

50376
1993
171

UNIVERSITE DES SCIENCES ET TECHNOLOGIES
DE LILLE

50376
1993
171

N° D'ORDRE: 1113

THESE

Présentée

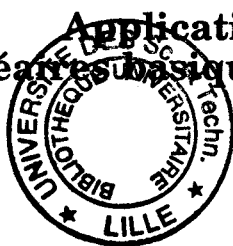
Pour l'obtention du grade de
Docteur en Sciences de la Vie et de la Santé

Par

Mostafa KOUACH

**CONTRIBUTION
DE LA SPECTROMETRIE DE MASSE
A L'ETUDE STRUCTURALE DES PROTEINES**

**Application
aux protéines nucléaires basiques de la Spermatogenèse**



Soutenue le 23 Avril 1993 devant la Commission d'Examen:

Président:	Professeur A. VERBERT
Rapporteurs:	Docteur J.C. PROMÉ Docteur A. VAN DORSSELAER
Examineurs:	Docteur P. SAUTIERE Docteur G. RICART Docteur G. BRIAND

A la mémoire de Monsieur le
Professeur Bernard FOURNET

A mes parents

A mon épouse

Aux familles:

KOUACH

ALIX

SIMON

A mes amis

Ce travail a été réalisé sous la direction scientifique du Docteur Gilbert BRIAND, Maître de Conférences à la Faculté de Médecine de Lille, dans l'Unité de Recherche Associée 409 du CNRS, "Biochimie et Biologie de la spermatogenèse", dirigée par le Docteur Pierre SAUTIERE, Directeur de recherche au CNRS, au sein de l'Institut de Recherches sur le Cancer de Lille, dirigé par le Professeur Jacques VANLERENBERGHE.

Les spectres de masse ont été réalisés dans :

- Le laboratoire de spectrométrie de masse dirigé par le Professeur Bernard FOURNET et le Docteur Guy RICART à Lille.
- Le laboratoire de spectrométrie de masse dirigé par le Docteur Alain VAN DORSSELAER à Strasbourg.

En outre, ce travail n'aurait pu être réalisé sans l'attribution d'une allocation co-financée par l'IRCL et la Région Nord Pas-de-Calais. A ce propos, je tiens à remercier du fond du coeur, toutes les personnes qui ont contribué à cette attribution.

Je tiens à remercier très vivement :

- Monsieur le Professeur A. VERBERT pour l'honneur qu'il m'accorde en présidant le jury de ma thèse.

- Messieurs les Docteurs J.C. PROME et A. VAN DORSSELAER d'avoir accepté d'être les rapporteurs de cette thèse.

- Messieurs les Docteurs P. SAUTIERE, G. RICART et G. BRIAND qui ont bien voulu examiner ce travail.

Je tiens à exprimer ma reconnaissance à :

Monsieur le Docteur Pierre SAUTIERE, qui m'a accueilli au sein de son Unité de Recherche, pour toute l'aide et la confiance qu'il m'a accordées.

Monsieur le Docteur Gilbert BRIAND d'avoir bien voulu m'encadrer et guider mes premiers pas dans la spectrométrie de masse. Je le remercie vivement pour la confiance qu'il m'a toujours accordée.

Madame le Docteur Denise BELAICHE pour toute l'aide qu'elle m'a prodiguée et tous les services qu'elle m'a rendus depuis mon arrivée au laboratoire.

Messieurs les Docteurs Guy RICART et Yves LEROY de m'avoir formé au fonctionnement de la FAB.

Monsieur le Docteur Alain VAN DORSSELAER et son équipe qui m'ont toujours bien reçu dans leur laboratoire et qui ont bien voulu m'initier à la technique d'électrospray.

Monsieur le Professeur VANLERENBERGHE et Madame BARBILLON pour la contribution financière qu'ils m'ont apportée tout au long de ce travail.

Je tiens également à remercier vivement :

Madame le Docteur LOUCHEUX et son équipe de m'avoir accepté dans leur Unité pour me permettre de mener à terme ce travail.

Messieurs les Professeurs DAUTREVAUX et FORMSTECHEUR ainsi que tous les membres de l'Interface Pharmacochimie-Biologie Cellulaire qui m'ont impliqué dans l'achat de l'appareil du service commun de spectrométrie de masse de l'Université de Lille II. Encore merci pour toute la confiance que vous m'accordez.

Tout le personnel des Unités 409 CNRS, 124 INSERM et de l'IRCL pour leur aide, leur soutien et l'amitié qu'ils m'ont témoignée.

Publications

- 1) Nuclear transition protein 1 from ram elongating spermatids. Mass spectrometry characterization, primary structure and phosphorylation sites of two variants.
Chirat F., Martinage A., Briand G., Kouach M., Van Dorsselaer A., Loir M. and Sautière P. (1991)
Eur. J. Biochem. 198, 13-20
- 2) A corrected primary structure for dog-fish *Scylliorhinus caniculus* protamine Z3.
Kouach M., Jaquinod M., Belaïche D., Sautière P., Van Dorsselaer A., Chevaillier P. and Briand G. (1993)
Biochim. Biophys. Acta 1162, 99-104
- 3) Application of Electrospray and Fast Atom Bombardment Mass Spectrometry to the identification of chemical and post-translational modifications of proteins and peptides.
Kouach M., Belaïche D., Jaquinod M., Couppez M., Kmiecik D., Ricart G., Van Dorsselaer A., Sautière P. and Briand G. (1993)
Biol. Mass Spectrom. soumis pour publication
- 4) Human bronchial mucus proteinase inhibitor and its amino- and carboxy- terminal domains. Electrospray mass spectrometry characterization and comparison with the recombinant protein.
Van Seuning I., Jaquinod M., Van Dorsselaer A., Briand G., Kouach M. and Davril M. (1993)
J. Biochem. soumis pour publication

Communications orales ou par affiches

- 1) Application de la spectrométrie de masse à l'étude structurale des peptides et des protéines.
Kouach M., Belaïche D., Chirat F., Briand G. and Sautière P.
NFWO-NFRS CONTACT GROUPS
"Structure and function of proteins"
"Structure, Conformation and Synthesis of Biological Active Peptides"
Meeting du 11 Janvier 1991- Institut Pasteur de Lille

- 2) Apport de la spectrométrie de masse à l'étude structurale des protéines et des peptides basiques.

Briand G., Kouach M., Belaïche D., Martin-Ponthieu A., Chirat F., Van Dorsselaer A. and Sautière P.

7ème Réunion Peptide 18-22 Mars 1991 Aussois

- 3) Contribution de la spectrométrie de masse à l'étude structurale de la protéine intermédiaire TP1 de bélier.

Kouach M., Chirat F., Belaïche D., Bitsch F., Sautière P. and Briand G.

7ème Réunion Peptide 18-22 Mars 1991 Aussois

- 4) Analyse des peptides et des protéines basiques par spectrométrie de masse.

Kouach M., Belaïche D., Kmiecik D., Jaquinod M., Ricart G., Sautière P. and Briand G.

XVIIIème Forum des jeunes chercheurs 3-6 Septembre 1991 Tours

Abréviations utilisées

API	:	atmospheric pressure ionization
CAD	:	collision activated dissociation
cDNA	:	acide désoxyribonucléique complémentaire
CF-FAB	:	continuous flow fast atom bombardment
CLHP	:	chromatographie liquide haute performance
Da	:	dalton
DTT	:	dithiothréitol
EDTA	:	acide éthylènediaminetétraacétique
ES-MS	:	electrospray mass spectrometry
eV	:	électron Volt
FAB	:	fast atom bombardment
GuCl	:	chlorhydrate de guanidine
IBB	:	indice de Bull et Breese
IE-TFA	:	N-(Iodoéthyl)-trifluoroacétamide
IS-MS	:	ion spray mass spectrometry
kDa	:	kilo Dalton
keV	:	kilo électron Volt
kV	:	kilo Volt
LC-MS	:	liquid chromatography mass spectrometry
LD-MS	:	laser desorption mass spectrometry
LSIMS	:	liquid secondary ion mass spectrometry
m-NBA	:	alcool méta-nitrobenzylique
m/z	:	masse sur charge
MALDI	:	matrix assisted laser desorption/ionization
MCA	:	multi channel analysis
MeV	:	méga électron Volt
MIKES	:	mass analysed ion kinetic energy spectrometry
MS	:	mass spectrometry
MS-MS	:	tandem mass spectrometry
PCA	:	acide perchlorique
PD-MS	:	plasma desorption mass spectrometry
PTH	:	phénylthiohydantoïne
TEA	:	triéthylamine
TFA	:	acide trifluoroacétique
TOF	:	time of flight
u m a	:	unité de masse atomique

Table des matières

	pages
Introduction générale	1
Chapitre 1: Généralités sur la spectrométrie de masse	4
I. Introduction	5
II. Définitions	7
II.1. Le spectromètre de masse	7
II.2. La résolution	7
II.3. Le dalton	9
II.4. Masse nominale, masse isotopique et masse chimique	9
II.5. La calibration	10
III. La technique FAB-MS	11
III.1. Introduction	11
III.2. Principe de la méthode FAB	11
III.3. Rôle de la matrice	13
III.4. Mécanisme d'ionisation et nature des ions formés	13
III.4.1. Mécanisme d'ionisation	13
III.4.2. Nature des ions formés	14
III.5. Phénomène de suppression	15
III.6. Continuous flow FAB	16
III.7. Utilisations et performances de la FAB	16
III.8. Limites de la FAB	17
IV. La désorption par Plasma ^{252}CF PD-MS	18
IV.1. Source au californium	18
IV.2. Mécanisme d'ionisation et nature des ions formés	18
IV.3. Matrices	20
IV.4. Utilisations et performances de la PD-MS	21
IV.5. Limites de la PD-MS	22
V. La désorption laser	23
V.1. Description de la technique	23
V.2. Utilisations et performances de la LD-MS	23
V.3. Limites de la LD-MS	25

VI. L'Electrospray-ionspray	26
VI.1. Introduction	26
VI.2. Principe	26
VI.3. Les facteurs influençant l'état de charge	28
VI.4. Etat de charge des ions	29
VI.5. Présentation d'un spectre obtenu en mode électrospray	30
VI.6. Performances et limites de la technique	32
VII. Les différents types d'analyseurs en spectrométrie de masse	34
VII.1. Analyseurs magnétiques et électrostatiques	34
VII.1.1. Analyseurs magnétiques	35
VII.1.2. Analyseurs électrostatiques	35
VII.2. Analyseurs en temps de vol	37
VII.3. Analyseurs quadripolaires	37
VII.4. Méthodes de visualisation	39
VII.5. Etalonnage	39
VIII. La spectrométrie de masse en tandem ou MS-MS	41
Chapitre 2: Rappels sur la spermatogenèse	46
Travaux personnels	52
Chapitre 3: Utilisation de la spectrométrie de masse pour la correction d'une séquence protéique.	53
Etude de la Protéine Z3 de Roussette	54
I. Introduction	54
II. Publication 1	57
III. Discussion et conclusion	63
Chapitre 4: Utilisation de la spectrométrie de masse pour la mise en évidence de variants structuraux.	65
Etude de la protéine de transition TP1 de bélier	66
I. Introduction	66
II. Publication 2	67

III. Conclusion	75
Chapitre 5: Contribution de la spectrométrie de masse à l'étude structurale d'une protéine.	81
Etude de la protéine de transition TP2 de béliet.	82
I. Introduction	82
II. Préparation de la protéine TP2	82
III. Réduction et alkylation des protéines	84
III.1. Par l'iodoacétamide	84
III.2. Par l'IE-TFA	84
IV. Contrôle électrophorétique	84
V. Purification de la protéine TP2	85
V.1. Purification sur colonne d'échange d'ions CM52	85
V.2. Séparation des protéines intermédiaires TP2 et TP3 en CLHP-phase inverse	85
VI. Composition en acides aminés de la protéine TP2	85
VII. Dégradation récurrente d'Edman	90
VIII. Analyse de la protéine TP2 en électrospray	90
IX. Hydrolyse de la protéine TP2 par la thermolysine	90
IX.1. Etude de la fraction B-1	92
IX.2. Etude de la fraction D-3	95
IX.3. Etude de la fraction D-2	95
IX.4. Etude de la fraction C-1	97
IX.5. Etude de la fraction A-6	99
X. Hydrolyse de la protéine TP2 par la protéase staphylococcique	99
XI. Hydrolyse du peptide N-terminal par l'endoprotéase Arg-C	102
XI.1. Etude de la fraction 2	102
XI.2. Etude de la fraction 4	102
XII. Coupure du peptide N-terminal par l'endoprotéase Lys-C	104

XIII. Coupure du peptide N-terminal par l'endoprotéase Asn-C	104
XIV. Conclusion	104
Chapitre 6: Utilisation de la spectrométrie de masse pour la détermination de modifications chimiques et post-traductionnelles	106
Etude de modifications chimiques et post-traductionnelles de protéines basiques.	107
I. Introduction	107
II. Publication 3	108
III. Discussion	148
III.1. Peptide NH ₂ -terminal de l'histone H5 de <i>Xenopus laevis</i>	148
III.2. Peptide de synthèse de la cholécystokinine	148
III.3. Peptide carboxamidométhylé de la protamine de bélier	148
III.4. Phosphorylation de la protamine de bélier	150
III.5. Acétylation et méthylation de l'histone H4 de thymus de veau	153
IV. Conclusion	158
Conclusion générale et perspectives	159
Bibliographie	163
Annexes	193

INTRODUCTION GENERALE

Ces dernières années, des progrès spectaculaires ont été réalisés en spectrométrie de masse, notamment dans le domaine des peptides et des protéines. En effet, l'introduction en 1981 de la technique FAB ou bombardement par atomes rapides a permis l'ionisation et la volatilisation des molécules polaires et thermolabiles. D'autres méthodes d'ionisation (Désorption au Californium 252 ou PD-MS, désorption laser ou LD-MS, électrospray ou ES-MS) ont également été développées, permettant ainsi de résoudre la plupart des problèmes auxquels sont confrontés les laboratoires spécialisés dans l'étude structurale des protéines. Ainsi, la spectrométrie de masse permet:

- de déterminer la masse moléculaire d'une protéine et d'estimer son degré de pureté,
- de vérifier la masse moléculaire d'une protéine après la détermination de sa séquence, soit par le cDNA, soit par la méthode récurrente d'Edman. Ce qui permet de voir:
 - *si la séquence établie est correcte,
 - *l'existence ou non de modifications post-traductionnelles,
 - *l'existence ou non de variants structuraux.
- de déterminer et de localiser les modifications post-traductionnelles ou chimiques,
- d'avoir des informations structurales sur une protéine après sa fragmentation,
- d'étudier l'existence et la localisation des ponts disulfure,
- d'étudier l'interaction d'une protéine avec son ligand ou des ions métalliques (par exemple le zinc dans la formation de doigts de zinc),
- de vérifier la masse et la structure d'un peptide de synthèse.

Ce constat a attiré notre laboratoire vers l'utilisation de cette technique. En effet, orienté vers l'étude structurale des protéines nucléaires basiques associées à la chromatine au cours de la spermatogenèse, notre laboratoire se trouvait confronté à un certain nombre de difficultés structurales inhérentes aux caractéristiques particulières de ces protéines: séquences répétitives, variants structuraux, modifications post-traductionnelles, NH₂-terminaux bloqués.

L'apport de la spectrométrie de masse a été fondamental dans la résolution de ces problèmes structuraux.

Notre laboratoire ne disposant pas de spectromètre de masse, nous avons travaillé en collaboration avec:

- le laboratoire de spectrométrie de masse dirigé par le Professeur Bernard FOURNET et le Docteur Guy RICART à Lille, pour la technique FAB-MS,

- le laboratoire de spectrométrie de masse dirigé par le Docteur Alain VAN DORSSELAER à Strasbourg pour la ES-MS.

C'est ainsi qu'au cours de notre DEA et de notre thèse, nous avons en étroite collaboration avec ces deux équipes, coopéré à l'utilisation et à la vulgarisation de cette technique au sein du laboratoire. Nous avons également montré son intérêt dans l'étude structurale des peptides et des protéines basiques.

Nos travaux, dans le cadre de la thématique de l'unité 409 CNRS, ont porté sur:

- l'étude de la protéine Z3 isolée du testicule de roussette (publication 1),
- l'étude de variants structuraux de la protéine TP1 isolée du testicule de bélier (publication 2),
- l'étude de modifications chimiques et post-traductionnelles de certaines protéines basiques (publication 3),
- l'étude partielle de la structure de la protéine TP2 isolée du testicule de bélier.

Nous nous proposons, tout d'abord, de décrire la technique de spectrométrie de masse, ses champs d'investigations, ses limites, puis après un bref rappel sur la spermatogenèse, de présenter nos résultats.

Chapitre 1

Généralités sur la spectrométrie de masse

I. Introduction:

La spectrométrie de masse, utilisée depuis longtemps pour l'analyse de composés organiques, était limitée aux molécules volatiles. Pour certaines substances dites "non volatiles", il était possible d'accroître, par un apport thermique, la tension de vapeur nécessaire à l'obtention d'un spectre de masse. Malheureusement, le produit thermolabile n'était pas identifié sous sa forme originelle à cause des transformations pyrolytiques qu'il subissait. D'autre part, les températures maximales atteintes par la sonde d'introduction directe (300-350°C) ne conduisaient pas toujours à la production de cette tension de vapeur (cas de sels inorganiques ou de composés organiques de haut poids moléculaire: peptides, protéines...). Pourtant, de telles substances se rencontrent très fréquemment en biochimie, chimie, pharmacologie. La faible volatilité de ces produits provient de l'existence d'interactions intermoléculaires de forte énergie comme par exemple des liaisons ioniques (sels), ou d'un grand nombre de liaisons hydrogènes comme dans le cas des substances polyhydroxylées (sucres), ou encore de leur masse moléculaire très élevée.

A l'époque, l'analyse des composés de masse inférieure à 1500 Da par spectrométrie de masse était possible en les rendant volatils. Pour cela, une modification des fonctions chimiques présentes était nécessaire. On remplaçait ainsi les interactions polaires à forte énergie par des interactions hydrophobes de plus faible énergie. C'est ainsi que Das et *al* (1) ont dérivé, en 1967, des peptides par N-acétylation et N-O-perméthylation. Un simple chauffage sous vide permettait alors de volatiliser les peptides ainsi modifiés. Cette dérivation avait en plus l'avantage de fragiliser les liaisons peptidiques. Les ions ainsi formés lors de l'ionisation par des électrons (ionisation par Impact Electronique), ou par un gaz interposé (ex: le méthane, Ionisation Chimique), donnaient ce qu'on appelle des pics de séquence. De la différence de masse entre les pics obtenus, on déduisait la séquence du peptide analysé. C'est ainsi que des séquences de décapeptides ont été effectuées.

En 1978, Biemann a introduit une technique de dérivation qui consistait à réduire le peptide en polyaminoalcools. Cette technique permettait de séquencer des peptides plus longs (2). Néanmoins, ces techniques avaient plusieurs inconvénients, à savoir : la réalisation délicate des réactions chimiques de dérivation sur des faibles quantités de peptides, la difficulté de l'interprétation des spectres quand les réactions n'étaient pas

complètes et enfin, l'impossibilité de volatiliser des peptides de masse supérieure à 500 Da, même après dérivation.

Ce n'est qu'en 1981, que la spectrométrie de masse a véritablement trouvé sa place dans le domaine des molécules non volatiles et thermolabiles par l'introduction, par Barber, de la technique de bombardement par atomes rapides ou **FAB-MS** (Fast Atom Bombardment-Mass Spectrometry) (3). Cette technique facile à mettre en oeuvre permet de travailler sur des peptides non modifiés. Depuis, d'autres techniques ont été développées dans le but de travailler d'une façon routinière sur des molécules polaires et thermosensibles en très faibles quantités. Elles ont permis de déplacer les frontières de la spectrométrie de masse des petites molécules vers les macromolécules.

Ce développement a abouti aux techniques: **LSIMS** (Liquid Secondary Ion Mass Spectrometry (4), désorption par plasma ou **PD-MS** (Plasma Desorption Mass Spectrometry) (5), désorption par Laser assistée par une matrice ou **LD-MS** (Laser Desorption Mass Spectrometry) (6), électrospray ou **ES-MS** (7) ou ion spray ou **IS-MS** (8).

Nous allons décrire ces six techniques qui sont, actuellement, les plus couramment utilisées dans le domaine des peptides et des protéines. Mais avant cela, rappelons d'abord quelques définitions:

II. Définitions:

II.1. Le spectromètre de masse:

Un spectromètre de masse est un appareil qui sert à mesurer avec une grande précision les masses moléculaires et qui permet d'avoir des informations structurales sur les substances étudiées.

Il est constitué de quatre parties (Fig.1):

- Un système d'introduction. Il peut s'agir d'un système d'introduction directe ou de couplage à une méthode chromatographique (CLHP) ou électrophorétique (électrophorèse capillaire).

- Une source d'ions où ont lieu la volatilisation et l'ionisation des composés. Elle se caractérise par son mode d'ionisation.

- Un analyseur qui sépare les ions formés dans la source suivant le rapport masse sur charge (m/z). Il se caractérise par son principe de séparation des ions.

- Un détecteur qui donne l'intensité du faisceau ionique correspondant à chaque espèce.

Dans le domaine des peptides et des protéines, la difficulté réside au niveau de la volatilisation et de l'ionisation de ces molécules ainsi qu'au niveau de la mesure de la masse (m/z) pour les molécules de masses élevées. En fonction de la nature de la molécule à analyser, de sa masse moléculaire et de la quantité disponible, nous serons amené à choisir une technique plutôt qu'une autre.

II.2. La résolution :

La résolution se définit comme étant la capacité de l'appareil à discerner 2 pics de masses très proches. Elle est égale au rapport de la masse d'un des pics sur la différence de masse entre les 2 pics: $R = m/\Delta m$ lorsque ceux-ci se recouvrent sur une hauteur h (pourcentage de vallée = $n\%$ de la hauteur des pics, n étant généralement égal à 5 ou 10) (Fig.2).

Si la résolution d'un appareil est fixée, par exemple à 1000, il pourra distinguer les pics de masses 1000 et 1001 Da ou encore 100,0 et 100,1 Da.

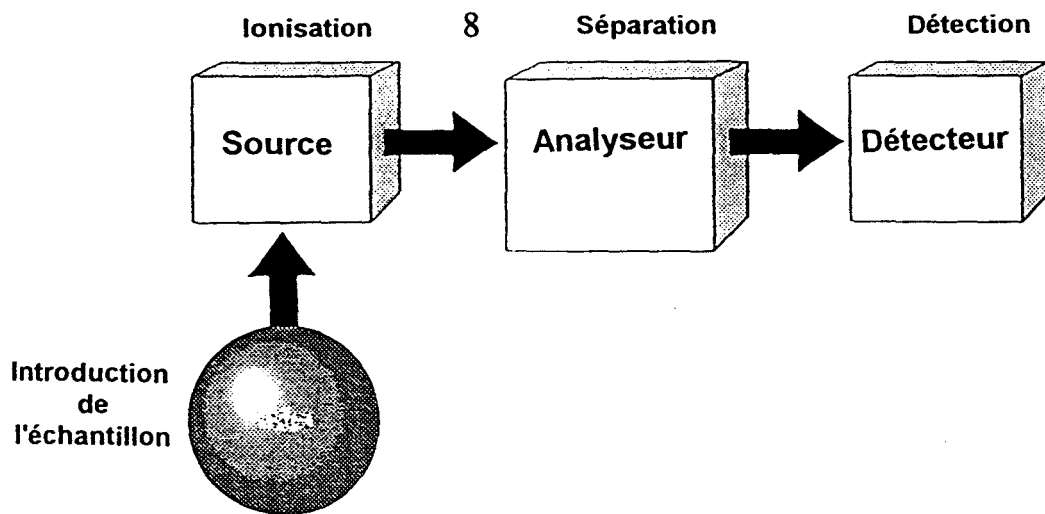


Figure 1: Les différentes parties d'un spectromètre de masse

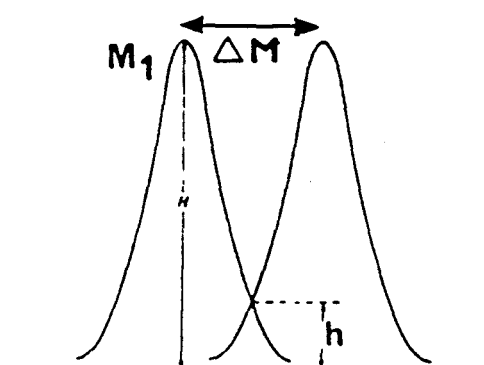


Figure 2: Définition de la résolution

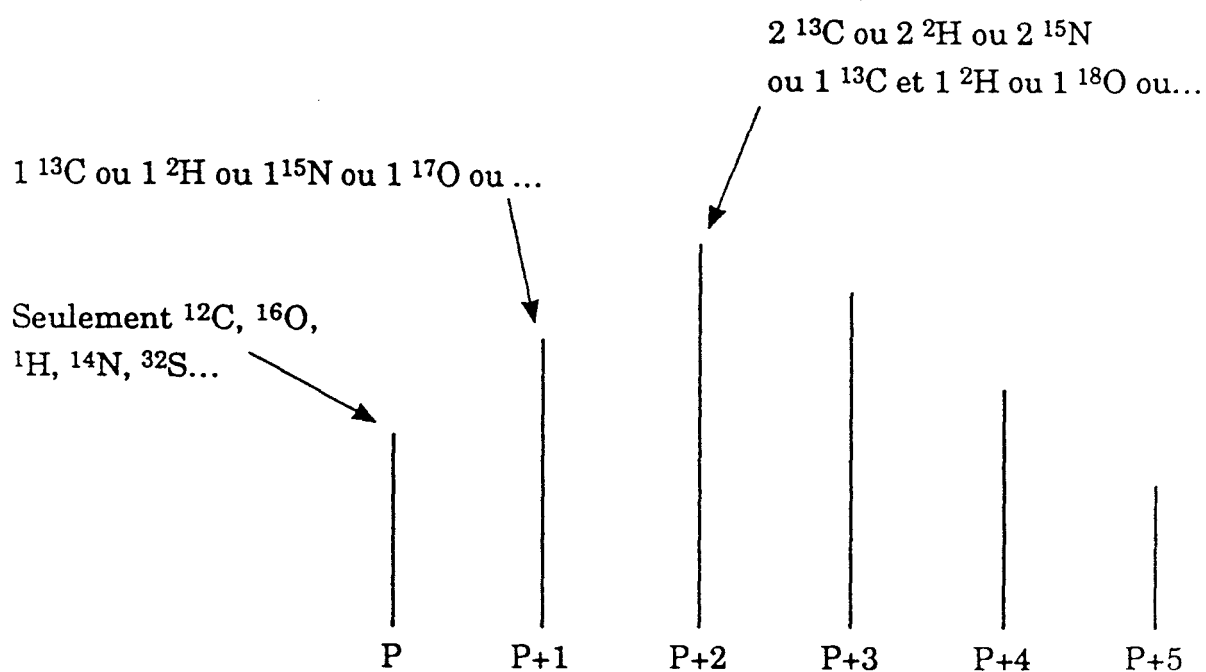


Figure 3: Profil isotopique d'un pic moléculaire à résolution unitaire

II.3. Le dalton :

Le dalton (Da) ou unité de masse atomique (u.m.a) représente le 1/12^{ème} de la masse d'un atome de carbone de masse 12.

II.4. Masse nominale, masse isotopique et masse chimique :

Dans la nature, les composés organiques et minéraux sont formés de mélanges d'isotopes stables des éléments. Le carbone, par exemple, contient 1,1% de l'isotope ^{13}C . La pesée d'un produit représente une masse moyenne (ou masse chimique). En réalité, cette masse correspond au barycentre de la masse des espèces moléculaires comprenant les différents isotopes des éléments.

En spectrométrie de masse, lorsqu'on travaille à une résolution unitaire, le signal obtenu est sous la forme d'une succession de pics séparés par une unité de masse. Ces pics correspondent à une distribution statistique des différents isotopes dans la molécule étudiée (9). (Fig.3 page8). Sur cette figure, le pic P correspond à la molécule ne contenant que les isotopes les plus légers (^{12}C , ^{14}N , ^1H , ^{16}O , ^{32}S etc...). Dans ce cas, il s'agit du **pic monoisotopique**. Quand la masse moléculaire est donnée en ne tenant compte que des isotopes les plus légers des éléments, on parlera de **masse monoisotopique**, le calcul étant fait à plusieurs décimales près (ex: $\text{H}=1,007825$). Quand celui-ci est fait au chiffre arrondi (ex: $\text{H}=1$, $\text{C}=12$ etc...), on parlera de **masse nominale**.

Les pics $\text{P}+1$, $\text{P}+2$,..., $\text{P}+n$ (Fig.3) représentent les combinaisons multiples des autres isotopes. Le pic $\text{P}+1$, par exemple, peut provenir des molécules contenant un ^{13}C ou un ^2H ou encore un ^{15}N etc.... Dans le cas de peptides, et d'une façon générale, plus la masse augmente, plus l'intensité des pics $\text{P}+1$, $\text{P}+2$,..., devient supérieure à celle du pic P et celui-ci sera alors difficilement identifiable dans le massif isotopique (9). Par conséquent, il est préférable de mesurer la masse chimique et non isotopique pour des peptides de masses élevées. D'autre part, lorsque la résolution de l'appareil est faible, les pics isotopiques sont confondus et l'appareil donne un pic correspondant à la **masse chimique** c'est à dire au barycentre des masses de tous les pics isotopiques.

En fonction de la résolution de l'appareil et/ou en fonction de la masse moléculaire du peptide ou de la protéine à analyser, on mesurera tantôt la masse monoisotopique, tantôt la masse chimique.

Un programme informatique, mis au point dans notre laboratoire, nous permet de calculer les masses isotopiques et chimiques des peptides et des protéines. Ce programme nous permet également de:

- calculer l'indice de Bull et Breese (I_{BB}) (cf. page 15), indicateur du caractère hydrophile ou hydrophobe d'un peptide ou d'une protéine,
- prévoir les différents fragments, N et C-terminaux et faciliter ainsi l'interprétation des spectres.

II.5. La calibration:

C'est l'étalonnage de l'appareil qui permet une bonne corrélation entre le temps de détection d'un ion et sa masse. La machine est calibrée en utilisant un produit de référence dont les masses sont connues.

III. La technique FAB-MS

III.1. Introduction:

En technique FAB, l'échantillon est solubilisé dans une matrice de faible tension de vapeur et bombardé par un jet d'atomes de gaz neutre de forte énergie cinétique. Sont employés généralement l'argon (Ar) ou le xénon (Xe) avec une énergie des particules du faisceau primaire allant de 2 à 8 keV (3). Un bombardement par un faisceau d'ions césium d'énergie de 8 à 30 keV est également utilisé. Dans ce cas, on parlera de LSIMS (4).

III.2. Principe de la méthode FAB:

Le produit à analyser est déposé sur un porte-échantillon métallique (en acier inoxydable ou en cuivre généralement) fixé au bout d'une canne dite d'introduction directe. Cette canne est introduite dans la source au travers d'un sas (10). Les atomes d'argon (ou de xénon) bombardent l'échantillon, ce qui aboutit à la formation d'ions. Ces ions sont extraits, accélérés, focalisés et analysés par l'analyseur de masse. En mode FAB, l'analyseur est constitué le plus souvent de secteurs électrostatique et magnétique qui permettent d'atteindre des masses élevées avec une bonne résolution et une bonne sensibilité. Un analyseur de type quadripolaire peut aussi être employé (11-13) mais celui-ci est surtout utilisé en mode Electrospray-Ionspray. Il existe d'autres configurations où sont utilisés, sur un même appareil, des secteurs électrostatique, magnétique et quadripolaire. Un tel appareil est dit hybride (14-15).

Lors du bombardement, des ions négatifs ou positifs ainsi que des fragments neutres sont expulsés (Fig.4). On peut donc choisir de détecter soit en mode positif (16), soit en mode négatif (17), choix qui élargit le champ d'investigation de la spectrométrie de masse. En effet, en fonction de la nature de la molécule étudiée et du but recherché, nous serons amené à choisir un mode plutôt qu'un autre.

En technique LSIMS, on remplace l'argon ou le xénon par du césium (Cs^+) de plus forte énergie (4). Ceci permet, d'une part, d'atteindre une sensibilité 10 à 20 fois supérieure par rapport à la FAB (18) et d'autre part, de désorber des ions plus lourds.

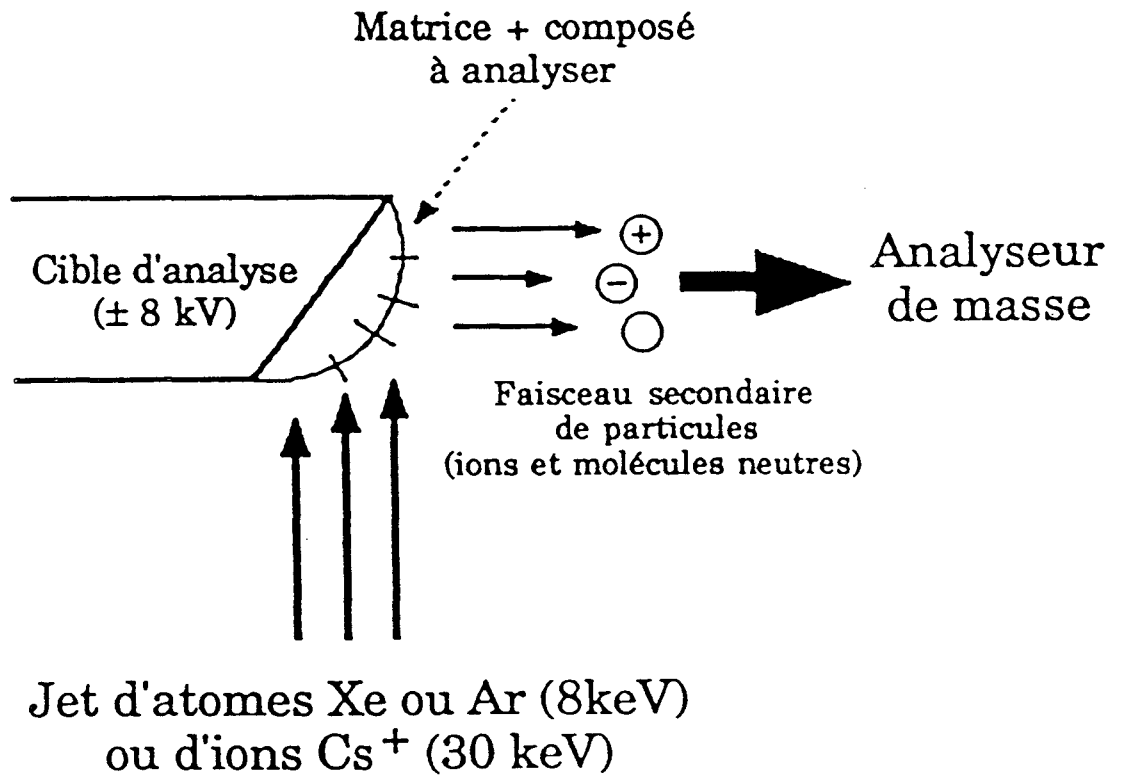


Figure 4: Principe de la désorption-ionisation en FAB et LSIMS

III.3. Rôle de la matrice:

En mode FAB, la matrice liquide utilisée joue un rôle crucial dans le phénomène d'ionisation (19). De manière générale, la matrice utilisée est une solution visqueuse qui doit satisfaire aux conditions suivantes:

- être un bon solvant du composé analysé et lui permettre une bonne diffusion à la surface,
- être peu volatile sous vide pour avoir un signal de longue durée,
- ne pas réagir chimiquement avec les produits analysés,
- avoir un spectre d'intensité suffisamment faible pour ne pas masquer celui du produit étudié,
- ne pas avoir la même masse que celle du produit étudié.

