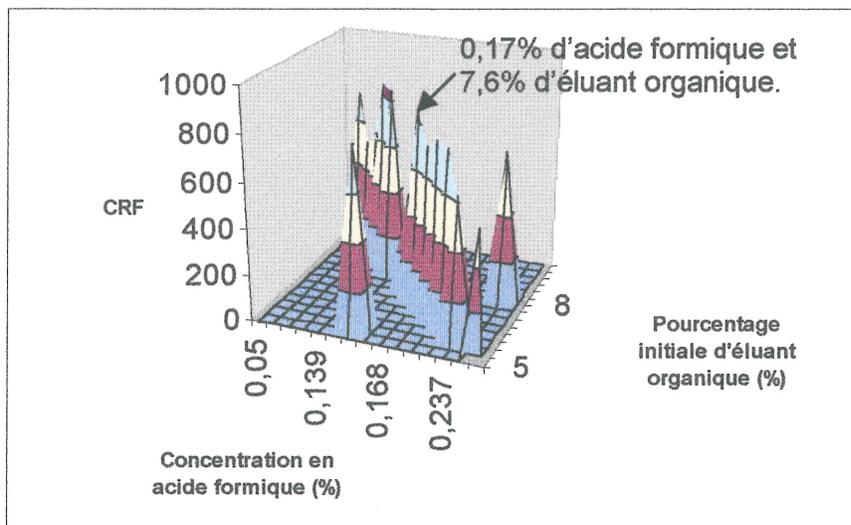


**Figure 5.9 Détermination du débit et du temps d'analyse optimum.**

Puisque les logiciels informatiques ne permettent pas la visualisation de formes de plus de trois dimensions, les graphiques ont été effectués à partir de paires de paramètres tout en gardant les deux autres constants. Ainsi, il sera possible de mieux visualiser pourquoi les conditions choisies par le logiciel sont considérées comme les meilleures.

Il est facile de voir que le débit et le temps d'analyse n'ont pas beaucoup d'effet sur la séparation chromatographique puisqu'il y a une dispersion aléatoire de sommets. (figure 5.9) Cependant, les valeurs optimales choisies sont celles correspondant au plus haut sommet. Ces deux paramètres étant des paramètres mécaniques de la chromatographie liquide à haute performance, ils ont un moindre effet sur la séparation des peptides.

La paire de paramètres que font la concentration en acide formique et la concentration de départ de l'éluant organique montrent qu'il y a une variation organisée de la fonction de réponse face à ceux-ci (figure 5.10). En effet, les fonctions de réponses obtenues sont centrées sur le graphique et ne varient pas de façon aléatoire. Ceci démontre que ces paramètres chimiques ont une plus grande influence sur la séparation des peptides.



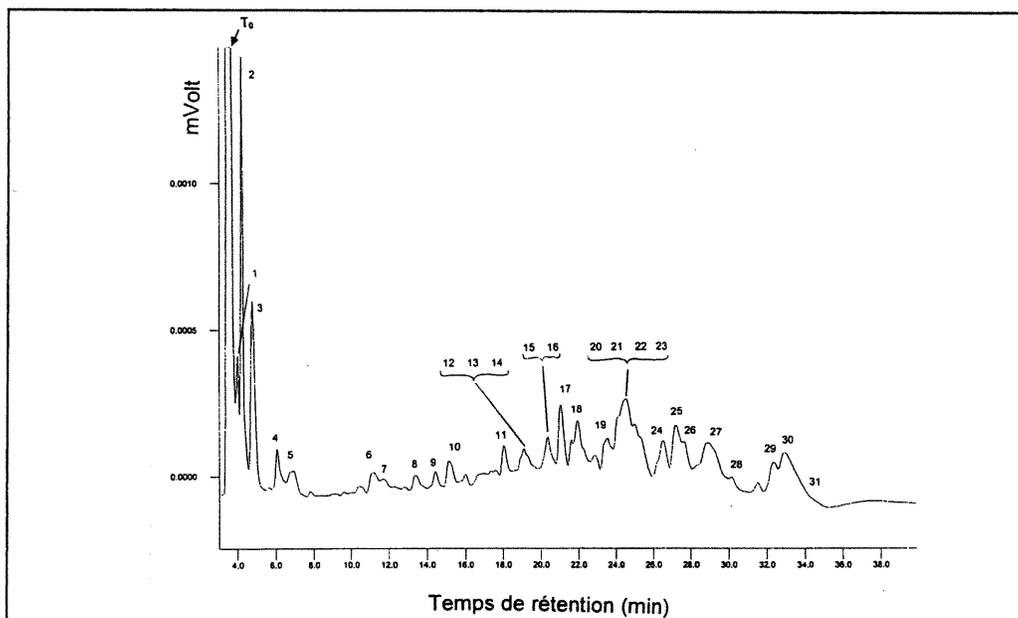
**Figure 5.10 Détermination de la concentration en acide formique et la concentration initiale de l'éluant organique optimum.**

Ces graphiques montrent aussi la capacité du logiciel utilisé à tenir compte du fait qu'il peut y avoir plusieurs sommets. Cela est très important, car le maximum d'un sommet n'est pas nécessairement le plus haut maximum possible.

Alors, en voyant les graphiques en trois dimensions et le tableau des valeurs de fonction de réponse chromatographiques, il est évident que le logiciel est efficace et que les conditions optimales obtenues avec celui-ci sont les meilleures pour le mélange analysé. Les figures 5.11 et 5.12 montrent des chromatogrammes obtenus dans des conditions intermédiaires, ayant des valeurs de CRF plus faibles que celle des conditions optimales. Il est visible que ces conditions ne sont pas optimales car plusieurs pics y sont non résolus.

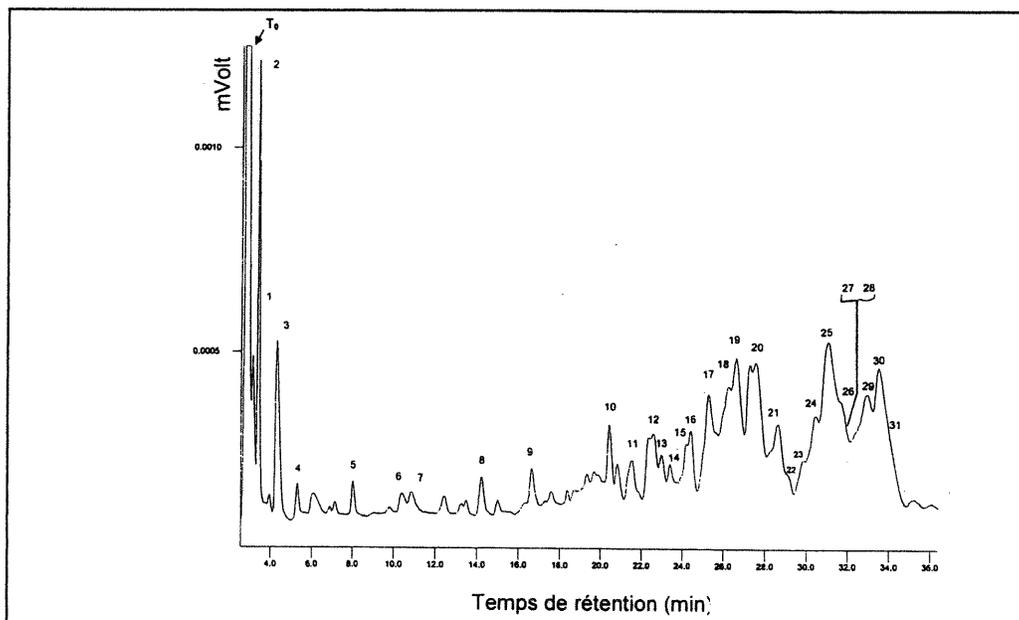
En effet, dans la figure 5.11, il y a plusieurs groupes de pics chromatographiques non résolus ; 12,13,14 ; 15,16 et 20,21,22,23. Cela explique la valeur de CRF faible car la fonction de réponse chromatographique tient compte de la résolution entre les paires de pics. La figure 5.12 montre une CRF plus élevée car les groupes de pics 12,13,14 et 20,21,22,23 sont mieux résolus entre eux. Cependant, les conditions

chromatographiques ne peuvent obtenir la meilleure valeur de CRF car les pics 15 et 16 ainsi que 27 et 28 ne sont pas bien résolus.



**Figure 5.11 Chromatogramme du mélange de peptides tryptiques de la  $\beta$ -caséine donnant une CRF = 568,8 .**

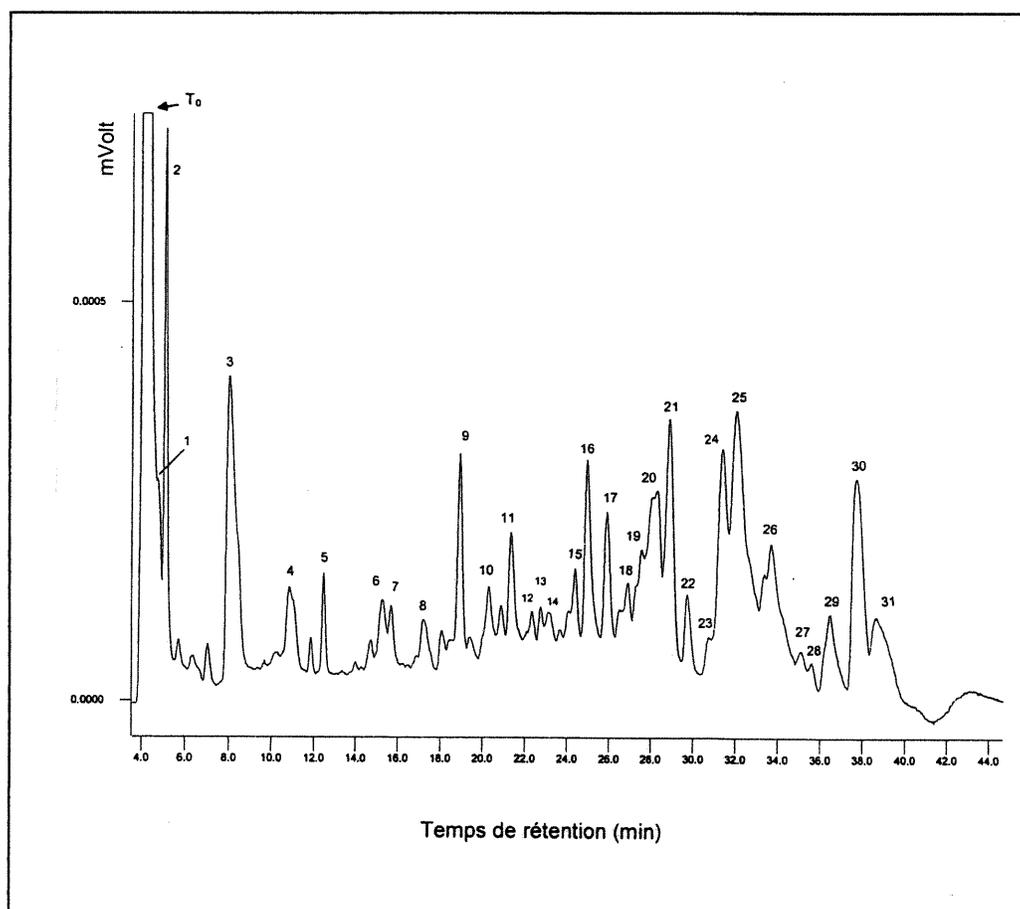
(Conditions: Éluant A : 0,1% Acide formique/5%ACN/H<sub>2</sub>O ,  
Éluant B : 0,03% Acide formique /5%H<sub>2</sub>O /ACN,  
gradient de 9% à 65% B en 40 minutes, débit : 0,8mL/min)



**Figure 5.12 Chromatogramme du mélange de peptides tryptiques de la  $\beta$ -caséine donnant une CRF = 723,7 .**

(Conditions: Éluant A : 0,18% Acide formique/5%ACN/H<sub>2</sub>O ,  
Éluant B : 0,06% Acide formique /5%H<sub>2</sub>O /ACN,  
gradient de 7,3% à 65% B en 48,3 minutes, débit : 1,03 mL/min)

Il est visible que le chromatogramme obtenu dans ces conditions déterminées comme optimales (figure 5.13) est le même que celui fait en utilisant les conditions optimum avec du TFA (figure 5.7). Cependant, les pics sont plus larges et cela est dû au pairage d'ions qui est plus faible avec l'acide formique qu'avec l'acide trifluoroacétique. En effet, puisque le pairage d'ions est plus faible, les peptides à séparer ne sont pas tous neutres ce qui fait que la séparation ne sera pas uniquement due au mode de partage. Alors, il y aura un élargissement des pics dû aux peptides qui seront séparés par d'autres modes de séparation.



**Figure 5.13 Chromatogramme du mélange de peptides protéolytiques de la  $\beta$ -caséine dans les conditions optimales.**

(Conditions: Éluant A : 0,17% Acide formique/5%ACN/H<sub>2</sub>O ,  
Éluant B : 0,06% Acide formique /5%H<sub>2</sub>O /ACN,  
gradient de 7,6% à 65% B en 46,9 minutes, débit : 0,7mL/min)

En conclusion, puisque les conditions optimales trouvées pour la  $\beta$ -caséine donnent d'excellents résultats, il semble correct de penser que ces conditions chromatographiques peuvent être utilisées pour toutes les grosseurs de peptides et protéines en affinant un peu les conditions. Ainsi, ce sont les conditions qui sont utilisées lors de toutes les analyses LC-MS et LC-MS/MS qui suivront dans les prochains chapitres.

Ces conditions chromatographiques permettent de séparer les peptides protéolytiques d'un mélange complexe, ceci est la première étape du séquençage des peptides. En effet, il faudra maintenant déterminer la masse moléculaire de ces fragments tryptiques par LC-MS et obtenir les spectres de masse en tandem pour avoir assez d'informations significatives pour la reconstruction de la protéine native ou le peptide de départ.

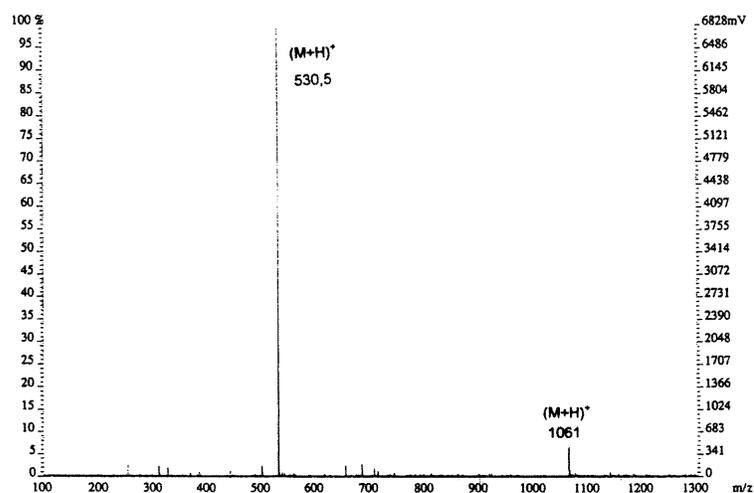
**CHAPITRE 6**  
**RÉSULTATS DES ANALYSES LC-MS**

## 6.1 LES ANALYSES PAR LC-MS

Les analyses d'une solution résultant de protéolyse, par le couplage de la chromatographie en phase liquide et la spectrométrie de masse, permettent l'obtention du spectre de masse de chaque pic chromatographique correspondant à un peptide tryptique. L'identification de la masse moléculaire de chaque peptide protéolytique est alors possible. Pour cette partie de la recherche, les solutions utilisées sont celles provenant de la digestion enzymatique faite en microréacteur pendant trois heures, puisque cela suffit à la digestion complète des peptides utilisés et que cela est plus rapide que d'attendre les 24 heures de la digestion en solution.

Comme il a été discuté à la section 3.2.3 du présent document, l'ionisation par électronébulisation produit des ions multichargés. Donc pour un même ion, le spectre de masse peut présenter plusieurs pics. Cependant, le nombre de pics attendus pour un peptide dépend de sa masse moléculaire. Bien que la limite de l'ionisation par électronébulisation est d'une charge par kDa, il arrive de voir jusqu'au plus deux charges par kDa. Ainsi, il faut avoir à faire à une protéine ou un peptide de masse élevée pour obtenir un spectre de masse avec une distribution gaussienne de pics de masse, telle qu'à la figure 3.5. Or, les peptides protéolytiques sont de petites tailles puisqu'il s'agit là d'un avantage de l'utilisation de la digestion enzymatique préalable à la chromatographie. Alors, les spectres de masse obtenus seront affublés d'au plus deux pics de masse uniquement pour un même ion, comme celui obtenu pour la bradykinine ( $MM = 1060,23 \text{ g/mol}$ ).

Comme le démontre la figure 6.1, le spectre de masse de la bradykinine comporte deux ions, l'ion monochargé  $(M+H)^+$  au  $m/z = 1061$  et l'ion doublement chargé  $(M+H)^{2+}$  au  $m/z = 530,5$ . La présence d'ions multichargés pourrait compliquer les spectres de masse obtenus. Cependant, les peptides à analyser donnent des fragments protéolytiques de masse inférieure ou égale à  $700 \text{ g/mol}$ , alors la présence d'ions multichargés sera évitée.



**Figure 6.1 Spectre de masse de la bradykinine.**

## 6.2 ANALYSES LC-MS DES DIFFÉRENTS PEPTIDES

Les expériences de couplage de la chromatographie en phase liquide et la spectrométrie de masse ont été faites sur des peptides de petites tailles comme l'angiotensine II (MM = 1045,53 g/mol) et la ranatachykinine (MM = 1309,65 g/mol) afin de faciliter les analyses. En effet, l'utilisation de petits peptides assure l'obtention d'un petit nombre de peptides protéolytiques, ce qui limite le nombre de pics chromatographiques et donc la complexité de l'analyse.