La basicité ou l'acidité relative (20) d'une matrice est essentielle à la production des ions. Ainsi, une meilleure protonation des protéines est obtenue avec le thioglycérol, sa fonction thiol étant acide. La durée d'observation est cependant plus courte qu'avec le glycérol car le thioglycérol est plus volatil sous vide. Le glycérol peut alors être mélangé avec le thioglycérol pour augmenter la durée d'observation des spectres.

Le m-NBA (alcool méta-nitrobenzylique), le mélange dithiothréitol: dithioérythritol 5:1, ou d'autres molécules (19,21) peuvent également servir de matrice. Il est généralement admis que l'utilisation d'un cosolvant comme l'eau, le méthanol ou le chloroforme, augmente l'intensité du signal et sa durée. Par contre, pour les composés anioniques (sulfates, phosphates, glycanes riches en acides sialiques), l'utilisation d'une matrice acceptrice de protons telle que la diéthanolamine amplifiera le signal. Pour accentuer l'ionisation de la molécule, des additifs modifiants le pH tels que l'HCl, le NH₄Cl, ou l'acide paratoluène sulfonique peuvent être ajoutés.

III.4. Mécanisme d'ionisation et nature des ions formés:

III.4.1. Mécanisme d'ionisation:

En mode FAB, le mécanisme de désorption-ionisation reste encore mal connu et sujet à controverse. D'après Sunner et *al* (22), les collisions en cascade résultant de l'impact des atomes ou des ions utilisés lors du bombardement, provoquent l'ionisation dans la matrice et induisent une

transition localisée d'une phase liquide vers une phase gazeuse par augmentation de la température. Les ions de la surface de la matrice (situation amphiphile) se retrouvent dans un état proche de la phase gazeuse. Les réactions ions/molécules qui se produisent sont celles plus généralement observées en phase gazeuse: transfert de protons, transfert d'électrons, recombinaisons, formation de clusters. Seule une faible fraction des ions subsistera dans le vide. Cette transition très rapide (10^{-9} sec.), comparée à une phase d'explosion qui se caractérise par une augmentation brutale de la température, est suivie d'une expansion avec baisse de température. Ce temps très court expliquerait que, même chauffées à forte température, les grosses molécules généralement thermolabiles seraient intactes.

II.4.2. Nature des ions formés:

En mode FAB positive, lors du bombardement, il y a formation de l'ion monochargé $(M+H)^+$ appelé ion pseudomoléculaire. M correspond à la masse de la molécule étudiée. Par contre, en mode négatif, l'ion pseudomoléculaire correspond à $(M-H)^-$. L'ion pseudomoléculaire est en général l'ion le plus intense. Il y a parfois formation également des ions doublement chargés $(M+2H)^{2+}$ ou $(M-2H)^{2-}$, des ions fragments, des ions d'addition avec le solvant ou la matrice $(M+matrice+H)^+$ ou $(M+matrice-H)^-$ ou des ions de cationisation $(M+Na)^+$, $(M+K)^+$, $(M+Na+K-H)^+$, $(M+2Na-H)^+$ etc.... Les ions de cationisation en faible quantité, permettent de confirmer la nature de l'ion pseudomoléculaire. Par contre, la présence de sels, en grande quantité dans le milieu, inhibe la réponse lors de l'analyse peptidique. On peut même avoir une disparition complète du pic $(M+H)^+$. Seuls les clusters de matrice additionnés de sels sont alors observés. Pour éviter ce phénomène, il est conseillé de conserver les peptides dans des tubes eppendorf et non dans des tubes en verre. Il a été observé que l'addition de tétraméthylurée dans la matrice diminuait les effets de cationisation. D'autres composés tels que les agents de dénaturation des protéines peuvent avoir un effet supprimeur de spectre (23) c'est pourquoi il faut absolument éviter leur présence dans le milieu lors de l'analyse en masse.

III.5. Phénomène de suppression:

Lors d'une analyse en FAB d'un mélange de peptides, l'intensité des signaux des pics pseudomoléculaires est fonction directe de l'hydrophilie de ces peptides (24-26). Un peptide de caractère hydrophobe va avoir une activité de surface plus importante qu'un peptide de caractère hydrophile et aura donc une probabilité plus grande d'occuper la surface de la matrice que le peptide hydrophile (24). Par conséquent, il sera plus facilement décapé de la surface lors du bombardement. Le phénomène est également dépendant de la quantité de peptide présente dans la matrice. En effet, il existe une compétition pour l'occupation de la surface de la matrice et par conséquent, plus il y a de molécules hydrophobes, plus le phénomène de suppression est important (27). L'intensité des pics pseudomoléculaires $(M+H)^+$ en mode FAB n'est donc pas proportionnelle à la quantité des peptides présents dans un mélange. La proportionnalité peut néanmoins être discutée dans le cas de mélanges de peptides de caractère identique (28).

Afin d'augmenter l'activité de surface des peptides hydrophiles, il est possible de les dériver en leur fixant des groupements plus hydrophobes. Une estérification des fonctions acides, par exemple, (29) augmente l'intensité du pic pseudomoléculaire. De la même façon, nous avons montré que l'intensité du signal d'un peptide hydrophile peut être amplifiée en augmentant le rapport concentration de peptide hydrophile/concentration de peptide hydrophobe (27). Généralement, pour les peptides ou les protéines de caractère très hydrophile, la concentration doit être relativement élevée (20 à 30 $\mu\text{g}/\mu\text{l}$) (30).

En 1974, Bull et Breese (31) ont calculé, pour les acides aminés, des indices basés sur l'enthalpie libre de passage de l'acide aminé de la solution à la surface. Le résidu de glycine est pris comme acide aminé de référence avec une enthalpie libre ΔF de 810 cal/mole et un indice nul. **L'indice de Bull et Breese** attribué au peptide étudié correspond à la moyenne de tous les indices des acides aminés qui le constituent. Un **indice positif** indiquera un **caractère hydrophile** du peptide tandis qu'un **indice négatif** indiquera un **caractère hydrophobe**. Cet indice permet une meilleure prévision de la réponse en FAB.

Le phénomène de suppression est un des handicaps de la technique FAB. Il est moins important lors de l'analyse en technique de "Continuous Flow FAB" (CF-FAB) (25,32).

III.6. Continuous Flow FAB:

Nous avons vu que l'analyse en FAB se fait en déposant l'échantillon (1µl), auquel on ajoute 1µl de matrice, sur une cible fixée au bout d'une canne d'introduction. Dans ce cas, il s'agit de la **FAB statique**. Il existe une autre technique dite de débit continu ou continuous flow-FAB (**CF-FAB**) ou encore **FAB dynamique**. Elle consiste à injecter le produit à analyser dans un capillaire où passe un tampon mélangé avec la matrice (ex: eau/glycérol 90:10). Cette technique a eu beaucoup de succès depuis les travaux de Caprioli (25, 32,33). Elle s'est avérée plus sensible que la FAB statique (32-34). Elle permet de réduire le bruit de fond (32,33) ainsi que le phénomène de suppression (32,35). Elle a été utilisée avec succès pour les peptides (32, 36-43), les glycopeptides (40,44) ainsi que pour les oligonucléotides (34) et les acides biliaires (45). Des résultats intéressants ont été également obtenus lors de l'analyse en MS/MS (42, 46-48).

On peut coupler, à cette technique, la chromatographie liquide haute performance (32,34,37,40,45) ou l'électrophorèse capillaire (38).

Moseley et ses collaborateurs (49,50) ont couplé la chromatographie liquide à la spectrométrie de masse par un système dit coaxial, où la matrice et le tampon passent séparément dans 2 capillaires couplés. Le mélange se fait à la sortie des capillaires au niveau de la source. Ce système a l'avantage, par rapport à la CF-FAB classique, de permettre une très bonne séparation chromatographique en évitant l'apport de la matrice au niveau du système CLHP et l'utilisation d'un faible débit de la phase mobile. Ce faible débit réduit les problèmes de couplage au spectromètre de masse d'une part et permet l'analyse de très faibles quantités de l'échantillon d'autre part. Moseley et *al* ont ainsi détecté un tripeptide à 500 attomoles (49).

Cependant, la réalisation de ces techniques de FAB dynamique est assez délicate, c'est pourquoi, l'orientation se fait actuellement plus vers l'électrospray, technique plus sensible encore que la FAB en flux continu (40, 51).

III.7. Utilisations et performances de la FAB:

Les performances de la FAB peuvent être résumées de la façon suivante:

-Avec de l'argon ou du xénon (8 keV), la détection des masses est possible jusqu'à 10 000 Da.

-Avec du césium (30 keV), la masse détectée la plus élevée est de 24 000 Da (52).

-La sensibilité est de 500 pmoles. (Pour un peptide de synthèse, notre record est de 4 pmoles (27)).

-En LSIMS, il est possible de détecter quelques centaines de pmoles.

-La précision est d'environ 0,01% (1 unité à 10 000 Da). Couplée à un secteur magnétique, la FAB permet en effet de travailler à une résolution unitaire jusqu'à 6000. La comparaison entre le massif isotopique théorique et celui observé permet de vérifier la pureté du produit. Une désamidation partielle (53) ou une réduction partielle des ponts disulfure peuvent être mises en évidence.

L'analyse directe d'un mélange de peptides issu de l'hydrolyse chimique ou enzymatique d'une protéine (**FAB mapping**) a été largement utilisée (54-59).

La FAB, avec sa méthode d'ionisation douce et avec sa très grande précision, a permis de résoudre beaucoup de problèmes rencontrés dans les études structurales des peptides, que ce soit pour vérifier la séquence d'un peptide naturel ou de synthèse, ou encore pour déterminer et localiser une modification (55,60) post-traductionnelle ou chimique.

Biemann (61), rappelle dans sa revue, l'importance de la FAB. La revue de Burlingame et *al* (62) rassemble la plupart des domaines dans lesquels la spectrométrie de masse en mode FAB peut intervenir.

III.8. Limites de la FAB:

Malgré ses multiples avantages, la FAB reste tout de même limitée:

- par le phénomène de suppression,
- par la présence de sels dans le milieu,
- pour la détection des hauts poids moléculaires,
- pour l'analyse de peptides très hydrophiles ou très hydrophobes,
- pour l'analyse de peptides très hydrophiles et de petites masses (< 300 Da), zone où le bruit de fond dû à la matrice est très important.

IV. La désorption par plasma ^{252}Cf PD-MS

McFarlane et ses collaborateurs (63,64) ont mis en évidence, lors de leurs premiers travaux, que l'impact des produits de fission spontanée du californium 252 avec un film mince d'une substance organique polaire permet d'obtenir des ions positifs et négatifs renfermant divers types d'ions pseudomoléculaires. C'est la méthode de désorption au californium ou " $^{252}\text{Californium Plasma Desorption Mass Spectrometry} : ^{252}\text{Cf PD-MS}$ ".

IV.1. Source au californium:

Le ^{252}Cf est un isotope ayant une demi-vie de 2,65 années. Il se décompose selon deux voies:

- par rayonnement alpha en proportion de 96.9%
- par fission spontanée (3.1%)

Les fissions donnent naissance à des paires de particules chargées de masses et d'énergies inégales dont les plus éloignées en masse sont le ^{142}Ba et le ^{106}Tc (Barium et Technétium)



Les 2 particules de la fission chargées positivement, se repoussent et l'angle entre les trajectoires est d'environ 180° . L'une de ces particules se dirige vers l'échantillon et provoque par collision la désorption et l'ionisation des molécules se trouvant dans la zone d'impact. La seconde particule est utilisée pour déclencher l'acquisition du signal (Fig.5).

Les ions formés sont extraits de la source par accélération sous l'effet d'une tension de 10 à 20 kV. Ils sont ensuite séparés suivant leur temps de vol dans un analyseur dit à temps de vol (cf. page37) d'où le nom de PD-MS-TOF (Time Of Flight).

IV.2. Mécanisme d'ionisation et nature des ions formés:

Comme en mode FAB, le mécanisme intime de la désorption-ionisation en PD-MS n'est pas connu mais il semble dépendre de la masse moléculaire (65,66). Pour des masses supérieures à 5000 Da, il y a désorption de molécules sous forme de clusters qui donnent naissance, après quelques nanosecondes, aux ions pseudomoléculaires $(\text{M}+\text{H})^+$ ou $(\text{M}+\text{nH})^{n+}$. Des ions de cationisation $(\text{M}+\text{Na})^+$, $(\text{M}+\text{K})^+$ ainsi que des ions fragments sont

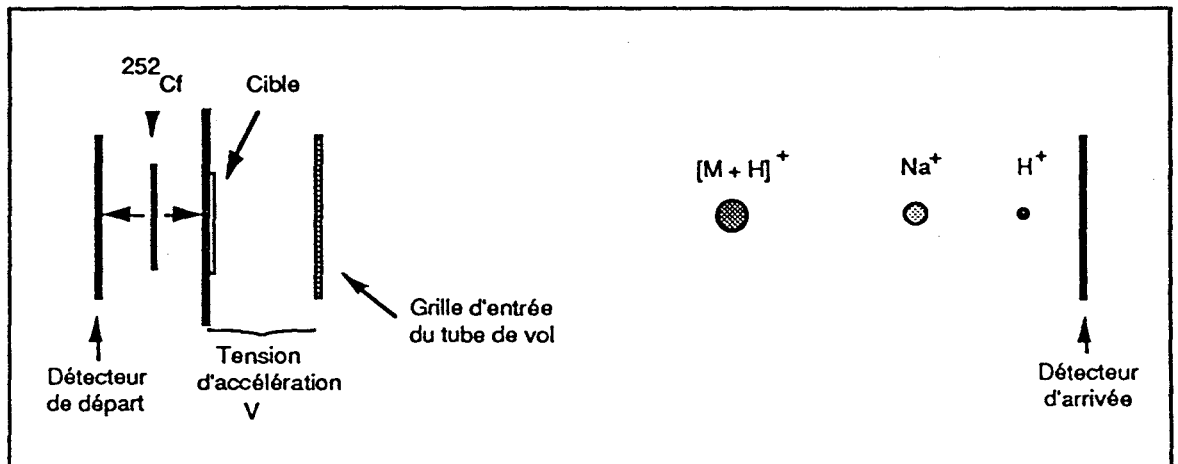


Figure 5: Principe de la désorption-ionisation en ^{252}Cf PD-MS

également formés (67). Ces fragments peuvent jouer un rôle déterminant dans l'élucidation de séquences d'oligopeptides (68,69). On peut choisir de détecter en mode positif ou en mode négatif. La nature de la surface, sur laquelle les molécules à analyser sont déposées, a généralement une très grande importance.

IV.3. Matrice:

En PD-MS, l'échantillon est déposé sous la forme d'un film solide sur une cible placée en face de la source au californium. Les cibles les plus couramment utilisées sont le Nylon, le Naphron, l'Aluminium ou le Mylar aluminisé.

Dans le cas des protéines, l'utilisation de ces surfaces donne un faible rendement de désorption (70). Ce rendement est nettement amélioré quand la cible est recouverte d'une fine couche de nitrocellulose (71,72). En effet, cette couche ou matrice de nitrocellulose subit un échauffement dû à l'impact des produits de fission du californium, ce qui provoque des microexplosions permettant la désorption des peptides et des protéines. Le glutathion a été utilisé avec succès pour l'analyse des protamines (73) et d'autres protéines (74). L'utilisation de la nitrocellulose et du glutathion, plus particulièrement sous sa forme réduite, augmente le rapport signal/bruit, l'intensité des ions multichargés et la résolution des pics (75).

Le mode de dépôt des polypeptides sur la cible a une influence sur le spectre. En PD-MS, il existe deux techniques pour déposer les échantillons sur la cible:

- l'"electrospray" (76) ou
- le "spin-drying" (77)

En technique "d'electrospray", le produit à étudier est en solution dans un capillaire. Sous l'action d'une différence de potentiel, la solution est expulsée du capillaire sous forme d'un jet de microgouttelettes qui vont s'évaporer et seule la protéine atteint la cible.

La technique de "spin-drying" consiste à déposer une goutte de solution, contenant la protéine à étudier, sur la cible recouverte de nitrocellulose puis à sécher cette goutte en faisant tourner la cible à grande vitesse. Cette technique est plus sensible que celle d'"electrospray" car sous l'action de la vitesse de rotation de la cible, une monocouche de molécules se forme sur la nitrocellulose. De plus, il est possible d'éliminer la plupart

des contaminants de faibles masses comme les sels, qui diminuent le rendement de désorption, en lavant la cible par une solution aqueuse.

IV.4. Utilisations et performances de la PD-MS:

Cette technique est généralement utilisée dans le domaine des peptides et des protéines pour la détermination de la masse moléculaire. Ce poids moléculaire est obtenu avec beaucoup plus de précision que par les techniques couramment utilisées dans la chimie des protéines comme l'électrophorèse en gel de polyacrylamide.

Le record de masse pour une protéine est celui de la pepsine porcine 34,6 kDa (78).

Cette technique a l'avantage de consommer très peu d'échantillon au cours de la mesure. Il est également possible de récupérer l'échantillon en solubilisant la nitrocellulose, ou de réaliser directement sur la cible des réactions chimiques ou enzymatiques. Klarskov et *al.* ont ainsi déterminé des séquences peptidiques par hydrolyse par des carboxypeptidases (79).

L'analyse d'hydrolysat de peptides (**PD mapping**) est également possible (80-83). Mais, comme en FAB, l'étude de peptides en mélange peut poser quelques problèmes dus au phénomène de suppression. En PD-MS, la suppression n'est pas due à l'hydrophobicité relative du peptide, mais elle est reliée à sa charge relative (84). En mode positif, le signal des peptides de charge négative nette est partiellement ou complètement supprimé en présence de peptides de charge positive nette. La situation est inverse en mode négatif. Cette suppression disparaît presque totalement après dérivation des peptides (estérification par exemple) (85). Schmitter décrit dans son article les performances et les limites de la PD-MS dans la détermination de la structure primaire des protéines (85).

La PD-MS est utilisée avec succès pour l'analyse d'oligosaccharides (86), de glycopeptides (87), de glycolipides (88), de peptides naturels ou de synthèse (89,90), de peptides recombinants (91), d'antibiotiques aminoglycosylés (92) ou encore de désoxyribonucléotides (93).

La précision est d'environ 0,1% jusqu'à la masse 10 000 Da (70,72,94). Pour des masses supérieures, elle varie de 1 à 5% (70,94). La sensibilité est de l'ordre de quelques dizaines de picomoles.

IV.5. Limites de la PD-MS:

La PD-MS est limitée par:

- le phénomène de suppression,
- la présence de sels dans le milieu,
- le manque de résolution. Elle varie de 200 à 500 pour les appareils classiques. Elle est améliorée par l'utilisation d'un réflectron ou miroir électrostatique qui compense la différence de temps des trajectoires des ions,
- la durée de l'acquisition des spectres (parfois jusqu'à 6 à 9h d'accumulation pour des masses > 10 000 Da),
- l'impossibilité de couplage à la chromatographie liquide,
- l'impossibilité d'analyser les hauts poids moléculaires.

Toutefois, le remplacement de la source au californium par une source de désorption laser permet d'accéder aux hautes masses.

V. La désorption laser

V.1. Description de la technique:

En 1987, Karas et *al.* ont développé la technique de désorption U.V. Laser assistée par une matrice (LD-MS) (95) ("Matrix Assisted Laser Desorption/Ionization" ou MALDI).

Des lasers émettant dans l'ultraviolet (U.V) ou l'infra-rouge (I.R) sont utilisés pour désorber et ioniser des molécules enrobées dans une matrice. Le mélange protéine-matrice forme, lors du séchage de l'échantillon sur la cible, une couche monocristalline. La matrice a une double fonction: absorber l'énergie du laser et isoler les molécules de l'échantillon les unes des autres pour éviter la formation de complexes de taille trop importante. Le rapport molaire matrice:échantillon sera compris entre 100:1 et 5000:1 (96). Les matrices couramment utilisées sont: l'acide nicotinique (6) et l'acide sinapinique (97). Dans l'analyse de protéines après transfert sur membrane de PVDF, Eckerskorn et *al.*(98) utilisent l'acide malique comme matrice.

Le laser produit un bref rayonnement U.V. pulsé de 5 à 20 ns qui va irradier l'échantillon. L'énergie déposée par ce rayonnement est absorbée par la matrice. Elle est ensuite restituée en partie et transférée à l'échantillon qui sera alors volatilisé et ionisé par des processus photochimiques. Il y a, en général, formation d'ions pseudomoléculaires mono et doublement chargés. Les lasers à U.V. sont généralement à azote ($\lambda = 337\text{nm}$), ceux à I.R. à CO_2 ($\lambda = 10,6\text{ }\mu\text{m}$). Les avantages du laser U.V. sont de focaliser sur une plus petite surface et donc de favoriser la désorption des molécules.

L'analyseur utilisé est de type temps de vol comme pour la source au californium. (cf. Fig.5 page19) (96).

V.2. Utilisations et performances de la LD-MS:

La technique de désorption par le laser est facile d'utilisation. Elle permet l'analyse de protéines de haut poids moléculaire (98-100), d'oligosaccharides (101), de glycoprotéines (102), d'oligonucléotides (103) ou encore de peptides de synthèse pour le contrôle de leur pureté (104,105). Les protéines riches en glycannes sont souvent réfractaires à l'analyse en spectrométrie de masse. De telles protéines donnent cependant un spectre en désorption laser. Les pics sont très larges et reflètent l'hétérogénéité de l'échantillon. La résolution généralement de 300, est insuffisante pour

mesurer individuellement les différentes molécules. Cependant, la différence entre la moyenne déduite des centroïdes des pics et la masse peptidique calculée d'après la séquence, donne une masse moyenne des glycanes fixés sur l'axe peptidique (96). L'analyse d'hydrolysats de peptides (**LD mapping**) est souvent proposée dans la littérature (106, 194,195). Beavis et Chait (107) ont montré la sensibilité de cette technique pour l'analyse de mélange de protéines. L'acquisition des spectres est rapide par rapport à la PD-MS et leur interprétation est facilitée par l'absence de fragments.

La LD-MS permet également l'analyse de petites molécules. En effet, elle s'est avérée plus sensible que la FAB pour l'analyse de mélange d'acides aminés ou de tripeptides (108). L'analyse de ces molécules est faite avec un excès de NaCl. Les pics sont observés à $(M-H+2Na)^+$.

L'électrophorèse capillaire peut être couplée à la LD-MS soit directement, soit après récupération de l'échantillon dans un tube de polypropylène. Dans ce cas, l'échantillon est ensuite mélangé avec la matrice et déposé sur la cible. Des résultats intéressants sont obtenus avec des sensibilités de l'ordre de $1/10^{\text{ème}}$ de picomole (109,110).

L'adaptation de la transformée de Fourier et l'addition du D-ribose ou du D-fructose (comme co-matrice) à la matrice (acide 2,5-dihydroxybenzoïque DHB ou l'acide nicotinique) ont permis l'analyse de protéines (masses 5000 - 12 000 Da) avec une très bonne précision, de l'ordre de 0,001%. (111).

La comatrice, relativement volatile en se décomposant en eau et CO_2 , crée localement un environnement gazeux favorable à l'ionisation et la désorption des molécules (111). Les proportions idéales de mélange échantillon/matrice/comatrice sont 1/5000/5000.

Par rapport à la FAB et à la PD-MS, la technique de désorption par le laser a l'avantage :

- de permettre l'analyse de protéines de haut poids moléculaire (>250 kDa),
- de permettre l'analyse de molécules complexes (glycoprotéines etc...),
- d'être plus simple d'emploi,
- d'être moins sensible à la présence de sels dans le milieu. Cette technique a, en effet, l'avantage d'être insensible aux tampons et aux sels aux concentrations généralement utilisées en biochimie. L'utilisation de

l'acide 2,5 dihydroxybenzoïque comme matrice permet de tolérer une contamination par le SDS de l'ordre de 0.1g/ml,

-d'analyser plus facilement des mélanges de peptides.

V.3. Limites de la LD-MS:

Cette technique par désorption laser est limitée par:

-la présence d'adduits constitués de fragments de matrice photochimiquement générés et fixés à la protéine. La présence de ces adduits est préjudiciable à la résolution et responsable des erreurs de mesure. L'intensité de ces adduits, dépendant de la matrice utilisée, est plus intense dans le cas de l'utilisation de l'acide nicotinique.

-sa faible résolution,

-son inadaptation au couplage à la chromatographie liquide,

-l'absence d'ions fragments porteurs d'informations structurales,

-le bruit de fond important dû à la matrice (jusqu'à m/z 500 pour l'acide nicotinique et m/z 1000 pour l'acide sinapinique).

VI. Electrospray-Ionspray

VI.1. Introduction:

Le développement des techniques de spectrométrie de masse dans le but de gagner en sensibilité et d'augmenter en masse a permis à la technique d'électrospray de s'imposer dans le domaine des peptides et des protéines. Connue depuis les expériences de Dole et *al.* (112), l'électrospray ne s'est véritablement développée que depuis cinq ans grâce aux travaux de Fenn et *al.* (113) et Covey et *al.* (114).

La particularité des techniques de spectrométrie de masse en mode électrospray (ES-MS) ou ionspray (IS-MS) réside dans l'utilisation de sources d'ions à la pression atmosphérique (Atmospheric Pressure Ionisation ou API).

VI.2. Principe:

L'injection dans la source, de la solution contenant le produit à étudier, se fait par un capillaire d'environ 0.1 mm de diamètre avec un débit constant réglé entre 1 et 20 $\mu\text{l}/\text{min}$. Une différence de potentiel élevée (4 à 8 kV), appliquée entre le capillaire et la chambre de nébulisation, provoque la formation d'un faisceau de gouttelettes (ou "spray") très chargées. Selon la polarité du capillaire, leur charge est positive ou négative. Le solvant des gouttelettes est évaporé par le passage d'un flux d'azote maintenu à une température de 60°C. Les ions formés sont alors introduits par un orifice dans l'analyseur. La formation des gouttelettes ainsi que leur évaporation peuvent être assistées par un jet d'azote coaxial ou par un léger chauffage de la chambre de nébulisation (115). L'assistance est dite pneumatique dans le premier cas et thermique dans le second. Ces deux types d'assistances permettent de travailler avec des débits plus forts allant jusqu'à 200 $\mu\text{l}/\text{min}$. Dans le cas de l'assistance pneumatique, on parle d'ionspray (114). La figure 6 montre les schémas des sources développées par la société VG BioTech (A) et par la société Perkin-Elmer Sciex (B).

Plusieurs mécanismes sont avancés pour expliquer la formation d'ions à partir de gouttelettes très chargées. Les gouttelettes explosent à un rayon critique (ou limite de Rayleigh), au moment où les forces de répulsion coulombiennes sont du même ordre que la tension de surface (forces

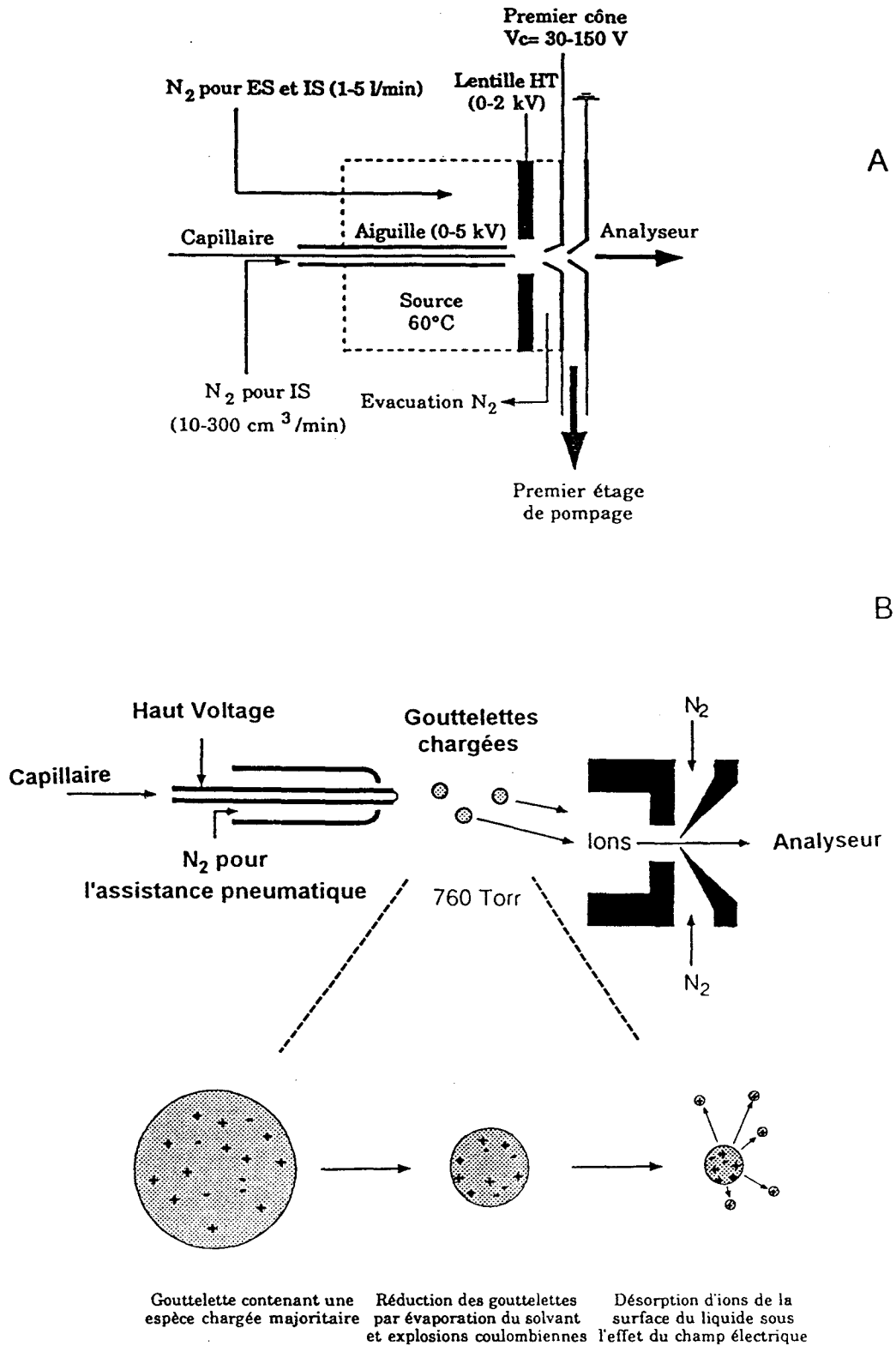


Figure 6: Schémas des sources électrospray-ion spray développées par VG Bio Tech (A) et Perkin-Elmer Sciex (B)

cohésives) (Fig. 6). Cette séquence d'événements se répète jusqu'à ce que la gouttelette ait atteint la taille pour laquelle le champ électrique dû aux charges de la surface soit suffisamment important pour désorber les ions de la gouttelette dans la phase gazeuse. Cette dernière étape, basée sur la désorption d'ions sous l'effet d'un champ électrique présent dans les gouttelettes multichargées ("évaporation d'ions"), dépend de son énergie de solvation (modèle de Iribarne et Thomson (116,117)). L'évaporation d'ions est un phénomène sélectif, puisque l'énergie de solvation augmente de façon quadratique avec le nombre de charges et de façon inversement proportionnelle au rayon de la goutte.

Roellgen et *al.* (118) contestent ce modèle en montrant que le champ électrique atteint la valeur critique de Rayleigh avant d'atteindre la valeur nécessaire à la désorption. Ils proposent simplement que de très petites gouttelettes, contenant chacune une seule molécule, sont éjectées durant, ou peu avant, l'explosion de Coulomb.

Fenn et *al.* ont montré que ce modèle n'était pas applicable à des grosses molécules polychargées (115) et proposent un modèle dit "actif", conciliant les mécanismes précédemment cités. La formation des gouttelettes se fait par une succession d'explosions coulombiennes, au cours desquelles, au moment où le diamètre est faible ($< 1\mu\text{m}$), il y a éjection de fines gouttelettes contenant chacune une seule molécule.

VI.3. Les facteurs influençant l'état de charge:

Le débit, la conductivité et la tension de surface de la solution contenant la molécule à étudier, sont des paramètres importants (119). La sensibilité et la stabilité du signal diminuent avec l'augmentation du pourcentage d'eau dans le mélange eau/méthanol. Cette perte de sensibilité (facteur 5 entre 100% d'eau et 100% de méthanol) est attribuée au plus faible rendement de désorption des ions des gouttelettes d'eau par rapport au gouttelettes de méthanol (120). Cette observation est en accord avec le modèle "évaporation d'ions" d'Iribarne et Thomson.

Dans le but de favoriser la protonation, il est préférable d'abaisser le pH par addition d'une certaine quantité d'acide. Selon Hiraoka et Kudaka (121), le choix de l'acide et de sa concentration optimum sont fonction de l'échantillon. Ainsi, dans le cas d'un oligopeptide Arg-Arg-Arg, des trois acides (formique, acétique et trifluoroacétique) étudiés, l'acide acétique est

le plus approprié car il est possible de l'utiliser dans une très grande plage de concentration (10^{-6} - 4M). Le mélange acide acétique à 5% dans l'eau-méthanol 50:50, permettant une solubilité, une ionisation et une désorption optimale de l'échantillon, est généralement préconisé.

VI.4. L'état de charge des ions:

Le principal avantage de ces techniques (ES-MS et IS-MS), est la formation d'ions multichargés $(M+nH)^{n+}$ en mode positif ou $(M-nH)^{n-}$ en mode négatif. En effet, comme le spectromètre de masse mesure les rapports m/z , une protéine de masse 30 000 Da et chargée 30 fois, par exemple, apparaîtra dans le spectre à une valeur de 1001 uma ce qui est à la portée d'un analyseur de type quadripolaire. On peut également observer des ions de cationisation $(M+aH+bX)^{(a+b)+}$ avec $X = Na, K, \dots$. Mais ils sont de faible intensité sur le spectre. Les molécules analysées sont parfois accompagnées d'adduits non covalents, qui se manifestent sur le spectre par la présence de pics correspondant à la masse attendue majorée de celle de l'adduit (souvent acide phosphorique H_3PO_4 ou sulfurique H_2SO_4 de masse 98 Da). Ces adduits réduisent l'intensité du signal de la molécule étudiée et compliquent le spectre de masse. La fixation de ces adduits se fait sur les sites basiques mais non protonés (122). L'élimination totale ou partielle de ces adduits se fait en augmentant la tension de l'orifice ou du cône d'extraction.

Ces techniques de nébulisation sont considérées douces dans les conditions normales d'utilisation. En effet, on n'observe pas d'ions fragments. Néanmoins, la fragmentation peut être provoquée par une augmentation de la tension de l'orifice (123)(cf. chapitre MS/MS page 41).

Smith et *al.* (119) ont montré une relation quasi-linéaire entre la masse moléculaire et le maximum de charges observées sur les polypeptides et les protéines. Pour les protéines de masse moléculaire inférieure à 40 000 Da et ne contenant pas de ponts disulfure, une corrélation similaire existe entre le nombre total de résidus basiques (lysine, arginine, histidine) et le nombre maximum de charges positives (124). Dans certains cas, le nombre maximum de charges observées est supérieur à celui attendu. Cela voudrait dire que d'autres acides aminés sont des sites potentiels de protonation. D'après l'étude de Bojensen sur l'affinité protonique en phase gazeuse des acides aminés (125), il s'agit très probablement d'une protonation des

résidus de glutamine, qui ont une affinité pour les protons similaire aux résidus de lysine (124). Le plus souvent, le nombre maximum de charges observées est inférieur au nombre maximum de charges possibles. En effet, nous avons vu que lors de l'évaporation du solvant, il y a augmentation du champ électrique et de la charge par ion à la surface de la gouttelette. La vitesse de désorption augmente, et l'ion en passant moins de temps dans la gouttelette n'atteindra pas son état de charge maximum. De plus, la présence de ponts disulfure dans les molécules réduit l'accessibilité aux sites protonables et, en diminuant la distance entre les charges, augmente le nombre de répulsions électrostatiques. Après réduction des ponts disulfure, le nombre maximum de charges observées se rapproche du nombre maximum de charges prévues. Dans le cas de protéines de taille importante comme l'albumine bovine (66600 Da), après réduction, le nombre de charges observées est toujours inférieur au nombre théorique (124). Une protonation maximale, pour ce type de protéine, devient impossible à cause des répulsions électrostatiques. De plus, leur formation nécessite une densité de charges trop importante à la surface des gouttelettes (115). Ces phénomènes pourraient également expliquer le comportement des protamines. Pour ces protéines très basiques, pouvant contenir jusqu'à 65% de résidus d'arginine, le nombre maximum de charges observées est de 50% par rapport au nombre prévu (28,126).

Récemment, Loo et *al.* (127) ont décrit les conditions d'utilisation de l'électrospray en mode négatif. Les protéines et peptides, contenant des fonctions acides tels que les groupements carboxyliques, phosphoriques ou sulfoniques, les oligosaccharides, les oligonucléotides, peuvent être analysés à pH basique (ex: NH_4OH 1%, pH 11.3).

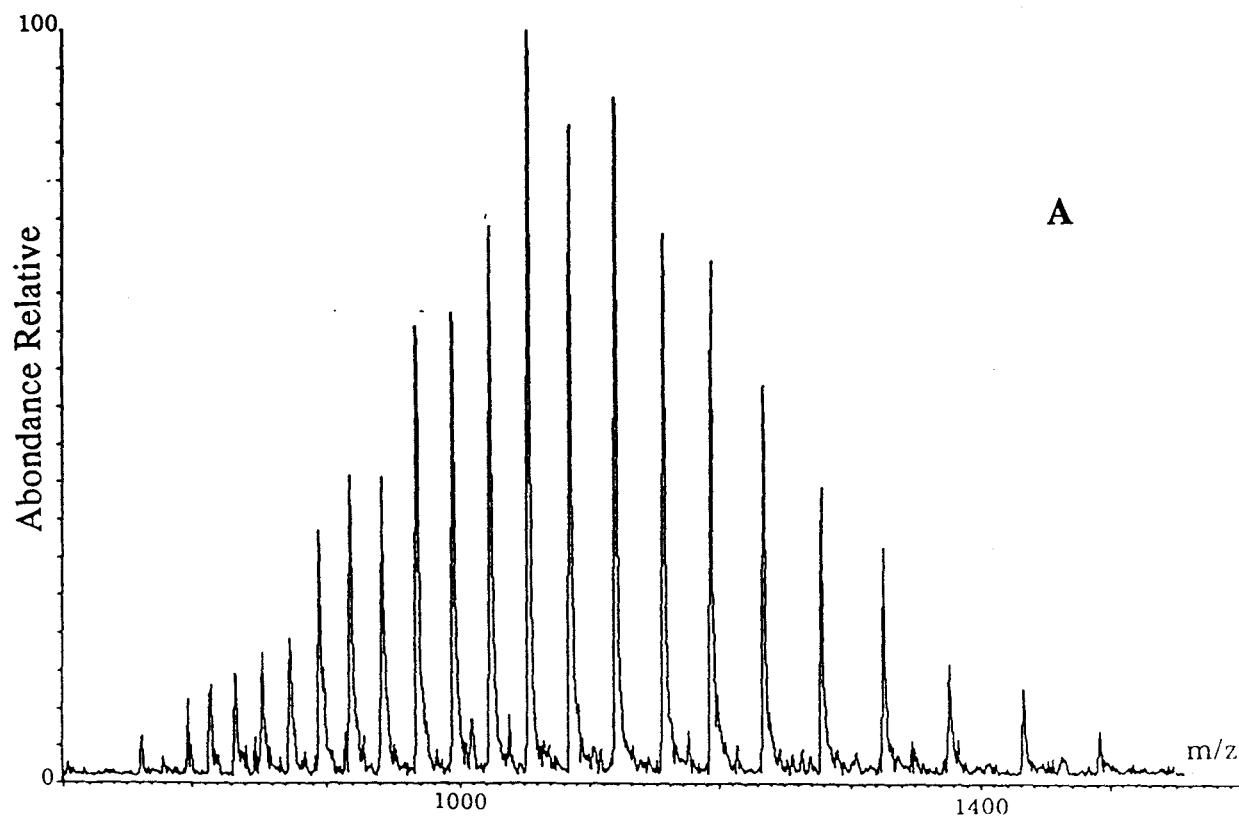
VI.5. Présentation d'un spectre obtenu en mode électrospray:

Les spectres obtenus en électrospray sont constitués en général, d'une série de pics correspondant à différents états de charge de la molécule étudiée. (Fig.7-A). Chaque pic correspond à la masse m/z . La masse moléculaire est calculée à partir de deux pics multichargés (protonations) adjacents différents d'un état de charge (Fig.7B).

En effet, si:

$$m_1 = M + n_1/n_1$$

$$m_2 = M + n_2/n_2$$



B

$$\begin{aligned}
 \text{si: } n_2 &= n_1 + 1 \\
 \text{alors: } n_2 &= \frac{m_1 - 1}{m_2 - m_1} \\
 \text{et: } M &= n_2 \cdot (m_2 - 1)
 \end{aligned}$$

$$\begin{aligned}
 m_1 &= \frac{M + n_1}{n_1} \\
 m_2 &= \frac{M + n_2}{n_2}
 \end{aligned}$$

$m_2 < m_1$
 $n_1 < n_2$

Figure 7: Exemple d'un spectre obtenu en électrospray (mode positif) (A) et calcul de la masse (B)

et $n_2 = n_1 + 1$

on a donc: $n_2 = (m_1 - 1) / (m_2 - m_1)$

Connaissant les différents états de charge, la masse moléculaire peut alors être calculée pour chaque pic. Un écart-type est ensuite calculé par ordinateur à partir de ces masses moléculaires. Deux types d'algorithmes de calcul interprètent les spectres obtenus. Le premier estime la masse moléculaire à l'aide des formules précédemment décrites, le deuxième appelé algorithme de déconvolution transforme mathématiquement le spectre des différents pics dus aux différentes charges en un seul pic correspondant à l'ion simplement chargé (128). Cet algorithme est complémentaire du premier en particulier dans l'interprétation de mélange.

VI.6. Performances et limites de la technique:

Le nombre d'articles, traitant de l'électrospray depuis les travaux de Fenn et *al.* (7) va croissant (129). Actuellement, l'électrospray s'impose par ses possibilités de mesurer, avec une bonne précision (de l'ordre de 0.01%), de hautes masses moléculaires notamment dans le domaine des protéines (119, 130-132). Feng et Konishi (133) ont déterminé les masses moléculaires (> 150 000 Da) d'un grand nombre de glycoprotéines d'importance immunologique et en particulier celle du composant C4 du Complément (196 863 Da).

La technique est très sensible. Smith et *al.* (119) ont mesuré la masse moléculaire du cytochrome C (12 360 Da) en n'utilisant que 23 femtomoles. Mais en général, les quantités de protéines nécessaires varient de 10 à 200 pmoles (132).

L'électrospray peut être utilisée pour:

- la vérification de la masse de peptides de synthèse (134) ou de protéines recombinantes (132,135),

- la caractérisation et la quantification de neuropeptides (136),

- l'analyse de variants structuraux (28, 137-140),

- l'analyse d'hydrolysats de protéines. Le phénomène de suppression est, dans ce cas, moins important qu'en FAB et PD-MS (141).

La fixation d'ions métalliques comme le cuivre ou le zinc, par des protéines, a été directement observé en ES-MS (196,197). Cette technique joue également un rôle important dans l'identification des modifications post-traductionnelles (198,199).

Le développement de la spectrométrie de masse en mode électrospray (ES) ou en mode ionspray (IS) est également lié à ses possibilités de couplage soit à la chromatographie liquide, soit à l'électrophorèse capillaire. De nombreux auteurs ont proposé des interfaces permettant le couplage des techniques ES et IS avec les méthodes de chromatographie liquide (8,119, 142,143). Dans le cas de l'électrospray proprement dit, il est nécessaire d'utiliser un diviseur de débit ("splitter") à la sortie de la chromatographie liquide haute performance. Par contre, la source ion spray est directement connectable à des microcolonnes CLHP ou colonnes microbore (5-200 μ l/min.). Les développements actuels sont axés vers le couplage de l'électrospray avec la technique d'électrophorèse capillaire. Les résultats publiés sont prometteurs (144-146).

L'absence d'ions fragments dans le spectre de peptides et de protéines, dans les conditions normales d'utilisation et la difficulté de la MS-MS sur un appareil triple quadripoles, peuvent être considérées actuellement comme les limites de cette technique.

VII. Les différents types d'analyseurs utilisés en spectrométrie de masse

Après avoir vu les différents mode d'ionisation, nous allons développer dans ce paragraphe les différents types d'analyseurs utilisés. Nous rappelons, qu'en dehors de certains cas particuliers, chaque technique de spectrométrie de masse utilise un type précis d'analyseur. La distribution peut se résumer ainsi:

- FAB ou LSIMS ---> analyseurs électrostatique et magnétique
- PD-MS ou LD-MS ---> analyseur à temps de vol ou TOF
- ES-MS ou IS-MS ---> analyseur quadripolaire

VII.1. Analyseurs magnétique et électrostatique:

Au cours de notre thèse, nous avons travaillé en technique FAB-MS sur deux spectromètres différents:

-un ZAB-HF à 2 secteurs (VG Analytical Ltd, Manchester, U.K.) à Strasbourg,

-un Kratos 4 secteurs Kratos Concept II HH (Kratos Analytical Ltd, Urmston, Manchester, U.K.) à Lille.

Le ZAB a été conçu à géométrie inverse c'est à dire, que le secteur magnétique précède le secteur électrostatique. Cette configuration permet de faire de la MS-MS en technique MIKES (cf. chap. MS-MS page41) en plus de la FAB classique. Le Kratos 4 secteurs électrostatique-magnétique-électrostatique-magnétique ou E-B-E-B permet de faire de la FAB-MS-MS (spectrométrie de masse en tandem) en plus de la FAB-MS.

Nous avons vu dans le chapitre FAB-MS (cf. page11) que les ions formés dans la source sont accélérés sous une tension V allant de 2 à 8 keV en général. En fin d'accélération, l'énergie cinétique acquise est donc égale à (147):

$$E_c = 1/2.m.v^2 = z.V \quad (1)$$

z = nombre de charges élémentaires,

V = tension d'accélération,

m = masse moléculaire,

v = vitesse

VII.1.1. Secteur magnétique:

Dans un champ magnétique, les ions sont soumis à 2 forces : (Fig.8-A)

-une force magnétique $F_B = B.z.v$

-une force centrifuge de valeur $m.v^2/r_B$, r_B étant le rayon de la trajectoire dans le secteur magnétique.

A l'équilibre, nous avons :

$$B.z.v = m.v^2/r_B \text{ -----> } v = B.z.r_B/m \quad (2)$$

De (1), on a: $m/z = 2V/v^2 \quad (3)$

de (2) on a: $v = B.z.r_B/m$

En remplaçant v par sa valeur dans (3) on arrive à :

$$m/z = B^2 r_B^2 / 2V \quad (4)$$

Cette équation montre que la masse (ou le rapport m/z) dépend uniquement du champ magnétique B (r_B et V étant des constantes). Autrement dit, en faisant un balayage de champ magnétique (ou "scan"), on fera parvenir au détecteur des ions de masses différentes.

VII.1.2. Secteur électrostatique:

Dans un champ électrostatique, des ions de vitesse v sont soumis à une force d'intensité $z.E$ (Fig. 8-B). Pour qu'un ion de masse m parcourt une trajectoire circulaire de rayon r_E , il faut qu'il y ait équilibre entre la force électrostatique et la force centrifuge, c'est à dire:

$$z.E = m.v^2/r_E \text{ -----> } r_E = m.v^2/z.E$$

d'après (1) $E_c = 1/2 m.v^2 = z.V$

$$m.v^2 = 2.z.V \text{ -----> } r_E = 2.V/E$$

Cette formule montre que la trajectoire des ions est indépendante de leur charge et de leur masse. Elle ne dépend que du rapport V/E (tension d'accélération/champ électrostatique).

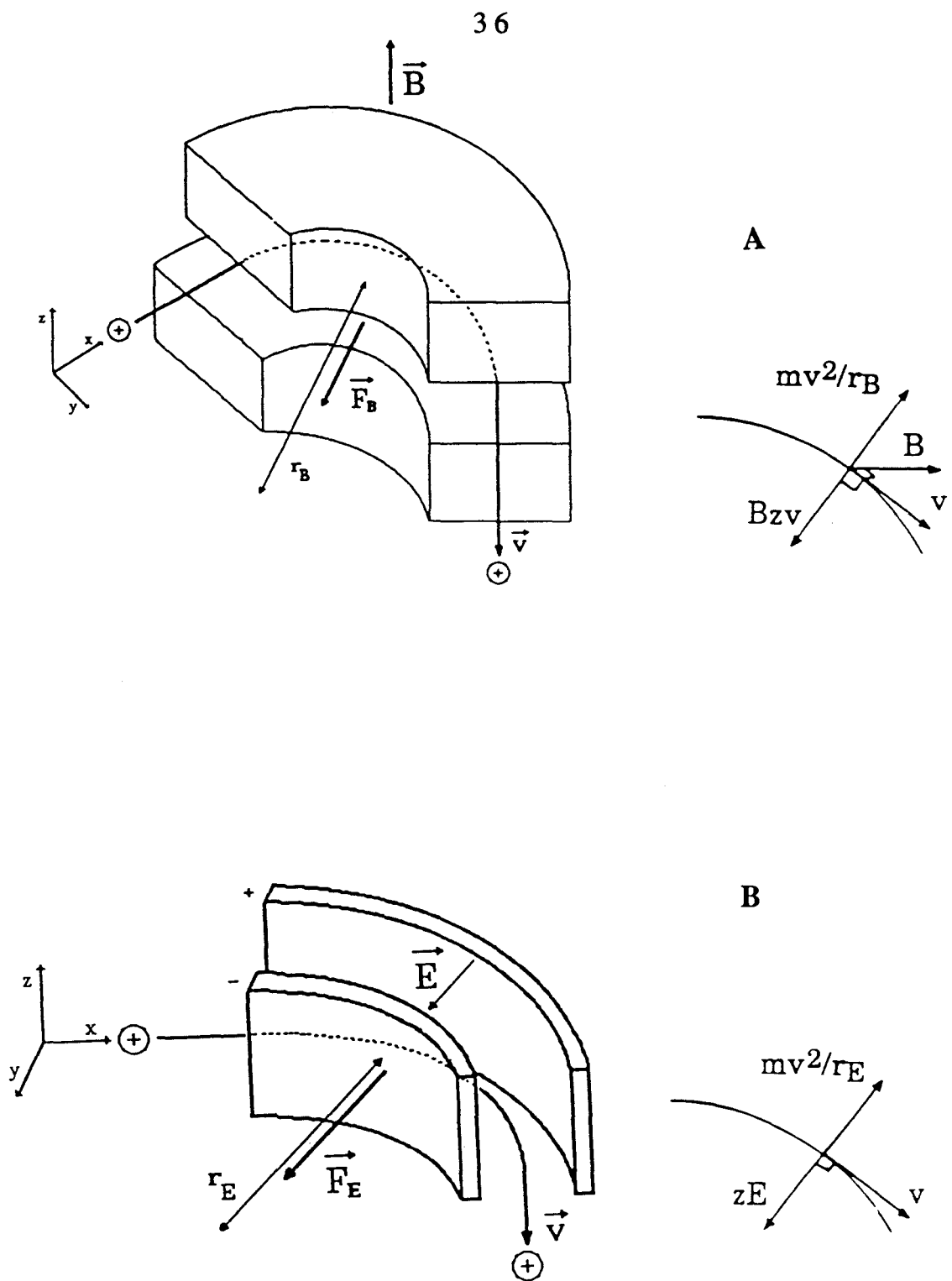


Figure 8: Trajectoire des ions positifs dans un champ magnétique (A) et électrostatique (B)

VII.2. Analyseurs en temps de vol (Time Of Flight):

L'analyseur est constitué d'un tube de "vol" dans lequel ne règne aucun champ électrique. La longueur du tube varie de 15 cm à 1,5 m environ. L'analyse consiste à calculer le temps que mettent les ions pour parcourir un tube de longueur L . Ce temps est comparé au signal de départ déclenché par une des particules des produits de la fission du californium (cf. page 18). Cotter (148) donne une description détaillée de la méthode de mesure.

Nous avons vu que les ions formés dans la source sont accélérés sous une tension V de 10 à 20 kV. Leur énergie cinétique E_c est donc:

$$E_c = 1/2.m.v^2 = z.V \text{ -----> } v = (2.z.V/m)^{1/2}$$

Cette formule montre, qu'à charge égale, plus la masse est grande, plus la vitesse est faible et que pour une masse donnée, plus l'état de charge z est élevé, plus la vitesse des ions sera grande. A l'intérieur du tube les ions se déplaçant d'un mouvement rectiligne uniforme, nous avons l'équation:

$L = t.v$ avec L = longueur du tube et v = vitesse de la particule.

L'équation définissant le temps de vol est donc:

$$t = L.(m/2zv)^{1/2}$$

VII.3. Analyseurs quadripolaires:

Un analyseur (ou filtre) quadripolaire est constitué par un assemblage de 4 électrodes (ou barres) parallèles de section hyperbolique ou cylindrique (Fig. 9). Les électrodes opposées sont au même potentiel. Généralement, le potentiel appliqué aux électrodes est la somme de 2 composantes: une composante continue (U) et une sinusoïdale (V) à haute fréquence (ω). Il est donc de la forme $\phi_0 = U + V \cos \omega t$. Ces électrodes sont connectées électriquement par paires. Ces paires sont de polarités opposées.

Les ions se trouvant dans l'analyseur vont osciller entre les barres. Ainsi, pour des valeurs données de U et V (avec ω fixe), seuls des ions ayant un m/z donné vont traverser le quadripole. Les autres ions vont être absorbés par les barres. Ce qui veut dire qu'en faisant varier U et V (en maintenant leur rapport U/V constant), on fait passer successivement entre les barres les ions ayant des rapports m/z différents (Fig.9).

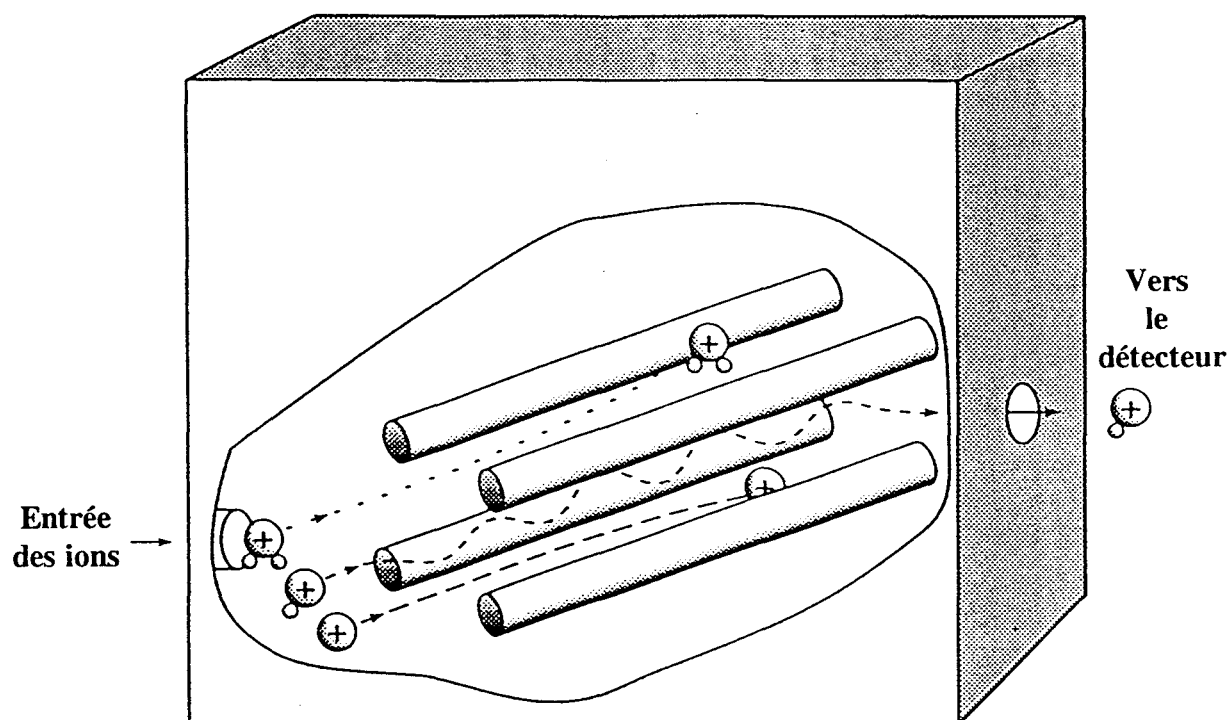


Figure 9: Schéma d'un analyseur quadripolaire

VII.4. Méthodes de visualisation:

Les ions, formés dans la source et séparés dans l'analyseur, arrivent au collecteur où ils sont détectés. La détection des ions consiste en la mise en évidence du signal ionique suite à l'interaction avec un "convertisseur" sensible à l'impact des ions. Le convertisseur est souvent un multiplicateur d'électrons. Son principe est le suivant: l'émission des électrons secondaires a lieu suite à l'impact des ions positifs sur une plaque dite cathode de conversion (ou dynode de conversion). Ces électrons sont accélérés et dirigés à leur tour sur une deuxième électrode de forme spécifique (dynode) ce qui provoque une nouvelle émission électronique et ainsi de suite jusqu'à la dernière dynode. Le courant est collecté par l'anode.

L'acquisition des données se fait par un microprocesseur avec une interface qui permet de transcrire les signaux analogiques sous forme digitale. Les résultats sont stockés dans la mémoire de l'ordinateur pour un éventuel traitement.

VII.5. Etalonnage:

L'étalonnage permet une bonne corrélation entre le moment où est détecté un ion et sa masse. Cet étalonnage peut être externe (fichier d'acquisition distinct de l'analyse) ou interne (pics d'étalonnage et composé à analyser sur le même fichier d'acquisition).

En FAB-MS, le produit de référence le plus souvent utilisé est l'iodure de césium (CsI) aussi bien en mode positif qu'en mode négatif. En effet, l'iodure de césium donne en mode positif des clusters de type $(\text{CsI})_n\text{Cs}^+$, alors qu'en mode négatif, les clusters sont de type $(\text{CsI})_n\text{I}^-$ (10, 16,17). Sont également utilisés, l'iodure de sodium (NaI) dans le glycérol, le polyéthylèneglycol seul ou en solution de 1 à 5% dans le thioglycérol.

En PD-MS ou en LD-MS, un couple d'ions suffit pour calibrer l'analyseur à temps de vol. En ^{252}Cf PD-MS, les couples les plus souvent utilisés sont $(\text{H}^+, \text{Na}^+)$ ou $(\text{H}^+, \text{NO}^+)$ en mode positif et $(\text{H}^-, \text{CN}^-)$ en mode négatif. (54).

En LD-MS, l'insuline porcine (5 778 Da) et la ribonucléase A (13 682 Da) sont généralement utilisées.

En électrospray-ionspray, l'appareil est calibré aux basses masses (< 1000 Da) avec le polyéthylèneglycol (PEG 400 ou 1000). Pour des m/z compris entre 500 et 1500 Da, la myoglobine de coeur de cheval (16 951,5 Da (149)) ou l'ubiquitine (8 564,9 Da) peuvent être utilisées. Pour des masses m/z autour de 2000 Da, on peut calibrer avec le lysozyme, le trypsinogène ou la β -lactoglobuline A.

VIII. La spectrométrie de masse en tandem

La technique MS-MS (150) utilise deux analyseurs en série. Le premier analyseur dit MS1 sert à sélectionner un ion ou **ion parent** et le deuxième analyseur MS2 va permettre la séparation des ions secondaires ou **ions fils** issus de la décomposition de l'ion sélectionné (Fig.10).

L'ion sélectionné dans le premier analyseur va se fragmenter dans une zone sans champ soit d'une façon spontanée, soit aidé par un gaz de collision (He, Ar ou Xe), auquel cas on parlera de CAD ou CID c'est à dire "Collision Activated Decomposition" ou "Collision Induced Dissociation". Les ions fils issus de cette fragmentation sont ensuite séparés et analysés dans le deuxième spectromètre de masse.

La technique MS-MS se pratique soit sur des appareils ayant une source FAB et on parlera de FAB-MS-MS, soit sur des appareils ayant une source électrospray et on parlera de ES-MS-MS.

En technique FAB, il existe deux types de géométrie: une géométrie normale et une géométrie inversée. Dans le premier cas, l'analyseur électrostatique précède l'analyseur magnétique. Pour faire de la MS-MS il est nécessaire dans ce cas d'avoir deux spectromètres en série ou tandem. Par contre, en technique MIKES (Mass Analysed Ion Kinetic Energy Spectrometry) (151), un simple appareil à géométrie inversée est utilisé.

En MIKES, l'analyseur magnétique sélectionne l'ion parent qui se fragmente ensuite en ions fils entre les deux analyseurs dans la région sans champ. Les ions fils sont ensuite analysés dans le secteur électrostatique (152). La technique MIKES a une résolution d'environ 100 et sa précision est également faible (environ 0.5).

En technique électrospray, l'analyse MS-MS est effectuée sur des analyseurs triple quadripoles. Le premier quadripole sélectionne l'ion parent, le second sert de cellule de collision et le troisième analyse les ions fils. Les collisions sont induites à basse énergie (1 à 200 eV) et le rendement de fragmentation est de l'ordre de 90%. Il est supérieur aux rendements observés sur des appareils à secteurs multiples à des énergies de collisions élevées (4 à 8 keV). La technique ES-MS-MS est particulièrement délicate car elle utilise un ion multichargé (153,207).

La fragmentation des peptides se fait surtout au niveau de l'axe peptidique et principalement au niveau de la liaison amide (154-156,204,205).

Roepstorff et Fohlman (154) ont proposé une nomenclature des différents types d'ions de fragmentations peptidiques (Fig.11). Les ions de

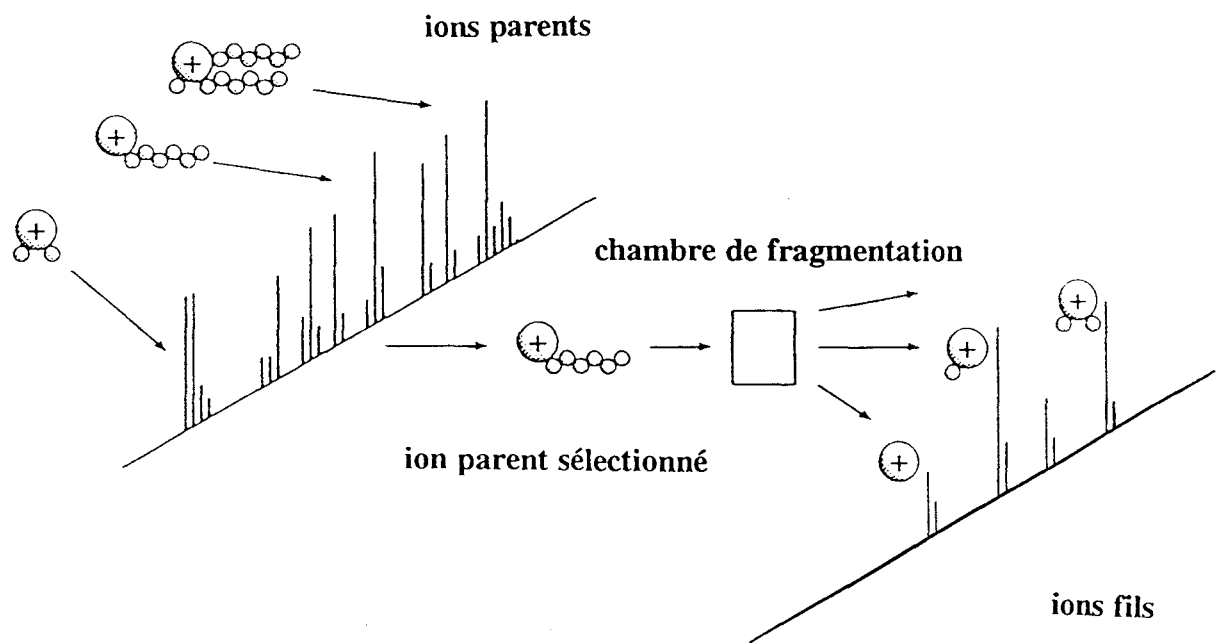


Figure 10: La spectrométrie de masse en tandem ou MS-MS

type A, B et C correspondent aux fragments chargés du côté N-terminal. Les fragments X, Y et Z emportent la charge du côté C-terminal. Des fragmentations au niveau des chaînes latérales sont également observées (157), elles correspondent dans la nomenclature aux fragments D, V et W. C'est cette fragmentation qui a permis de distinguer la leucine de l'isoleucine en MS-MS. (158,203).

L'utilisation d'un couple de lettres signifie l'existence de fragments internes. $(B_3Z_4)_2$, par exemple, signifie l'existence d'un fragment interne contenant 2 résidus et correspondant à une fragmentation en position 3 à partir du N-terminal et en position 4 à partir du C-terminal.

Nous avons vu que, dans les conditions normales d'utilisation, l'ionisation en ES-MS ou en IS-MS est douce et la molécule analysée reste souvent intacte. Il est cependant possible de provoquer sa fragmentation en augmentant la tension électrique d'extraction des ions dans la source (123). En Allemagne, dans le laboratoire de Perkin-Elmer Sciex, nous avons obtenu des résultats très satisfaisants en provoquant la fragmentation d'un médicament, le piroxicam (masse 332 Da).

Actuellement, toutes les techniques de MS-MS sont limitées aux masses inférieures à 1500 Da. En général, le peptide sélectionné doit être d'intensité suffisamment importante pour espérer des résultats dans le deuxième analyseur.

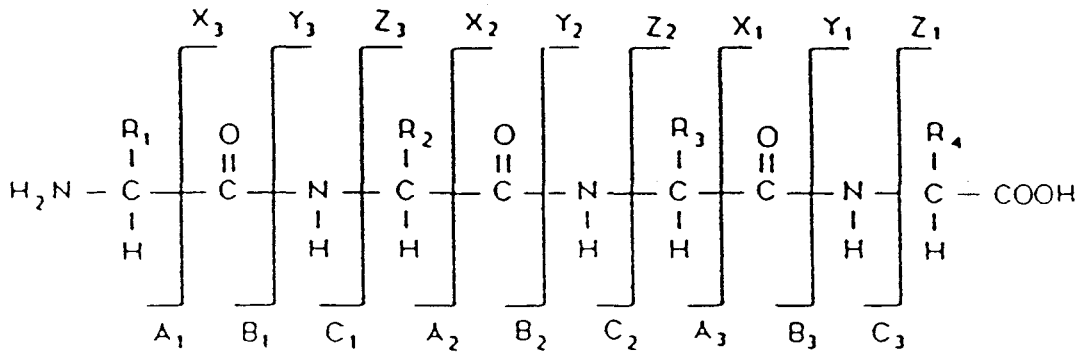
Au cours de notre thèse, les essais que nous avons réalisés en MS-MS ont été un échec à l'exception des peptides de synthèse disponibles en grande quantité et ceci pour deux raisons principales:

- les peptides obtenus à partir de protéines nucléaires basiques sont souvent de caractère très hydrophile et par conséquent, le signal obtenu en FAB est très faible,

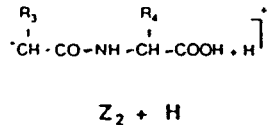
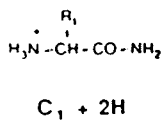
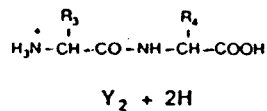
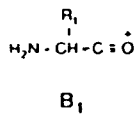
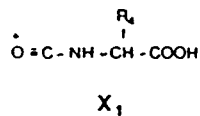
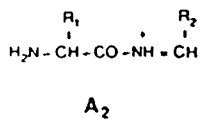
- d'autre part, nous avons toujours disposé de très faible quantité de produit.