### 6.2.1 L'ANALYSE LC-MS DE L'ANGIOTENSINE II

La digestion enzymatique par la trypsine de l'angiotensine II mène à la production de deux fragments protéolytiques (figure 6.2). Alors, comme il y a deux fragments protéolytiques attendus, il n'y a que deux pics chromatographiques importants (figure 6.3). Le chromatogramme du blanc de la digestion en microréacteur (figure 6.4) montre une absence de pics. Alors, les pics ne correspondant pas aux fragments tryptiques sur les chromatogrammes des digestions enzymatiques en microréacteur, ne

peuvent provenir de l'autoprotéolyse de la trypsine.<sup>(81)</sup> Il s'agit donc d'impuretés dans les standards de peptides utilisés et de peptides dus aux clivages incomplets ou non spécifiques par la trypsine du peptide analysé.<sup>(3)</sup>

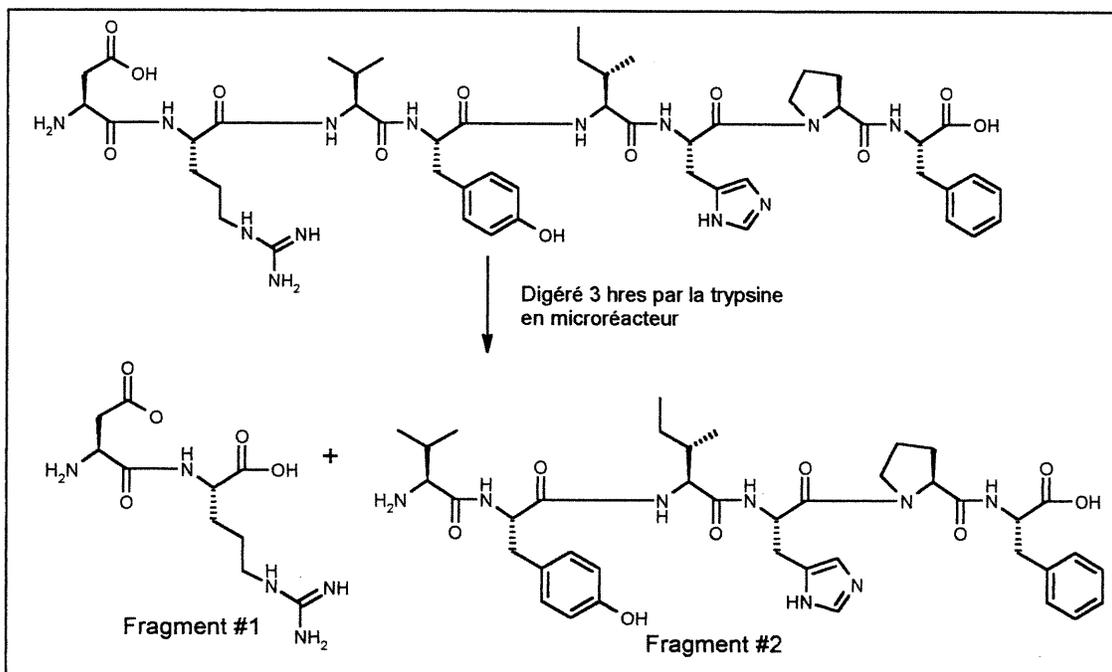


Figure 6.2 Peptides protéolytiques obtenus pour l'angiotensine II.

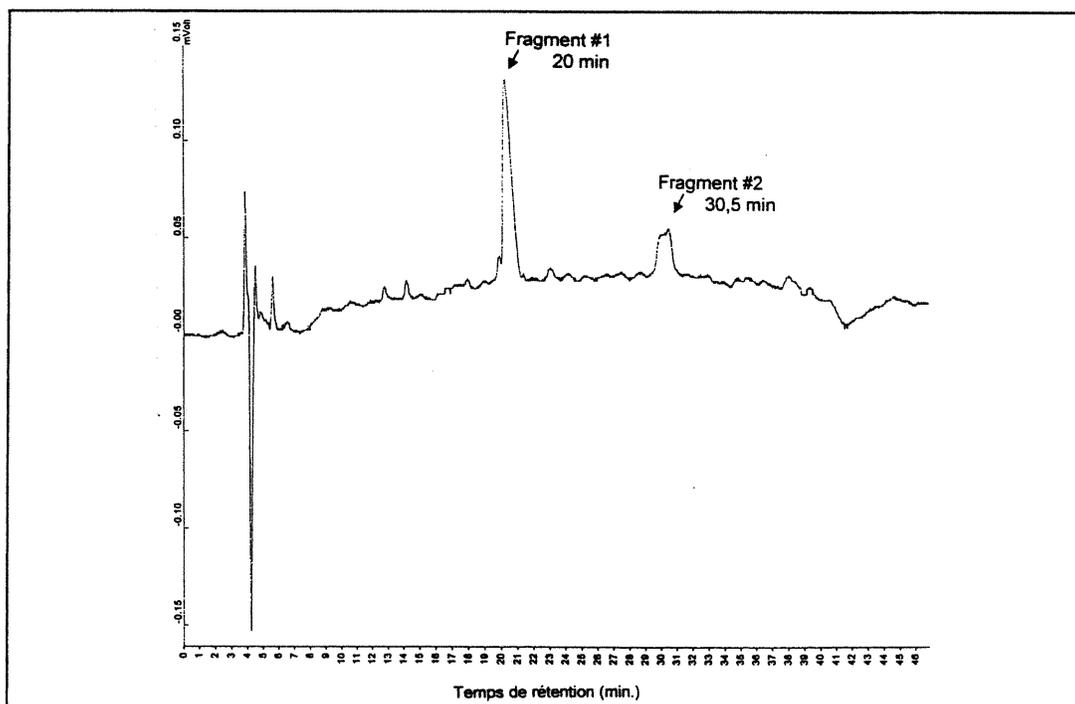
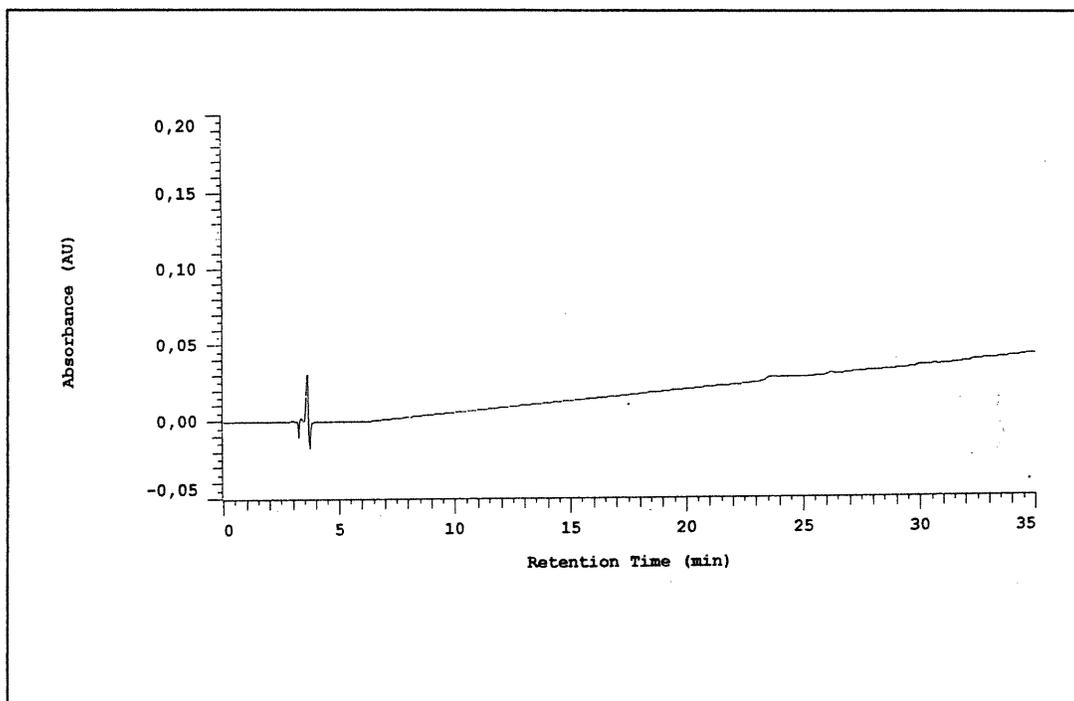


Figure 6.3 Chromatogramme du mélange protéolytique de l'angiotensine II, obtenu par digestion de trois heures en microréacteur.

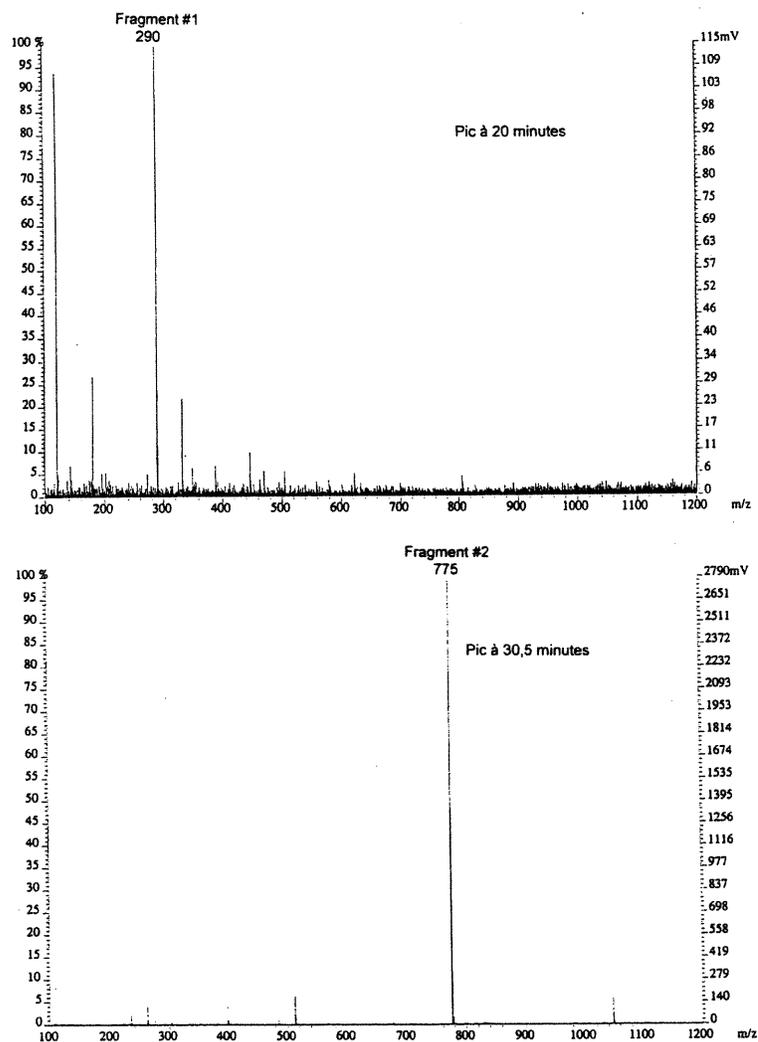


**Figure 6.4 Chromatogramme du blanc de la digestion en microréacteur.**

Ce sont les spectres de masse suivants (figure 6.5) qui permettent d'identifier les pics chromatographiques comme étant les peptides protéolytiques #1 et #2. Le spectre de masse du pic chromatographique au temps de rétention de 20 minutes montre un ion moléculaire à  $m/z = 290$ , ce qui correspond au  $(M+H)^+$  du peptide Asp-Arg et l'ion doublement chargé  $(M+H)^{2+}$  à  $m/z = 180$ . De même l'ion moléculaire de  $m/z = 775$  du spectre de masse du pic chromatographique à 30,5 minutes est le  $(M+H)^+$  du peptide tryptique Val-Tyr-Ile-His-Pro-Phe.

Les spectres de masse des pics chromatographiques, correspondant aux peptides tryptiques de l'angiotensine II, montrent des pics significatifs qui ne correspondent pas à ces peptides. L'ionisation par électronébulisation ne permettant pas la fragmentation étant une méthode d'ionisation ultra-douce, il est évident que les pics de masse ne sont pas des fragments dus à l'ionisation. Il y a quatre hypothèses pour ces pics inconnus : la contamination de la colonne chromatographique, l'effet de mémoire de la

source du spectromètre de masse, les produits chimiques constituant les éluants ou les impuretés provenant du standard d'angiotensine II utilisé.



**Figure 6.5 Spectres de masse de chaque pics du mélange protéolytique de l'angiotensine II.**

Il se peut que les pics de masse proviennent de produits présents dans la colonne chromatographique avant l'injection et qu'ils soient élués simultanément avec le peptide protéolytique. Cependant, cette hypothèse est peu plausible vu que des pics indésirables se retrouvent dans tous les spectres de masse. De plus, les conditions chromatographiques utilisées sont celles permettant la séparation d'un mélange de plusieurs peptides, même si rien indique que les produits contaminants sont des peptides, tout

porte à croire qu'ils devraient être séparés des analytes par la chromatographie utilisée, ce qui élimine cette hypothèse.

L'hypothèse de l'effet de mémoire de la source du spectromètre de masse est aussi peu vraisemblable. Des produits présents sur l'écrèmeur dans le spectromètre de masse (figure 3.2), avant l'injection dans le chromatographe, peuvent apporter des interférences dans les spectres de masse. En effet, l'appareil étant utilisé par les étudiants et le Centre Régional de Spectrométrie de Masse, il arrive parfois qu'il y ait un effet de mémoire pour certains produits. Cependant, la source a toujours été nettoyée avant chaque utilisation pour cette étude, ce qui élimine d'emblée cette hypothèse.

Puisque l'angiotensine II, provenant de la compagnie Sigma, est pure à 97%, alors il est évident qu'il y a des impuretés qui peuvent entraîner des pics significatifs dans les spectres de masse.

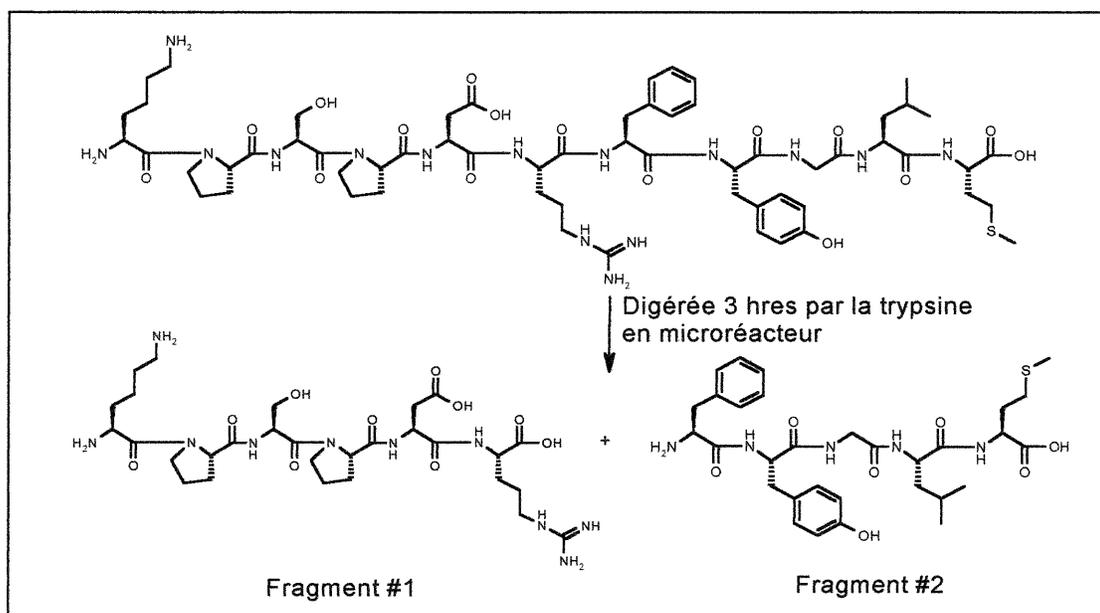
Les éluants utilisés lors de l'analyse LC-MS ne peuvent donner des pics de masse à  $m/z=120$  ou plus haut, car les constituants ; l'acétonitrile, l'eau et l'acide formique ont des masses moléculaires plus petites que 100. Les pics inconnus ne peuvent donc pas s'expliquer par cette hypothèse.

La seule explication possible est la présence d'impuretés qui sont élués en même temps que les fragments tryptiques qui apportent ces pics de masse. Mais, il est impossible de faire mieux puisqu'il s'agit de standards achetés. Bien que ces pics de masse inconnus soient d'intensité non négligeable, ils ne sont pas vraiment nuisibles, puisque les pics correspondant au peptide tryptique montrent une intensité beaucoup plus importante. L'ionisation par électrobulbion ne donnant que des ions multichargés, il est facile d'identifier le peptide protéolytique grâce à l'unique pic d'importance sur le spectre de masse. De plus, il faut toujours garder en tête que ce projet se fait en complémentarité avec celui de l'étude cinétique de la réaction de

digestion enzymatique. Cela apporte un autre support quant à l'identification des pics de masse correspondant aux fragments protéolytiques. C'est l'analyse par spectrométrie de masse en tandem de ces peptides tryptiques, après leur séparation chromatographique, qui permettra d'en connaître la séquence en acides aminés.

### 6.2.2 L'ANALYSE LC-MS DE LA RANATACHYKININE

La digestion enzymatique par la trypsine de la ranatachykinine mène, tout comme celle de l'angiotensine II, à la production de deux fragments protéolytiques.