Toutefois, les fragmentations spontanées au niveau de la source, observées sur ce type de molécules, ont été déterminantes dans l'élucidation de leur structure primaire (30,53).

Cette technique est particulièrement intéressante car elle permet de déterminer la structure de peptides de masse connue se trouvant en mélange et ceci sans purification préalable (61,202). Bean et Carr (159) l'ont



Fragmentation selon l'axe peptidique

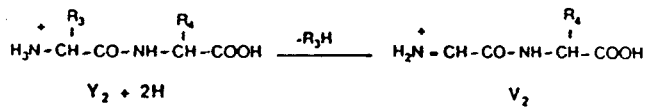


Fragmentation selon la chaîne latérale

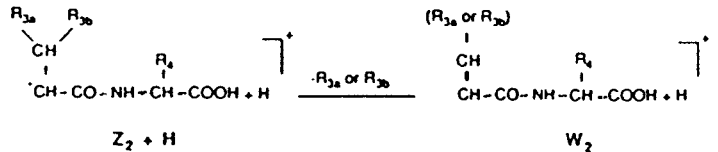
D IONS



Y IONS



W IONS



where R_{3a} = H except for
 Ile: R_{3a} = CH₃, R_{3b} = C₂H₅
 Val: R_{3a} = R_{3b} = CH₃
 Thr: R_{3a} = OH, R_{3b} = CH₃

Figure11: Nomenclature des types de fragmentations de peptides selon Roepstorff et Fohlman (154)

appliquée à l'analyse des ponts disulfure et au positionnement de ces derniers. Elle a été également utilisée dans la vérification de structure de peptide de synthèse (206) ou de structure établie par le séquençage du cDNA (160). Elle a également permis de vérifier la nature de modifications chimiques (189) ou post-traductionnelles (200,201) et de les localiser. L'utilisation de la CF-FAB, en réduisant la quantité de matrice injectée, augmente le rapport signal/bruit. Les performances de la MS-MS s'en trouvent améliorées et permettent l'analyse séquentielle de plusieurs petits peptides en faible quantité (48).

En conclusion, on peut dire que la spectrométrie de masse en tandem est une technique difficilement maîtrisable. Son utilisation n'est réelle que dans très peu de laboratoires. Elle doit faire appel à des programmes informatiques qui permettent la prévision de tous les fragments à partir d'une structure partiellement ou entièrement connue. L'élucidation complète d'une structure inconnue, uniquement par cette technique, est actuellement impossible. Elle nécessite la complémentarité des autres techniques habituellement utilisées dans la détermination d'une structure primaire: dégradation récurrente d'Edman, hydrolyses chimiques et/ou enzymatiques.

Chapitre 2

Rappels sur la spermatogenèse

A partir des années 80, notre laboratoire, qui était spécialisé dans l'étude des protéines nucléaires associées au DNA dans les cellules somatiques, s'est orienté vers l'étude structurale des protéines nucléaires basiques associées à la chromatine au cours de la spermatogenèse dans différents modèles (mammifères, invertébrés, poissons sélaciens).

Dans les différentes espèces animales, la spermatogenèse se caractérise par un ensemble de phénomènes complexes qui conduisent, à partir d'une cellule somatique diploïde, à la formation d'une cellule haploïde ayant une morphologie caractéristique et une fonction hautement spécialisée: le spermatozoïde. Parallèlement aux changements morphologiques de la cellule, le noyau subit de profondes modifications. Ainsi, au cours de la spermatogenèse, la chromatine passe progressivement d'un état diffus, (où elle est transcriptionnellement active), à un état fortement condensé dans le spermatozoïde mûr, et devient transcriptionnellement inactive. Parallèlement, cette transition chromatine active/chromatine inactive est liée à une modification de la nature des protéines basiques associées au DNA. Deux cas de figure sont possibles : soit que les histones de type somatique sont remplacées directement par des protéines hautement basiques et de taille plus petite, appelées protamines, comme cela est le cas chez les poissons téléostéens, soit que ces histones de type somatique sont d'abord remplacées par des protéines dites intermédiaires, présentes de façon transitoire (d'où le nom également de protéines de transition) dans les spermatides, avant de laisser place dans le spermatozoïde mûr aux protamines comme cela se rencontre chez les mammifères (ex: bélier) (Fig.12).

Chez le bélier, les noyaux des spermatocytes et des spermatides renferment les cinq histones de type somatique H1, H2A, H2B, H3 et H4 ainsi que deux histones spécifiques du testicule TH2A et TH2B. Toutes ces protéines disparaissent en phase d'élongation. Une histone particulière H1⁰, dont la présence a été corrélée avec la terminaison de la maturation cellulaire durant le développement ou la régénération (161,162), apparaît en abondance dans les spermatides en voie d'élongation (stade 9-11) dont la chromatine se condense, alors qu'elle est absente dans les noyaux des spermatocytes et des spermatides rondes. Elle disparaît ensuite dès le stade 12 (spermatides allongées). Au cours de la différenciation des spermatides (stade 9-12), quatre protéines spécifiques apparaissent: P1, 3, 7 et T (163,164). Ces protéines, qui sont en fait les protéines TP2, TP3, TP4 et TP1 respectivement, sont de taille moléculaire plus petite que les histones, sont

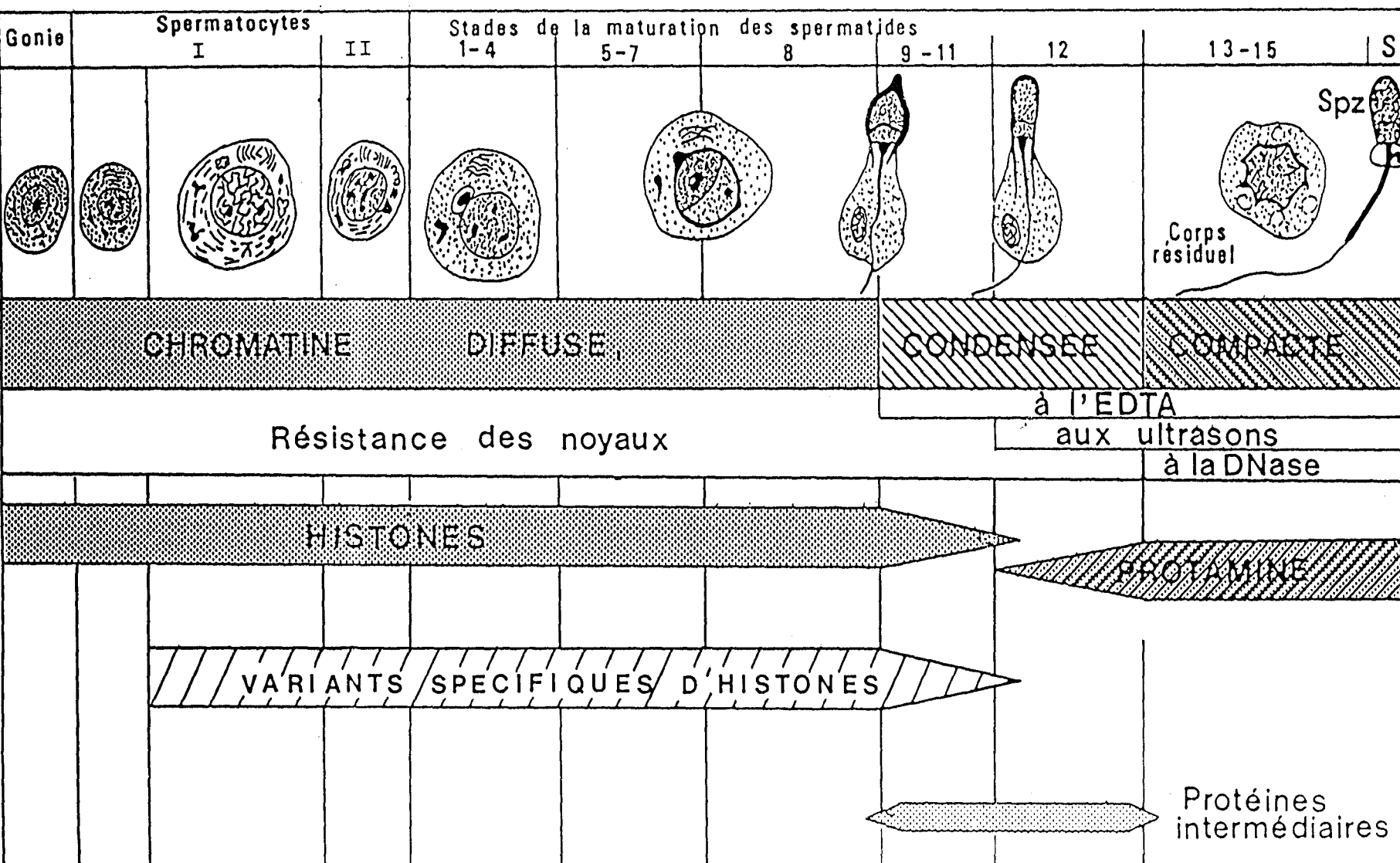


Figure 12 : Les différentes étapes de la spermatogenèse chez le Bélier

beaucoup plus basiques et contrairement aux histones, elles renferment de la cystéine. Enfin, à partir des spermatides allongées (stade 12), les protéines intermédiaires sont remplacées par une seule protamine. La protamine de bœuf est de taille moléculaire plus petite que les protéines intermédiaires, elle renferme également de la cystéine et est beaucoup plus basique.

Ces protéines qui se mettent en place au cours de la spermatogenèse subissent des **modifications post-traductionnelles** (méthylation, acétylation, phosphorylation).

Les histones encore présentes en début de phase d'élongation sont liées de manière plus lâche au DNA. Selon Le Gac-Jegou (165), et en accord avec les résultats obtenus chez la truite (166,167) et chez le rat (168,169), les histones H2A, H2B, H3 et H4 incorporent rapidement l'acétate marqué au ^{14}C . L'analyse électrophorétique à haute résolution confirme pour l'histone H4, une des histones les plus étudiées, qu'elle est beaucoup plus acétylée dans les spermatides en élongation que dans les spermatocytes et les spermatides rondes. Ces observations sont en accord avec le fait établi par ailleurs, que l'acétylation des histones réduit leur interaction avec le DNA (170).

Au niveau des protéines intermédiaires et des protamines, la modification post-traductionnelle la plus connue est la phosphorylation. Ainsi, la protéine TP1 (protéine T d'après Loir et Lanneau (163)) qui existe sous forme de **2 variants structuraux** (28), est présente sous 3 formes: non phosphorylée, mono et diphosphorylée. Cette phosphorylation pourrait intervenir avant et/ou après le dépôt de la protéine TP1 sur le DNA.

Dans le premier cas, la phosphorylation permettrait un dépôt correct des protéines sur le DNA en modulant les interactions DNA-protéine TP1 ou en changeant la conformation de la protéine. Dans le second cas, cette modification pourrait entraîner une dissociation du complexe DNA-protéine TP1 pour la mise en place de la protamine. La protamine qui remplace les protéines intermédiaires apparaît, quant à elle, sous la forme monophosphorylée dans le testicule, puis déphosphorylée dans le spermatozoïde éjaculé. D'après Marushige et Marushige (171,172), la phosphorylation de la protamine pourrait intervenir après le dépôt de celle-ci sur la chromatine. On observe également une légère diminution du point de fusion du DNA complexé avec la protamine phosphorylée. Cette

diminution est d'autant plus importante que le degré de phosphorylation est élevé (173,174).

L'ensemble de ces résultats montre donc que la phosphorylation des molécules de protamines entraîne une diminution des interactions entre le DNA chargé négativement et la protamine chargée positivement.

Plusieurs fonctions biologiques de la phosphorylation ont également été proposées. Elle pourrait être impliquée soit dans le transport des protéines basiques de leur site de synthèse vers le noyau (175), soit dans le positionnement correct des protamines sur le DNA (sur le petit ou le grand sillon) afin de permettre la formation correcte des ponts disulfure (171,172).

En résumé, et de manière générale, que l'on soit dans le cas d'une transition protéique simple (histones ---> protamine (s)), ou double (histones ---> protéines intermédiaires ---> protamine (s)), les protéines qui remplacent les histones se caractérisent par leur **faible diversité en acides aminés** (126) (4 sortes d'acides aminés pour la protamine Sp de seiche (176)), ce qui limite grandement les sites de coupures enzymatiques ou chimiques. Elles renferment parfois des variants structuraux difficilement séparables par les méthodes chromatographiques actuelles (28). Quelques unes d'entre elles ont leur extrémité **NH₂-terminale bloquée** ce qui les rend inaccessibles à la dégradation d'Edman (ex: protéine TP2 de bélier). De plus, elles sont très souvent **présentes en très faible quantité** dans les spermatides (chez le bélier le stade 9-12 représentatif des protéines TP dure environ 4 jours) et c'est une des raisons pour laquelle leur structure primaire n'est souvent établie qu'à partir du cDNA (177-180). Toutefois, comme nous venons de le voir, ces protéines subissent souvent des modifications post-traductionnelles et dans ce cas, la séquence établie par le cDNA ne nous donne aucune information sur ces modifications. Si ces modifications sont de type acétylation ou méthylation (hormis l'acétylation ou la méthylation en N-terminal de la protéine), elles sont détectables par des méthodes classiques de détermination de structure primaire des protéines. Cependant, dans le cas de la phosphorylation, la méthode récurrente d'Edman ne nous renseigne pas sur ces modifications car les anilinothiazolinones phosphates ne sont pas solubles dans le solvant d'extraction utilisé au cours de la dégradation. On peut toutefois détecter la sérine phosphorylée et uniquement celle-ci, si on fait subir à la protéine une modification qui transforme ce résidu en S-éthylcystéine (181). Cependant, cette méthode a l'inconvénient de transformer également les résidus de S-

carboxamidométhylcystéine en S-éthylcystéine ce qui complique l'interprétation des résultats obtenus par dégradation d'Edman.

Dans la résolution de ces différents problèmes structuraux, la spectrométrie de masse a été d'un intérêt capital.

Travaux Personnels

Chapitre 3

Utilisation de la spectrométrie
de masse pour la correction
d'une séquence protéique

Etude de la protéine Z3 de Roussette

I. Introduction:

Nous avons vu qu'au cours de la spermatogenèse, la transition chromatine active-chromatine inactive est liée à une modification de la nature des protéines basiques associées au DNA. Ainsi, les histones sont soit directement remplacées par des protéines très basiques riches en arginine appelées protamines, soit remplacées de façon transitoire par des protéines spécifiques des spermatides, qui sont ensuite elles-mêmes remplacées par des protamines.

La protéine Z3 est une des quatre protamines Z1, Z2, Z3 et S4 associées à l'ADN chez la roussette *Scylliorhinus caniculus*. En effet, le remaniement qui intervient dans les protéines nucléaires basiques au cours de la spermatogenèse de la roussette comporte deux étapes:

- une première étape correspondant à l'élongation des spermatides au cours de laquelle les histones sont remplacées par les protéines de transition S1 et S2 (182),

- une seconde étape correspondant à la maturation des spermatozoïdes et au cours de laquelle les protéines de transition disparaissent pour laisser place aux quatre protamines Z1, Z2, Z3 et S4 (183).

La composition en acides aminés de ces quatre protamines est présentée dans le tableau I.

La protamine Z3 a été étudiée dans notre laboratoire et sa séquence a été élucidée en 1981 (184). Elle renferme 31 résidus et ne comporte que 5 types d'acides aminés: Ala, Ser, Tyr, Gly et Arg.

L'étude des protéines basiques par spectrométrie de masse étant très peu développée, nous nous sommes attaché à étudier la protamine Z3, de structure connue par cette technique et ceci dans le but de mettre au point les conditions de travail, pour ce type de protéine.

Un spectre de masse en mode électrospray a été réalisé en vue de la vérification et de la confirmation de la séquence établie en 1981. La masse moléculaire calculée correspondant à la séquence de Z3 sans modification est de 3979 Da. La masse observée en ES-MS est de 4748 +/- 0,1 Da, soit 769 Da de différence. Nous avons analysé la protéine en mode électrospray sur 3 appareils différents:

Tableau I

Composition en acides aminés des scylliorhines de la roussette *Scylliorhinus caniculus* d'après Gusse et al. (183).

Acides aminés	Z2	Z1	Z3	S4
Asx	-	3.1 (2)*	-	-
Thr	2.1 (1)*	3.5 (2)	-	-
Ser	3.9 (2)	3.7 (2)	10.0 (3)	-
Glx	2.4 (1)	2.5 (1)	-	-
Pro	2.4 (1)	5.0 (3)	-	6.5 (2)
Gly	2.5 (1)	2.5 (1)	19.1 (6)	3.4 (1)
Ala	10.4 (5)	4.6 (1)	3.4 (1)	9.3 (3)
Cys	11.2 (4)	8.9 (4)	-	13.4 (4)
Val	3.9 (2)	3.6 (2)	-	3.2 (1)
Met	1.7 (1)	-	-	-
Ile	1.8 (1)	-	-	-
Leu	2.4 (1)	6.4 (4)	-	-
Tyr	-	1.4 (1)	2.7 (1)	-
Phe	-	1.5 (1)	-	-
His	5.9 (3)	3.8 (2)	-	-
Lys	13.8 (6)	17.4 (8)	-	41.4 (14)
Arg	35.9 (17)	32.1 (16)	64.8 (20)	22.9 (7)
Total	46	50	31	32
% basiques	55.6	53.3	64.8	64.3

* Les résultats sont exprimés en moles/100 moles

Les chiffres entre parenthèses correspondent aux résidus effectivement présents dans la séquence.

- à Strasbourg sur un Bio Q (VG),
- en Allemagne sur un API III (Perkin-Elmer Sciex),
- en Angleterre sur un TSQ 700 (Finnigan Mat).

Nous avons obtenu la même masse dans les trois cas. Cette même masse de 4748 Da a également été obtenue en désorption laser.

Deux hypothèses ont alors été émises:

- la séquence établie est incomplète,
- la protéine s'accompagne de modifications chimiques ou post-traductionnelles qui n'ont pas été détectées par les méthodes classiques.

Le travail que nous avons réalisé, nous a permis de confirmer la première hypothèse. En effet, la protamine Z3 renferme 37 résidus d'acides aminés. L'essentiel des résultats obtenus est présenté dans la publication 1:

A corrected primary structure for dog-fish *Scylliorhinus caniculus* protamine Z3

Kouach M., Jaquinod M., Belaïche D., Sautière P., Van Dorsselaer A., Chevaillier P. and Briand G. (1993)

Biochim. Biophys. Acta 1162, 99-104

BBAPRO 34411

A corrected primary structure for dog-fish *Scylliorhinus caniculus* protamine Z3

Mostafa Kouach ^a, Michel Jaquinod ^b, Denise Belaïche ^a, Pierre Sautière ^a,
Alain van Dorsselaer ^b, Philippe Chevaillier ^c and Gilbert Briand ^a

^a URA 409 CNRS, Université de Lille II, Institut de Recherches sur le Cancer, Lille (France), ^b Laboratoire de Chimie Organique des Substances Naturelles, URA 31 CNRS, Strasbourg (France) and ^c Laboratoire de Biologie Cellulaire, Université Paris-Val-de-Marne, Créteil (France)

(Received 14 July 1992)

(Revised manuscript received 28 September 1992)

Key words: Protamine; α -Chymotrypsin; Trypsin; Primary structure; Mass spectrometry; (*S. caniculus*)

We have redetermined the primary structure for dog-fish protamine using automated amino-acid sequencing associated to mass spectrometry techniques and report, on the basis of these findings, that the previously published amino-acid sequence is incorrect. The correct protamine sequence is 37 amino acids long and differs from the original published sequence by the C-terminal hexapeptide Arg³²-Gly-Arg-Arg-Ser-Arg³⁷.

Introduction

Spermatogenesis involves a progressive differentiation of the germ cells towards a terminal stage represented by the mature spermatozoa. During spermatogenesis, the histones are replaced by more basic proteins: in early spermatids of the dog-fish, histones are lost and replaced by two intermediary proteins which are themselves eliminated and replaced by four protamines in late spermatids [1,2]. These protamines were called protamines Z1, Z2, Z3 and S4 and their primary structure was determined [3–6]. At the end of the process, the DNA-protamine complex has acquired a highly compact and transcriptionally inactive form. These protamines are characterized by a very high arginine content (more than 50%).

Conventional methodologies, such as Edman degradation and amino-acid analysis, have serious limitations with regard to characterizing structurally basic proteins. During Edman degradation, the repetitive yield of the phenylthiohydantoin of arginine falls and background rises with each degradation when arginines are consecutive. Amino-acid analysis yields poor infor-

mation because of the disproportion between the arginines and the other residues.

Mass spectrometry, based on precise measurement on the mass-to-charge ratio, is now considered to be a significant aid in protein structure determination. In the past several years, new ionization technologies have been developed: Fast atom bombardment (FAB) [7–9], ²⁵²Cf plasma desorption (PD) [10], matrix-assisted laser desorption (LD) [11–14] and electrospray ionization (ESI) [15–17]. These methods show extremely high ionization efficiencies and, for FAB and ESI, the capacity for very precise molecular mass measurements.

The electrospray mass spectrum of protamine Z3 shows a molecular mass of 4748.8 Da, inconsistent with the published sequence. In this report, we describe the redetermination of the primary structure for dog-fish *Scylliorhinus caniculus* protamine Z3 using electrospray and FAB mass spectrometry associated with automated amino-acid sequencing.

Materials and Methods

Preparation and purification of protamine Z3

Protamine Z3 was isolated and purified as previously described [5].

Sequence analysis

Protamine Z3 (2.6 nmol) was sequenced on an Applied Biosystems 470A sequencer using 03 RARG program. The phenylthiohydantoin derivatives of amino

Correspondence to: G. Briand, URA 409 CNRS, Université de Lille II, Institut de Recherches sur le Cancer, Place de Verdun, 59045 Lille Cedex, France.

Enzymes: alpha-chymotrypsin (EC 3.4.21.1), trypsin TPCK (EC 3.4.21.4).

acids were identified on an on-line phenylthiohydantoin analyser, Applied Biosystems 120A, equipped with an CR4A Shimadzu integrator-calculator. For the identification of the modified tyrosine PTH, the thiazolinone was collected from the reaction cartridge and lyophilized at the ninth cycle. Conversion was accomplished by heating the thiazolinone in 0.2 ml of 1 M HCl at 80°C for 10 mn. The PTH was dissolved in 20% acetonitrile and identified by HPLC 120A.

Cleavage with trypsin

66 nmol of Z3 dissolved in 200 μ l of 0.1 M ammonium bicarbonate (pH 8.0) were digested with trypsin at 37°C for 4 h using an enzyme-to-substrate ratio of 1:50 (w/w). The digestion was separated by reverse-phase HPLC on a C₁₈ μ Bondapak Waters column (220 $\times$ 4.6 mm, particle size 10 μ m) using the gradient system indicated in the legend of Fig. 3A.

Cleavage with chymotrypsin

50 nmol of protamine Z3, dissolved in 300 μ l of 0.1 M ammonium acetate (pH 5.0) were digested with chymotrypsin at 30°C for 2 h using an enzyme-to-substrate ratio of 1:100 (w/w). The digestion was stopped by lowering the pH with formic acid. The chymotryptic peptides were subsequently separated by reverse-phase HPLC on a C₁₈ μ Bondapak Waters column (220 $\times$ 4.6 mm, particle size 10 μ m) using the gradient system indicated in the legend of Fig. 3B.

Mass spectrometry

ESMS. ESMS was performed on a VG BIO-Q mass spectrometer [18], an instrument which consists essentially of an electrostatic spray ion source operating at atmospheric pressure followed by a quadrupole mass analyser with a mass range of 3000. The voltage of the extracting cone (V_c) was adjusted to 150 V. The mass spectrometer was scanned from m/z 100 to m/z 1000 in 15 s at unit resolution. Calibration was performed using charged ion from poly(ethylene glycol) 800 obtained from a separate introduction. The protein sample was first dissolved in water containing 5% acetic acid and an equivalent volume of methanol was added. The concentration used was between 30–100 pmol/ μ l. This solution (2–10 μ l) was introduced into the ion source at a flow rate of 2 μ l/min.

FAB. A Concept II HH (Kratos Analytical, Manchester, UK) four-sector tandem mass spectrometer [19], which consists of two double-focusing forward-geometry instruments joined back-to-back (EBEB) was used to acquire spectra. It is equipped with a commercial Kratos FAB source, an Ion Tech B11 NF saddle-field fast atom gun (Ion Tech, Teddington, UK) and a Kratos DS-90 data system. A beam of Xe atoms of approx. 7 keV impact energy and equivalent to 1 mA emission current was employed to ionize peptides dis-

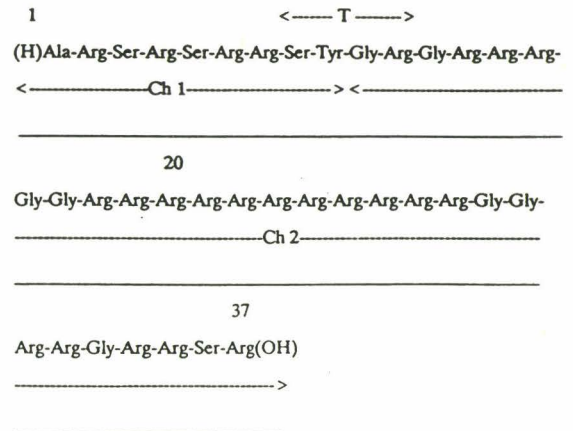


Fig. 1. Complete amino-acid sequence of dog-fish Z3 protein. (T) and (Ch) indicate peptides derived from cleavage with trypsin and chymotrypsin, respectively. (—), Automated Edman degradation of the dog-fish Z3.

solved in the matrix (glycerol/water/trifluoroacetic acid in ratio 10:88:2 in positive ion mode or triethanolamine/tetramethylurea in ratio 50:50 in negative ion mode). The FAB-produced ions were then accelerated through a potential of 8 kV and mass-selected by using MS-I (ESA-I and the magnetic sector) at a resolution of about 1500 (full width at 5% height). Cesium iodide was used as standard compound for mass calibration of samples. The peptides were dissolved in deionized water at a concentration of 1 nmol/ μ l. 1 μ l of peptide solution was deposited on a stainless steel target, and 1 μ l of matrix was added.

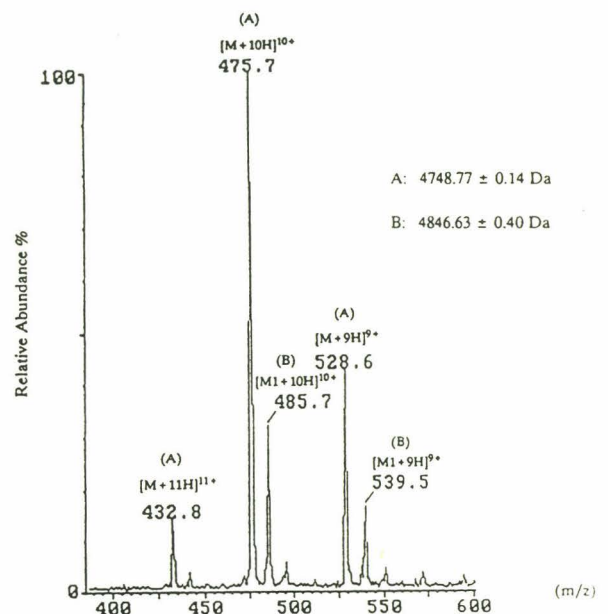


Fig. 2. Section of the ESMS spectrum of dog-fish Z3. Two series (A) and (B) of multicharged ions with 9 to 11 charges were detected. The major series (A) corresponds to the expected protamine Z3 and the minor series (B) corresponds to protamine Z3 with a molecular mass increase of 98.0 Da due to non-covalently bound adducts.

Results and Discussion

The strategy and sequence data which allowed us to elucidate the complete sequence of dog-fish Z3 are presented in Fig. 1. On two different preparations of protamine Z3, automated Edman degradation was de-

termined up to C-terminal residue 37. For one sample only, it was observed in the course of protamine Z3 sequence determination that the tyrosine residue at position nine, was modified. The modification was observed with 1 M HCl conversion as described above. The retention time in HPLC of the PTH of this substi-

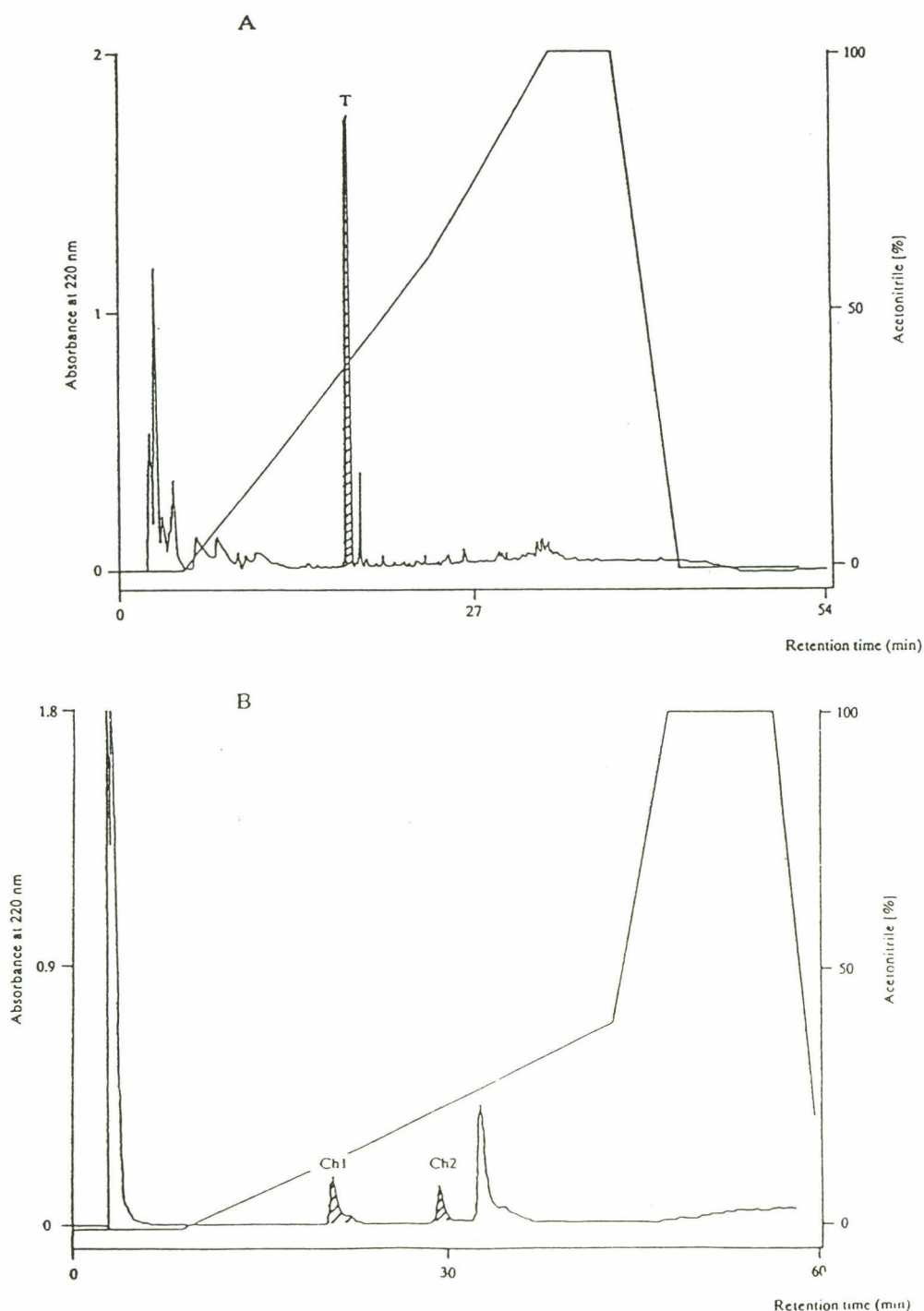


Fig. 3. Fractionation by reverse-phase HPLC of peptides generated from tryptic cleavage (A) and from chymotryptic cleavage (B) of dog-fish Z3. The peptides (66 nmol and 50 nmol, respectively) were separated on a C_{18} column, pore size $10\ \mu\text{m}$ (Waters $220\times 4.6\ \text{mm}$). The elution was performed with a 0–100% linear gradient of 50% acetonitrile in 0.1% trifluoroacetic acid, at a flow rate of 1 ml/min. The absorbance was measured at 220 nm. The hatched peptides peaks correspond to peptides submitted to mass spectrometry.

TABLE I

Amino-acid sequence of protamine Z3

Phenylthiohydantoin derivatives were identified and quantified by reverse-phase HPLC.