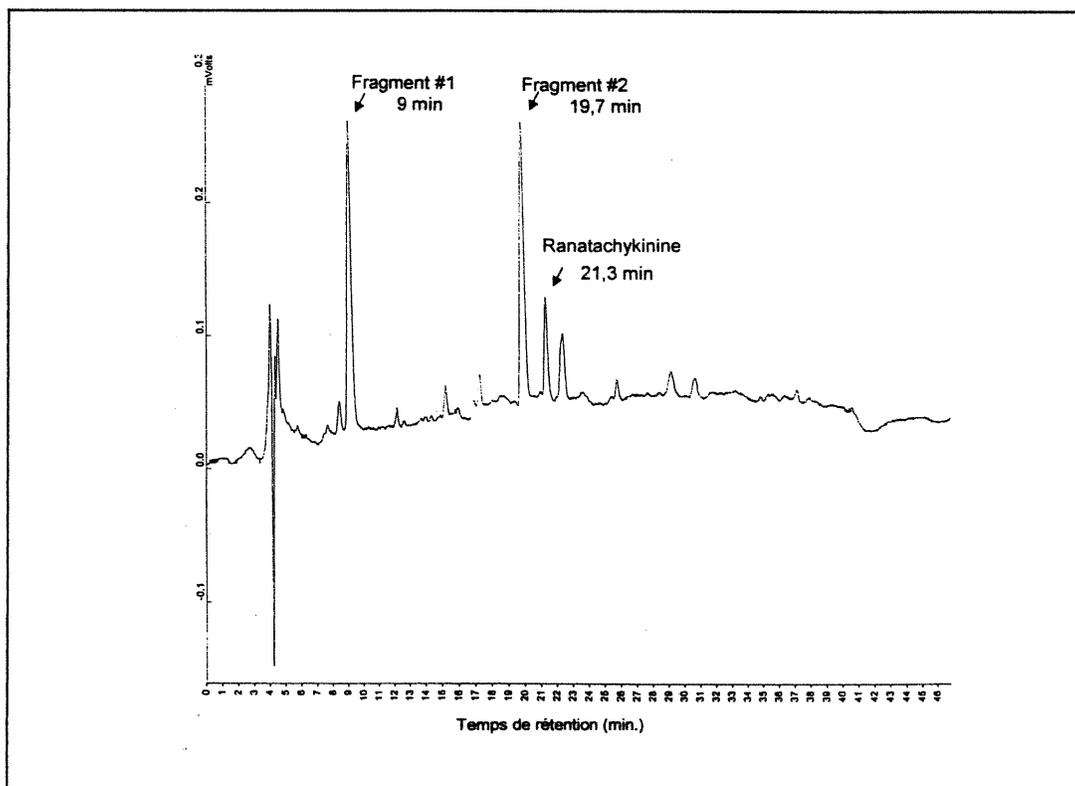


**Figure 6.6** Fragments protéolytiques obtenus pour la ranatachykinine.

Cependant, le chromatogramme (figure 6.7) du mélange protéolytique de la ranatachykinine montre plus de pics qu'il y a de fragments protéolytiques attendus. Les pics chromatographiques qui ne correspondent pas aux peptides tryptiques proviennent sans contredit des impuretés contenues dans le standard de ranatachykinine ayant une pureté de 98%. Comme précédemment discuté pour l'angiotensine II, la présence d'impuretés dans

le standard utilisé pour la digestion enzymatique est la seule hypothèse plausible. Cependant, puisque les pics interférents sont séparés par chromatographie, il est donc facile de les différencier des peptides protéolytiques d'intérêt par l'utilisation de la spectrométrie de masse.

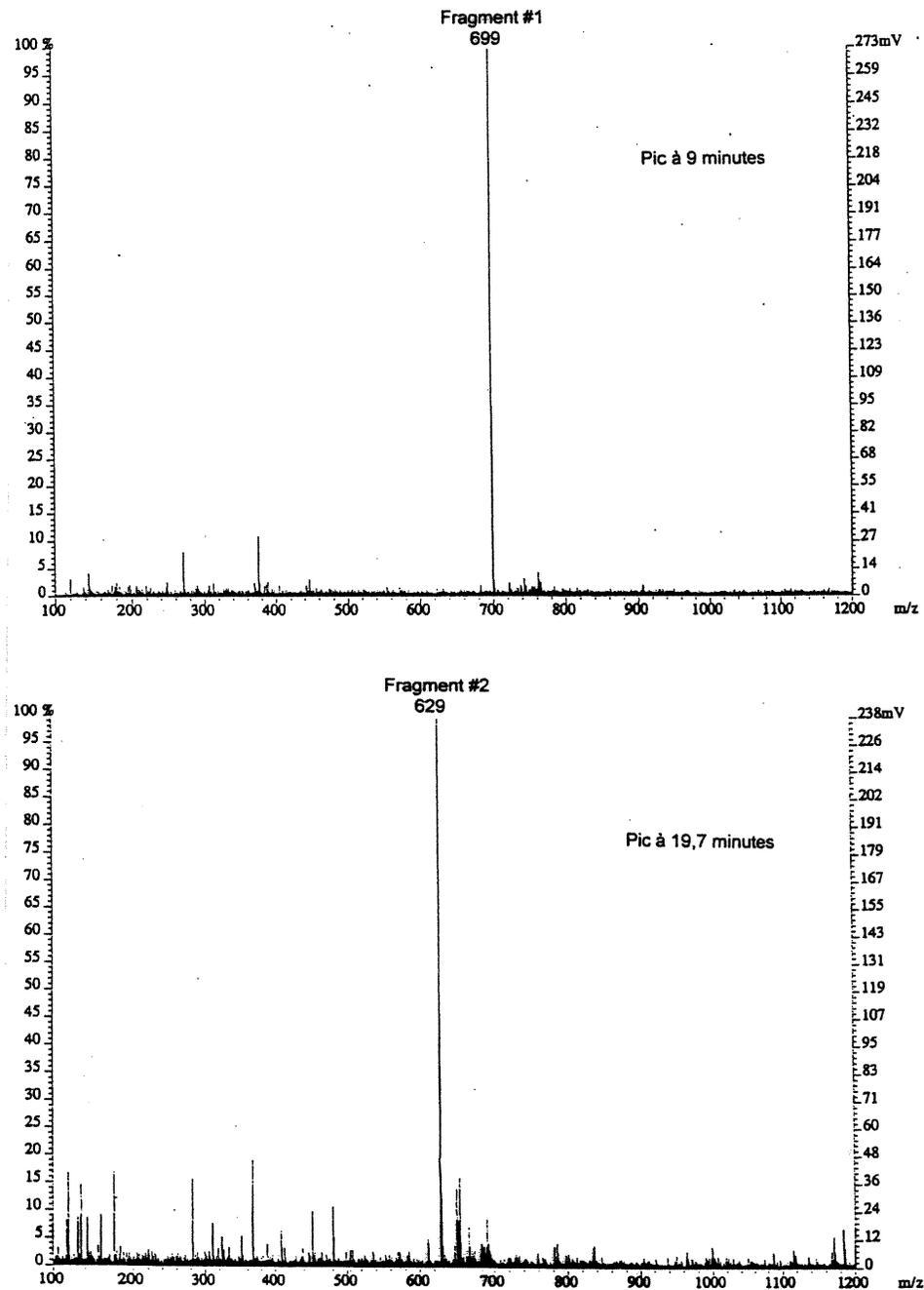
Les spectres de masse de chaque pics chromatographiques (figure 6.8) permettent l'identification des pics correspondants aux peptides protéolytiques. Ainsi, le pic chromatographique ayant un temps de rétention de 9 minutes donne un spectre de masse comportant un ion moléculaire à  $m/z = 699$ . Cet ion moléculaire correspond à l'ion  $(M+H)^+$  du peptide Lys-Pro-Ser-Pro-Asp-Arg ( $MM= 698,37$  g/mol).



**Figure 6.7** Chromatogramme du mélange protéolytique de la ranatachykinine, obtenu par digestion de trois heures en microréacteur.

Il en est de même pour le pic chromatographique à 19,7 minutes qui correspond au fragment #2, le Phe-Tyr-Gly-Leu-Met-NH<sub>2</sub>. En effet, l'ion moléculaire à  $m/z = 629$  du spectre de masse de ce pic chromatographique est l'ion  $(M+H)^+$  du peptide Phe-Tyr-Gly-Leu-Met-NH<sub>2</sub> ( $MM=628,30$  g/mol).

Pour ce qui est des autres pics chromatographiques, celui à 21,3 minutes est celui de la ranatachykinine de départ, ce qui signifie que la digestion enzymatique de trois heures effectuée dans le microréacteur n'était pas tout à fait complète.



**Figure 6.8** Spectres de masse de chaque pics du mélange protéolytique de la ranatachykinine.

En conclusion, ces analyses LC-MS des mélanges tryptiques, de l'angiotensine II et de la ranatachykinine, obtenus par une digestion de trois heures dans le microréacteur montrent que les 24 heures de digestion en solution ne sont pas nécessaires pour l'angiotensine. Pour la ranatachykinine, la digestion enzymatique n'est pas complète après trois heures en microréacteur. Cependant, puisque après trois heures, le pic de ranatachykinine restante est le tiers de ceux des fragments, il n'est pas faux de penser que la digestion sera complète en moins de 24 heures.

L'obtention de la masse moléculaire des fragments tryptiques est très utile pour l'analyse subséquente par spectrométrie de masse en tandem. En effet, la masse moléculaire des peptides protéolytiques ainsi obtenue, permettra d'effectuer des analyses par spectrométrie de masse en tandem sur ces peptides tryptiques en fixant la masse de l'ion moléculaire de ces peptides, comme la masse de l'ion parent. Ce sont ces analyses MS/MS qui donneront les dernières informations (fragments de séries A, B et Y) nécessaires à la reconstruction de la séquence en acides aminés des peptides tryptiques. Cependant, l'ordre dans lequel se trouvait les fragments protéolytiques avant la digestion enzymatique demeurera inconnue. Dans le cadre de notre recherche, c'est par l'étude cinétique de la protéolyse de la protéine native qui nous fournira ce renseignement.

**CHAPITRE 7**

**RÉSULTATS DES ANALYSES PAR**

**SPECTROMÉTRIE DE MASSE EN TANDEM**

## 7.1 ANALYSES PAR SPECTROMÉTRIE DE MASSE EN TANDEM

C'est grâce à l'analyse par spectrométrie de masse en tandem que la séquence en acides aminés des peptides protéolytiques peut être déterminée. En effet, après la digestion enzymatique, les peptides tryptidiques obtenus sont séparés par la chromatographie liquide et leur masse moléculaire est déterminée par l'utilisation du spectromètre de masse EBE, comme détecteur. Cependant, la spectrométrie de masse en tandem est la seule façon d'obtenir de l'information sur la composition en acides aminés de ces fragments protéolytiques. En se basant sur la nomenclature de Roepstorff préalablement discutée au chapitre 3, il est possible de connaître les acides aminés qui les composent ainsi que l'ordre dans lequel ils s'y retrouvent. La séquence en acides aminés des peptides tryptidiques est alors disponible. Les analyses par spectrométrie de masse en tandem ont été séparées en deux parties ; premièrement, le choix du gaz de collisions et deuxièmement, les analyses de peptides protéolytiques de deux peptides analytes.

### 7.1.1 CHOIX DU GAZ DE COLLISION

Lors des analyses par spectrométrie de masse en tandem, le choix du gaz de collisions qui sera utilisé pour permettre la fragmentation des ions parents, qui sont les peptides protéolytiques, est très important. C'est pour cette raison que l'étude des informations disponibles en spectrométrie de masse en tandem selon le gaz de collisions utilisé a été faite en premier lieu. En effet, il est normal de s'attendre à une fragmentation différente pour un même peptide avec des différents gaz et pour différents peptides avec le même gaz de collisions. Comme le démontre l'équation 3.10, la masse du gaz de collisions affecte la fragmentation des ions parents en gardant tous les autres paramètres constants, tels que la pression de gaz et l'énergie appliquée à la cellule de collisions ( $E_{lab}$ ). Ainsi, l'étude a été réalisée sur quatre peptides de masses moléculaires différentes (tableau VII) avec trois gaz de collisions différents (tableau VIII). Alors, pour chacun des peptides mentionnés dans ce tableau, les spectres de masse en tandem ont été

obtenus avec chacun des trois gaz de collisions du même tableau. Lors de ces expériences, les solutions de peptides standards ont été injectés en continu dans le spectromètre de masse. Ces analyses démontreront l'effet du gaz de collisions sur la dissociation induite par collisions.

**Tableau VII: Masse moléculaire des peptides analysés**

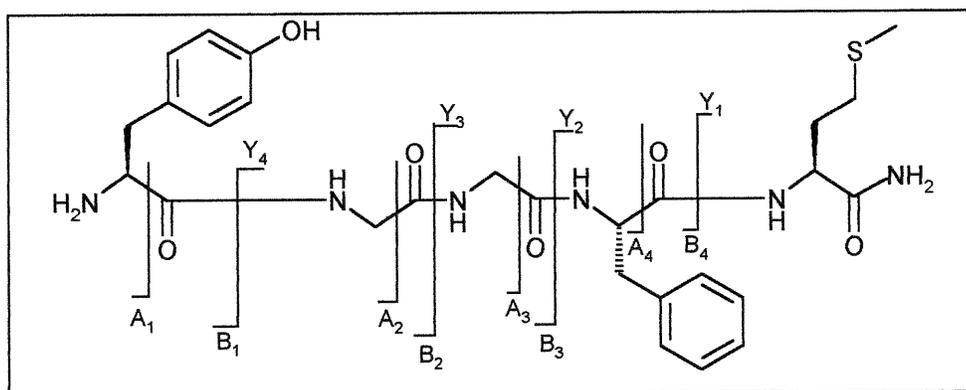
Peptides	Masse moléculaire (g/mol)
Méthionine enképhalinamide	572,24
Pentaphénylalanine	753,35
Bradykinine	1060,22
Substance P	1346,73

**Tableau VIII: Masse moléculaire des gaz de fragmentation utilisés dans la cellule à collisions**

Gaz de fragmentation	Masse moléculaire (g/mol)
Méthane (CH <sub>4</sub> )	16
Argon (Ar)	40
Xénon (Xe)	131,9

#### 7.1.1.1 Résultats de la MS/MS pour la méthionine enképhalinamide.

Puisque les fragments A, B et Y doivent se retrouver en majorité dans les spectres de masse en tandem des peptides protéolytiques, la première étape est la détermination des rapports m/z des fragments attendus pour la méthionine enképhalinamide.<sup>(75)</sup> C'est en déterminant la présence des fragments attendus sur les spectres de masse obtenus qu'il est possible de quantifier la qualité des analyses de spectrométrie de masse en tandem.



**Figure 7.1 Fragmentation de la méthionine enképhalinamide.**

**Tableau IX: Rapport m/z des fragments A, B et Y attendus de la méthionine enképhalinamide (M+H)<sup>+</sup> = 573.**

Fragments	m/z	Fragments	m/z	Fragments	m/z
A <sub>1</sub>	136,08	B <sub>1</sub>	164,07	Y <sub>1</sub>	149,07
A <sub>2</sub>	193,10	B <sub>2</sub>	221,09	Y <sub>2</sub>	296,14
A <sub>3</sub>	250,12	B <sub>3</sub>	278,11	Y <sub>3</sub>	353,16
A <sub>4</sub>	397,19	B <sub>4</sub>	425,18	Y <sub>4</sub>	410,19

Alors, avec ces valeurs de m/z pour les fragments, il est possible d'assigner le bon fragment au bon pic de masse du spectre de masse MS/MS. Ainsi, il sera possible de déterminer quel gaz de collisions convient le plus à cette taille de peptide, soit environ 570 g/mol. Pour comparer les gaz entre eux, il a fallu déterminer une fonction de réponse qui nous permettrait de quantifier le mieux possible la qualité de la fragmentation. En tenant compte du nombre de fragments obtenus et l'intensité relative (IR) du signal obtenu pour ceux-ci, on utilise deux informations importantes que nous procure un spectre de masse en tandem. Ainsi, la fonction de réponse choisie est celle de l'équation 7.1.

$$\text{Fonction de réponse} = \sum IR_i \times \text{nb. fragments } i \text{ obtenus} \quad \text{Équation 7.1}$$

Le tableau X donne les résultats obtenus pour la méthionine enképhalinamide avec le méthane, l'argon et le xénon. Il est évident, selon les données de ce tableau que le gaz le plus favorable pour des peptides de la taille de la méthionine enképhalinamide est le méthane. Bien que l'argon pourrait aussi donner de bons résultats.

**Tableau X: Détermination du bon gaz de collisions pour la méthionine enképhalinamide.**

Gaz	Fonction de réponse obtenue (%)
Méthane	3190
Argon	2430
Xénon	1062

Dans cette recherche, on se base sur les principes premiers pour l'interprétation des spectres de masse MS/MS et non sur la recherche de l'empreinte digitale MS/MS dans une base de données. Il faut donc avoir au minimum une série entière de fragments A, B, C, X, Y ou Z afin de déterminer la structure primaire du peptide protéolytique. Par la suite, plus le nombre de fragments de même série est élevée, plus il devient facile d'obtenir la séquence.

Les spectres de masse de la méthionine enképhalinamide obtenus par spectrométrie de masse en tandem avec le méthane, l'argon et le xénon démontrent aussi que le nombre de fragments ainsi que l'intensité relative de ceux-ci diminuent avec l'augmentation de la masse moléculaire du gaz de fragmentation. La figure 7.2 permet aussi d'identifier le méthane comme meilleur gaz à collisions pour la fragmentation de la méthionine enképhalinamide.

En effet, le spectre de masse en tandem obtenu grâce à la dissociation induite par collisions avec le méthane présente la série entière des fragments A attendus. Cette série entière permet l'identification de la séquence. Pour ce qui est des fragments B et Y, ils sont tous présents exception faite des fragments B<sub>1</sub> et Y<sub>4</sub>. Le grand nombre de fragments de ces deux autres séries permet la reconstruction facile de la séquence complète en acides aminés de la méthionine enképhalinamide, car ils permettent de confirmer rapidement l'hypothèse obtenue par la série de fragments A.

Sur le spectre de masse avec l'argon, il est possible de voir que les fragments de m/z élevé perdent de leur intensité face aux fragments de rapport m/z plus faible. Il en est de même dans le spectre de masse en tandem obtenu par la dissociation induite par collisions avec le xénon. En effet, les fragments de rapport m/z plus grand que 350 sont pratiquement invisibles sur le spectre de masse en tandem avec le xénon.

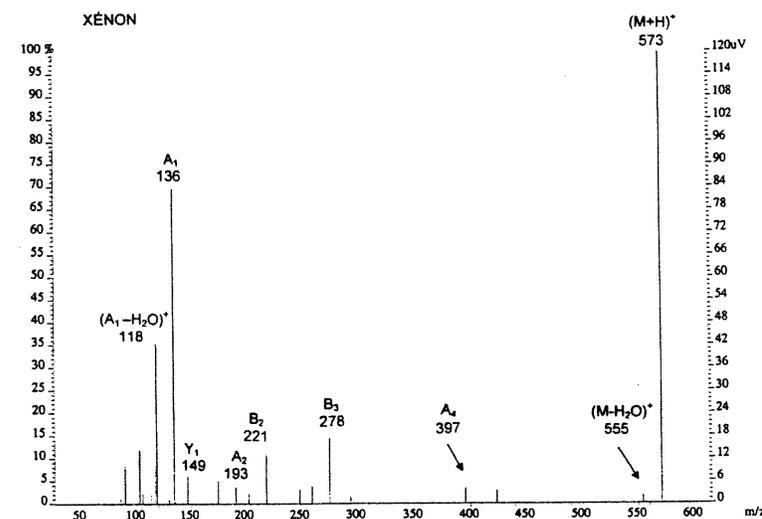
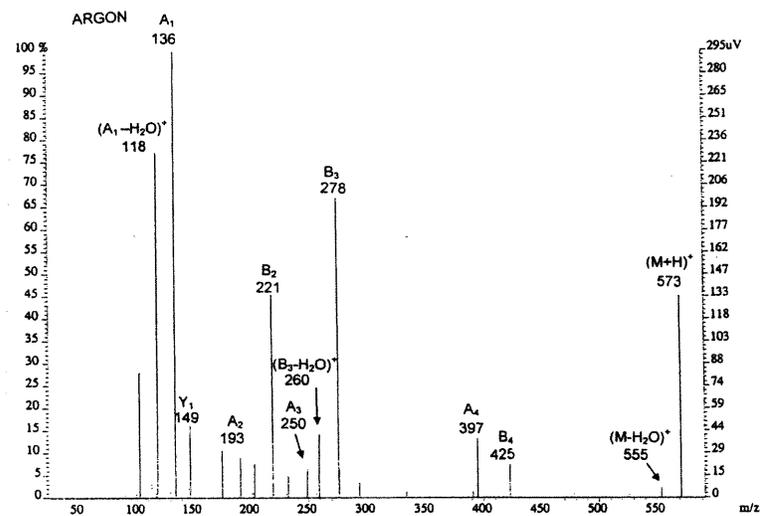
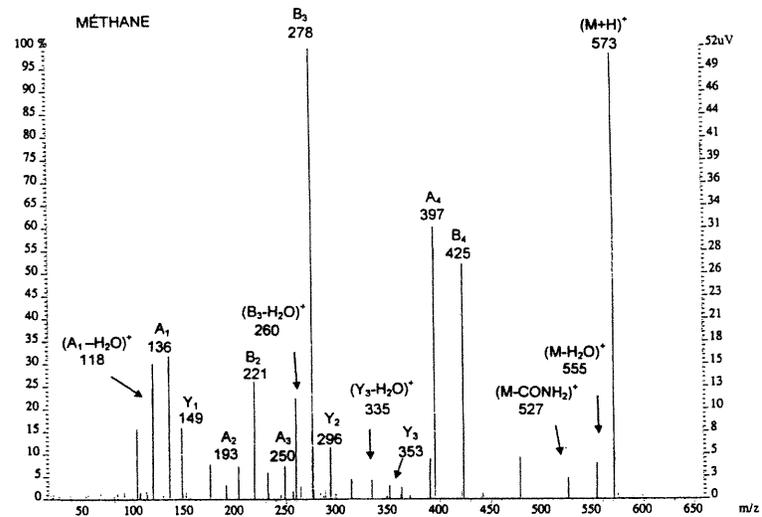
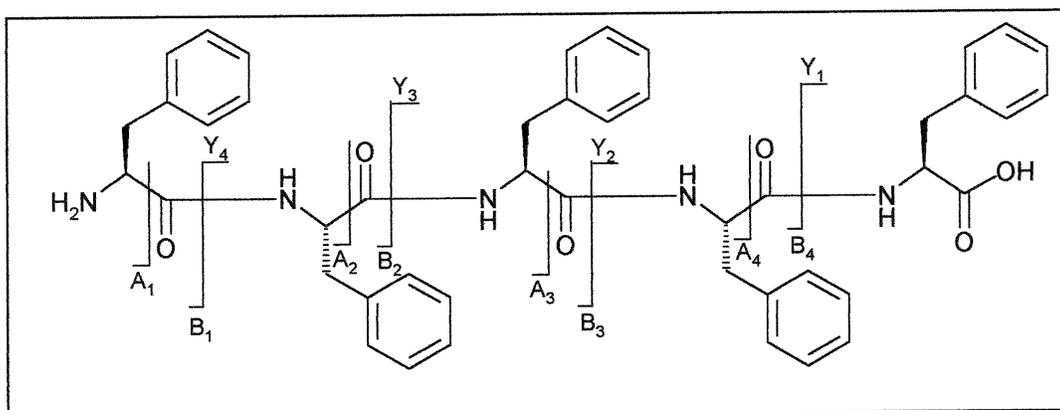


Figure 7.2 Spectres de masse MS/MS de la méthionine enképhalinamide pour les trois gaz à collisions.

Les spectres de masse en tandem obtenus avec l'argon et le xénon démontrent l'effet de la masse moléculaire du gaz de fragmentation sur la dissociation induite par collisions. En effet, plus le gaz de fragmentation est de masse moléculaire élevée, plus sa taille est élevée. Lorsque la taille du gaz est grande, les collisions avec l'analyte sont nombreuses dans la cellule à collisions. La fragmentation sera d'autant plus forte qu'il y aura de collision. Les fragments obtenus seront petits lorsqu'il y aura une fragmentation élevée. Cela explique la perte d'intensité des pics de masse des fragments les plus gros ( $A_4$  et  $B_4$ ).

#### 7.1.1.2 Résultats de la MS/MS pour la pentaphénylalanine.

Comme pour la méthionine enképhalinamide, la détermination du rapport  $m/z$  des fragments A, B et Y attendus pour la pentaphénylalanine, selon la nomenclature de Roepstorff, est la première étape à effectuer. (Tableau XI)



**Figure 7.3 Fragmentation de la pentaphénylalanine**

Ainsi, comme précédemment avec la méthionine enképhalinamide, le choix du gaz de collisions pour la fragmentation optimale d'un peptide d'environ 750 g/mol, sera possible. Le tableau XII donne la fonction de réponse obtenue pour la pentaphénylalanine avec le méthane, l'argon et le xénon. Les résultats de la pentaphénylalanine suivent la même tendance que ceux obtenus pour la méthionine enképhalinamide.

**Tableau XI : Rapport m/z des fragments A, B et Y attendus de la pentaphénylalanine (M+H)<sup>+</sup> = 754.**

Fragments	m/z	Fragments	m/z	Fragments	m/z
A <sub>1</sub>	120,08	B <sub>1</sub>	148,08	Y <sub>1</sub>	166,09
A <sub>2</sub>	267,15	B <sub>2</sub>	295,14	Y <sub>2</sub>	313,16
A <sub>3</sub>	414,22	B <sub>3</sub>	442,21	Y <sub>3</sub>	460,22
A <sub>4</sub>	561,29	B <sub>4</sub>	589,28	Y <sub>4</sub>	607,29

**Tableau XII: Détermination du bon gaz de collisions pour la pentaphénylalanine.**

Gaz	Fonction de réponse obtenue (%)
Méthane	1010
Argon	745
Xénon	168

Tout comme les spectres de masse le démontrent (figure 7.4), le méthane semble le meilleur gaz de fragmentation alors que l'argon pourrait aussi être un choix acceptable.

Dans le cas de la pentaphénylalanine, les fragments A<sub>3</sub>, B<sub>1</sub>, Y<sub>1</sub> et Y<sub>4</sub> sont absents du spectre de masse en tandem obtenu avec le méthane, le gaz de fragmentation considéré comme le meilleur des trois gaz utilisés. En ayant les fragments A<sub>1</sub>, A<sub>2</sub> et A<sub>4</sub>, il est possible d'identifier l'acide aminé en N-terminal et le suivant. Il manque A<sub>3</sub> pour avoir la séquence entière. C'est là que les fragments de la série B entrent en jeu. En effet, il est possible d'obtenir la valeur de A<sub>3</sub> à partir du fragment B<sub>3</sub> puisqu'il n'y a qu'un lien carbonyle (C=O) qui les différencie. Par la suite, les fragments Y<sub>2</sub> et les fragments A<sub>4</sub> et B<sub>4</sub> permettent de terminer et confirmer la séquence. Ainsi, bien que le nombre de fragments attendus absents soit plus grand que pour la méthionine enképhalinamide, il n'en demeure pas moins facile de déterminer l'entière séquence en acides aminés de la pentaphénylalanine.

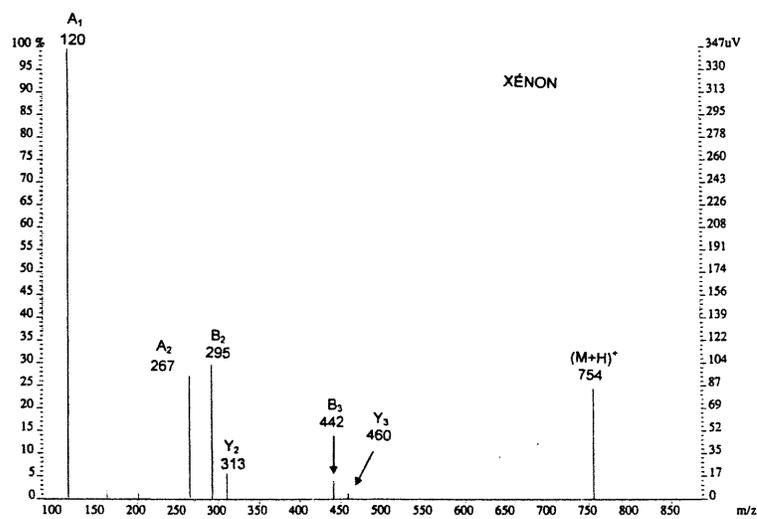
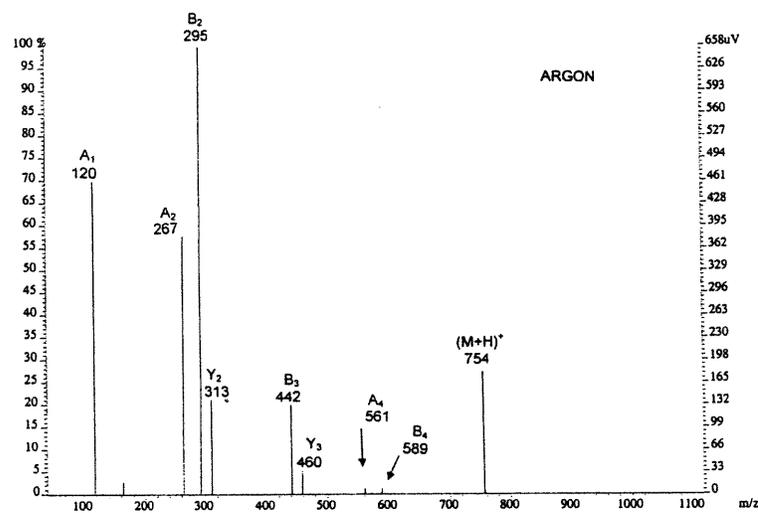
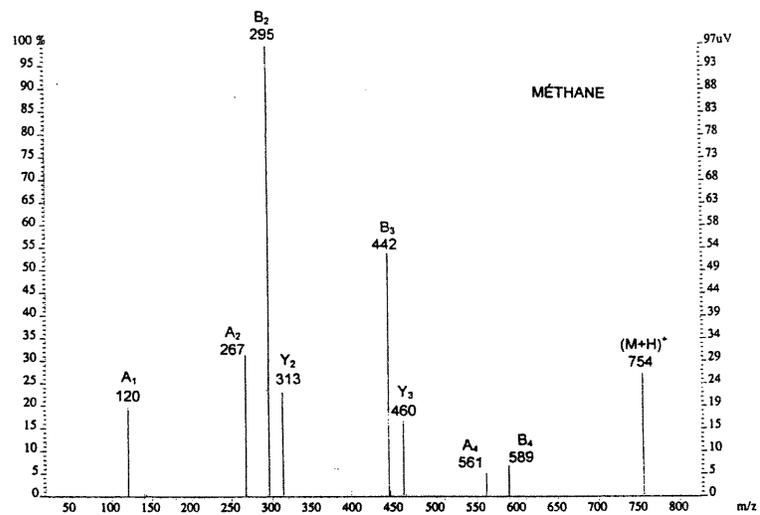


Figure 7.4 Spectres de masse MS/MS de la pentaphénylalanine pour les trois gaz à collisions.

Pour la pentaphénylalanine, comme pour la méthionine enképhalinamide, plus le gaz de collisions est de taille élevée, plus l'intensité des fragments contenant le plus d'acides aminés se trouve affaiblie ( $B_3$ ,  $Y_3$ ,  $A_4$  et  $B_4$ ). Cependant, dans le cas de la pentaphénylalanine on perd même les fragments  $A_4$  et  $B_4$ .

Alors, il est évident que le facteur de réponse choisi pour la quantification de la qualité de la fragmentation est satisfaisant puisqu'il mène aux mêmes choix de gaz de collisions pour les deux premiers peptides que l'analyse visuelle des spectres de masse obtenus. Ainsi, pour les autres peptides, l'étude se contentera des facteurs de réponse obtenus pour chaque gaz dans la détermination du meilleur gaz de fragmentation.

### 7.1.1.3 Résultats de la MS/MS pour la bradykinine.

Comme pour les deux premiers peptides analysés, trouver le rapport  $m/z$  des fragments A, B et Y attendus pour la bradykinine, selon la nomenclature de Roepstorff, est la première partie des résultats obtenus.

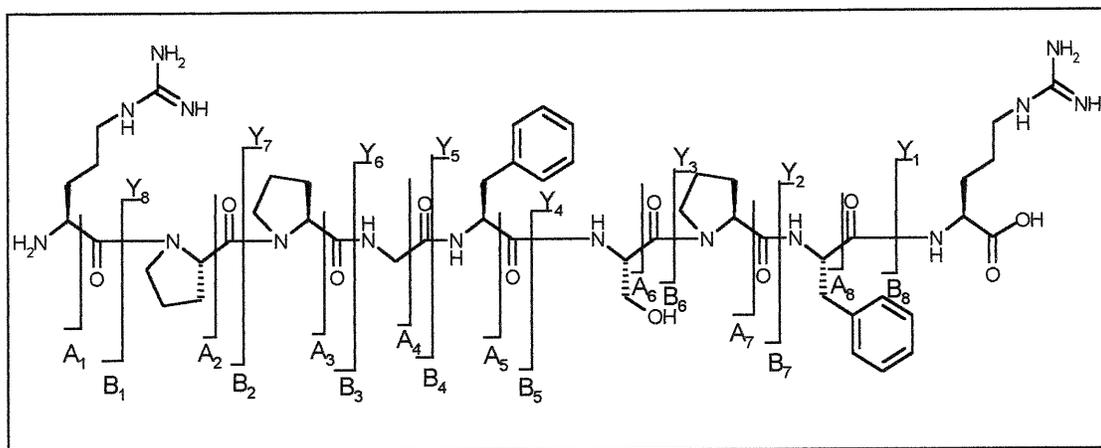


Figure 7.5 Fragmentation de la bradykinine.

**Tableau XIII : Rapport m/z des fragments A, B et Y attendus de la bradykinine (M+H)<sup>+</sup> = 1061.**

Fragments	m/z	Fragments	m/z	Fragments	m/z
A <sub>1</sub>	129,19	B <sub>1</sub>	157,11	Y <sub>1</sub>	175,12
A <sub>2</sub>	226,17	B <sub>2</sub>	254,16	Y <sub>2</sub>	322,19
A <sub>3</sub>	323,22	B <sub>3</sub>	351,21	Y <sub>3</sub>	419,24
A <sub>4</sub>	380,24	B <sub>4</sub>	408,24	Y <sub>4</sub>	506,27
A <sub>5</sub>	527,31	B <sub>5</sub>	555,30	Y <sub>5</sub>	653,34
A <sub>6</sub>	614,34	B <sub>6</sub>	642,34	Y <sub>6</sub>	710,36
A <sub>7</sub>	711,39	B <sub>7</sub>	739,39	Y <sub>7</sub>	807,42
A <sub>8</sub>	858,46	B <sub>8</sub>	886,46	Y <sub>8</sub>	904,47

Le tableau XIV donne les résultats obtenus pour la bradykinine avec le méthane, l'argon et le xénon, et permet de déterminer le meilleur gaz de fragmentation pour un peptide d'environ 1060 g/mol. Tout comme pour les deux peptides analysés dans les sections antérieures, les résultats de fragmentation pour la bradykinine mènent encore à l'identification du méthane comme gaz de choix alors que l'argon pourrait encore être utilisé comme remplaçant sans trop de changement dans l'apparence du spectre de masse. Cependant, il apparaît que le changement de gaz de fragmentation affecte beaucoup moins les peptides de m/z > 1000 (Tableaux XIV et XVI) que ceux de m/z < 1000 (Tableaux X et XII).