Cycle number	Residue	Amount (pmol)
1	Ala	532
2	Arg	4.7
3	Ser	41.2
4	Arg	5.7
5	Ser	28.1
6	Arg	n.q.
7	Arg	n.q.
8	Ser	16.9
9	Tyr	113.9
10	Gly	140.2
11	Arg	n.q.
12	Gly	187
13	Arg	n.q.
14	Arg	30.6
15	Arg	n.q.
16	Gly	144.7
17	Gly	149.6
18	Arg	20.8
19	Arg	32.6
20	Arg	n.q.
21	Arg	29.1
22	Arg	29.6
23	Arg	29.8
24	Arg	36.4
25	Arg	35.3
26	Arg	39.7
27	Arg	34.3
28	Arg	44.5
29	Gly	94.3
30	Gly	106.6
31	Arg	23.3
32	Arg	32
33	Gly	77.1
34	Arg	22.3
35	Arg	20.3
36	Ser	n.q.
37	Arg	12

n.q., not quantified.

tuted tyrosine was found to be shorter than that of the standard tyrosine PTH. The sequence data are given in Table I. In the original published sequence [5], it was concluded that the tyrosine derivative present in the protamine was tyrosine-*o*-sulfate. Sulfated amino-acid residues are difficult to detect directly by conventional approaches but they are readily detected by mass spectrometry [20,21]. But, in the case of suspected tyrosine sulfate, it is essential that the FAB mass spectra be acquired in the negative ion mode. In contrast to the FAB results, sulfotyrosine-containing peptides do not fragment easily in electrospray under normal operating conditions [22].

The ESMS of protamine Z3 showed two individual series of peaks, marked respectively A and B in Fig. 2.

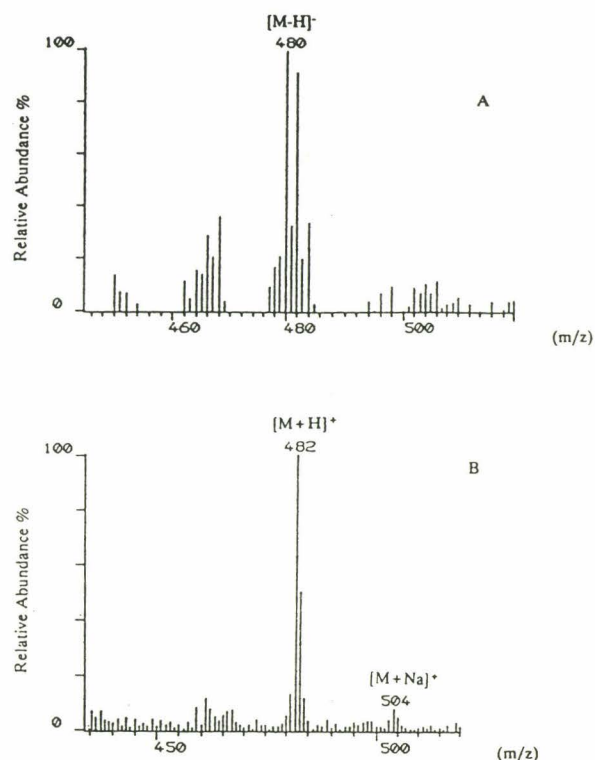


Fig. 4. Section of the FABMS spectra in negative ion mode (A) and positive ion mode (B) of peptide (T) Ser-Tyr-Gly-Arg derived from cleavage of dog-fish Z3 with trypsin.

The molecular mass measured (4748.8 ± 0.1 Da) for the major peaks (A) corresponds to that of Z3 (expected mass: 4748.4 Da). The minor series of peaks (B) yielded a mass of 4846.6 ± 0.4 . The 98 Da difference

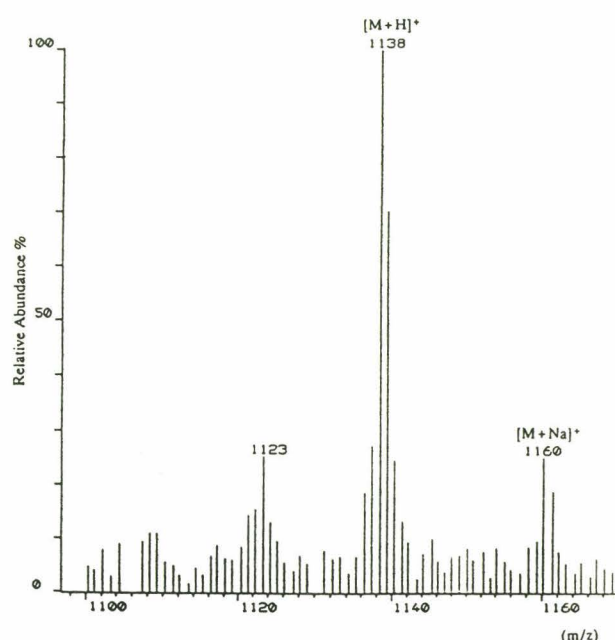


Fig. 5. Section of the positive ion FABMS spectrum of peptide Ch1 derived from cleavage of dog-fish Z3 with chymotrypsin.

can be attributed to adduct ions. Sulfated and/or phosphate impurities present in the protein samples produce an amount of sulfuric acid and/or phosphoric acid that bind strongly to the accessible basic sites of the protein ions [23]. These adducts are non-covalently bound to the protonated protein molecule ions. For the two samples, the measured molecular masses are identical and in agreement with the Edman degradation data without modification.

In order to confirm sequence data, protamine was cleaved either with trypsin or with chymotrypsin. The resulting fragments were separated by means of reverse-phase HPLC (Fig. 3). Most peptides, very hydrophilic, were eluted at the beginning of the gradient. For the tryptic peptide (T) Ser-Tyr-Gly-Arg, the molecular mass determined by FABMS, in the negative and positive ion modes (Fig. 4), agreed with the sequence without covalent modification. HPLC of the chymotryptic digest produced two peptide peaks (Ch1 and Ch2) and a peak corresponding to the intact protein (Fig. 3B). Fig. 5 shows the positive ion FABMS spectrum, of the N-terminal chymotryptic peptide Ch1 with the sequence from Ala at the N-terminus to Tyr at position 9 of protamine Z3. The mass of this peptide was observed at m/z 1138 (theoretical value: 1138). The same result was obtained in ESMS (data not shown). Fig. 6 shows the ESMS spectrum obtained on the C-terminal chymotryptic peptide Ch2 with the sequence from Gly at position 10 to Arg at position 37 of

protamine Z3. The mass of the peptide was observed at 3629.6 ± 0.7 Da (theoretical value: 3628.2 Da).

From these data, we conclude the structure of the protamine Z3 does not contain a sulfated group on tyrosine in position 9. The modification detected by Edman degradation on one sample could be the result of artefact during preparation of the protein and not the result of a post-translational modification. We use three observations to support this affirmation; firstly, the extreme lability of this substituent upon enzymatic hydrolysis, mass spectrometry or standard Edman degradation. Secondly, spermatozoa are the terminal stage of spermatogenesis. The presence of a posttranslational modification would be permanent. Thirdly, in other protamines no identical modification has been detected or suspected.

The corrected protein sequence is 37 amino-acids long and differs from that originally published by the hexapeptide Arg³²-Gly-Arg-Arg-Ser-Arg³⁷.

One may explain the difference on the basis that this hexapeptide was missed when the protein was originally sequenced. Edman degradation was performed with a Beckman 890C sequencer and manual conversion with 1 M HCl. In the spinning cup, even in presence of carriers such as polybrene, the extractive losses are very significant for hydrophilic peptides. The repetitive yield of phenylthiohydantoin of arginine decreases rapidly and overlaps increase when arginines are consecutive. So it is difficult to identify the C-

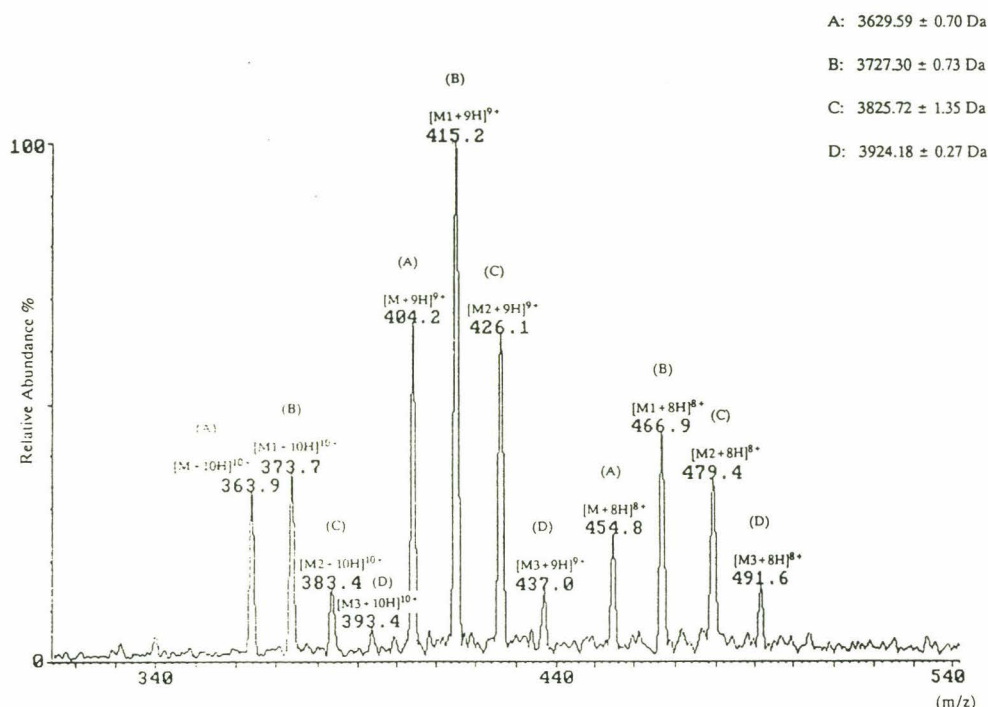


Fig. 6. Section of the ESMS spectrum of peptide Ch2 derived from the cleavage of dog-fish Z3 with chymotrypsin. Four series (A, B, C and D) of multiply-charged ions with 8 to 10 charges were detected. Series A correspond to the peptide Ch2. The other series (B, C and D) correspond to the peptide Ch2 with a molecular mass increase of a multiply of 98 due to non-covalently bound adducts.

terminal residue with absolute certainty. Primary structure elucidation of the C-terminal amino acid of protamine by carboxypeptidase digestion is always hazardous.

The ability of the mass spectrometry to determine the molecular masses of proteins can be used to confirm and correct protein sequence. This study demonstrates that mass spectrometry is now a necessary adjunct to the Edman degradation for establishing the structure of very hydrophilic proteins, such as protamines.

Acknowledgements

We would like to express our gratitude for the scientific collaboration of the Service Commun de Spectrométrie de Masse, Lille I 59655 Villeneuve d'Ascq cedex, France, including Prof. B. Fournet and Dr G. Ricart. We thank A. Hemez for the separation in HPLC and O. Hannedouche-Moreau for sequence analysis.

References

- Gusse, M. and Chevaillier, P. (1978) *Cytobiologie* 16, 421–443.
- Gusse, M. and Chevaillier, P. (1980) *Chromosoma* 77, 57–68.
- Chevaillier, P., Martinage, A., Gusse, M. and Sautière, P. (1987) *Biochim. Biophys. Acta* 914, 19–27.
- Martinage, A., Gusse, M., Belaïche, D., Sautière, P. and Chevaillier, P. (1985) *Biochim. Biophys. Acta* 831, 172–178.
- Sautière, P., Briand, G., Gusse, M. and Chevaillier, P. (1981) *J. Biochem.* 119, 251–255.
- Sautière, P., Gusse, M., Briand, G., Martinage, A. and Chevaillier, P. (1984) *Biochim. Biophys. Acta* 791, 82–86.
- Barber, M., Bordoli, R.S., Sedgwick, R.D. and Tyler, A.N. (1981) *J. Chem. Soc. Chem. Commun.* 7, 325–327.
- Barber, M., Bordoli, R.S., Sedgwick, R.D. and Tyler, A.N. (1981) *Nature* 293, 270–275.
- Barber, M., Bordoli, R.S., Elliot, G.J., Sedgwick, R.D. and Tyler, A.N. (1982) *Anal. Chem.* 54, 645A–657A.
- MacFarlane, R.D. and Torgenson, D.F. (1976) *Science* 191, 920–923.
- Karas, M. and Hillenkamp, F. (1988) *Anal. Chem.* 60, 2299–2301.
- Karas, M., Bachmann, D., Bahr, U. and Hillenkamp, F. (1987) *Int. J. Mass Spectrom. Ion Processes* 78, 53–68.
- Karas, M., Bahr, U. and Hillenkamp, F. (1989) *Int. J. Mass Spectrom. Ion Processes* 92, 231–242.
- Tanaka, K., Waki, H., Ido, Y., Akita, S., Yoshida, Y. and Yoshida, T. (1988) *Rapid Commun. Mass Spectrom.* 2, 151–153.
- Fenn, J.B., Mann, M., Meng, C.K., Wong, S.F. and Whitehouse, C.M. (1990) *Mass Spectrom. Rev.* 9, 37–70.
- Smith, R.D., Loo, J.A., Edmonds, C.G., Barinaga, C.J. and Udseth, H.R. (1990) *Anal. Chem.* 62, 882–899.
- Covey, T.R., Bonner, R.F., Shushan, B.I. and Henion, J. (1988) *Rapid Commun. Mass Spectrom.* 2, 249–256.
- Van Dorsselaer, A., Bitsch, F., Green, B., Jarvis, S., Lepage, P., Bischoff, R., Kolbe, H.V.J. and Roitsch, C. (1990) *Biomed. Environ. Mass Spectrom.* 19, 692–704.
- Despeyroux, D., Bordas-Nagy, J. and Jennings, K.R. (1991) *Rapid Commun. Mass Spectrom.* 5, 156–159.
- Gibson, B.W. (1990) in *Biological Mass Spectrometry*, (Burlingame, A.L., McCloskey, J.A., eds.), pp. 315–336, Elsevier, Amsterdam.
- Bateman, A., Solomon, S. and Bennett, H.P.J. (1990) *J. Biol. Chem.* 265, 22130–22136.
- Covey, T., Shushan, B., Bonner, R., Schroder, W. and Hucho, F. (1991) in *Methods in Protein Sequence Analysis*, (Jornvall, H., Hoog, J.O. and Gustavsson, A.M., eds.), pp. 249–256. Birkhauser Press, Basel.
- Chowdhury, S.K., Katta, V., Beavis, R.C. and Chait, B.T. (1990) *J. Am. Soc. Mass Spectrom.* 1, 382–388.

II: Discussion et Conclusion:

L'hexapeptide C-terminal supplémentaire R-G-R-R-S-R a été confirmé par la méthode récurrente d'Edman sur la protéine. La séquence de la protamine Z3 est donc de 37 résidus d'acides aminés au lieu de 31 préalablement publiés. La masse du peptide chymotrypsique 10-37 a été vérifiée en électrospray.

Nous n'avons jamais détecté la présence d'un groupement sulfate au niveau de la tyrosine en position 9. En effet, le peptide chymotrypsique 1-9 a été analysé en FAB-MS et en ES-MS. La masse obtenue correspond à celle calculée sans aucune modification.

Nous avons ensuite réalisé une coupure de la protamine Z3 par la trypsine pour isoler le tétrapeptide S-Y-G-R et l'étudier en FAB positive et négative. La masse obtenue dans les deux modes est celle du peptide non modifié.

Nous avons étudié le comportement d'une tyrosine sulfate en spectrométrie de masse. Ces études ont été réalisées sur un peptide de synthèse de la cholécystokinine (CCK) de séquence:

D-Y(SO₃H)-M-G-W-M-D-F(NH₂) (cf. publication 3, page 118).

Les meilleurs résultats de stabilité du groupement sulfate sont obtenus en mode négatif aussi bien en FAB qu'en ES-MS. En mode positif, le sulfate est plus stable en ES-MS qu'en FAB (publication 3).

Le peptide sulfaté de la CCK a été traité par la trypsine dans les mêmes conditions que la protamine Z3 et la masse de ce peptide après traitement montre la stabilité du groupement sulfate.

L'ensemble de ces résultats nous permet donc d'exclure la possibilité de l'existence d'un groupement sulfate au niveau de la tyrosine en position 9. Nous pensons fortement que la modification détectée par la dégradation d'Edman dans une seule préparation de protamine Z3 est le résultat d'un artéfact de préparation et non d'une modification post-traductionnelle.

Trois observations viennent appuyer ceci:

- l'extrême labilité de ce groupement modifiant,
- la protéine Z3 est isolée du spermatozoïde qui représente le stade final de la spermatogenèse. Une modification post-traductionnelle devrait donc être permanente,
- dans les protamines des autres espèces un tel groupement n'a jamais été détecté.

En conclusion, cet exemple montre le rôle que peut jouer la spectrométrie de masse dans l'élucidation d'une structure primaire de protéine. En effet, un simple spectre de vérification de la masse à l'origine de ce travail, nous a permis de compléter la séquence de la protamine Z3. Toutefois, il faut préciser que les performances du séquenceur en phase liquide, utilisé en 1981, étaient moindres que celles des séquenceurs en phase gazeuse actuels. Les pertes dues aux lavages malgré la présence de polybrène étaient abondantes. De plus la conversion se faisait manuellement. Enfin, l'interprétation des hydrolyses par les exopeptidases est toujours délicate.

Chapitre 4

Utilisation de la spectrométrie
de masse pour la mise en
évidence de variants
structuraux

Etude de la protéine de transition TP1 de bélier

I. Introduction:

Chez le bélier, au cours de la spermatogenèse, les histones sont remplacées par 4 protéines intermédiaires TP1, TP2, TP3 et TP4, qui elles mêmes sont remplacées par une protamine.

Ces protéines intermédiaires sont présentes en très faible quantité dans les spermatides. En effet, le stade 9-12 représentatif des protéines de transition (TP), dure environ 4 jours (cf. Fig.12 page 48). C'est pourquoi, les séquences de ces protéines actuellement connues ont été déduites à partir des séquences nucléotidiques des gènes correspondant.

Kistler et *al* ont établi la séquence de la protéine TP1 isolée du testicule de rat (185).

La protéine TP1 isolée du testicule de bélier *Ovis aries* a été étudiée dans notre laboratoire. Notre apport dans ce travail consistait en l'analyse en spectrométrie de masse de la protéine et des différents peptides thermolysiques.

Ce travail nous a permis de mettre en évidence l'existence de 2 variants structuraux qui ne diffèrent que par un seul résidu d'acide aminé à la position 27 (cystéine dans le variant majeur et glycine dans le variant mineur).

Les résultats sont présentés dans la publication 2:

Nuclear transition protein 1 from ram elongating spermatids. Mass spectrometric characterization, primary structure and phosphorylation sites of two variants.

Chirat F., Martinage A., Briand G., Kouach M., Van Dorsselaer A., Loir M. and Sautière P. (1991)

Eur. J. Biochem. 198, 13-20

Nuclear transition protein 1 from ram elongating spermatids

Mass spectrometric characterization, primary structure and phosphorylation sites of two variants

Frédéric CHIRAT¹, Arlette MARTINAGE¹, Gilbert BRIAND¹, Mostafa KOUACH¹, Alain VAN DORSSELAER², Maurice LOIR³ and Pierre SAUTIÈRE¹

¹ Unité de Recherche Associée n° 409 au Centre National de la Recherche Scientifique, Université de Lille II, Institut de Recherches sur le Cancer, Lille, France

² Laboratoire de Chimie Organique des Substances Naturelles,

Unité de Recherche Associée n° 31 au Centre National de la Recherche Scientifique, Strasbourg, France

³ Laboratoire de Physiologie des Poissons, Institut National de la Recherche Agronomique, Campus de Beaulieu, Rennes, France

The ram transition protein 1 (TP1) is present in spermatid cell nuclei in the nonphosphorylated, monophosphorylated and diphosphorylated forms. Its primary structure was determined by automated Edman degradation of S-carboxamidomethylated protein and of peptides generated by cleavage with thermolysin and endoproteinase Lys-C.

The ram TP1 is a small basic protein of 54 residues and structurally very close to other mammalian TP1. The mass spectrometric data obtained from the protein and its fragments reveal that ram TP1 is indeed a mixture (approximately 5:1) of two structural variants (M_r 6346 and 6300). These variants differ only by the nature of the residue at position 27 (Cys in the major variant and Gly in the minor variant). The study of phosphorylation sites has shown that four different serine residues could be phosphorylated in the monophosphorylated TP1, at positions 8, 35, 36 or 39.

From previous physical studies, it has been postulated that the Tyr32 surrounded by two highly conserved basic clusters was responsible for the destabilization of chromatin by intercalation of its phenol ring between the bases of double-stranded DNA.

The presence of three phosphorylatable serine residues in the very conserved sequence 29–42 is another argument for the involvement of this region in the interaction with DNA.

Spermatogenesis involves a progressive differentiation of the germ cells towards a terminal stage represented by the mature spermatozoa.

In mammals, the first part of spermiogenesis where the spermatid development takes place, is quiescent relative to changes in chromatin structure, while the second part of this stage is characterized by a progressive condensation of the diffuse chromatin to a highly compact and transcriptionally inactive structure [1–6].

This condensation is correlated with a double-protein transition. The first transition corresponds to the complete removal and replacement of histones (somatic and testis-specific variants) by several spermatid-specific basic proteins [7–9] which are themselves eliminated and replaced by one [10–16] or two protamines [17–23] during the second transition.

While histones and protamines have been extensively studied, little is known about spermatid-specific proteins also called transition proteins.

In rat and mouse four transition proteins [7, 8] named TP1, TP2, TP3 and TP4 have been identified. Among them, only TP1 and TP2 have been fully characterized.

Rat and mouse TP1 are low molecular mass basic proteins (54 residues), rich in arginine (20%), lysine (19%) and serine (14%) and devoid of cysteine [24–27].

TP2 is larger than TP1 (116 and 117 residues in rat and mouse, respectively) [28, 29]. This protein is characterized by a high amount of serine (22%) and proline (19%) and by the presence of cysteine (about 5%).

Little is known about transition proteins in other mammalian species. In man, bull and boar, a transition protein related to TP1 has been identified and sequenced [30, 31]. Like rat and mouse TP1, human TP1 lacks of cysteine, while boar and bull TP1 contain one cysteine residue.

In ram, four transition proteins called P1, P3, P7 and T have been purified and characterized [9]. The amino acid composition and the C-terminal sequence (27 residues) of ram protein P1 (F. Chirat, unpublished results) clearly indicate a close structural relationship between this protein and rat and mouse TP2, whilst the amino acid composition of ram protein

Correspondence to P. Sautière, URA 409 CNRS, Université de Lille II, Institut de Recherches sur le Cancer, Place de Verdun, F-59045 Lille Cédex, France

Abbreviations. ESMS, electrospray mass spectrometry; FAB-MS, fast-atom-bombardment mass spectrometry; PDMS, plasma-desorption mass spectrometry; TP, transition protein.

Enzymes. Carboxypeptidase A (EC 3.4.17.1); endoproteinase Lys-C (EC 3.4.21.50); thermolysin (EC 3.4.24.4).

T [9] is very similar to those of the other mammalian TP1. Moreover, this protein was shown to be monophosphorylated and diphosphorylated to a small extent [9].

In this paper, we report the complete primary structure of the ram protein T and the different sites of phosphorylation identified in the monophosphorylated form of the protein. Since the sequence analysis of ram protein T has shown a close structural relationship with known mammalian TP1 proteins [31], the ram protein T will be named from now, for the sake of clarity, ram transition protein 1 or ram TP1.

MATERIALS AND METHODS

See Supplementary section.

RESULTS

Ram TP1 is present in elongating spermatid nuclei in three forms: nonphosphorylated, monophosphorylated and diphosphorylated [9]. These forms were separated by fractionation of the 0.75-M-perchloric-acid-soluble cysteine-containing proteins extracted from EDTA-resistant spermatid nuclei, on a CM-Sephadex C25 column. Prior to ion-exchange chromatography, the spermatid proteins were reduced and alkylated with iodoacetamide.

The nonphosphorylated and monophosphorylated forms of TP1 were obtained in a highly pure state, while the diphosphorylated form coeluted with the spermatid protein 7.

Amino acid sequence of ram TP1

The strategy and sequence data which allowed us to elucidate the complete sequence of ram TP1 are presented in Fig. 1. The elution diagrams of the peptides derived from the protein and the table providing sequence data are given in the Supplementary section (Figs S1 and S2, Table S1).

The amino-terminal sequence of TP1 was determined up to residue 38.

The digestion of the S-carboxamidomethylated protein with carboxypeptidase A for 15 min released 0.9 nmol leucine/nmol protein and 0.5 nmol histidine/nmol protein.

The remainder of the sequence was unambiguously established by sequencing peptides derived from hydrolysis of the S-carboxamidomethylated protein with thermolysin and endoproteinase Lys-C.

Cleavage of the protein with thermolysin generated six peptides (Th1 to Th6) which altogether account for the total number of amino acids of the TP1 molecule (54 residues) (Table 1).

Among the peptides derived from the cleavage of the protein with endoproteinase Lys-C, only three peptides K1, K2 and K3 were useful in providing the necessary overlaps for the alignment of the thermolysin peptides Th4, Th5 and Th6 (Table S1). The peptides K1 and K2 were obtained as a mixture in equal amounts as shown by the amino acid composition. The correct assignment of the Edman data to each peptide was possible only because the sequence was known at that point.

Mass spectrometry of the S-carboxamidomethylated protein

Electrospray mass spectrometry (ESMS) of the non-phosphorylated form of TP1 yields two fragments having

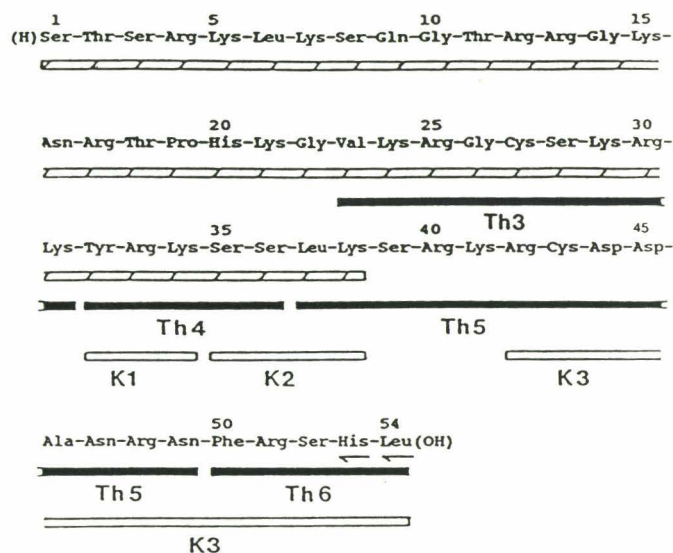


Fig. 1. Complete amino acid sequence of ram TP1 protein. (Th) and (K) indicate peptides derived from cleavage with thermolysin and endoproteinase Lys-C, respectively. (◻) Automated Edman degradation of the S-carboxamidomethylated ram TP1; (■) automated Edman degradation of thermolysin peptides (Th); (□) automated Edman degradation of endoproteinase Lys-C peptides (K); (—) carboxypeptidase A digestion

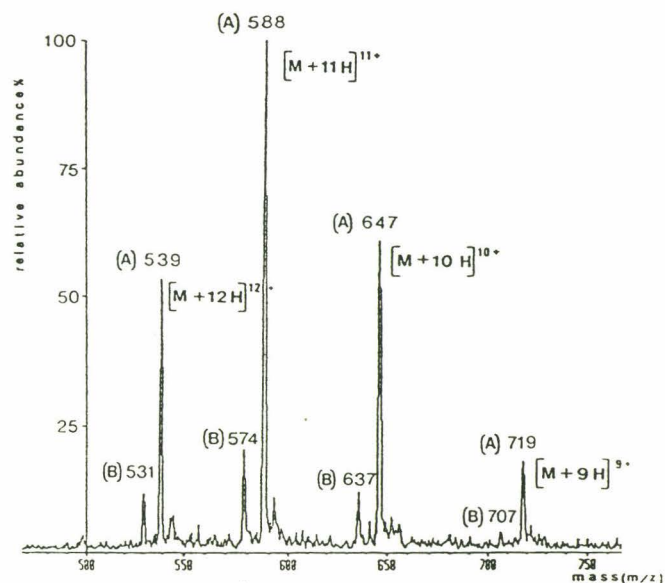


Fig. 2. Section of the ESMS spectrum of ram TP1. Two series (A) and (B) of multicharged ions with 9 to 12 charges were detected. The major series (A) corresponds to the major variant (6460 Da), and the minor series (B) corresponds to the minor variant (6357 Da)

molecular masses of 6460 Da and 6357 Da (Fig. 2 and Table 1). The fragment of molecular mass 6460 Da corresponds to the presence of two S-carboxamidomethylated cysteine residues. The other fragment, with a molecular mass of 6357 Da, is in agreement with the replacement of one cysteine residue by a glycine residue.

From these data, the two components were estimated to be present in a 5:1 ratio.

Table 1. Amino acid composition of ram TP1 and peptides generated by cleavage with thermolysin (Th) and with endoproteinase Lys-C (K). Values in parentheses are the number of residues/molecule of protein or peptide derived from the sequence. CmCys, S-carboxymethyl cysteine; n.d., not determined

Amino acid	TP1	Amino acid content of TP1 fragment (residue nos)							
		Th1 (1–5)	Th2 (6–22)	Th3 (23–31)	Th4 (32–36)	Th5 (37–49)	Th6 (50–54)	K1/2 (32–34/ 35–38)	K3 (42–54)
	mol/mol								
CmCys	2.1(2)	—	—	0.8(1)	—	0.9(1)	—	—	0.8(1)
Asx	5.0(5)	—	1.2(1)	—	—	4.0(4)	—	—	4.1(4)
Thr	2.9(3)	1.0(1)	2.0(2)	—	—	—	—	—	—
Ser	7.6(8)	2.0(2)	1.5(1)	1.4(1)	2.1(2)	1.5(1)	1.3(1)	1.6(2)	1.5(1)
Glx	1.3(1)	—	1.1(1)	—	—	—	—	—	—
Pro	1.2(1)	—	1.0(1)	—	—	—	—	—	—
Gly	4.3(4)	—	3.0(3)	1.3(1)	—	—	—	—	—
Ala	1.3(1)	—	—	—	—	1.1(1)	—	—	1.2(1)
Val	1.1(1)	—	—	1.0(1)	—	—	—	—	—
Met	—	—	—	—	—	—	—	—	—
Ile	—	—	—	—	—	—	—	—	—
Leu	3.0(3)	—	1.0(1)	—	—	1.0(1)	1.0(1)	1.0(1)	1.2(1)
Tyr	1.1(1)	—	—	—	0.9(1)	—	—	0.7(1)	—
Phe	1.0(1)	—	—	—	—	—	1.0(1)	—	1.0(1)
His	1.9(2)	—	1.0(1)	—	—	—	1.1(1)	—	1.0(1)
Lys	10.3(10)	1.0(1)	3.1(3)	3.0(3)	1.0(1)	2.1(2)	—	2.0(2)	—
Arg	11.3(11)	0.9(1)	3.2(3)	2.0(2)	1.0(1)	3.2(3)	0.9(1)	1.1(1)	3.0(3)
Mass ^a	6462	—	—	1117.6	—	1632.8	—	—	—
Mass ^b	6460	n.d.	n.d.	1117.6	n.d.	1633.8	n.d.	n.d.	n.d.
	6357			1014.6					
Total ^c	54	5	17	9	5	13	5	3/4	13

^a Calculated masses in Da.

^b Masses in Da, measured by mass spectrometry (ESMS for TP1, PDMS for Th5 and FAB-MS for Th3).

^c Total number of amino acid residues.

Mass spectrometry of

S-carboxamidomethylated thermolysin peptides Th3 and Th5

When subjected to fast-atom-bombardment mass spectrometry (FAB-MS), peptide Th3 (residues 23–31) yielded two fragments having molecular masses of 1117.6 Da and 1014.6 Da (Fig. 3 and Table 1).

The difference (103 Da) between these two masses is identical to that observed between the two molecular masses determined for the S-carboxamidomethylated protein TP1.

The analysis of peptide Th5 (residues 37–49) by ²⁵²Cf plasma-desorption mass spectrometry (PDMS) gave a single mass of 1633.8 Da which corroborates the sequence determined by the automated Edman degradation. It must be emphasized that FAB-MS failed to give the mass of this peptide.

Identification of the phosphorylation sites in ram TP1

The phosphorylation sites were identified in the monophosphorylated form of ram TP1. Due to the low amount of protein available, no mass determination of phosphorylated TP1 by ESMS was carried out.

The S-carboxamidomethylated derivative of the monophosphorylated protein was cleaved by thermolysin. The elution pattern of the digest presented in Fig. S1B shows an

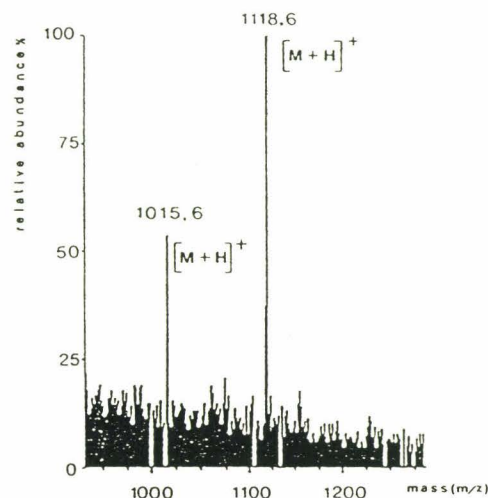


Fig. 3. Section of the FAB-MS spectrum of peptides Th3 derived from cleavage of S-carboxamidomethylated ram TP1 with thermolysin

additional peptide, Thx, by comparison with the elution profile of thermolysin peptides generated from the nonphosphorylated protein (Fig. S1A). Among the seven peptides obtained, only peptides Th2 and Thx were found to contain

phosphoamino acid(s) upon amino acid analysis after a 2-h hydrolysis in 6 M HCl at 110°C.

These two peptides were treated according to the procedure described by Meyer et al. [33] in order to convert the phosphoserine residues into S-ethyl cysteine residues.

The sequence data presented in Table S2 led us to conclude that serine residues at positions 8, 35, 36 and 39 are partially phosphorylated in the monophosphorylated form of the protein. The data also indicate that peptides Th2 and Thx were cross contaminated and that peptide Thx corresponds to the sequence 32–49 of ram TP1.