**Tableau XIV : Détermination du bon gaz de collisions pour la bradykinine.**

Gaz	Fonction de réponse obtenue (%)
Méthane	6090
Argon	5684,5
Xénon	467,5

Le spectre de masse de la bradykinine obtenu par spectrométrie de masse en tandem avec le méthane comme gaz de collisions est celui de la figure

7.6. Il est à noter que plus le peptide à analyser est grand, plus le nombre de fragments étant élevé, plus les spectres de masse deviennent complexes à analyser au point de vue du nombre de pics de masse présents. Ainsi, il est possible de voir que la bradykinine, tout comme la substance-P (section 7.1.1.4), donnent des spectres de masse en tandem beaucoup plus peuplés en pics que la méthionine enképhalinamide et la pentaphénylalanine. Cependant, le nombre de fragments attendus obtenus est limité et permet de déterminer qu'une partie de la séquence en acides aminés des peptides et non la séquence en totalité.

Les fragments attendus ne sont pas nombreux dans le spectre de masse en tandem de la bradykinine. Il faudra donc travailler beaucoup plus fort pour arriver à une séquence en acides aminés satisfaisante. D'abord, il n'y a pas de fragments  $A_1$  ce qui déjà nous complique beaucoup la tâche. En effet, sans fragments  $A_1$ , il devient presque impossible d'identifier l'acide aminé en N-terminale (R). Cependant, en ayant le fragment  $Y_8$  et la masse moléculaire du peptide protéolytique analysé il est possible d'y arriver. Par la suite, les fragments  $A_2$  et  $B_2$  nous permettent d'obtenir le deuxième acide aminé à partir du N-terminal (P). Avec le fragment  $A_8$  et la masse moléculaire du peptide tryptique d'intérêt, il est possible d'identifier l'acide aminé en C-terminale (R). Par la suite, les fragments  $A_5$ ,  $A_6$  et  $A_7$  donnent l'identité des deux acides aminés suivants (P-F) à partir du C-terminal, alors que  $Y_3$  et  $Y_4$  les confirment. La séquence est alors connue en partie (R-P-?-?-?-S-P-F-R) mais il est impossible de l'obtenir entièrement car il nous manque toujours trois acides aminés. Il faudrait envisager l'utilisation d'un gaz de collision, autre que les trois testés, afin d'obtenir une meilleure fragmentation ou une fragmentation qui apporterait les informations complémentaires à celles obtenues par le méthane.

L'étude MS/MS de la substance P permettra de voir si le phénomène est général pour les peptides de  $m/z > 1000$  ou s'il s'agit d'un cas isolé.

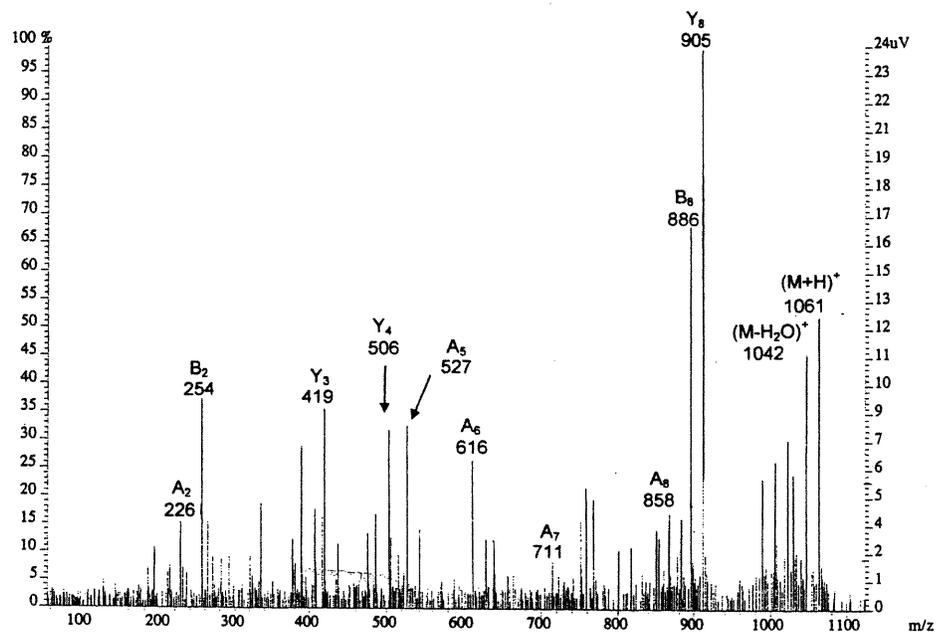


Figure 7.6 Spectre de masse MS/MS de la bradykinine avec le méthane.

#### 7.1.1.4 Résultats de la MS/MS pour la substance P.

La détermination du rapport  $m/z$  des fragments A, B et Y attendus pour la substance P, selon la nomenclature de Roepstorff, est réalisée en premier lieu lors de l'analyse par spectrométrie de masse en tandem de la substance P.

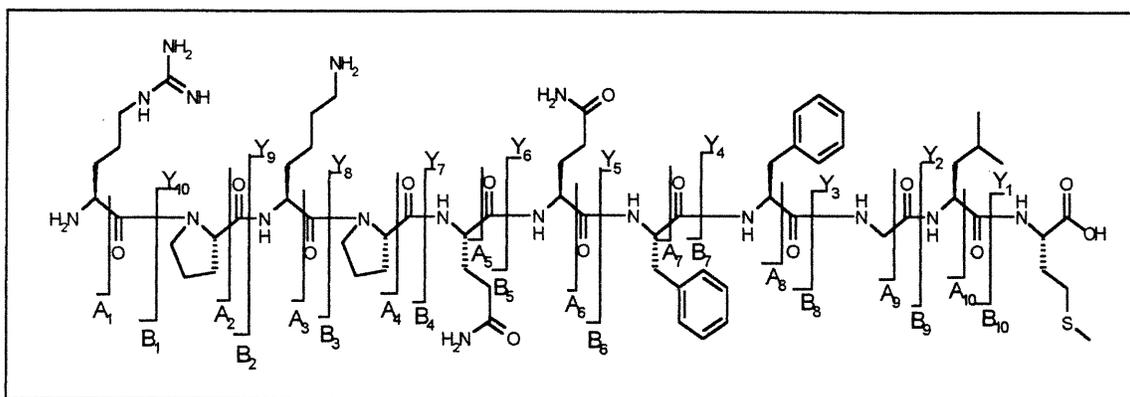


Figure 7.7 Fragmentation de la substance P.

Le tableau XVI donne les résultats obtenus pour la substance P avec le méthane, l'argon et le xénon, et permet de déterminer le meilleur gaz de fragmentation pour un peptide d'environ 1350 g/mol. Les résultats de fragmentation pour la substance P mènent à la conclusion que le xénon est le gaz de choix lorsque les peptides à fragmenter commencent à dépasser les masses de 1300 g/mol.

**Tableau XV : Rapport m/z des fragments A, B et Y attendus de la substance P  $(M+H)^+ = 1347$ .**

Fragments	m/z	Fragments	m/z	Fragments	m/z
A <sub>1</sub>	129,11	B <sub>1</sub>	157,11	Y <sub>1</sub>	149,07
A <sub>2</sub>	226,17	B <sub>2</sub>	254,16	Y <sub>2</sub>	262,16
A <sub>3</sub>	354,26	B <sub>3</sub>	382,26	Y <sub>3</sub>	319,18
A <sub>4</sub>	451,31	B <sub>4</sub>	479,31	Y <sub>4</sub>	466,25
A <sub>5</sub>	579,37	B <sub>5</sub>	607,37	Y <sub>5</sub>	613,32
A <sub>6</sub>	707,43	B <sub>6</sub>	735,43	Y <sub>6</sub>	741,38
A <sub>7</sub>	854,50	B <sub>7</sub>	882,49	Y <sub>7</sub>	869,43
A <sub>8</sub>	1001,57	B <sub>8</sub>	1029,56	Y <sub>8</sub>	966,49
A <sub>9</sub>	1058,59	B <sub>9</sub>	1086,66	Y <sub>9</sub>	1094,58
A <sub>10</sub>	1171,67	B <sub>10</sub>	1199,67	Y <sub>10</sub>	1191,63

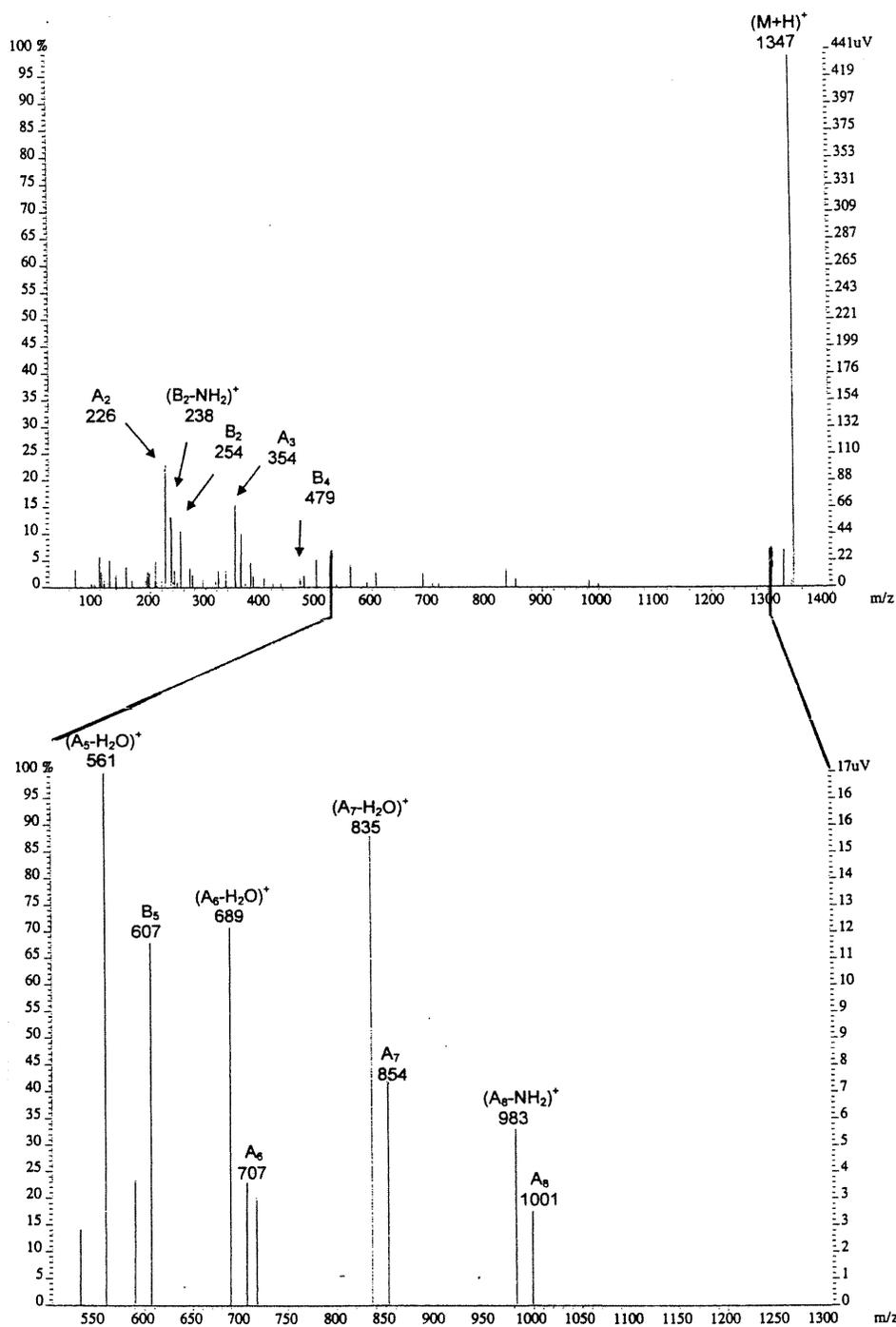
**Tableau XVI : Détermination du bon gaz de collisions pour la substance P.**

Gaz	Fonction de réponse obtenue (%)
Méthane	198
Argon	473,6
Xénon	550,7

Le spectre de masse de la substance P obtenu par spectrométrie de masse en tandem avec le xénon comme gaz de collisions est celui de la figure 7.8. Les fragments des séries A et B sont prioritaires sur ce spectre de masse en tandem. Ce sont ces fragments qui mèneront à la détermination de la

séquence inachevée de la substance P. En effet, même avec la masse moléculaire du peptide analysé, il est impossible de déterminer les acides aminés N-terminal et C-terminal car les fragments  $A_1$  et  $Y_{10}$  sont absents. Le groupe de fragments  $B_5$ ,  $B_4$ ,  $A_3$  et  $B_2$  permettent d'identifier les acides aminés trois, quatre et cinq à partir du N-terminal (K-P-Q). Alors, que les fragments  $A_6$ ,  $A_7$  et  $A_8$  (F-F) mènent aux acides aminés quatre et cinq à partir du C-terminal. La séquence ainsi obtenue ( ?-?-K-P-Q-?-F-F-?-?-?) comporte de grands manques.

Les résultats pour la substance P confirment l'hypothèse selon laquelle la séquence des peptides de  $m/z > 1000$  ne peut être obtenue avec l'un ou l'autre des trois différents gaz de collisions. Puisqu'il est de plus en plus difficile d'obtenir une séquence satisfaisante pour des peptides de  $m/z > 1000$ , il est évident que la digestion enzymatique facilitera les analyses par spectrométrie de masse en tandem. En effet, la protéolyse permettant de cliver le peptide à analyser en plus petits peptides tryptiques, la détermination de la séquence en acides aminés de ceux-ci sera plus facile vu la simplicité des spectres de masse en tandem obtenus. Cependant, pour des fragments tryptiques de  $m/z > 1000$ , il faudrait, comme dit précédemment, trouver une façon d'obtenir une meilleure fragmentation ou les informations manquantes à la bonne identification de la séquence.



**Figure 7.8** Spectre de masse MS/MS de la substance P avec le xénon.

À l'aide des données expérimentales obtenues dans les tableaux X, XII, XIV et XVI, il a été possible de tracer un graphique (figure 7.9). En se fiant à ce graphique, il est possible de choisir le bon gaz de collisions (parmi les trois

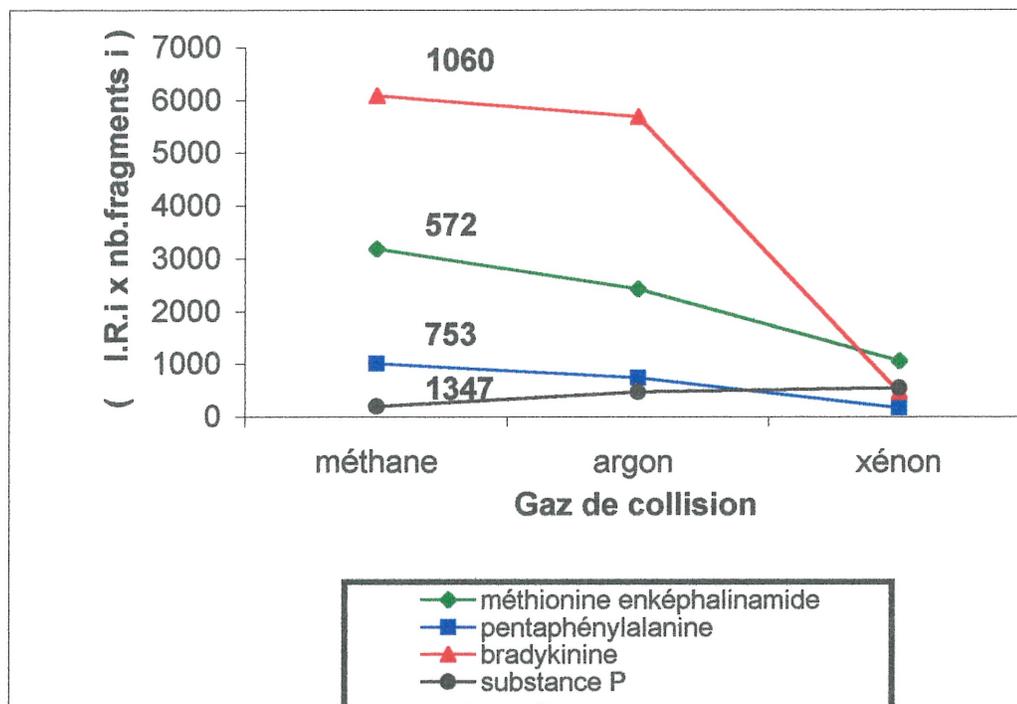


Figure 7.9 Détermination du gaz de collisions.

La tendance est la même pour les peptides de masse moléculaire variant de 500 à 1100g/mol. La fragmentation devient de moins bonne qualité lorsque la taille du gaz de collisions utilisée est plus élevée. Ainsi, les spectres de masse en tandem de la méthionine enképhalinamide, la pentaphénylalanine et la bradykinine présenteront plus de fragments significatifs, pour la détermination de la séquence en acides aminés, avec le méthane qu'avec l'argon et le xénon. Cependant, la substance P demande l'utilisation du xénon pour obtenir un spectre de masse en tandem contenant le plus grand nombre de fragments de série A,B et Y. Mais comme discuté pour la bradykinine et la substance P, les fragments obtenus pour les peptides de  $m/z > 1000$  sont insuffisants à l'obtention de la séquence entière de ceux-ci.

## 7.2 ANALYSE DES PEPTIDES TRYPTIQUES

Les analyses par spectrométrie de masse en tandem des peptides protéolytiques sont réalisées de la manière suivante. D'abord, la solution résultant de la digestion enzymatique du peptide à analyser est injectée en LC/MS selon les conditions optimales du chapitre 6, afin de connaître la masse moléculaire des peptides tryptiques dans la solution. Par la suite, l'analyse MS/MS est effectuée, par diffusion (sans colonne), à la masse fixée de l'ion parent. Alors, selon la grosseur des peptides protéolytiques obtenus lors de la digestion enzymatique, le gaz de collisions choisi sera différent. Dans les deux cas pertinents de cette recherche, c'est-à-dire l'angiotensine II et la ranatachkinine, les peptides tryptiques obtenus sont de masse inférieure à 1000 g/mol. Cela permet donc d'utiliser uniquement le méthane comme gaz de fragmentation lors des analyses de spectrométrie de masse en tandem de ces peptides. De plus cela élimine le problème rencontré lors de l'analyse MS/MS de peptides ayant un  $m/z > 1000$ .

### 7.2.1 SPECTROMÉTRIE DE MASSE EN TANDEM DE L'ANGIOTENSINE II

Tout comme lors de l'étude des différents gaz de fragmentation avec les différentes tailles de peptides à analyser, il faut débiter l'analyse des peptides protéolytiques par la détermination des rapports  $m/z$  des fragments attendus par la spectrométrie de masse en tandem selon Roepstorff. Les tableaux XVII, XVIII et la figure 7.10, décrivent bien les fragments A, B et Y provenant de la fragmentation des peptides tryptiques #1 et #2.

**Tableau XVII : Rapport  $m/z$  des fragments A, B et Y attendus du peptide tryptique #1 de l'angiotensine II  $(M+H)^+ = 290$ .**

Fragments	$m/z$
A <sub>1</sub>	88,04
B <sub>1</sub>	116,04
Y <sub>1</sub>	175,12

Tableau XVIII : Rapport m/z des fragments A, B et Y attendus du peptide tryptique #2 de l'angiotensine II  $(M+H)^+ = 776$ .

Fragments	m/z	Fragments	m/z	Fragments	m/z
A <sub>1</sub>	72,08	B <sub>1</sub>	100,08	Y <sub>1</sub>	166,09
A <sub>2</sub>	235,14	B <sub>2</sub>	263,14	Y <sub>2</sub>	263,14
A <sub>3</sub>	348,23	B <sub>3</sub>	376,22	Y <sub>3</sub>	400,20
A <sub>4</sub>	485,29	B <sub>4</sub>	513,28	Y <sub>4</sub>	513,28
A <sub>5</sub>	582,34	B <sub>5</sub>	610,34	Y <sub>5</sub>	676,35

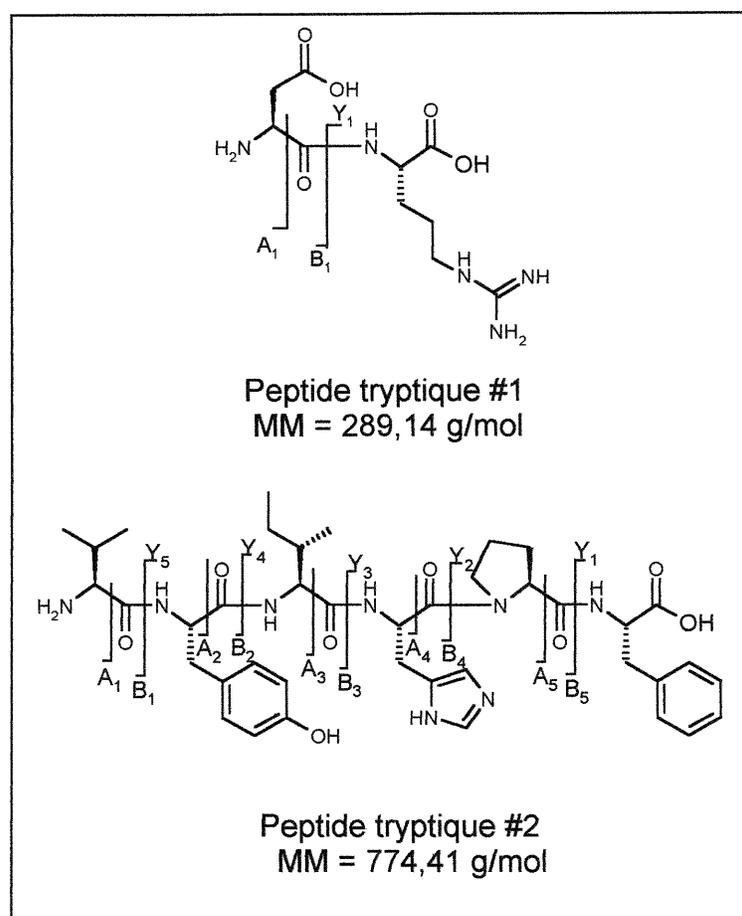


Figure 7.10 Fragmentation des peptides protéolytiques de l'angiotensine II

Les spectres de masse en tandem des peptides protéolytiques de l'angiotensine II se retrouvent à la figure 7.11. Le spectre de masse en

tandem du premier peptide protéolytique permet bien entendu d'en déterminer la séquence en acides aminés car tous les fragments attendus sont présents.

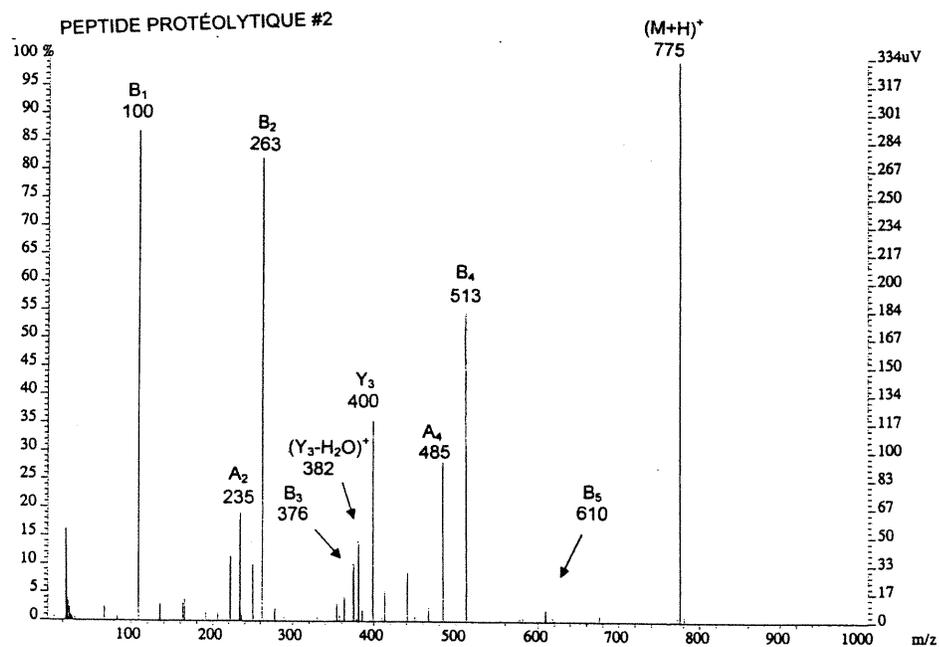
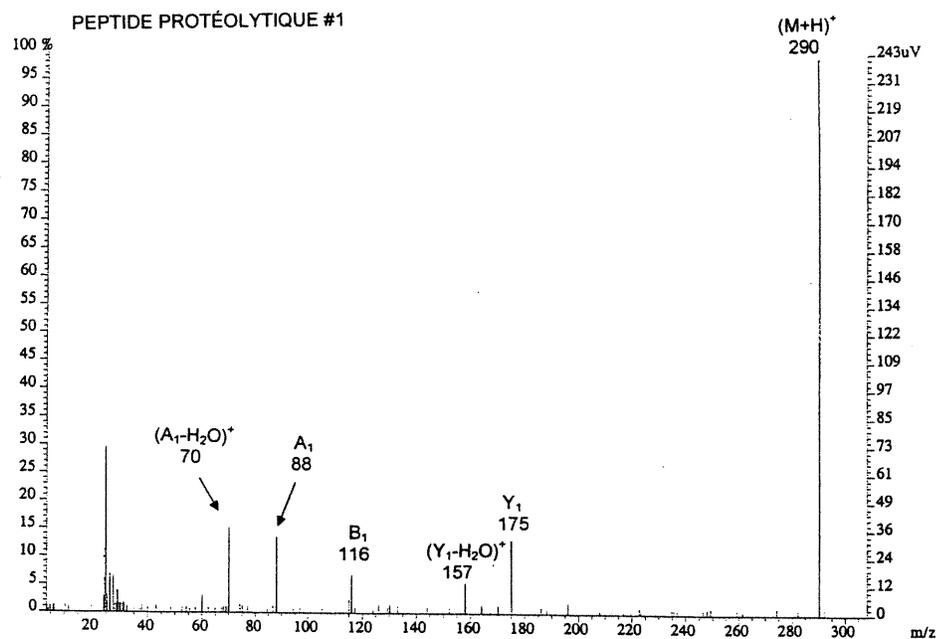
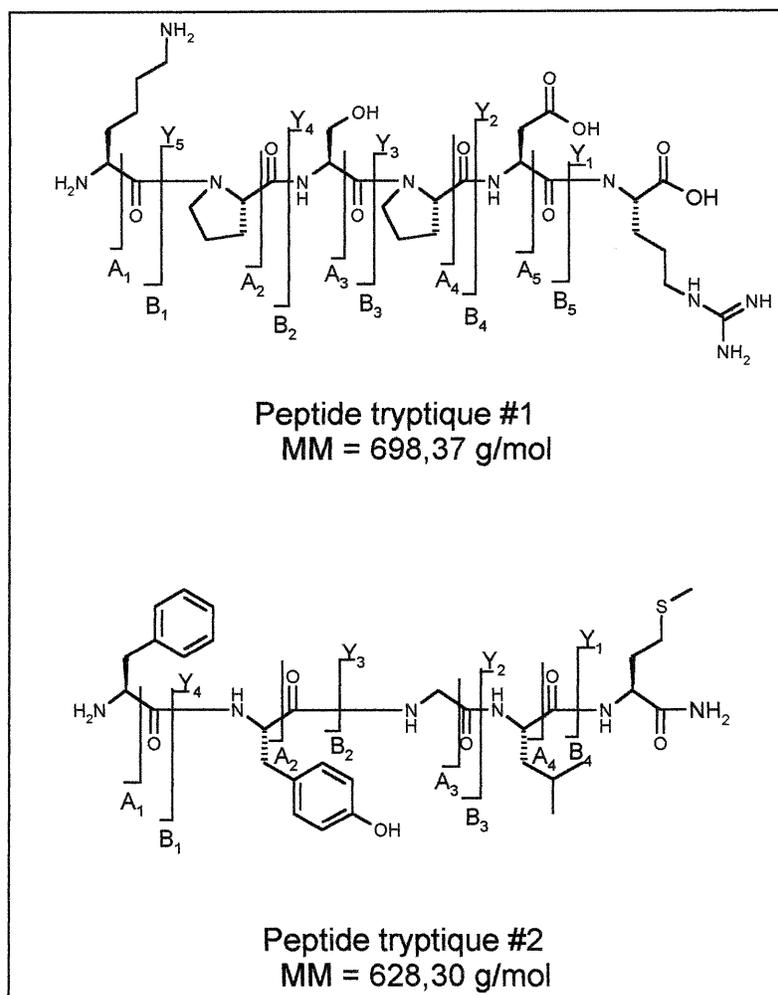


Figure 7.11 Spectres de masse en tandem des peptides tryptiques de l'angiotensine II.

La séquence en acides aminés du deuxième peptide protéolytique s'avère trouvable puisque la série de fragments B est entièrement présente. De plus les trois autres fragments de type A et Y permettront de confirmer quelques acides aminés trouvés.

### 7.2.2 SPECTROMÉTRIE DE MASSE EN TANDEM DE LA RANATACHYKININE

Les tableaux XIX, XX et la figure 7.12, décrivent bien les fragments A, B et Y attendus de la fragmentation des peptides tryptiques de la ranatachykinine, tout comme fait pour l'angiotensine II, dans la section précédente.



**Figure 7.12** Fragmentation des peptides protéolytiques de la ranatachykinine.

**Tableau XIX : Rapport m/z des fragments A, B et Y attendus du peptide tryptique #1 de la ranatachykinine (M+H)<sup>+</sup> = 699.**

Fragments	m/z	Fragments	m/z	Fragments	m/z
A <sub>1</sub>	101,11	B <sub>1</sub>	129,10	Y <sub>1</sub>	175,12
A <sub>2</sub>	198,16	B <sub>2</sub>	226,16	Y <sub>2</sub>	290,15
A <sub>3</sub>	285,19	B <sub>3</sub>	313,19	Y <sub>3</sub>	387,20
A <sub>4</sub>	382,24	B <sub>4</sub>	410,24	Y <sub>4</sub>	474,23
A <sub>5</sub>	497,27	B <sub>5</sub>	525,27	Y <sub>5</sub>	571,28

**Tableau XX : Rapport m/z des fragments A, B et Y attendus du peptide tryptique #2 de la ranatachykinine (M+H)<sup>+</sup> = 629.**

Fragments	m/z	Fragments	m/z	Fragments	m/z
A <sub>1</sub>	120,08	B <sub>1</sub>	148,08	Y <sub>1</sub>	149,07
A <sub>2</sub>	283,14	B <sub>2</sub>	311,14	Y <sub>2</sub>	262,16
A <sub>3</sub>	340,17	B <sub>3</sub>	368,16	Y <sub>3</sub>	319,18
A <sub>4</sub>	453,25	B <sub>4</sub>	481,25	Y <sub>4</sub>	482,24

Les spectres de masse en tandem des peptides protéolytiques de la ranatachykinine se retrouvent à la figure 7.13. La détermination de la séquence en acides aminés du premier peptide tryptique se fait en partie grâce à la présence dans le spectre de masse en tandem des quelques fragments de la série Y. En effet, il faut débiter par le N-terminal (K) que l'on obtient avec le fragment A<sub>1</sub> et qui sera confirmé par le fragment Y<sub>5</sub>. Par le groupe de fragments Y<sub>3</sub>, Y<sub>4</sub> et Y<sub>5</sub>, il est possible d'identifier les acides aminés deux et trois à partir du N-terminal (P-S). L'absence des fragments Y<sub>1</sub> et Y<sub>2</sub> rendent impossible l'obtention de la séquence complète pour ce peptide.

Les fragments de la série Y, tous présents, du deuxième peptide tryptique mènent directement à la connaissance de la séquence complète en acides aminés de ce dernier peptide protéolytique. Les autres fragments permettent la confirmation partielle de la séquence obtenue.

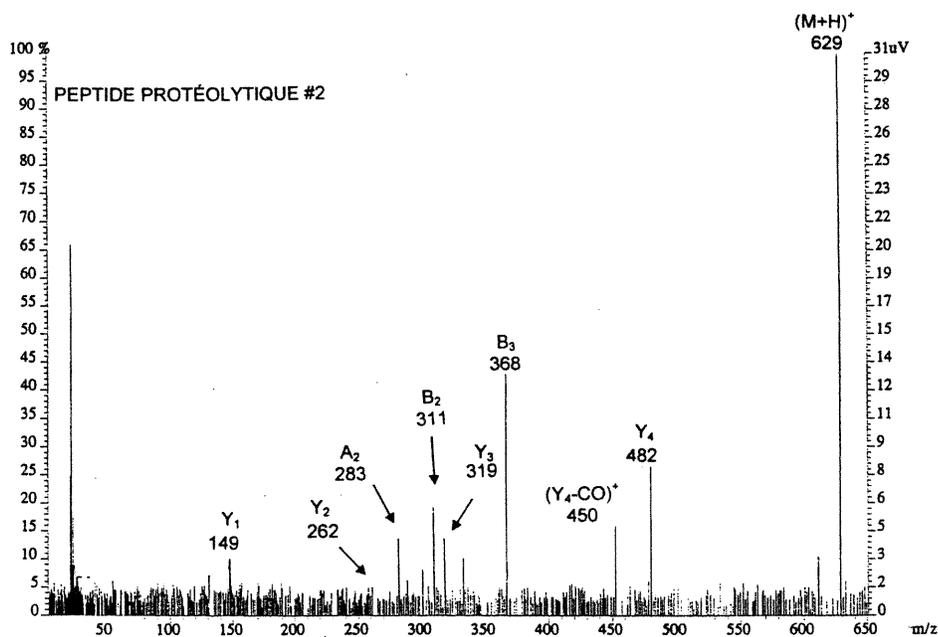
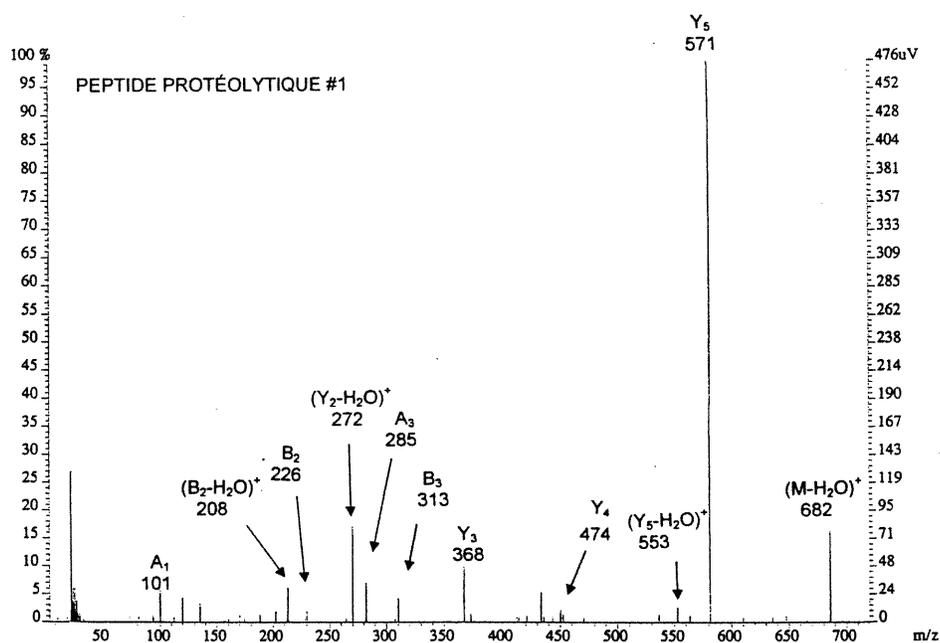


Figure 7.13 Spectres de masse en tandem des peptides tryptiques de la ranatychinine.