In the course of the sequence analysis, we have not observed the two peaks characteristic of the phenylthiohydantoin derivative of the carboxamidomethylated cysteine expected at position 12 (Cys43 in TP1), but another peak with a retention time identical to that of the phenylthiohydantoin derivative of S-ethyl cysteine. To check the possibility that the procedure of Meyer et al. [33], could also affect the S-carboxamidomethylated cysteine, we submitted peptide Th3, VKRGCSKRK (Table 1 in Supplementary section) (residues 23–31) generated from the non-phosphorylated TP1, which contains one S-carboxamidomethylated cysteine at position 27, to an automated Edman degradation. This peptide was sequenced twice, before and after derivatization according to the procedure of Meyer et al. [33]. The sequence data, presented in the Fig. S3 obviously show that the S-carboxamidomethylated cysteine is also transformed by this procedure.

A complete study of this transformation, the mechanism of which does not appear to be a β -elimination, will be published elsewhere.

DISCUSSION

From mass spectrometric data, TP1 was found to be a mixture of two structural variants which only differ by a single change: the cysteine residue at position 27 in the major variant is replaced by a glycine residue in the minor one.

The contribution of mass spectrometry was essential to reveal the existence of the minor variant. Indeed, during the course of the automated Edman degradation of the protein, the identification of glycine at position 27 was hindered, both the low amount of this variant and the carry-over of glycine from the previous cycle.

With 43% of basic residues, ram TP1 has a basicity intermediate between that of histone H4 (26%) [35] and that of ram protamine (56%) [13]. The basic amino acids are distributed all along the polypeptide chain, most of them arranged in clusters of two or three residues. This protein is also characterized by a relative abundance of hydroxylated residues (22%) which accumulate in the first 75% of the molecule. The occurrence of two basic clusters interspersed with one tyrosine residue at position 32 and three serine residues at positions 35, 36 and 39, constitutes the most interesting feature of ram TP1.

These residues of serine were found to be phosphorylated in the monophosphorylated form of TP1. Two of them (Ser35 and Ser36) are present in a sequence Arg/Lys-Xaa-Ser specifically recognized by the cyclic AMP-dependent protein kinase, where Xaa represents any amino acid except proline [36, 37]. The two other sites of phosphorylation identified in ram TP1, are located in the sequence LKS, which occurs twice in the polypeptide chain, at positions 6–8 and 37–39 respectively. Thus, the presence of four phosphorylation sites in the mono-

	1	10	20	30
Ram (1)	STSRK	LK	SG	SKR
Ram (2)	STSRK	LK	SG	SKR
Boar	STSRK	LK	SG	SKR
Bull	STSRK	LK	SG	SKR
Rat	STSRK	LK	SG	SKR
Mouse	STSRK	LK	SG	SKR
Human	STSRK	LK	SG	SKR

	40	50
Ram (1)	KYRK	SS
Ram (2)	KYRK	SS
Boar	KYRK	GS
Bull	KYRK	SS
Rat	KYRK	SS
Mouse	KYRK	SV
Human	KYRK	GN

Fig. 4. Comparison of the amino acid sequence of TP1 from ram, bull [31], boar [31], rat [25, 26], mouse [27] and man [28]. The boxes correspond to sequence identity between the different TP1

phosphorylated form of ram TP1 indicates a lack of specificity in the phosphorylation of this protein.

All the known mammalian TP1 exhibit a high degree of similarity (Fig. 4) particularly in the very basic region 29–42 which contains one residue of tyrosine at position 32 and at least one phosphorylatable residue of serine at position 35, 36 or 39. In this region the tyrosine residue, likely to be involved in interactions with DNA, could be essential for the destabilization of the chromatin through its intercalation between the bases of DNA as postulated by Singh and Rao [38]. Moreover, the phosphorylation of serine residues which is known to modulate interaction of basic proteins (i.e. histones and protamines) with DNA [39–42] would allow the correct positioning of TP1 on DNA. Later on, the dephosphorylation of TP1 would bring together the region 29–42 and DNA, with as a consequence, the intercalation of Tyr32 between the DNA bases, thus inducing the destabilization of the chromatin necessary to the complete removal of histones before the deposition of protamines.

We are indebted to K. Klarskov and P. Roepstorff (University of Odense, Denmark) for the PDMS analysis of peptide Th5. We thank M. J. Dupire for amino acid analysis, the *Service commun de séquence peptidique* (URA 409 Centre National de la Recherche Scientifique, U 16 Institut National de la Santé et de la Recherche Médicale) and T. Ernout for the preparation of this manuscript. This work was supported by grants from the *Centre National de la Recherche Scientifique*, from the *Université de Lille II* and from the *Fondation pour la Recherche Médicale*.

REFERENCES

1. Kierszenbaum, A. L. & Tres, L. L. (1975) *J. Cell Biol.* 65, 258–270.

2. Meistrich, M. L., Reid, B. O. & Barcellona, W. J. (1976) *Exp. Cell Res.* **99**, 72–78.
3. Loir, M. & Lanneau, M. (1978) *Exp. Cell Res.* **115**, 231–243.
4. Meistrich, M. L., Trostle, P. K. & Brock, W. A. (1981) in *Bioregulators of development* (Jagiello, G. & Vogel, H. J., eds) pp. 151–166, Academic Press, New York.
5. Loir, M. & Lanneau, M. (1984) *J. Ultrastruct. Res.* **86**, 262–276.
6. Loir, M., Bouvier, D., Fornells, M., Lanneau, M. & Subirana, J. A. (1985) *Chromosoma (Berl.)* **92**, 304–312.
7. Meistrich, M. L., Brock, W. A., Grimes, S. R., Platz, R. D. & Hnilica, L. S. (1978) *Fed. Proc., FASEB.* **37**, 2522–2525.
8. Balhorn, R., Weston, S., Thomas, C. & Wyrobeck, A. J. (1984) *Exp. Cell Res.* **150**, 298–308.
9. Dupressoir, T., Sautière, P., Lanneau, M. & Loir, M. (1985) *Exp. Cell Res.* **161**, 63–74.
10. Coelingh, J. P., Monfoort, C. H., Rozijn, T. H., Gevers Leuven, J. A., Schiphof, R., Steyn-Parvé, E. P., Braunitzer, G., Schrank, B. & Ruhfus, A. (1972) *Biochim. Biophys. Acta* **285**, 1–14.
11. Calvin, H. I. (1976) *Biochim. Biophys. Acta* **434**, 377–389.
12. Tobita, T., Tsutsumi, H., Kato, A., Suzuki, H., Nomoto, M., Nakano, M. & Ando, T. (1983) *Biochim. Biophys. Acta* **744**, 141–146.
13. Sautière, P., Bélaïche, D., Martinage, A. & Loir, M. (1984) *Eur. J. Biochem.* **144**, 121–125.
14. Mazrimas, J. A., Corzett, M., Campos, C. & Balhorn, R. (1986) *Biochim. Biophys. Acta* **872**, 11–15.
15. Ammer, H. & Henschen, A. (1988) *FEBS Lett.* **242**, 111–116.
16. Ammer, H. & Henschen, A. (1988) *Biol. Chem. Hoppe-Seyler* **369**, 1301–1306.
17. McKay, D. J., Renaux, B. S. & Dixon, G. H. (1985) *Biosci. Rep.* **5**, 383–391.
18. Ammer, H., Henschen, A. & Lee, C. H. (1986) *Biol. Chem. Hoppe-Seyler* **367**, 515–522.
19. Gusse, M., Sautière, P., Bélaïche, D., Martinage, A., Roux, C., Dadoune, J. P. & Chevaillier, Ph. (1986) *Biochim. Biophys. Acta* **884**, 124–134.
20. McKay, D. J., Renaux, B. S. & Dixon, G. H. (1986) *Eur. J. Biochem.* **156**, 5–8.
21. Bélaïche, D., Loir, M., Kruggel, W. & Sautière, P. (1987) *Biochim. Biophys. Acta* **913**, 145–149.
22. Bellve, A. R. (1988) *Biochemistry* **27**, 2890–2897.
23. Pirhonen, A., Linnala-Kankkunen, A. & Mäenpää, P. H. (1989) *FEBS Lett.* **244**, 199–202.
24. Kistler, W. S., Geroch, M. E. & Williams-Ashman, H. G. (1973) *J. Biol. Chem.* **248**, 4532–4543.
25. Kistler, W. S., Noyes, C., Hsu, R., Heinrikson, R. L. (1975) *J. Biol. Chem.* **250**, 1847–1853.
26. Heidaran, M. A., Kozak, C. A. & Kistler, W. S. (1989) *Gene* **75**, 39–46.
27. Kleene, K. C., Borzorgzadeh, A., Flynn, J. F., Yelick, P. C. & Hecht, N. B. (1988) *Biochim. Biophys. Acta* **950**, 215–220.
28. Luerssen, H., Maier, W. M., Hoyer-Fender, S. & Engel, W. (1989) *Nucleic Acids Res.* **17**, 3585.
29. Kleene, K. C. & Flynn, J. F. (1987) *J. Biol. Chem.* **262**, 17272–17277.
30. Luerssen, H., Hoyer-Fender, S. & Engel, W. (1988) *Nucleic Acids Res.* **16**, 7723.
31. Kremling, H., Luerssen, H., Adham, I. M., Klemm, U., Tsousidou, S. & Engel, W. (1989) *Differentiation* **40**, 184–190.
32. Sundqvist, B., Kamensky, I., Hakansson, P., Kjellberg, J., Salehpour, M., Widdiyasekera, S., Fohlman, J., Petersson, P. A. & Roepstorff, P. (1984) *Biomed. Mass Spectrom.* **11**, 242–257.
33. Meyer, H. E., Hoffmann-Posorske, E., Korte, H. & Heilmeyer, L. M. G. (1986) *FEBS Lett.* **204**, 61–66.
34. Swiderek, K., Jaquet, K., Meyer, H. E. & Heilmeyer, L. M. G. (1988) *Eur. J. Biochem.* **176**, 335–342.
35. DeLange, R. J., Fambrough, D. M., Smith, E. L. & Bonner, J. (1969) *J. Biol. Chem.* **244**, 319–334.
36. Shlyapnikov, S. V., Arutyunyan, A. A., Kurochkin, S. N., Memelova, L. V., Nesterova, M. V., Sashchenko, L. P. & Severin, E. S. (1975) *FEBS Lett.* **53**, 316–319.
37. Martinage, A., Mangeat, P., Laine, B., Couppez, M., Sautière, P., Marchis-Mouren, G. & Biserte, G. (1980) *FEBS Lett.* **118**, 323–329.
38. Singh, J. & Rao, M. R. S. (1987) *J. Biol. Chem.* **262**, 734–740.
39. Marushige, Y. & Marushige, K. (1975) *J. Biol. Chem.* **250**, 39–45.
40. Balhorn, R. (1982) *J. Cell Biol.* **93**, 298–305.
41. Marushige, Y. & Marushige, K. (1978) *Biochim. Biophys. Acta* **518**, 440–449.
42. Pruslin, F. H., Imesh, E., Winston, R. & Rodman, T. C. (1987) *Gamete Res.* **18**, 179–190.

Supplementary material to

Nuclear transition protein 1 from ram elongating spermatids

Mass spectrometric characterization, primary structure and phosphorylation sites of two variants

Frédéric CHIRAT, Arlette MARTINAGE, Gilbert BRIAND, Mostafa KOUACH, Alain Van DORSSELAER, Maurice LOIR and Pierre SAUTIERE

EXPERIMENTAL PROCEDURES

Materials

Testes of sexually mature rams (*Ovis aries*) were collected at the *Station de la Fertilité mâle de Nouzilly* (France) and were used immediately or after freezing in liquid nitrogen and storage at –80 °C. Ram spermatidal nuclear proteins were extracted and purified as described in [9].

Carboxypeptidase A treated with phenylmethylsulfonyl fluoride was from Sigma. Thermolysin was from Merck. Endoproteinase Lys-C was purchased from Boehringer. Acetonitrile for reverse-phase HPLC was obtained from Rathburn. All solvents and reagents for gas-phase sequencing were from Applied Biosystems. All other reagents were of the highest purity available.

Carboxypeptidase A digestion

Ram TP1 (1 nmol) was digested with carboxypeptidase A in 0.1 M N-methylmorpholine acetate buffer, pH 8.0, with an enzyme/substrate ratio of 1:50 (by mass) for 15 min at 37 °C. The digestion was stopped by lowering the pH with glacial acetic acid. The amino acids released were identified on the amino acid analyzer.

Cleavage with thermolysin

10 nmol ram TP1, dissolved in 0.4 ml 0.1 M ammonium bicarbonate buffer, pH 8.0, was digested with thermolysin at 37 °C for 1 h, using an enzyme/substrate ratio of 1:50 (by mass). The digestion was stopped by lowering the pH with formic acid and the hydrolysate was freeze-dried prior to fractionation.

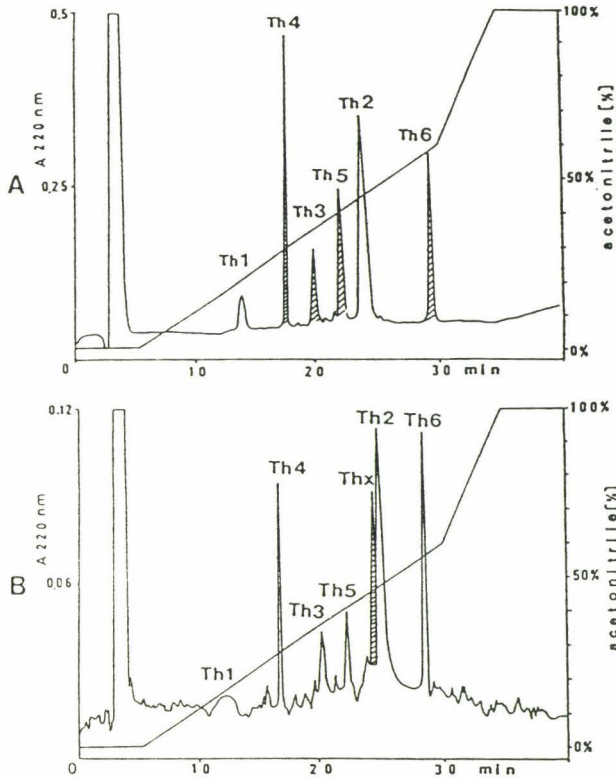


Fig. S1. Fractionation by reverse-phase HPLC of peptides generated from thermolysin cleavage of the nonphosphorylated ram TP1 (A) and monophosphorylated ram TP1 (B). The peptides (10 nmol) were separated on a C_{18} column, pore size 5 μ m (Brownlee 220 $\times$ 4.6 mm). The elution was performed with a 0–60% linear gradient of 50% acetonitrile in 0.1% trifluoroacetic acid, at a flow rate of 1 ml/min. The absorbance was measured at 220 nm. The hatched peaks correspond to peptides submitted to automated Edman degradation. The peptides were numbered according to their position in the sequence of the protein

Cleavage with endoproteinase Lys-C

10 nmol ram TP1 dissolved in 0.4 ml 0.1 M ammonium bicarbonate buffer, pH 8.0, was digested with endoproteinase Lys-C at 37°C for 2 h, using an enzyme/substrate ratio of 1:100 (by mass). The digestion was stopped by lowering the pH with formic acid and the hydrolysate was freeze-dried prior to fractionation.

Fractionation of peptides

Peptides generated by enzymatic cleavage were separated by reverse-phase HPLC on a C_{18} Brownlee column (220 $\times$ 4.6 mm) (particle size 5 μ m) using the gradient systems indicated in the figure legends.

Amino acid analyses

Peptide samples were hydrolyzed in evacuated sealed tubes at 110°C for 24 h in 6 M HCl. One drop of 1% phenol was added to prevent excessive degradation of tyrosine.

Sequence determination

Microsequencing was performed on a gas-phase sequencer, Applied Biosystems 470A, using the 03 RPTH Standard Program. The

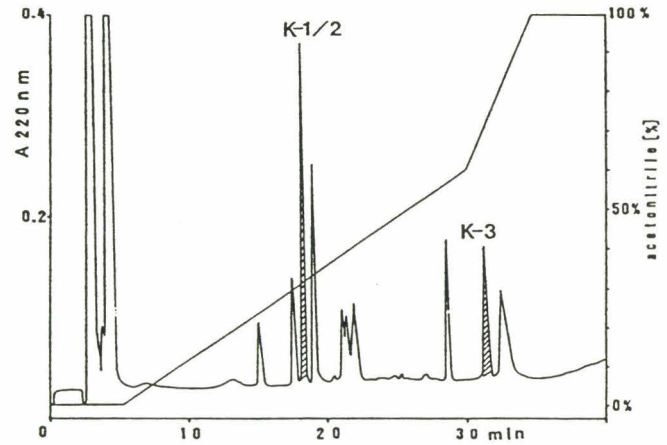


Fig. S2. Fractionation by reverse-phase HPLC of peptides generated from endoproteinase Lys-C cleavage of the nonphosphorylated ram TP1. The peptides (10 nmol) were separated and numbered as indicated in legend to Fig. S1. The hatched peaks correspond to peptides used to establish the complete amino acid sequence of ram TP1

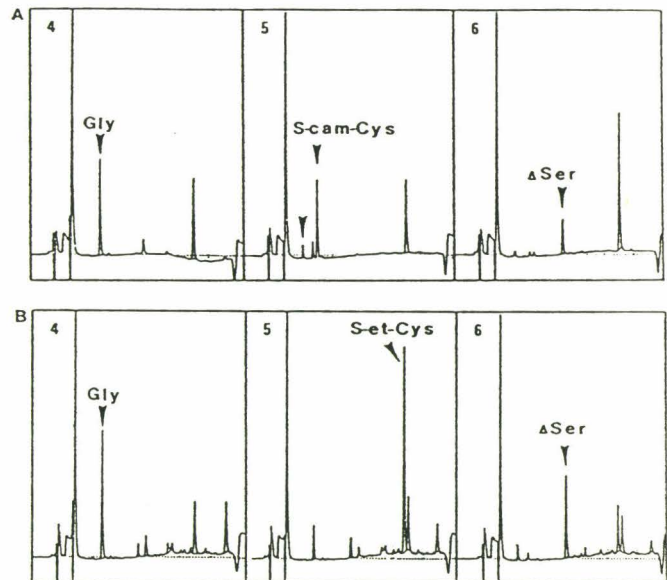


Fig. S3. Automated Edman degradation of peptide Th3 derived from the cleavage of ram TP1 with thermolysin. Only cycles 4, 5 and 6 are presented. Peptide Th3 (1 nmol) was sequenced (A) before and (B) after the treatment described by Meyer et al. [33] for the transformation of P-Ser in S-ethyl cysteine. At cycle 5, the phenylthiohydantoin derivative of S-carboxamidomethyl cysteine (S-cam-Cys) observed before treatment (A) was no longer observed after treatment (B), however, a phenylthiohydantoin derivative with a retention time identical to that of the phenylthiohydantoin of S-ethyl cysteine (S-et-Cys) was present in its place

phenylthiohydantoin derivatives of amino acids were identified on an on-line phenylthiohydantoin analyzer, Applied Biosystems 120A, equipped with an integrator-calculator, Shimadzu C-R4A. The amount of protein and of each peptide submitted to Edman degradation was approximately 1 nmol.

Mass spectrometry

Positive FAB-MS was carried out on a ZAB-HF double-focusing mass spectrometer (mass range 3200 Da at 8 keV) and recorded on

Table S1. Amino acid sequence of *S*-carboxamidomethylated ram TP1 and of peptides derived from cleavage with thermolysin (Th) and endoproteinase Lys-C (K). Phenylthiohydantoin derivatives were identified and quantified by reverse-phase HPLC. Peptides K1 and K2 could not be separated and were sequenced together. n.q., not quantified

Cycle	TP1		Protein fragment of TP1													
			Th3		Th4		Th5		Th6		K1/2				K3	
	residue	amount	residue	amount	residue	amount	residue	amount	residue	amount	residue	amount	residue	amount	residue	amount
		pmol		pmol		pmol		pmol		pmol		pmol		pmol		pmol
1	S	339	V	295	Y	68	L	270	F	157	S	n.q.	Y	23	R	21
2	T	267	K	88	R	28	K	318	R	n.q.	S	n.q.	R	15	C	n.q.
3	S	217	R	30	K	22	S	66	S	83	L	13	K	12	D	46
4	R	83	G	116	S	n.q.	R	56	H	8	K	10			D	48
5	K	341	C	n.q.	S	n.q.	K	171	L	23					A	46
6	L	294	S	n.q.			R	77							N	40
7	K	344	K	33			C	76							R	21
8	S	108	R	29			D	45							N	30
9	Q	194	K	n.q.			D	61							F	38
10	G	175					A	92							R	20
11	T	59					N	92							S	n.q.
12	R	38					R	49							H	12
13	R	53					N	51							L	n.q.
14	G	125														
15	K	168														
16	N	112														
17	R	34														
18	T	39														
19	P	67														
20	H	20														
21	K	96														
22	G	62														
23	V	50														
24	K	97														
25	R	23														
26	G	58														
27	C	n.q.														
28	S	21														
29	K	58														
30	R	17														
31	K	59														
32	Y	18														
33	R	18														
34	K	47														
35	S	15														
36	S	n.q.														
37	L	9														
38	K	34														

Table S2. *Edman degradation of thermolysin peptides Th2 and Thx from monophosphorylated ram TPI, performed on the gas-phase sequencer* These peptides were derivatized according to Meyer et al. [33]. Phenylthiohydantoin derivatives were identified and quantified by reverse-phase HPLC. n.q. not quantified

Cycle	Peptide			
	Th2		Thx	
	residue	amount	residue	amount
		pmol		pmol
1	Leu	433.2	Tyr	215.3
2	Lys	297.9	Arg	85.3
3	Ser/S-ethyl Cys	n.q.	Lys	n.q.
4	Gln	151.4	Ser/S-ethyl Cys	n.q.
5	Gly	140.9	Ser/S-ethyl Cys	n.q.
6	Thr	29.3	Leu	44.0
7	Arg	86.4	Lys	130.6
8	Arg	135.6	Ser/S-ethyl Cys	n.q.
9	Gly	58.5	Arg	n.q.
10	Lys	n.q.	Lys	n.q.
11	Asn	46.0	Arg	50.0
12	Arg	n.q.	S-ethyl Cys*	n.q.
13	Thr	n.q.	Asp	24.4
14	Pro	23.4	Asp	25.3
15	His	11.8	Ala	10.8

* At cycle 12, a phenylthiohydantoin derivative with a retention time identical to that of phenylthiohydantoin of S-ethyl cysteine was observed in place of phenylthiohydantoin of S-carboxamidomethyl cysteine.

a VG 11/250 data system (VG Analytical Ltd, Manchester, UK). Ionisation of the sample was performed with 1 mA of a 8-keV energy Xe atom beam. The peptides were dissolved in desionized water at a concentration of 1 nmol/ μ l. The matrix was thioglycerol/TFA 1%. 1 μ l matrix was deposited on a stainless steel target and 1 μ l peptide solution was added.

The wide-scan spectra were generated by classical exponential magnetic scanning at a resolution of 1000 and at an accelerating voltage of 8 keV.

ESMS. ESMS was performed on a VG BIO-Q quadrupole mass spectrometer with a mass range of 3000 Da. The mass spectrometer was scanned from $m/z = 100 - m/z = 1000$ in 15 s at unit resolution.

Calibration was performed using charged ions from poly(ethylene glycol) 800 obtained from a separate introduction.

The electrostatic spray ion source was operated at atmospheric pressure with an extraction cone voltage (V_c) value of 150 volts. The protein sample was first dissolved in water containing 5% acetic acid then an equivalent volume of methanol was added. The concentration used was between 30–100 pmol/ μ l. This solution (2–10 μ l) was introduced into the ion source at a flow rate of 2 μ l/min.

PDMS. The ^{252}Cf PDMS of peptide Th5 was obtained on a BIO-ION BIN 10 K Instrument (BIO-ION AB, Box 15045, Uppsala, Sweden). The spectrum was acquired for 10^6 fission events (15–20 min) or for 4×10^6 events. Data-handling procedures have been described by Sundqvist et al. [32]. The sample was dissolved (2.5 nmol/ μ l) in a water/methanol solution (4:1, by vol.) containing 0.1% TFA.

The sample (2 μ l) was deposited on the matrix and left to dry. Finally, the dry target was inserted into the mass spectrometer.

Nomenclature of peptides

Peptides obtained from cleavage with thermolysin and with endoproteinase Lys-C were designated as Th and K, respectively. Peptides were numbered according to their position in the amino acid sequence of the protein.

Conversion of phosphoserine to S-ethyl cysteine

The phosphoserine-containing thermolysin peptides were derivatized according to Meyer et al. [33] modified by Swiderek et al. [34]. 1 nmol peptide placed in a small glass tube was dissolved in 25 μ l derivatization mixture consisting of 0.4 ml water, 0.1 ml absolute ethanol, 65 μ l 5 M sodium hydroxide and 60 μ l ethanethiol.

The tube was flushed with argon, tightly closed and incubated at 50 °C for 1 h. After incubation the solution was evaporated to dryness. The derivatized peptide was then dissolved in 80 μ l water and applied directly on a polybrene-treated glass-fibre disk for sequence c' -termination.

The use of a small volume of derivatization mixture avoids clogging of the glass fibre disk by salts.

III. Conclusion:

L'ensemble de ces résultats montre clairement le rôle principal de la spectrométrie de masse dans la détermination de ces variants structuraux. En effet, lors du séquençage, par la méthode d'Edman, seul le résidu de cystéine à la position 27 a été retenu pour la séquence. Le résidu de glycine a été considéré comme le résiduel ("overlap") de la glycine à la position 26.

Le pic Th-3 obtenu après séparation des peptides thermolysiques en CLHP-phase inverse est parfaitement symétrique. Et c'est seulement lors de l'analyse en FAB-MS que l'existence de variants a été mise en évidence par l'obtention de 2 masses.

Le spectre de masse de la protéine TP1 obtenu en électrospray montre 2 séries de pics. Une famille de masse 6460 Da correspondant au variant majeur avec la cystéine et une famille de masse 6357 Da correspondant au variant mineur avec la glycine.

Les cystéines sont sous la forme de S-carboxamidométhylcystéine (160 Da) (Glycine: 57 Da). Ce qui explique la différence observée en électrospray et en FAB.

Toutes les masses des différents peptides thermolysiques ont été vérifiées en FAB (Fig.13-17) sauf celle du peptide Th-5. Le peptide Th-5 n'a pas donné de signal en FAB très probablement pour des raisons de quantité insuffisante. Il a été vérifié en PD-MS dans le laboratoire du Professeur Roepstorff à Odense au Danemark.

En conclusion, on peut dire que la protéine TP1 de bélier est une protéine de 54 résidus, qui ne contient pas de modification post-traductionnelle en dehors de la phosphorylation. Par contre, elle existe sous la forme de deux variants structuraux qui ne diffèrent que par un seul acide aminé. Elle présente une forte homologie structurale avec les protéines TP1 des autres espèces. Enfin, les sites de phosphorylation ont été localisés en réalisant sur la protéine phosphorylée la réaction de Meyer qui consiste en la conversion de la phosphosérine en S-éthylcystéine. La faible quantité de protéine disponible, sous la forme phosphorylée, ne nous a pas permis d'effectuer des spectres de masses.

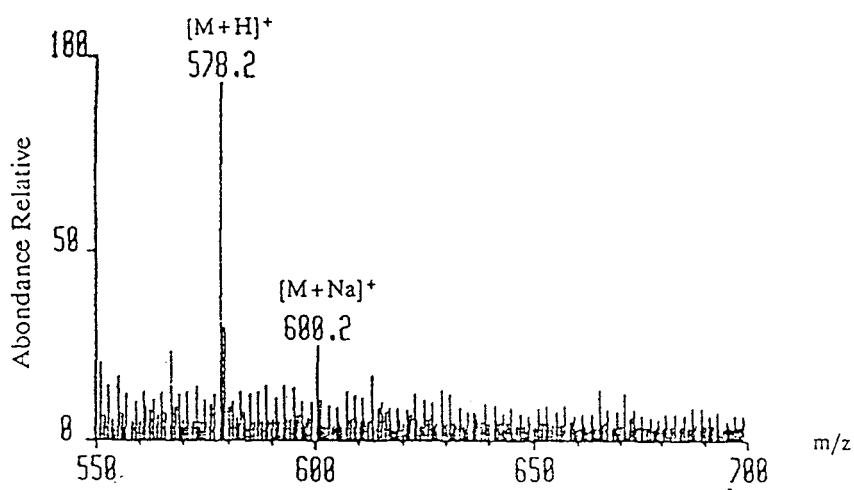
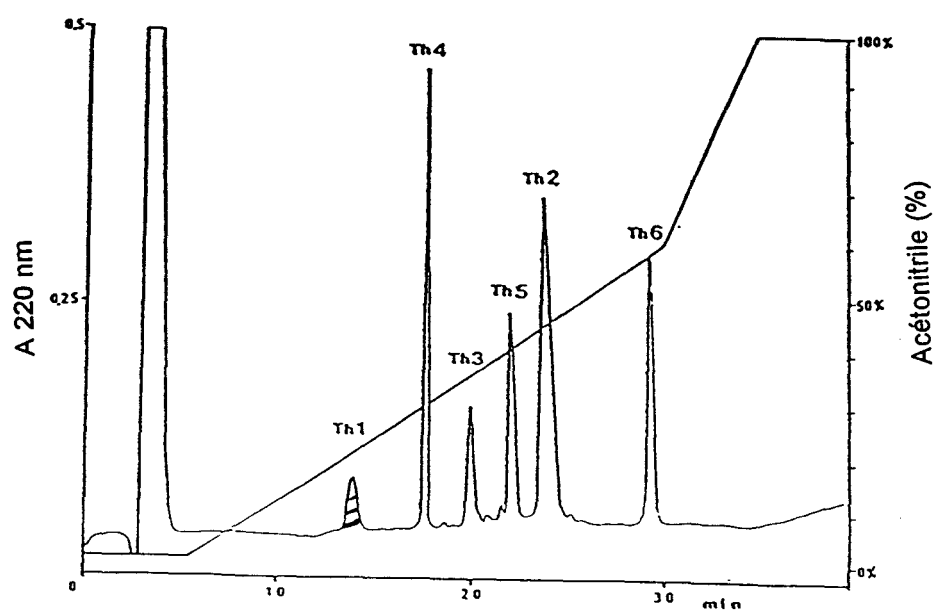


Figure 13: Spectre FAB + du peptide Th1 de séquence STSRK et de masse attendue $(M+H)^+ = 578.3$

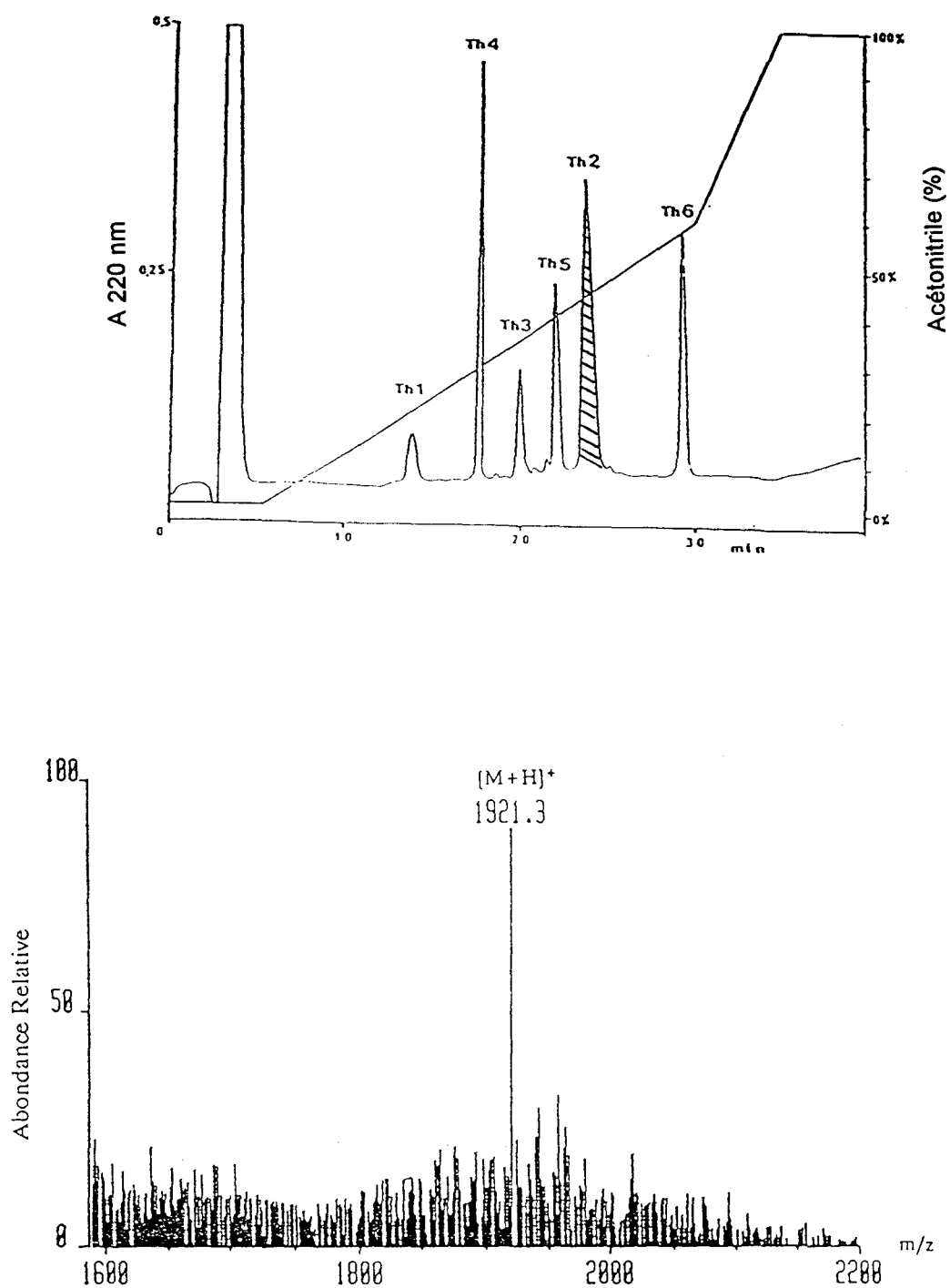


Figure 14: Spectre FAB + du peptide Th2 de séquence
 LKSQGTRRGKNRTPHKG
 et de masse attendue $(M+H)^+ = 1921.1$

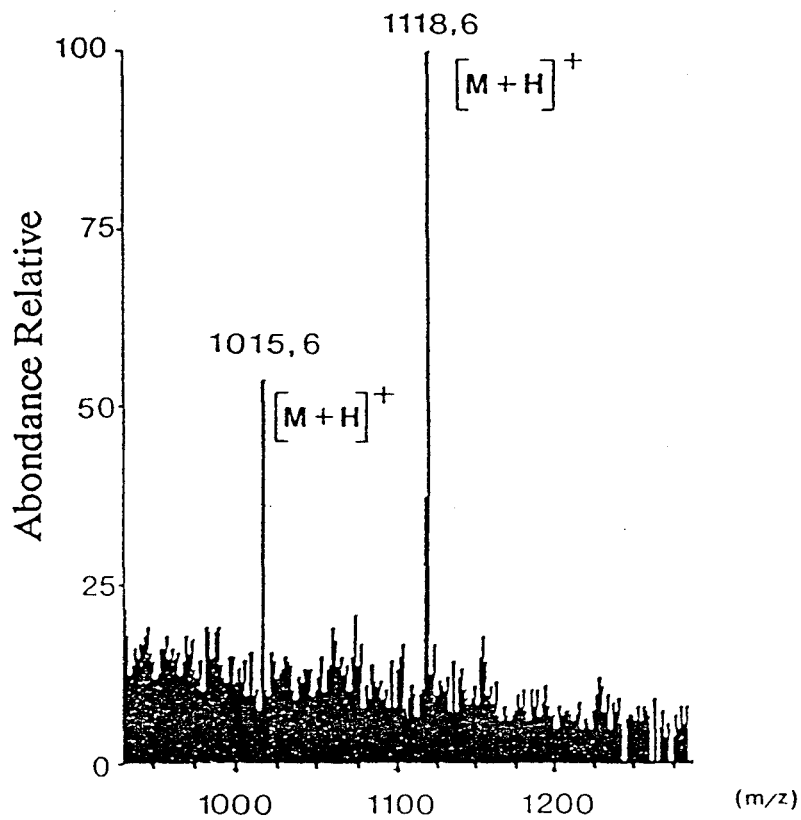
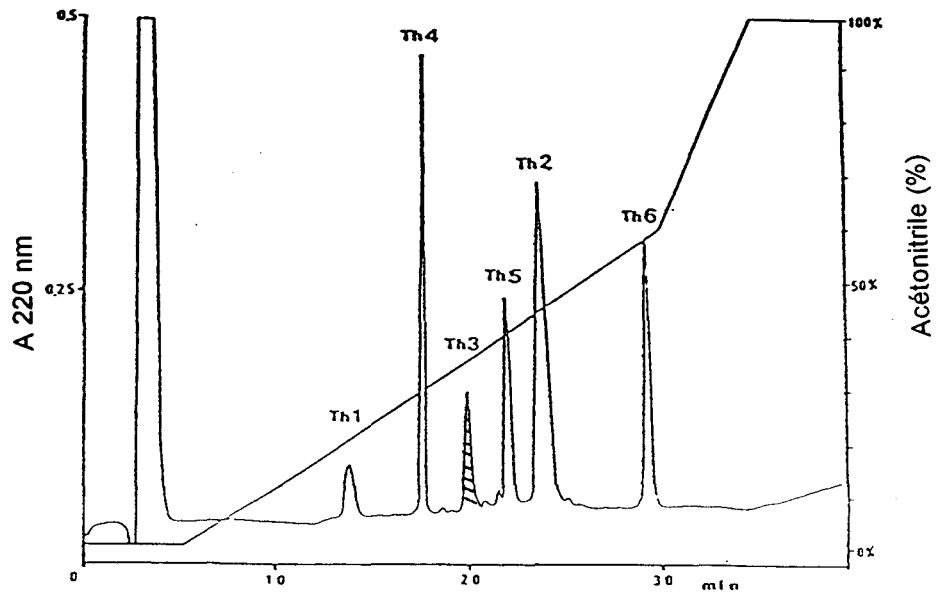


Figure 15: Spectre FAB + du peptide Th3 de séquence VKRGCSKRK et de masse attendue $(M+H)^+ = 1118.6$
 Masses observées: 1118.6 et 1015.6. Existence de 2 variants structuraux de séquences VKRGCSKRK et VKRGGSSKRK

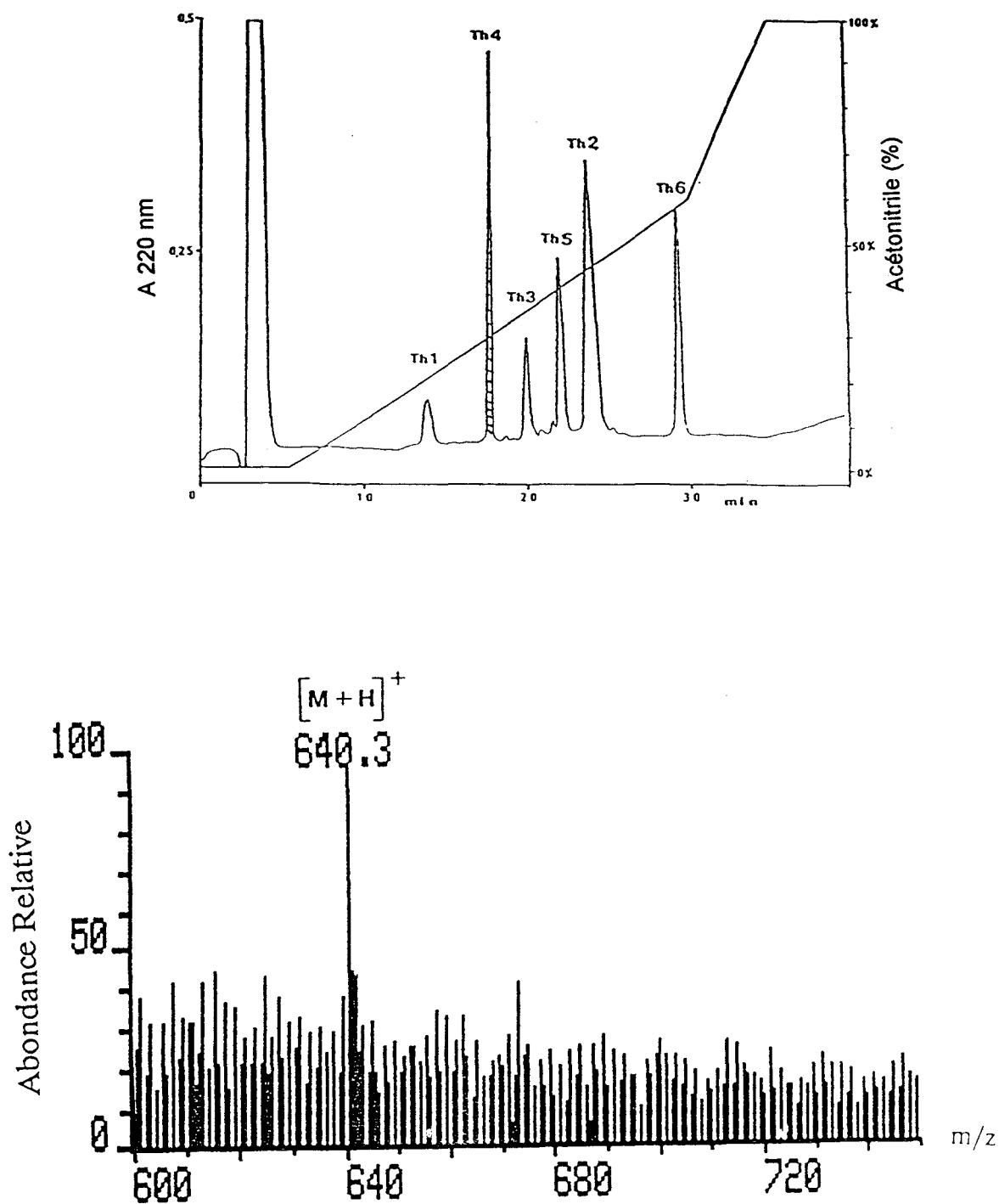


Figure 16: Spectre FAB + du peptide Th4 de séquence YRKSS et de masse attendue $(M+H)^+ = 640.3$

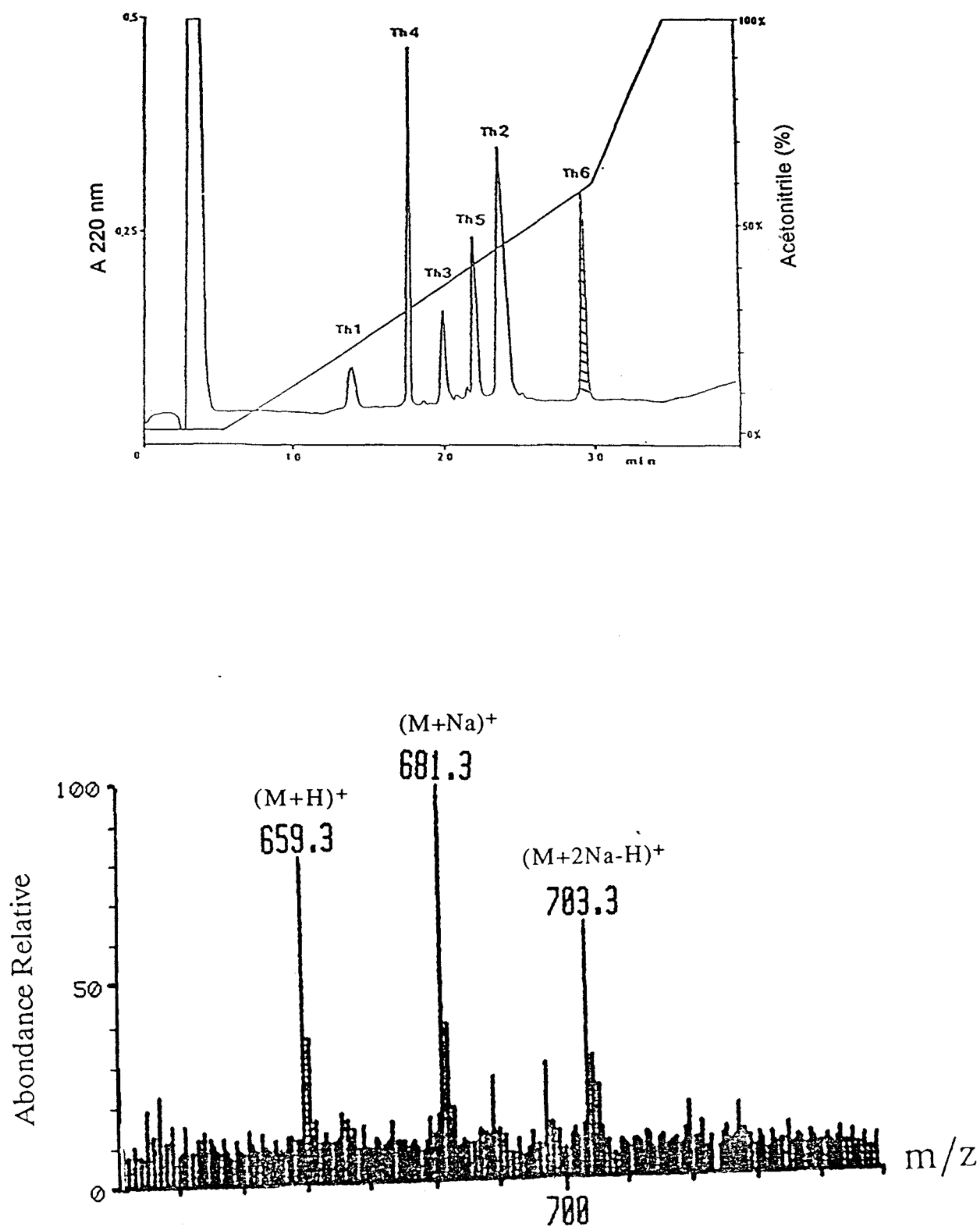


Figure 17: Spectre FAB + du peptide Th6 de séquence FRSHL et de masse attendue $(M+H)^+ = 659.4$

Chapitre 5

Contribution de la
spectrométrie de masse à
l'étude structurale d'une
protéine

Etude de la protéine de transition TP2 de bélier

I. Introduction:

Après l'élucidation de la structure primaire de la protéine intermédiaire TP1 (Chirat et *al.* (28)), nous avons entrepris les études structurales de la protéine intermédiaire TP2. Malheureusement, nous n'avons pas pu mener à terme ce travail et ceci pour 3 raisons principales:

- la faible quantité de protéine dans les spermatides,
- les difficultés de purification de la protéine,
- le blocage de l'extrémité NH₂-terminale de la protéine TP2.

Nous présenterons dans ce chapitre les résultats partiels obtenus sur la séquence de cette protéine.

II. Préparation de la protéine TP2:

Les testicules de bélier, congelés dans l'azote liquide et conservés à -80°C, nous parviennent de la station de fertilité mâle de l'INRA à Nouzilly-Tours.

Le protocole de préparation que nous avons utilisé comprend deux étapes: la préparation des noyaux et l'extraction des protéines nucléaires basiques (Fig.18). Toutes ces manipulations sont effectuées à +4°C. Les testicules décongelés dans un tampon, contenant de l'EDTA 7mM et des antiprotéases, sont broyés puis filtrés sur des filtres de nylon de porosités différentes (48µm puis 25µm). Par centrifugation, on sépare les noyaux non résistants (surnageant) des noyaux résistants (culot) à l'EDTA. Ces derniers correspondent au stade 9-15 de la spermiogenèse (Fig.12 page 48). Le culot subit ensuite plusieurs lavages dans des tampons contenant du sucrose puis du Triton X100 0.1%. Après chaque lavage, la pureté des noyaux est contrôlée au microscope en contraste de phase. Les histones sont ensuite extraites par l'acide sulfurique 0.2M. Les protéines intermédiaires ainsi que la protamine contenues dans le culot sont alors réduites. Le DNA est précipité à pH acide, puis éliminé après centrifugation. Les protéines intermédiaires et la protamine sont séparées selon leur solubilité dans l'acide perchlorique (PCA). La fraction soluble contient les protéines

Protocole de préparation des protéines intermédiaires
du testicule de bélier

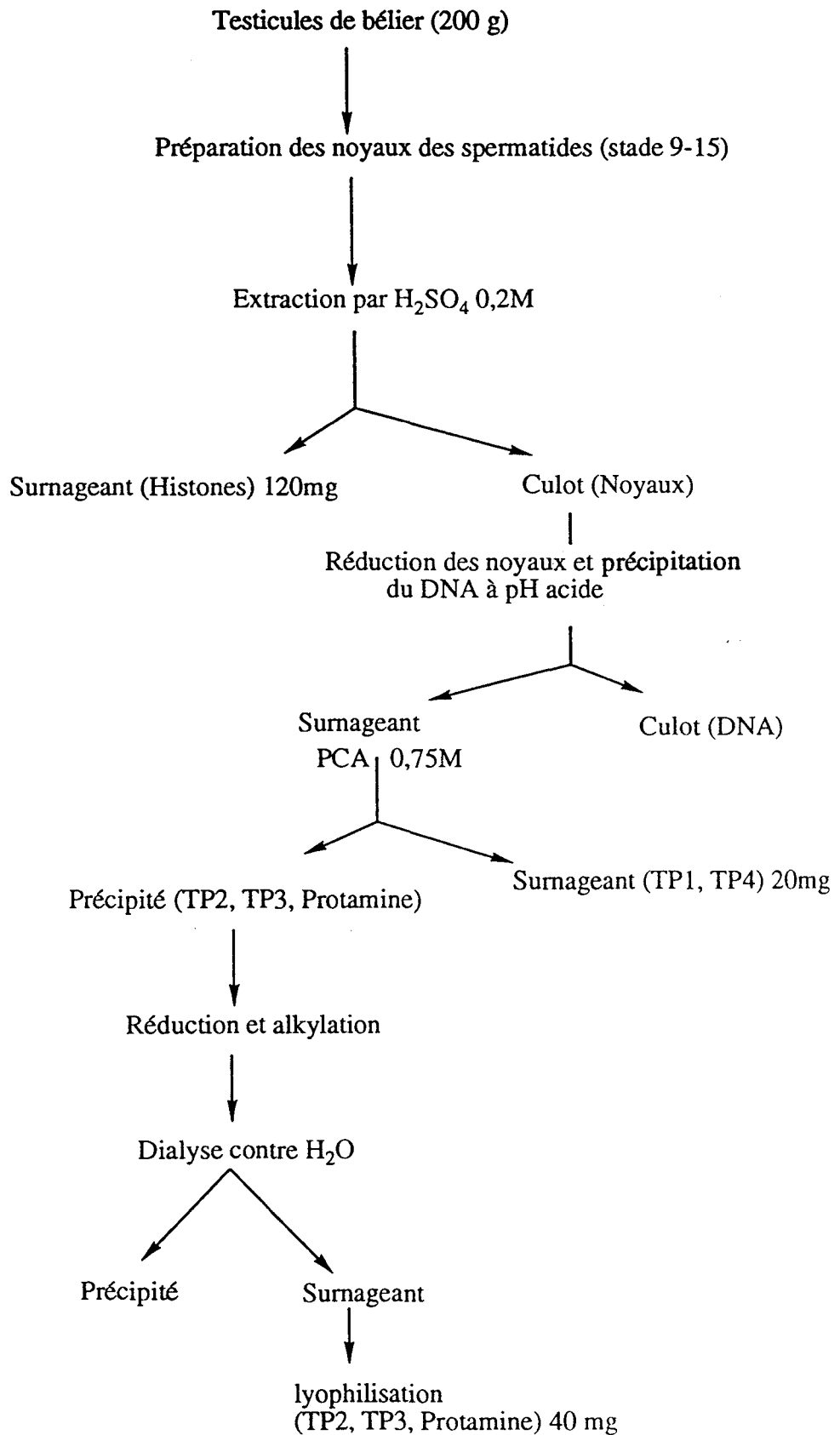


Figure 18

intermédiaires TP1 et TP4. La fraction insoluble renferme les protéines intermédiaires TP2, TP3 ainsi que la protamine (Fig.20 couloir 2). Les protéines PCA insolubles sont ensuite réduites et alkylées puis dialysées contre l'eau et enfin lyophilisées.

III. Réduction et alkylation des protéines:

III.1. Par l'iodoacétamide:

Les protéines sont dissoutes dans une solution contenant du Tris/HCl 0.1M, pH 8.0, GuCl 6M et 2-mercaptoéthanol 0.1M, puis incubées à 37°C, 4h sous agitation en atmosphère d'azote. L'iodoacétamide est ensuite ajouté à la solution pour atteindre la concentration de 0.5M. Le pH est réajusté à 8.0. La solution, incubée pendant 2h à 37°C sous agitation sous azote à l'obscurité, est ensuite dialysée puis lyophilisée.

III.2. Par le N-(Iodoéthyl)-trifluoroacétamide (IE-TFA):

Les protéines sont dissoutes dans une solution contenant du Tris/HCl 50mM pH 8.6, GuCl 5M, 2-mercaptoéthanol 0.1M, puis incubées à 50°C pendant 4h sous agitation en atmosphère d'argon. L'IE-TFA dissous dans le méthanol est ajouté à la solution pour atteindre la concentration de 0.25M. La réaction est effectuée pendant 2h30 sous agitation, en atmosphère inerte à l'obscurité. Les protéines réduites et alkylées sont ensuite dialysées contre de l'eau puis lyophilisées.

IV. Contrôle électrophorétique:

L'analyse des protéines a été réalisée en gel de polyacrylamide à 17%, urée 2.5M (186) (mini gels de 8x10cm). Une pré-électrophorèse de 1h est réalisée avant le dépôt des protéines. La migration, d'une durée comprise entre 1 et 2h, s'effectue en milieu acide acétique 0.9M sous un courant de 22mA. Pour la révélation, les gels sont ensuite laissés une nuit dans une solution de bleu de coomassie G250 à 4% dans un mélange eau/éthanol/acide acétique (45:45:10 v/v/v). La décoloration est assurée par des lavages successifs dans un mélange eau/éthanol/acide acétique (77.5:12.5:10 v/v/v).

V. Purification de la protéine TP2:

Nous avons réalisé la purification de la protéine TP2 en deux étapes:

- purification sur colonne d'échange d'ions,
- purification en CLHP phase inverse.

V.1. Purification sur colonne d'échange d'ions CM52:

Les protéines réduites et alkylées sont progressivement, après dialyses successives, amenées en milieu GuCl 0.3M dans l'acétate de sodium 50mM pH5.0. La solution protéique (70mg) est déposée sur une colonne de carboxyméthylcellulose CM52 équilibrée dans le même milieu. Les protéines sont séparées par un gradient discontinu de GuCl 0.3M, 0.52M et 0.8M. Cette première étape de purification consiste à séparer de la protamine les deux protéines intermédiaires TP2 et TP3. Le diagramme d'élution des protéines est présenté dans la figure 19. Un contrôle électrophorétique est effectué sur toutes les fractions (Fig. 20). Les protéines intermédiaires TP2 et TP3 présentes dans la fraction 3 (couloir 5) sont séparées en CLHP phase inverse.

V.2. Séparation des protéines intermédiaires TP2 et TP3 en CLHP-phase inverse:

0.2mg de mélange TP2 et TP3 sont déposés sur une colonne RP-18 "Encapped" (244x4mm, 5 μ m). L'élution est réalisée par un gradient d'acétonitrile de 0 à 50% dans l'acide trifluoroacétique (TFA) à 0.1% avec un débit de 1ml/min. (Fig.21). L'absorbance est mesurée à 220nm. Le contrôle électrophorétique montre que la fraction 1 correspond à la protéine TP2.

VI. Composition en acides aminés de la protéine TP2:

1 nmol. de la protéine TP2 a été hydrolysée par l'HCl 6M à 110°C pendant 24h dans un tube purgé à l'azote et scellé sous vide. Une goutte de phénol à 1% est ajoutée à l'acide pour éviter la destruction de la tyrosine. La séparation et la quantification des acides aminés sont réalisées sur un analyseur Beckman modèle 6300. Les résultats sont présentés dans le tableau II. La composition est caractérisée par sa richesse en résidus basiques (35%) et par la présence de nombreux résidus de sérine.

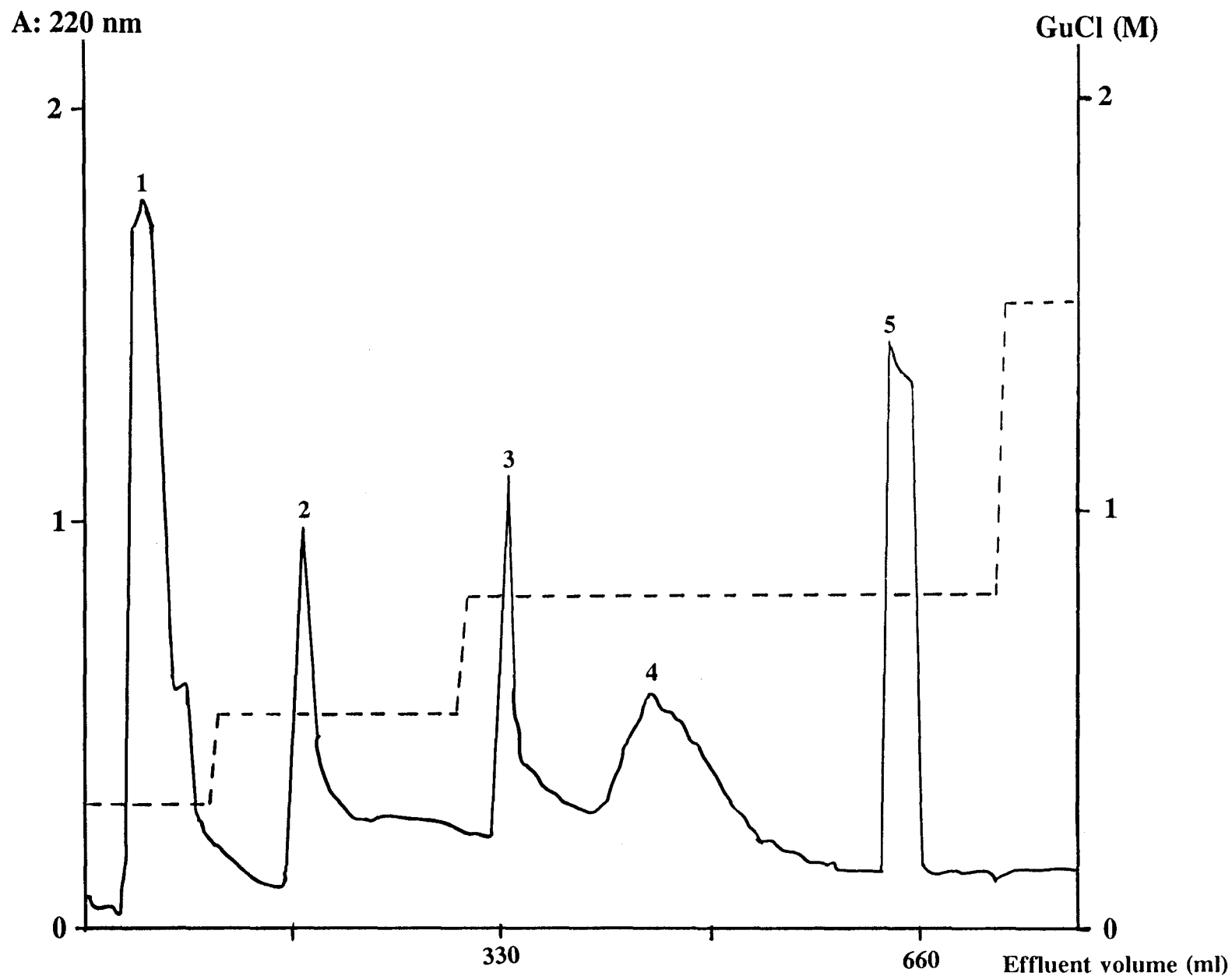


Figure 19: Séparation des protéines PCA insolubles (TP2, TP3, et protamine) sur colonne de CM-cellulose (CM 52).

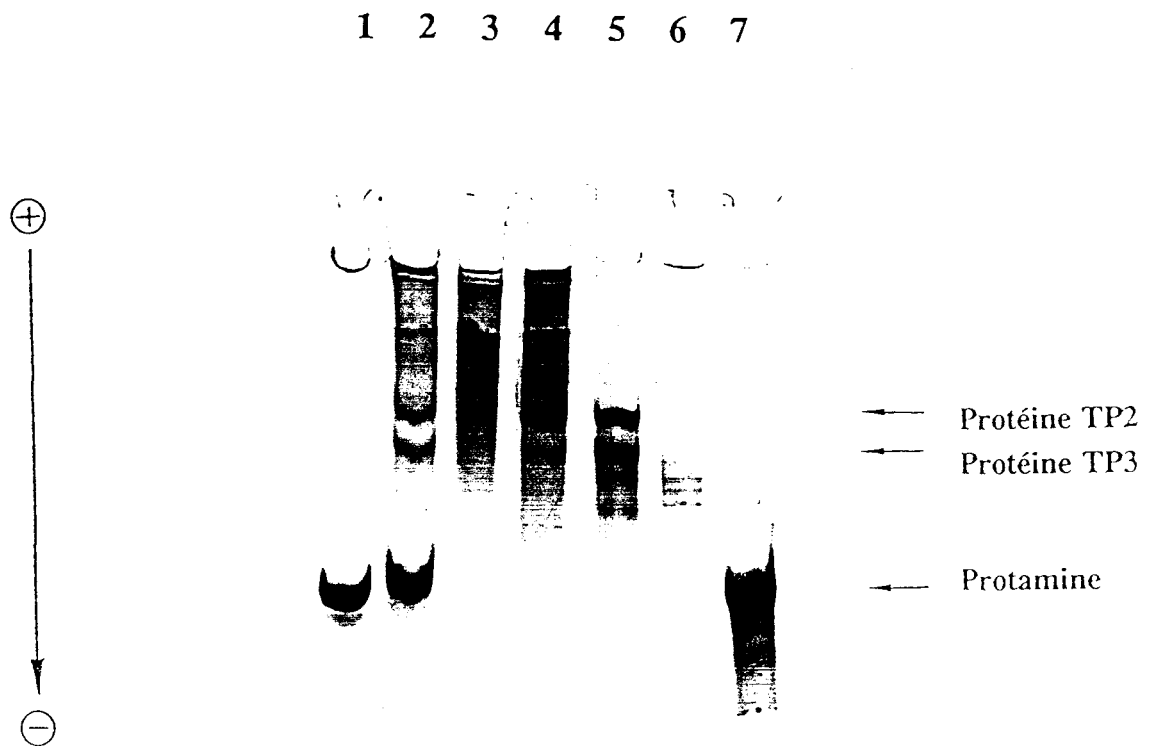


Figure 20: Contrôle électrophorétique en gel de polyacrylamide à 17% en milieu acide acétique 0.9M, urée 2.5M des protéines PCA insolubles (couloir 2) et des fractions éluées de la colonne de CM-cellulose. Les couloirs 3 à 7 correspondent respectivement aux fractions 1 à 5. Le couloir 1 correspond à la protamine isolée des spermatozoïdes éjaculés de bélier.

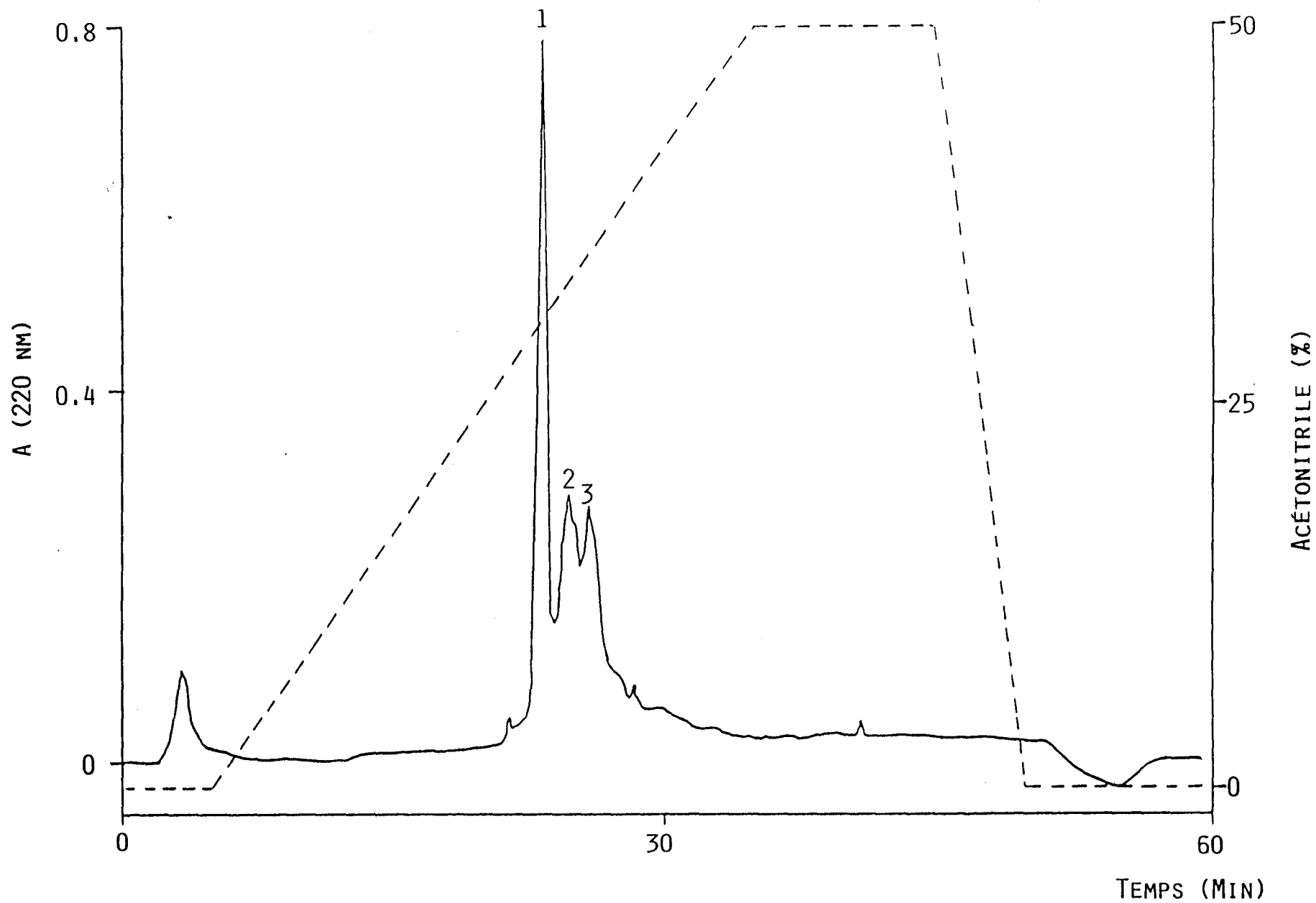


Figure21: Séparation des protéines intermédiaires TP2 et TP3 de bœuf par CLHP-phase inverse.

Tableau II

**Composition en acides aminés de la protéine de transition TP2
isolée de testicule de bélier.**

Acides aminés	TP2
Asp	5.49 ^a (7) ^b
Thr	3.83 (5)
Ser	19.06 (23)
Glu	8.71 (10-11)
Pro	5.62 (7)
Gly	5.84 (7)
Ala	2.47 (3)
Cys	4.52 (5-6)
Val	2.92 (3-4)
Met	0.94 (1)
Ile	0.58 (1)
Leu	2.38 (3)
Tyr	2.18 (2-3)
Phe	-
His	7.51 (9)
Lys	8.43 (10)
Arg	19.44 (24)
Total	120-124

a: les résultats sont exprimés en moles/100 moles.

b: le nombre de résidus entre les parenthèses a été déterminé sur la base d'un résidu d'isoleucine par molécule de protéine.

VII. Dégradation récurrente d'Edman:

1 nmol. de la protéine TP2 réduite et alkylée a été soumise à la dégradation récurrente d'Edman sur un séquenceur Applied Biosystems 470A couplé à un analyseur de PTH-aminoacides Applied Biosystems 120A. Aucun PTH-aminoacide n'a été détecté ce qui montre que l'extrémité NH₂-terminale de la protéine intermédiaire TP2 est bloquée.