En conclusion, le choix du gaz de fragmentation est très important pour obtenir un spectre de masse en tandem de qualité. Plus il y a un grand nombre de fragments attendus présents dans le spectre de masse en tandem, plus il est facile de reconstruire la séquence en acides aminés d'un peptide. Finalement, les analyses de spectrométrie de masse en tandem des peptides protéolytiques sont satisfaisantes pour obtenir la séquence en acides aminés de ceux de  $m/z < 1000$ . Ainsi, il a été possible de déterminer la structure en acides aminés de trois des quatre peptides protéolytiques résultants d'une digestion enzymatique, ce qui est plutôt bien. Cependant, l'ordre dans lequel se trouvaient ces peptides tryptiques dans le peptide de départ reste toujours inconnue. Elle pourra en complément être déterminée par l'étude cinétique de la protéolyse.

**CHAPITRE 8**

**CONCLUSION**

## 8.1 SOMMAIRE DES RÉSULTATS

Les nombreuses publications sur le séquençage des peptides et protéines par la LC-ESI-MS et la LC-ESI-MS/MS démontrent bien que ce sujet est en plein essor. La présente recherche a été réalisée afin d'établir une méthode de séquençage des peptides et protéines sur l'appareil utilisé. Cela en tenant compte que l'on n'utilise pas la recherche en base de données car on veut pouvoir analyser des protéines inconnues.

Le point de départ de cette recherche fût les conditions d'opération utilisées dans la littérature, par la suite ces conditions ont été ajustées afin de convenir au type d'appareil utilisé et au type d'analyse effectuée.

Le choix de l'utilisation de l'acide formique comme agent de pairage d'ions à éliminer les problèmes d'ionisation dus à l'utilisation de l'acide trifluoroacétique, dénotés dans la littérature<sup>(57)</sup>. Les premières expériences effectuées furent réalisées afin de trouver les nouvelles conditions chromatographiques pour un mélange de peptides protéolytiques, grâce à la méthode d'optimisation par simplex. L'optimisation des conditions chromatographiques, pour une solution résultant de la digestion de la  $\beta$ -caséine par la trypsine en solution, démontre bien que la méthode des simplex mène rapidement et de façon satisfaisante à la meilleure séparation des peptides protéolytiques. Puisque la solution utilisée lors de l'optimisation des conditions chromatographiques comportait 31 pics chromatographiques, il est évident que les conditions permettant de tous les séparés, permettront la séparation d'un plus petit nombre de fragments tryptiques. Ce sont donc ces conditions d'élution qui servirent par la suite pour les analyses LC-MS de cette recherche.

La deuxième partie de ce projet de recherche était en les analyses LC-MS. Les spectres de masse ainsi obtenus permirent d'identifier la masse moléculaire des fragments peptidiques obtenus par la digestion enzymatique.

En effet, il fût possible de déterminer la masse moléculaire des peptides protéolytiques de l'angiotensine II et de la ranatachykinine, grâce aux analyses LC-MS réalisées. Ces masses moléculaires sont les premières données importantes menant à l'obtention de la séquence d'acides aminés complète des peptides. C'est en fixant, comme masse de l'ion parent, celle de l'ion moléculaire du spectre de masse LC-MS du peptide protéolytique, qu'il sera possible d'obtenir le spectre de masse en tandem des peptides tryptiques.

Les expériences de spectrométrie de masse en tandem sur différents peptides ; la méthionine enképhalinamide, la pentaphénylalanine, la bradykinine et la substance P, ont rendu possible la détermination des gaz de fragmentation adéquats aux analyses MS/MS des peptides tryptiques. Ainsi, il fût conclu que pour des peptides protéolytiques de masse moléculaire < 1000, le gaz de fragmentation à favoriser est le méthane. Ainsi, les fragments obtenus ont permis l'identification de la séquence en acides aminés de la méthionine enképhalinamide et de la pentaphénylalanine. Cette étude est très importante pour les analyses de ce projet de recherche de l'angiotensine II et de la ranatachykinine. Tandis que pour des peptides tryptiques de masse moléculaire plus grande que 1000 (bradykinine et substance P), le gaz de collision donnant les plus d'informations dans le spectre de masse, ne donnait pas suffisamment d'indices pour l'obtention de la séquence en acides aminés entière des peptides analysés. Il faudrait donc utiliser un autre gaz de collisions pour obtenir un spectre de masse en tandem comportant des informations satisfaisantes ou bien des informations complémentaires à celles obtenues.

Les spectres de masse MS/MS des peptides tryptiques ont fourni les informations nécessaires à l'identification de la séquence en acides aminés les composant, dans le cas de trois des quatre fragments protéolytiques. La séquence des deux peptides protéolytiques de l'angiotensine II a été identifiée. Uniquement pour un des deux peptides tryptiques de la

ranatachykinine, la séquence entière fût obtenue. Cependant, il manquait toujours deux fragments sur le spectre de masse pour identifier les trois derniers acides aminés composant l'autre peptide tryptique.

Les étapes de la démarches expérimentales soient le clivage enzymatique, l'analyse LC-MS ainsi que l'analyse LC-MS/MS mènent donc à l'obtention de la séquence des peptides protéolytiques de la protéine native. Bien que le séquençage des peptides a mené dans 75% des cas à la séquence en acides aminés des peptides protéolytiques, il reste encore une étape afin d'obtenir l'entière séquence des peptides et protéines analytes.

En effet, les spectres de masse MS/MS permettent de connaître la séquence des peptides tryptiques mais pas celle de la protéine native. L'ordre dans lequel se trouvaient les fragments peptidiques dans la protéine de départ avant la digestion enzymatique est toujours inconnue. Bien entendu, l'utilisation d'une base de données pourrait permettre l'identification de la protéine native si celle-ci est connue et présente dans la base de données. Il faut toujours garder en tête que ce projet veut obtenir la séquence en acides aminé des protéines inconnues.

Il est possible avec le chevauchement des pics chromatographiques obtenus pour la digestion de deux différentes enzymes comme la trypsine et la chymotrypsine, d'obtenir l'ordre dans lequel se trouvait les peptides tryptiques avant le clivage enzymatique.

Puisque ce projet se fait en complémentarité avec un autre projet en cours dans notre groupe, l'étude cinétique de la réaction de potéolyse sera privilégiée pour l'identification de l'ordre des peptides tryptiques dans la protéin native.

## 8.2 PROJET DE RECHERCHE COMPLÉMENTAIRE

L'étude cinétique de la digestion enzymatique, dans le microréacteur contenant un gel d'enzyme, fournit un arbre de masse démontrant l'ordre dans lequel les fragments protéolytiques sont formés (figure 8.1). Ainsi, l'ordre de départ des fragments protéolytiques peut être connu. Cette étude permettra la confirmation de l'identification à partir de la base de données.

Comme démontré au chapitre 1 de ce document, l'étude cinétique de la réaction de clivage enzymatique apporte deux importants renseignements pour cette recherche. Premièrement, elle permet de déterminer après combien de temps la protéine (P) a subi une digestion enzymatique complète dans le microréacteur créé en laboratoire. Deuxièmement, le plus intéressant de cette étude est qu'elle permet de déterminer dans quel ordre se font les fragments peptidiques et donc donne une partie de la séquence en positionnant chaque peptide protéolytique ( $pp_i$ ) dans la protéine.

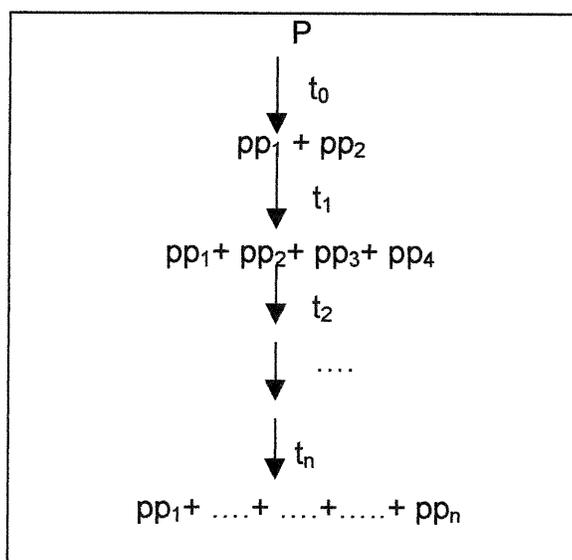


Figure 8.1 Allure de l'arbre de masse d'une protéine.

Ainsi, pour l'angiotensine II et la ranatychinine utilisées dans ce projet de recherche, il sera possible de déterminer l'ordre dans lequel se trouvaient les peptides tryptiques avant leur protéolyse par la trypsine.

Cette méthode a donc un potentiel intéressant pour le séquençage des protéines. Le microréacteur est toujours en étude dans le cadre d'un autre projet de recherche dans le groupe.

### **8.3 RECHERCHES FUTURES**

Pour les recherches futures il y a quelques détails qu'il faudrait considérés. Premièrement, il serait peut-être préférable d'utiliser une  $\mu$ -pompe et une  $\mu$ -colonne chromatographique afin de permettre l'introduction directe de l'éluant sortant du chromatographe dans le spectromètre de masse. Cela pourrait augmenter de façon considérable la quantité d'analyte entrant dans le spectromètre de masse et ainsi réduire la quantité de peptide nécessaire pour le séquençage. Dans une  $\mu$ -colonne, la proportion de l'analyte par rapport à l'éluant est plus grande que dans une colonne de diamètre interne de 4,60 mm. De plus, une  $\mu$ -pompe permet d'introduire directement l'éluant sortant de la colonne sans diviseur d'effluent, qui réduit considérablement la fraction entrant dans le MS. Par ces deux avantages, la quantité d'analyte entrant dans le MS se voit grandement accrue.

Deuxièmement, il serait aussi intéressant d'optimiser le temps de séjour du peptide dans le microréacteur de gel enzymatique pour obtenir une digestion complète de l'analyte. Ainsi, il serait sûrement possible de réduire le temps de préparation des peptides.

Troisièmement, il faudrait trouver un moyen d'obtenir une séquence entière et non partielle des peptides de  $m/z > 1000$ . Soit en utilisant un autre gaz de collisions que les trois gaz utilisés dans cette recherche pour avoir plus d'informations dans un spectre de masse ou obtenir des informations complémentaires à celles obtenues dans cet ouvrage.

Finalement, le plus intéressant pour le séquençage des peptides serait de réussir à coupler le microréacteur au système LC-MS/MS afin de permettre une digestion enzymatique en ligne. Ainsi, il y aurait beaucoup moins de chance de perte d'analyte et la préparation des peptides serait de beaucoup simplifiée.

Pour ce qui est de la méthode de séquençage proposée dans ce mémoire, elle convient à l'appareil utilisé et fournit les informations nécessaires à l'identification de la protéine native. Cette méthode répond donc aux objectifs fixés au départ pour ce projet de recherche, soit le séquençage des peptides par la spectrométrie de masse en tandem à ionisation par électronébulisation.

## **BIBLIOGRAPHIE**

- 1 G. Allan, « *Laboratory techniques in biochemistry and molecular biology vol.9 : sequencing of proteins and peptides* », North-Holland publishing company, NY, 1981, pp. 327
- 2 R.B. Cole (Ed), « *Electrospray ionization mass spectrometry : fundamentals, instrumentations and applications* », Wiley & sons, NY, 1997, pp.577
- 3 C.T. Mant & R.S. Hodges (Eds), « *High-performance liquid chromatography of peptides and proteins : separation, analysis and conformation* », CRC Press, Boston, 1991, pp.938
- 4 D. Voet & J.G. Voet, « *Biochemistry* », John Wiley & sons, NY,1990, pp.1223
- 5 J.A. Hunt & V.M. Ingram, *Biochimica et Biophysica Acta*, 1958, **28**, p.546
- 6 V.M. Ingram, *Biochimica et Biophysica Acta*, 1958, **28**, p.539
- 7 V.M. Ingram, *Biochimica et Biophysica Acta*, 1959, **36**, p.539
- 8 Hart & Conia, « *Introduction à la chimie organique* »,4<sup>e</sup>Ed, InterEditions, Paris, 1987, pp.567
- 9 P. Edman, *Acta Chemica Scandinavica*, 1950, **4**, p.283
- 10 J. Ozols, *Methods in Enzymology*, 1990, **182**, p.587
- 11 R. D. Voyksner, D.C. Chen & H.E. Swaisgood, *Analytical Biochemistry*, 1990, **188**, p.72
- 12 C.J.McNeal (Ed), « *The analysis of peptides and proteins by mass spectrometry* », Wiley & sons, NY, 1988, pp.322
- 13 L.R. Croft, « *Introduction to protein sequence analysis* », John Wiley & sons, Toronto, 1980, pp.157
- 14 S.B. Needleman,« *Protein sequence determination* », Springer-Verlag, NY, 1970 , pp.345
- 15 C.T. Mant & Robert S. Hodges, *Methods in Enzymology* , 1996, **271**, p.3
- 16 W.M.A. Niessen, « *Liquid chromatography - Mass spectrometry* », 2<sup>e</sup>Ed, Marcel Dekker, NY, 1999, pp.634

- 17 T.J. Yu, H.A. Schwartz, S.A. Cohen, P. Voutos & B.L. Karger, *Journal of Chromatography*, 1984, **301**, p.425
- 18 J. van der Greef, A.C. Tas, M.A.H. Rijk, M.C. ten Noever de Brauw, M. Höhn, G. Meyerhoff & U. Rapp, *Journal of Chromatography*, 1985, **323**, p.81
- 19 J.G. Stroh, J.C. Cook, R.M. Milberg, L. Brayton, T. Kihara, Z. Huang, K.L. Rinehart, Jr & I.A.S. Lewis, *Analytical Chemistry*, 1985, **57**, p.985
- 20 J.G. Stroh, K.L. Rinehart, Jr, J.C. Cook, T. Kihara, M. Suzuki & T. Arai, *Journal of the American Chemical Society*, 1986, **108**, p.858
- 21 C.N. Kenyon, *Biomed. Mass Spectrom.*, 1983, **10**, p.535
- 22 H. Milton & H. Bur, *Journal of Chromatography*, 1985, **350**, p.399
- 23 C.R. Blakley & M.L. Vestal, *Analytical Chemistry*, 1983, **55**, p.750
- 24 D. Pilosof, H.-Y. Kim, D.F. Dyckes & M.L. Vestal, *Analytical Chemistry*, 1984, **56**, p.1236
- 25 D. Pilosof, H.-Y. Kim, M.L. Vestal & D.F. Dyckes, *Biomedical Mass Spectrometry*, 1984, **11**, p.403
- 26 K. Stachowiak, C. Wilder, M.L. Vestal & D.F. Dyckes, *Journal of the American Chemical Society*, 1988, **110**, p. 1758
- 27 W.P. Blackstock, R.J. Dennis, S.J. Lane, J.I. Sparks & M.P. Weir, *Analytical Biochemistry*, 1988, **175**, p.319
- 28 D.S. Jones & S.T. Krolik, *Rapid Communications in Mass Spectrometry*, 1987, **1**, p.67
- 29 P.J. Rudewicz, *Biomedical and Environmental Mass Spectrometry*, 1988, **15**, p.461
- 30 R.M. Caprioli, T. Fan & J.S. Cottrell, *Organic Mass Spectrometry*, 1986, **58**, p.2949
- 31 K. Mock, J. Firth & J.S. Cottrell, *Organic Mass Spectrometry*, 1989, **24**, p.591
- 32 M.A. Moseley, L.J. Deterding, J.S.M. de Wit, K.B. Tomer, R.T. Kennedy, N. Bragg & J.W. Jorgenson, *Analytical Chemistry*, 1989, **61**, p.1577

- 33 L.J. Deterding, M.A. Moseley, K.B. Tomer, & J.W. Jorgenson, *Analytical Chemistry*, 1989, **61**, p.2504
- 34 M.A. Moseley, L.J. Deterding, K.B. Tomer, & J.W. Jorgenson, *Rapid Communications in Mass Spectrometry*, 1989, **3**, p.87
- 35 C.K. Meng, M. Mann & J.B. Fenn, « *Presented at the 36<sup>th</sup> ASMS Conference on Mass Spectrometry and Allied Topics* », San Francisco, June 5-10, 1988, p.771
- 36 M. Mann, C.K. Meng & J.B. Fenn, *Analytical Chemistry*, 1989, **61**, p.1702
- 37 M.S. Wilm & M. Mann, *Analytical Chemistry*, 1996, **68**, p.1
- 38 R. Feng, Y. Konishi & A.W. Bell, *Journal of the American Chemical Society*, 1991, **2**, p. 387
- 39 R. Feng & Y. Konishi , *Analytical Chemistry*, 1992, **64**, p.2090
- 40 A. Van Dorsselaer, F. Bitsch, B.N. Green, S.A. Jarvis, P. Lepage, R. Bischoff, H.V.J. Kolbe & C. Roitsch, *Biomedical and Environmental Mass Spectrometry*, 1990, **19**, p.692
- 41 L. Baczynskyj & G.E. Bronson, *Rapid Communications in Mass Spectrometry*, 1990, **4**, p.58
- 42 J.H. Bourell, K.P. Clauser, R. Kelley, P.Carter & J.T.Stults, *Analytical Chemistry*, 1994, **66**, p.2088
- 43 P. Roepstorff & J. Fohlmann, *Biomedical and Environmental Mass Spectrometry*, 1984, **11**, p.601
- 44 P. Roepstorff & J. Fohlmann, *Biomedical and Environmental Mass Spectrometry*, 1985, **12**, p.631
- 45 E.C. Huang & J.D. Henion, *Journal of the American Chemical Society*, 1990, **1**, p. 158
- 46 E.C. Huang & J.D. Henion, *Analytical Chemistry*, 1991, **63**, p.732
- 47 P.R. Griffin, J.A. Coffmann, L.E. Hood & J.R. Yates,III, *International Journal of Mass Spectrometry and Ion Processes*, 1991, **111**, p.131
- 48 P.Thibault, S. Pleasance, M.V. Laycock, R.M. Mackay & R.K. Boyd , *International Journal of Mass Spectrometry and Ion Processes*, 1991, **111**, p.317

- 49 V. Ling, A.W. Guzzetta, E. Canova-Davis, J.T. Stults, W.S. Hancock, T.R. Covey & B.I. Shushan, *Analytical Chemistry*, 1991, **63**, p.2909
- 50 M. Mann, P. Hojrup & P. Roepstroff, *Biological Mass Spectrometry*, 1993, **22**, p.338
- 51 S.K. Chowdhury, V. Katta & B.T. Chait, *Biochemical and Biophysical Research Communications*, 1990, **167**, p.686
- 52 J.K. Eng, A.L.M. McCormack & J.R. Yates,III, *Journal of the American Chemical Society*, 1994, **5**, p.976
- 53 J.R. Yates,III, J.K. Eng, A.L.M. McCormack & D. Schieltz, *Analytical Chemistry*, 1995, **67**, p.1426
- 54 A.L.M. McCormack, D. Schieltz, B. Goode, S. Yang, G. Barnes, D. Drubin & J.R. Yates,III, *Analytical Chemistry*, 1997, **69**, p.767
- 55 D.A. Skoog & J.J. Leary, « *Principles of instrumental analysis* », 4<sup>°</sup>Ed, NY, 1992,Saunders College Publishing, pp.700
- 56 M.-I. Aguilar & M.T.W.Hearn, *Methods in Enzymology* , 1996, **271**, p.3
- 57 A. Apffel, S.M. Fischer, G. Goldberg, P.C. Goodley & F.E. Kuhlmann, *Journal of Chromatography A*, 1995 , **712**, p.177
- 58 D.J. Poll & D.R.K. Harding, *Journal of Chromatography*, 1989, **469** , p.231
- 59 G.Thévenon & F.E. Regnier, *Journal of Chromatography* , 1989, **476**, p.499
- 60 F.E. Kuhlmann, A. Apffel, S.M. Fischer, G. Goldberg & P.C. Goodley, *Journal of the American Society in Mass Spectrometry*, 1995, **6** , p.1221
- 61 M.Bertrand, « *Notes de cours de CHM 6140* », UdeM, Mtl , 1999
- 62 M. Desiderio (Ed), « *Mass spectrometry : clinical and biomedical applications Vol.1* », Plenum Press, NY, 1992, pp.353
- 63 G.I. Taylor, « *Proc. R. Soc.* », 1964, **A280**, p.383
- 64 HP1100 LC/MS System Overview, Hewlett Packard, pp.42
- 65 M. Dole, « *The radiation chemistry of macromolecules* », Academic Press, NY, 1972, pp.369

- 66 J.V. Iribarne & B.A. Thomson, *The Journal of Chemical Physics*, 1976, **64**, p.2287
- 67 K.L. Busch, G.L. Glish & S.A. McLuckey, « *Mass spectrometry / mass spectrometry: techniques and applications of tandem mass spectrometry* », VCH, NY, 1988, pp.333
- 68 R.G. Cooks & G.L. Glish, *Chemical and Engineering News*, 1981, **Nov30**, p.40
- 69 M.E. Rose & R.A.W. Johnstone, « *Mass spectrometry for chemists and biochemists* », Cambridge University Press, NY, 1982, pp.307
- 70 R.S. Johnson, S.A. Martin, K. Biemann, J.T. Stults & J.T. Watson, *Analytical Chemistry*, 1987, **59**, p.2621
- 71 A.G. Craig, S.C. Koerber, J. Porter, C. Hoeger & J.E. Rivier, *Biological Mass Spectrometry*, 1993, **22**, p.31
- 72 J.T. Watson, « *Introduction to mass spectrometry* », Raven Press, NY, 1985, pp.351
- 73 P. Matsudaira (Ed), « *A practical guide to protein and peptide purification for microsequencing* », 2<sup>e</sup>Ed, Academic Press Inc, NY, 1993, pp.184
- 74 H. Abouchacra, « *Communications personnelles* »
- 75 B. Lönnerdal, S. Berström, Y. Andersson, K. Hjalmarsson, A.K. Sundqvist & O. Hernel, *FEBS Letters*, 1990, **269(1)**, p.153
- 76 S.N. Deming & L.R. Parker, Jr., *CRC Critical Reviews in Analytical Chemistry*, 1978, **7(3)**, p.187
- 77 Walters, Parker, Morgan & Deming (Ed), « *Sequential simplex optimization* », CRC Press LLC, Boca Raton (Florida), 1991, pp.403
- 78 L. de Galan & H.A.H. Billiet, *Advances in Chromatography*, 1986, **25**, p.63
- 79 J.C. Berridge, *Chromatographia*, 1982, **16**, p.172
- 80 J.C. Berridge & E.G. Morrissey, *Journal of Chromatography*, 1981, **316**, p.69
- 81 J.C. Berridge, *Journal of Chromatography*, 1982, **244**, p.1

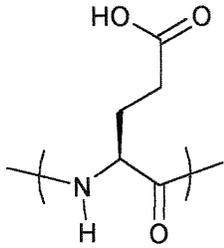
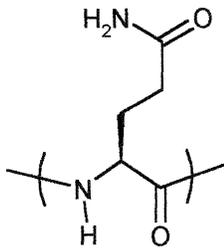
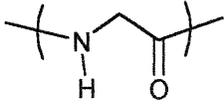
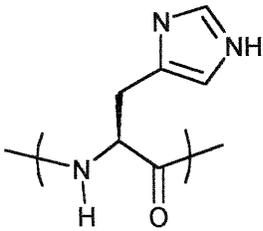
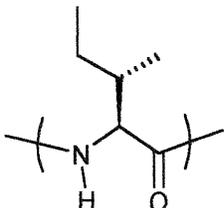
- 82 M. D'Amboise, « *Notes de cours de CHM6170* », UdeM, Mtl, 1999, pp.141
- 83 É. Bonneil, *Thèse de doctorat*, UdeM, Mtl, 2000, pp.196

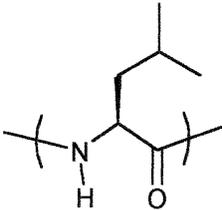
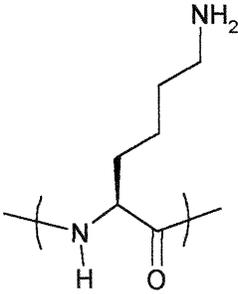
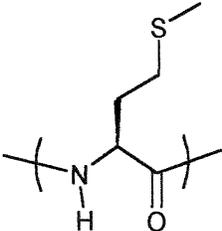
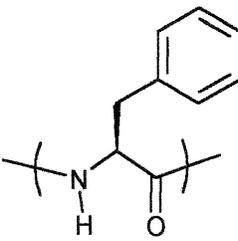
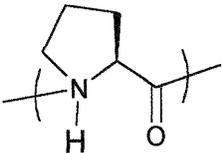
## **ANNEXE I**

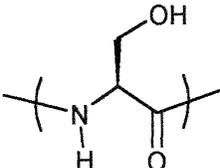
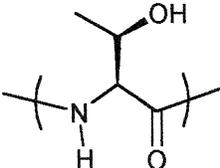
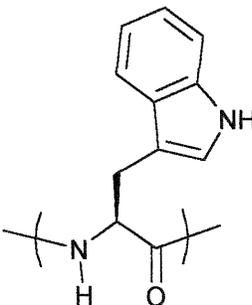
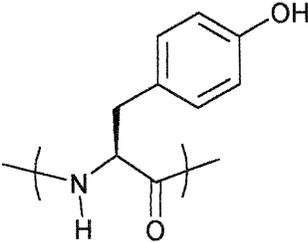
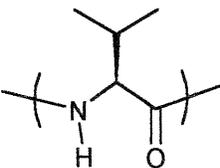
**STRUCTURES, MASSES MOLÉCULAIRES  
ET ABBRÉVIATIONS POUR LES VINGTS  
ACIDES AMINÉS PHYSIOLOGIQUES <sup>(12)</sup>**

### Table des vingt acides aminés physiologiques

Acide aminé	Abréviation en trois lettres de l'acide aminé	Abréviation en une lettre	Structure du résidu	Masse du résidu (g/mol)
Alanine	Ala	A		71,1
Arginine	Arg	R		156,2
Asparagine	Asn	N		114,1
Acide aspartique	Asp	D		115,1
Cystéine	Cys	C		103,1

Acide aminé	Abréviation en trois lettres de l'acide aminé	Abréviation en une lettre	Structure du résidu	Masse du résidu (g/mol)
Acide glutamique	Glu	E		129,1
Glutamine	Gln	Q		128,1
Glycine	Gly	G		57,1
Histidine	His	H		137,2
Isoleucine	Ile	I		113,2

Acide aminé	Abréviation en trois lettres de l'acide aminé	Abréviation en une lettre	Structure du résidu	Masse du résidu (g/mol)
Leucine	Leu	L		113,2
Lysine	Lys	K		128,2
Methionine	Met	M		131,2
Phénylalanine	Phe	F		147,2
Proline	Pro	P		97,1

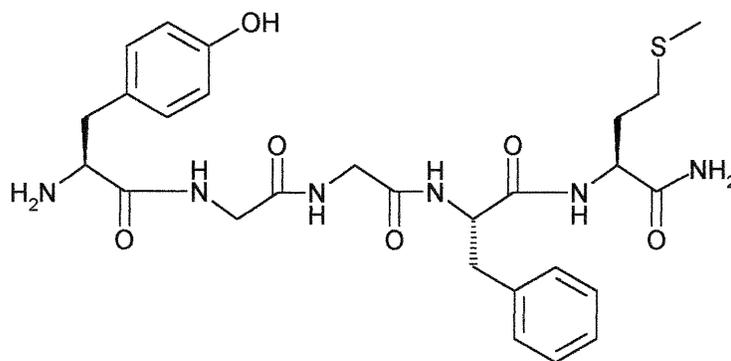
Acide aminé	Abréviation en trois lettres de l'acide aminé	Abréviation en une lettre	Structure du résidu	Masse du résidu (g/mol)
Sérine	Ser	S		87,1
Thréonine	Thr	T		101,1
Tryptophane	Trp	W		186,2
Tyrosine	Tyr	Y		163,2
Valine	Val	V		99,1

## **ANNEXE II**

### **STRUCTURES, MASSES ET FORMULES MOLÉCULAIRES DES PEPTIDES ANALYSÉS**

**MÉTHIONINE ENKÉPHALINAMIDE***Tyr-Gly-Gly-Phe-Met-NH<sub>2</sub>*Y-G-G-F-M-NH<sub>2</sub>C<sub>27</sub>H<sub>36</sub>N<sub>6</sub>O<sub>6</sub>S

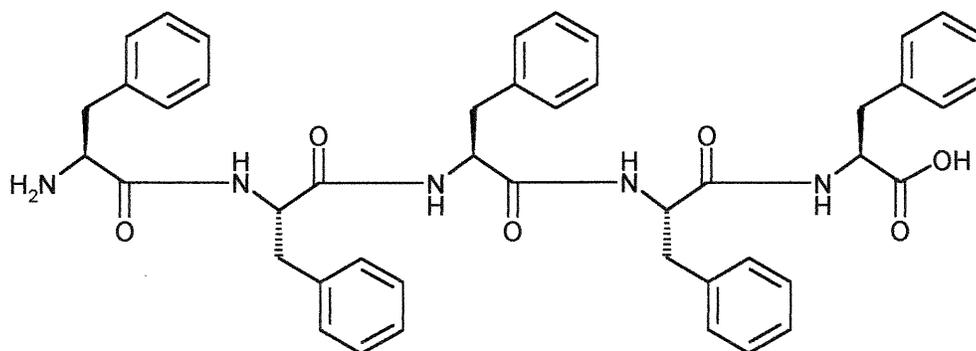
MM = 572,24 g/mol

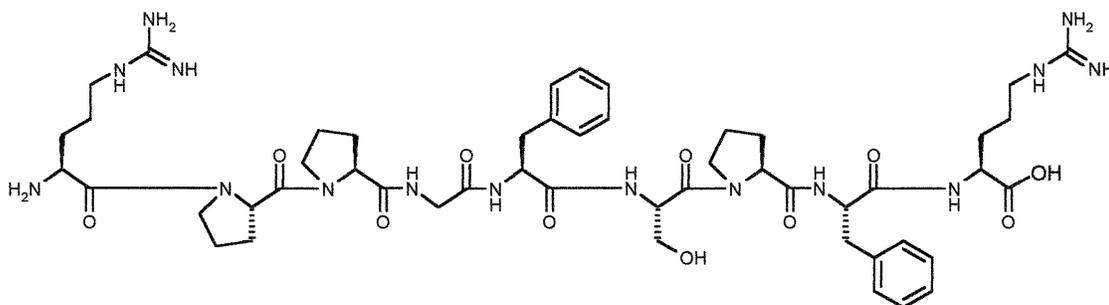
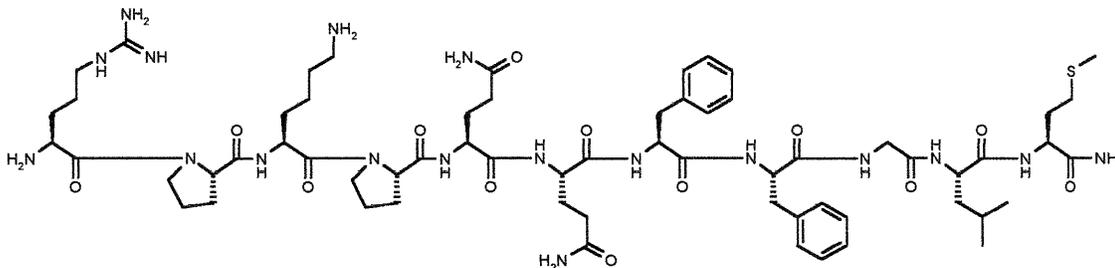
**PENTAPHÉNYLALANINE***Phe-Phe-Phe-Phe-Phe*

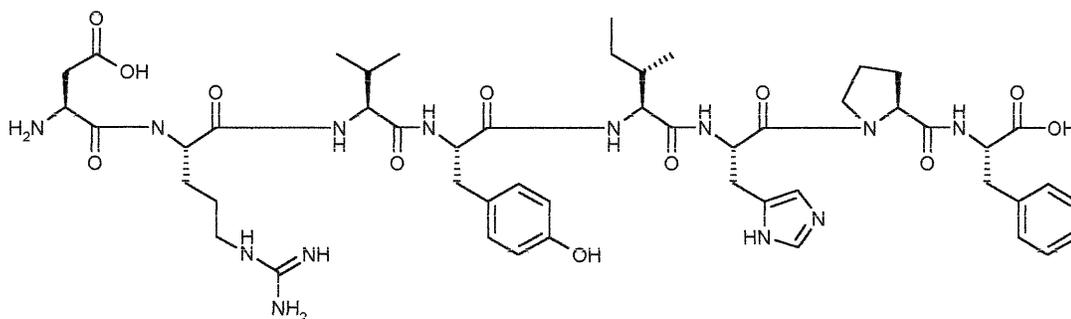
F-F-F-F-F

C<sub>45</sub>H<sub>47</sub>N<sub>5</sub>O<sub>6</sub>

MM = 753,35 g/mol



**BRADYKININE***Arg-Pro-Pro-Gly-Phe-Ser-Pro-Phe-Arg***R-P-P-G-F-S-P-F-R** $C_{50}H_{73}N_{15}O_{11}$ **MM = 1060,22 g/mol****SUBSTANCE P***Arg-Pro-Lys-Pro-Gln-Gln-Phe-Phe-Gly-Leu-Met-NH<sub>2</sub>***R-P-K-P-Q-Q-F-F-G-L-M-NH<sub>2</sub>** $C_{63}H_{98}N_{18}O_{13}S$ **MM = 1346,73 g/mol**

**ANGIOTENSINE II***Asp-Arg-Val-Tyr-Ile-His-Pro-Phe***D-R-V-Y-I-H-P-F****C<sub>50</sub>H<sub>71</sub>N<sub>13</sub>O<sub>12</sub>****MM = 1045,53 g/mol****RANATACHYKININE***Lys-Pro-Ser-Pro-Asp-Arg-Phe-Tyr-Gly-Leu-Met***K-P-S-P-D-R-F-Y-G-L-M****C<sub>60</sub>H<sub>91</sub>N<sub>15</sub>O<sub>16</sub>S****MM = 1309,65 g/mol**