VIII. Analyse de la protéine TP2 en électrospray:

La protéine TP2 est dissoute dans un mélange eau/méthanol (50:50, v/v) contenant 1% d'acide acétique, à la concentration de 20 pmoles/ μ l. 5 à 10 μ l de la solution sont introduits dans la source d'ions à un débit de 2 μ l/min. Le voltage du cône d'extraction des ions (55 volts) a été augmenté à 130V afin d'éliminer ou de diminuer les interactions non covalentes (adduits de phosphate ou sulfate etc...). Le spectre de masse obtenu est présenté dans la figure 22. Une famille d'ions chargés de 21 à 25 fois a été détectée donnant une masse moyenne de 16380 $\pm$ 9 uma. Cette masse est nettement supérieure à celle calculée à partir de la composition en acides aminés (autour de 14700 Da). De telles données ne permettent pas d'avancer l'existence ou non de modifications chimiques ou post-traductionnelles.

Dans le but d'obtenir des informations structurales sur cette protéine, nous l'avons soumise à différentes hydrolyses enzymatiques.

IX. Hydrolyse de la protéine TP2 par la thermolysine:

15 nmoles de protéine TP2 réduite et alkylée sont soumises à une hydrolyse par la thermolysine. La réaction est effectuée dans du bicarbonate d'ammonium 0.1M, pH 8.5 pendant 1h à 37°C avec un rapport enzyme/substrat = 1/50 (p/p). La réaction est arrêtée par addition d'acide formique. Le diagramme d'élution des peptides sur une colonne CLHP-phase inverse, est présenté dans la figure 23-A.

Les fractions 3, 4 et 5 (Fig.23-A) ont été purifiées en CLHP-phase inverse (Fig. 23-B, C et D). Différentes fractions ont été soumises à une composition en acides aminés, à la dégradation récurrente d'Edman et à l'analyse en

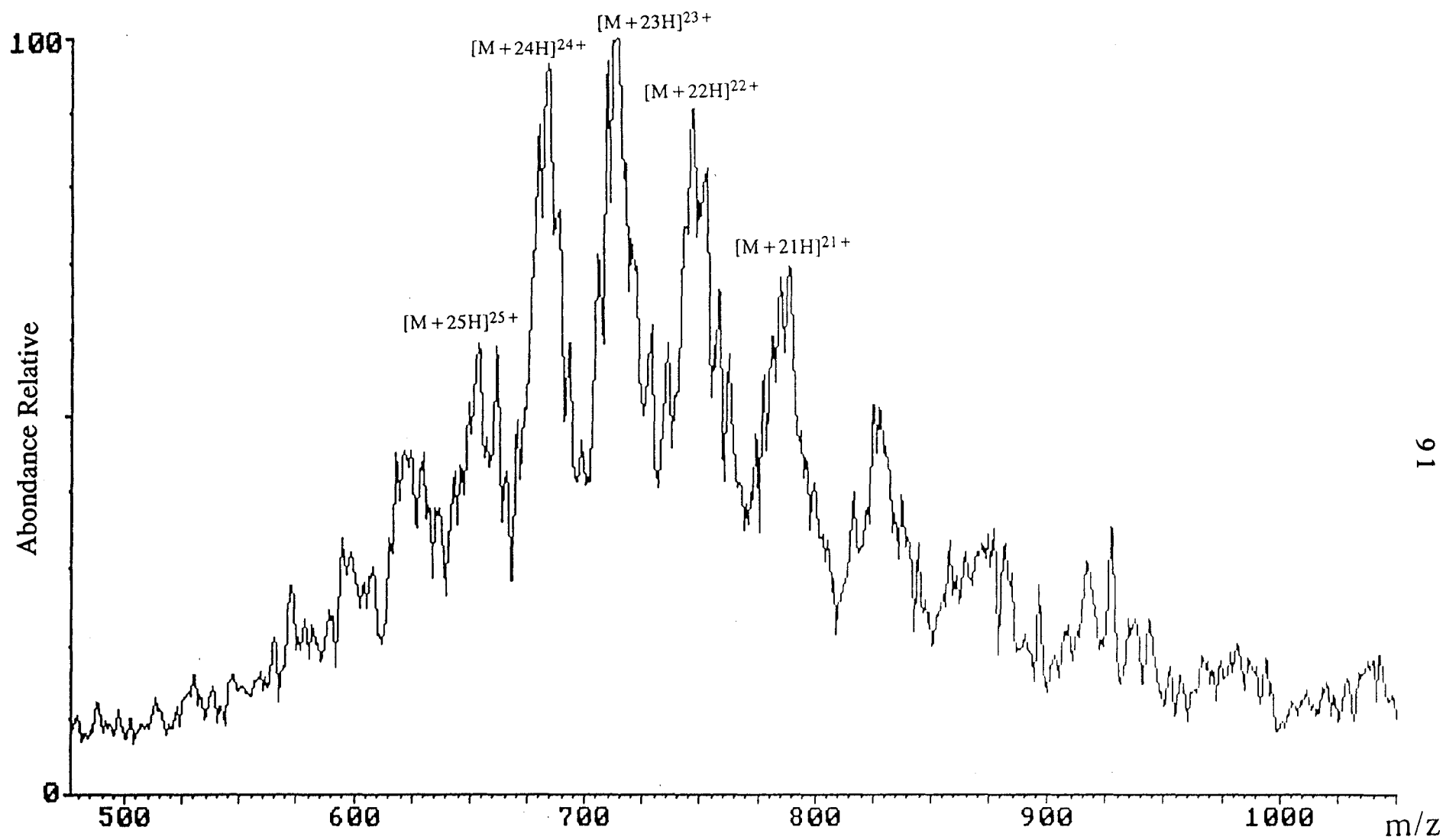


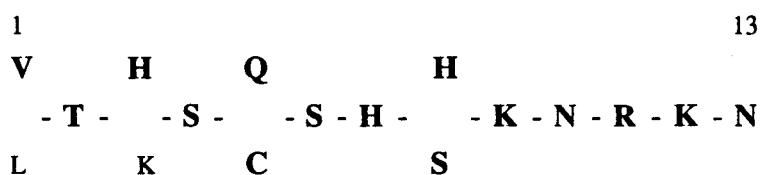
Figure 22: Spectre ES-MS de la protéine intermédiaire TP2 de bélier.

spectrométrie de masse (mode FAB). Les résultats exploitables que nous avons obtenus sont les suivants:

IX.1. Etude de la fraction B-1:

La dégradation d'Edman de la fraction B-1 (Fig.23) a donné les résultats suivants:

(les acides aminés dominants sont présentés en taille supérieure et en caractère **gras**)



La figure 24 montre le spectre de masse d'accumulation (MCA) obtenu sur ce peptide en mode FAB positif. L'analyse est faite sur 0.75 nmole de peptide. La matrice utilisée est le thioglycérol/TFA 1%. La masse 1554.8 uma observée en FAB, nous a permis de lever les ambiguïtés de séquence. En effet, seule la séquence:



correspond à la masse observée. La masse calculée est 1554.75 Da. La cystéine est sous la forme carboxamidométhylcystéine.

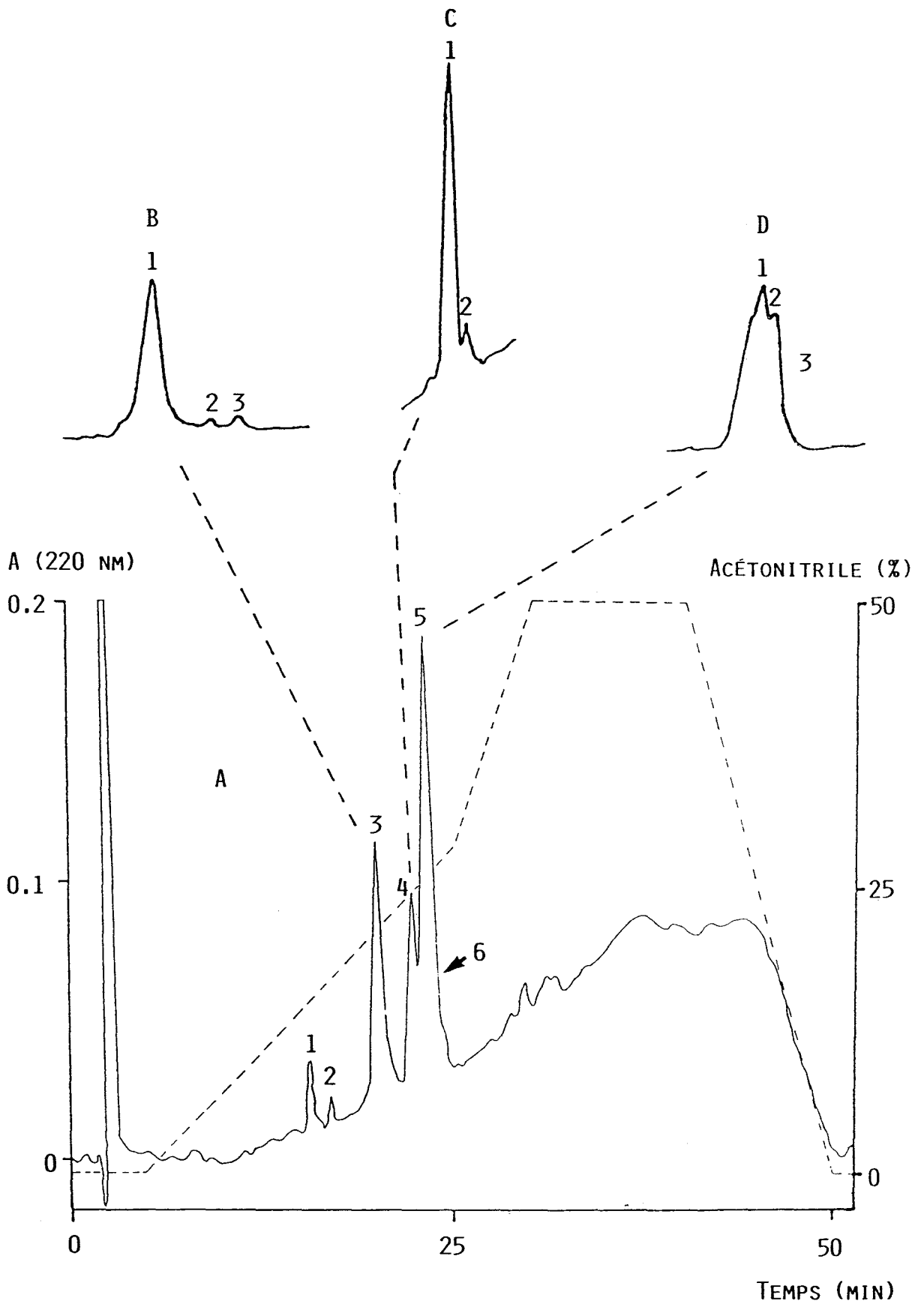


Figure 23: A: Séparation en CLHP-phase inverse des peptides thermolysiques de la protéine intermédiaire TP2 de bœuf.
 B, C et D représentent respectivement le refractionnement en CLHP-phase inverse des fractions 3, 4 et 5.

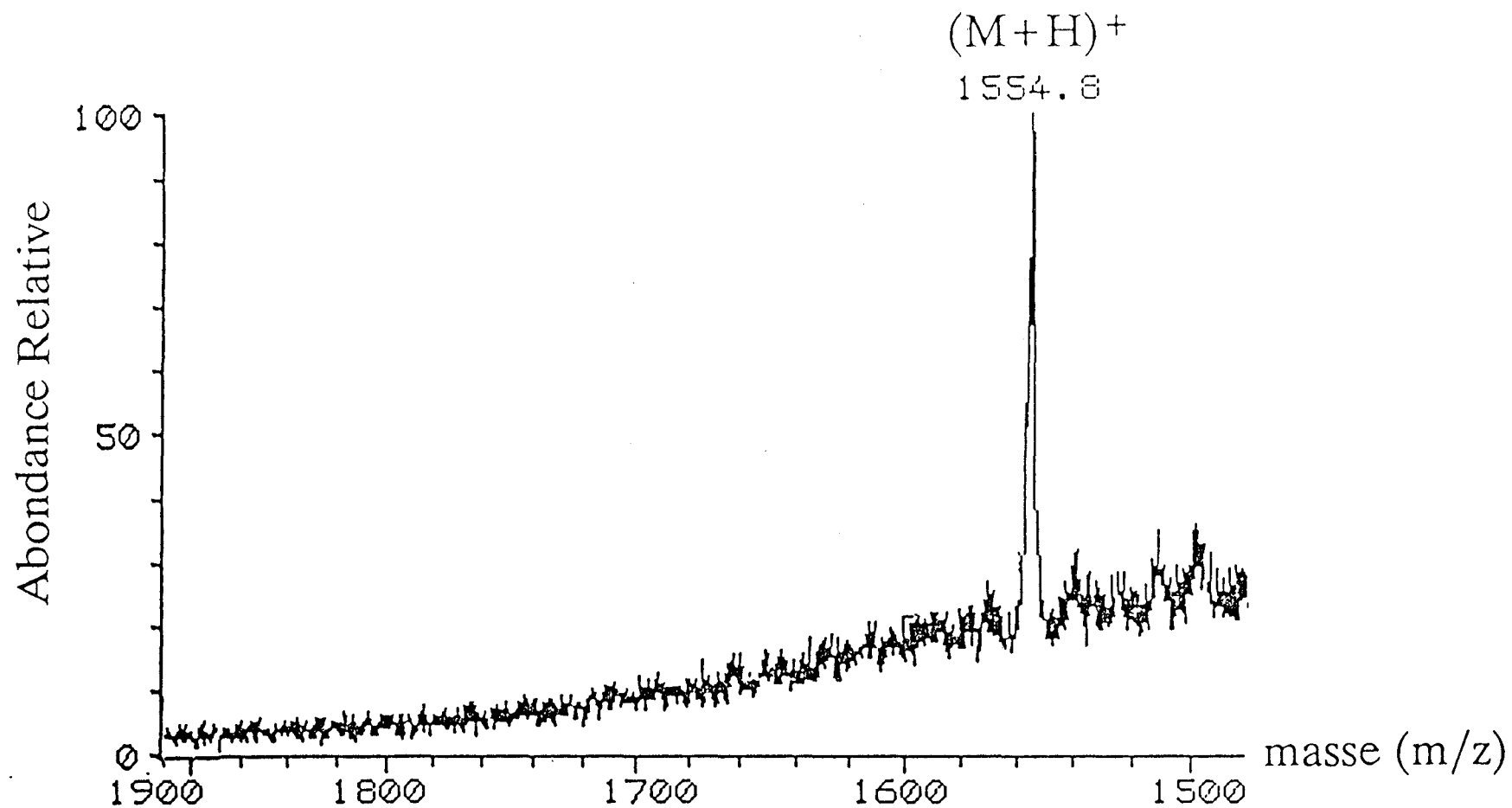


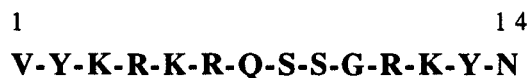
Figure 24: Spectre FAB-MS du peptide thermolysique (fraction B-1) de la protéine intermédiaire de bélier.

IX.2. Etude de la fraction D-3:

Le séquençage de la fraction D-3 (Fig.23 page 93) a donné les résultats suivants:



Ce résultat laisse penser qu'il peut s'agir de 2 peptides dont un majeur, de séquence :



et un mineur, de séquence :



La figure 25-A montre le spectre de masse d'accumulation (MCA) obtenu sur ce mélange peptidique en FAB positif. L'analyse est faite sur 0.5 nmole dans le glycérol/eau/TFA 10:88:2 (v/v/v). Le résultat obtenu confirme l'hypothèse avancée. Cependant, le pic à 1772 uma correspondant au peptide majeur ($I_{BB} + 239$) a une intensité relative inférieure au peptide mineur de masse 1198 uma ($I_{BB} + 51$). Cet exemple illustre bien le phénomène de suppression d'un peptide à caractère hydrophile par un peptide à caractère plus hydrophobe en mode FAB.

La figure 25-B montre les allures isotopiques de ces deux peptides. Les pics isotopiques du peptide le plus lourd sont moins résolus que ceux du peptide à 1198 Da. Il est donc préférable pour les masses élevées de mesurer la masse chimique qui correspond à la moyenne des pics isotopiques.

IX.3. Etude de la fraction D-2:

La présence de plusieurs phénylthiohydantoïnes d'acides aminés (PTH-AA) observés à chaque cycle de dégradation lors du séquençage de la fraction D-2 (Fig.23), montre que cette fraction contient plusieurs peptides.

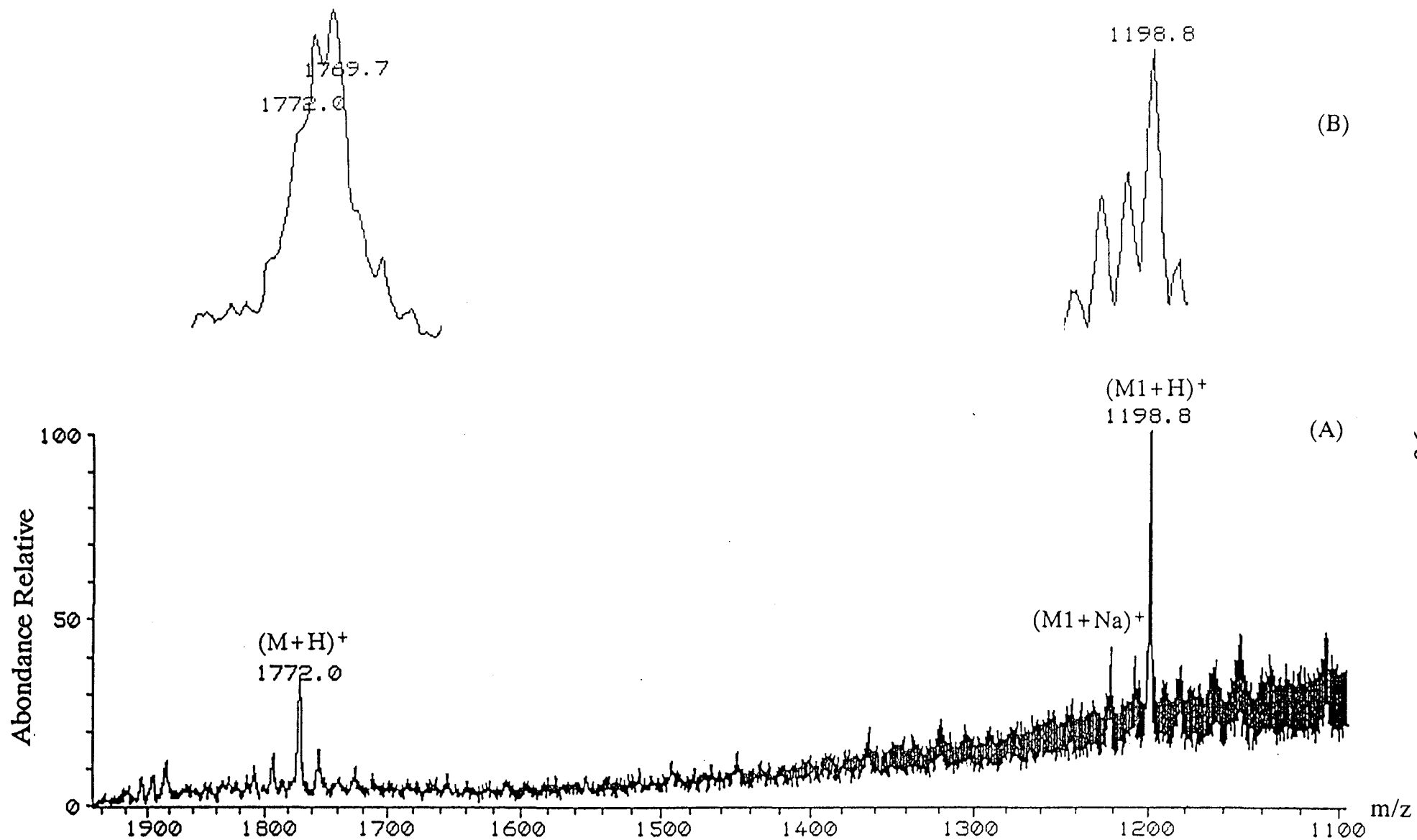
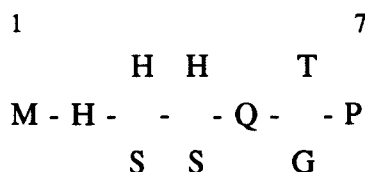


Figure 25: (A)-Spectre FAB-MS des peptides thermolysiques (fraction D-3) de la protéine intermédiaire de bœuf.
(B)-Allures isotopiques.

Dans ce mélange peptidique on retrouve les 2 peptides de la fraction D-3 de séquences: V-Y-K-R-K-R-Q-S-S-G-R-K-Y-N et L-E-G-K-V-I-K-R-K-Q.

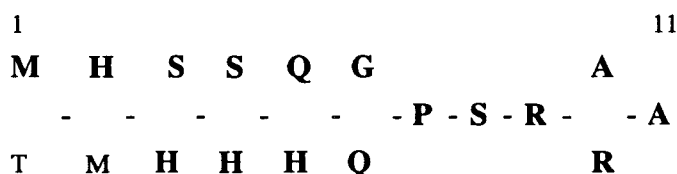
La séquence probable suivante a pu également être dégagée de ce mélange:



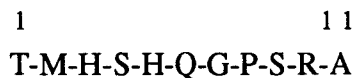
L'analyse en FAB-MS de ce mélange peptidique ne nous a pas permis de lever cette ambiguïté de séquence. En effet, seule la masse 1198 Da correspondant au peptide : L-E-G-K-V-I-K-R-K-Q a été observée (Fig.26). L'analyse est faite dans le glycérol/eau/TFA 10:88:2 sur une quantité de peptide inférieure à une nmole. On observe le pic de cationisation (M+Na)⁺ en plus du pic (M+H)⁺.

IX.4. Etude de la fraction C-1:

Le séquençage de la fraction C-1 (Fig.23 page 93) a donné les résultats suivants:



L'analyse de ces peptides en FAB-MS dans le glycérol/eau/TFA, n'a pas donné de spectre interprétable. Nous n'avons pas pu continuer cette étude à cause de la faible quantité de peptide disponible. Cependant, d'après les résultats de séquence, il peut s'agir d'un mélange de deux peptides qui ne diffèrent que par la présence du résidu de thréonine. Ceci explique le décalage des résultats de la dégradation d'Edman. La séquence probable serait alors:



D'autre part, les résultats obtenus sur la fraction D-2 corroborent cette proposition de séquence.

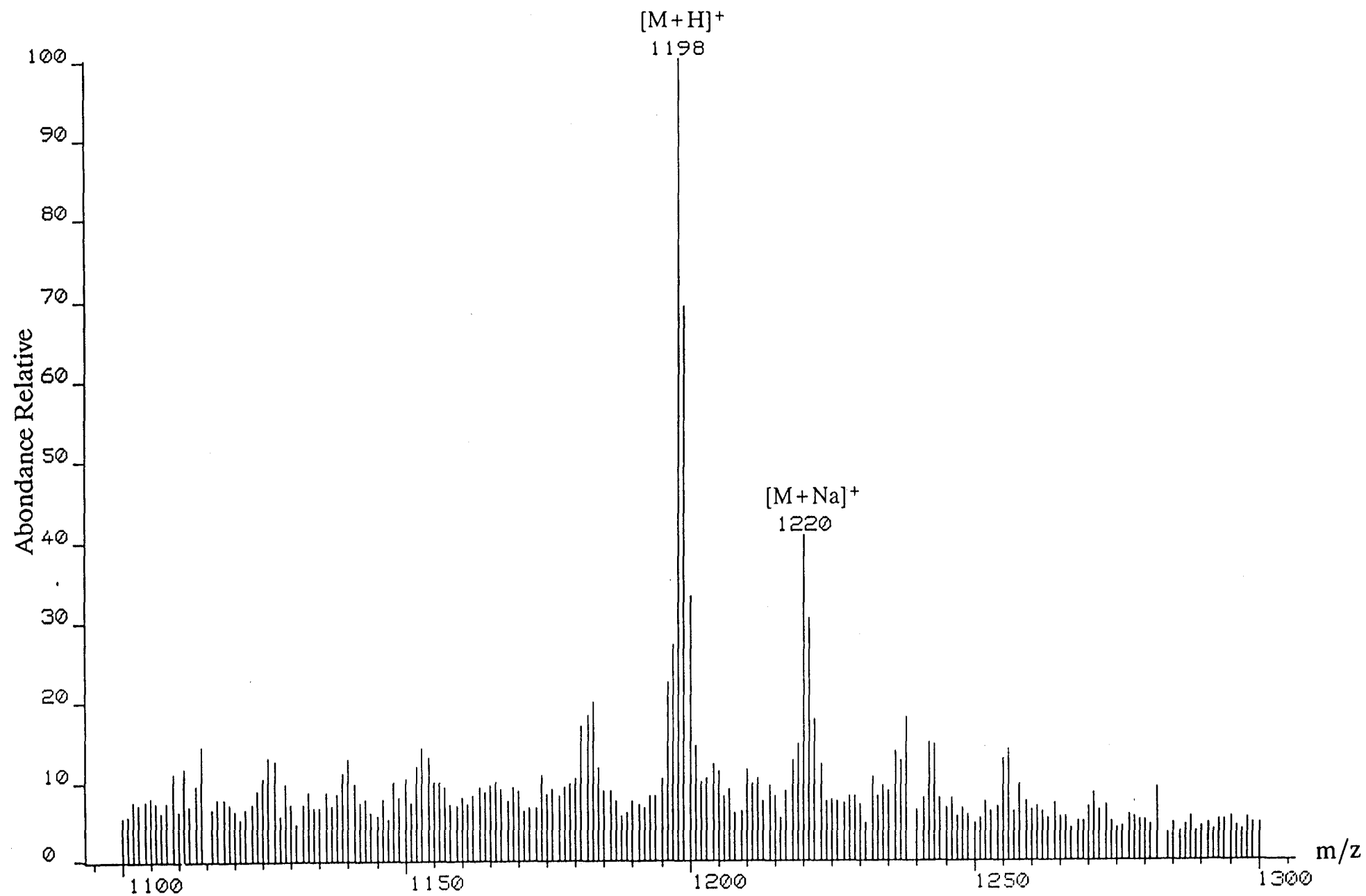


Figure 26: Spectre FAB-MS des peptides thermolysiques (fraction D-2) de la protéine intermédiaire de bélier. Seul le peptide de séquence LEGKVIK RKQ est observé.



IX.5. Etude de la fraction A-6:

Le séquençage de cette fraction A-6 (Fig.23 page 93) montre qu'elle est constituée d'un mélange peptidique avec les 2 peptides présents dans la fraction D-3, plus éventuellement un dipeptide de séquence T-K. Là encore, seule la masse 1198 Da a été observée en FAB.

Le séquençage, réalisé sur les autres fractions, a montré l'hétérogénéité de celles-ci et leur analyse en FAB-MS ne nous a pas donné de résultat interprétable. Il faut toutefois souligner que la quantité disponible pour la réalisation d'un spectre était le plus souvent inférieure à 1 nmole. Par conséquent, des essais avec plusieurs types de matrice étaient donc impossibles. La matrice glycérol/eau/TFA (10:88:2) a été le plus souvent utilisée.

X. Hydrolyse de la protéine TP2 par la protéase staphylococcique:

La protéine TP2, ne renfermant qu'un seul résidu d'acide glutamique (Tableau II, page 89), a été soumise à une coupure par la protéase staphylococcique V8 à pH 4.0 en milieu acétate d'ammonium 0.1M, pendant 16h à 37°C avec un rapport enzyme/substrat = 1/50 (p/p). Les peptides obtenus ont été séparés en CLHP-phase inverse (Fig.27). Les 2 fractions obtenues ont été analysées en composition en acides aminés et en gel de polyacrylamide à 17% en milieu acide/urée. La fraction 1 correspond, d'après les résultats obtenus, au peptide C-terminal alors que la fraction 2 renferme le peptide N-terminal ainsi que la protéine non hydrolysée. Nous avons ensuite soumis le peptide de la fraction 1 à la dégradation automatique d'Edman. Le résultat présenté dans la figure 28, est en accord avec celui obtenu par F. Chirat (187). Les résultats de séquence, obtenus sur ce peptide C-terminal, nous permettent de placer les peptides thermolysiques de la fraction D-3 et de prolonger ainsi la séquence du côté NH₂-terminal, du dipeptide L-E.

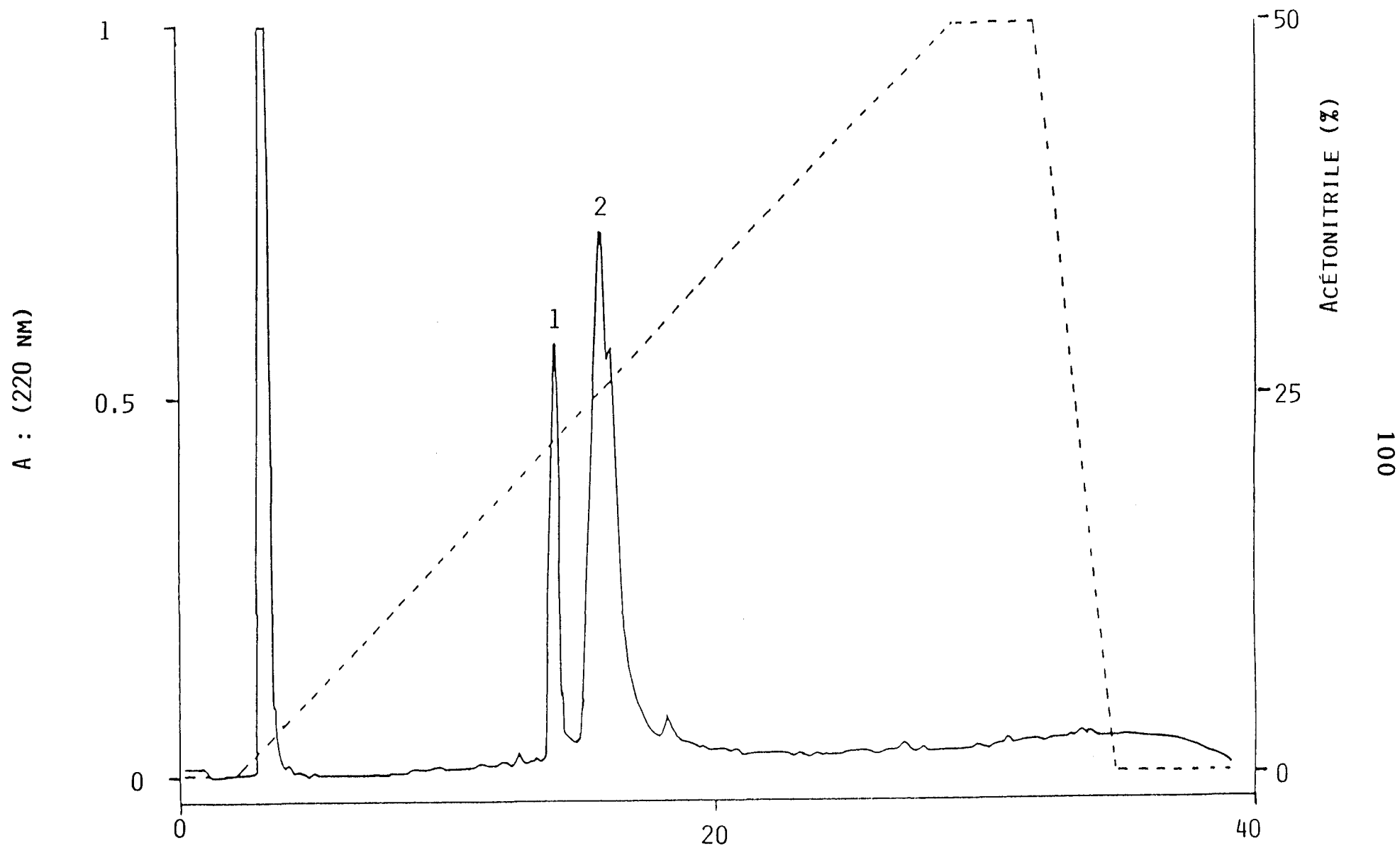


Figure27: Séparation en CLHP-phase inverse des peptides obtenus après clivage de la protéine intermédiaire TP2 de bœuf, par l'endoprotéinase Glu-C.

1 10
Gly-Lys-Val-Ile-Lys-Arg-Lys-Gln-Val-Lys-
20
Arg-Ser-Lys-Gln-Val-Tyr-Lys-Arg-Lys-Arg-
Gln-Ser-Ser-Gly-Arg-Lys-Tyr-Asn(OH)

Figure 28: Séquence en acides aminés du peptide
C-terminal (28 résidus) obtenu après clivage
de la protéine TP2 de bélier par
l'endoprotéinase Glu-C.

XI. Hydrolyse du peptide N-terminal par l'endoprotéase Arg-C:

Après avoir séparé les peptides C- et N-terminaux en CLHP, nous avons soumis le peptide N-terminal à une hydrolyse par l'endoprotéase Arg-C en milieu bicarbonate d'ammonium 0.1M à pH 8.5 pendant 2h à 37°C avec un rapport enzyme/substrat de 1/50 (p/p). La réaction est arrêtée par addition d'acide formique. Les peptides sont séparés en CLHP-phase inverse (Fig.29). Une composition en acides aminés a été réalisée sur les différentes fractions et certaines ont été soumises à la dégradation automatique d'Edman.

XI.1. Etude de la fraction 2:

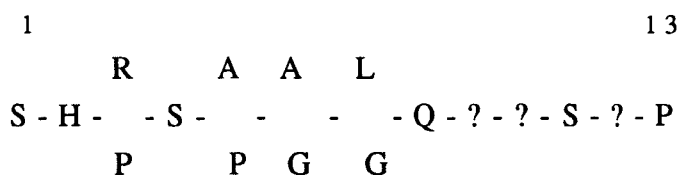
Le séquençage de la fraction 2 (Fig.29) a donné le résultat suivant:



Cette séquence confirme sans ambiguïté celle proposée dans la fraction thermolysique C-1, et permet de la prolonger de 4 résidus du côté N-terminal.

XI.1. Etude de la fraction 4:

Le séquençage de la fraction 4 (Fig.29) montre qu'elle est constituée d'un mélange de peptides, dont le peptide de la fraction 2. De ces résultats, la séquence probable suivante peut être proposée:



Une nanomole de cette fraction 4 a été analysée en FAB. Seule la masse 1171 uma a été observée, mais aucune séquence n'a pu lui être attribuée. La complexité des autres fractions ne nous a pas permis d'avoir plus d'informations structurales.

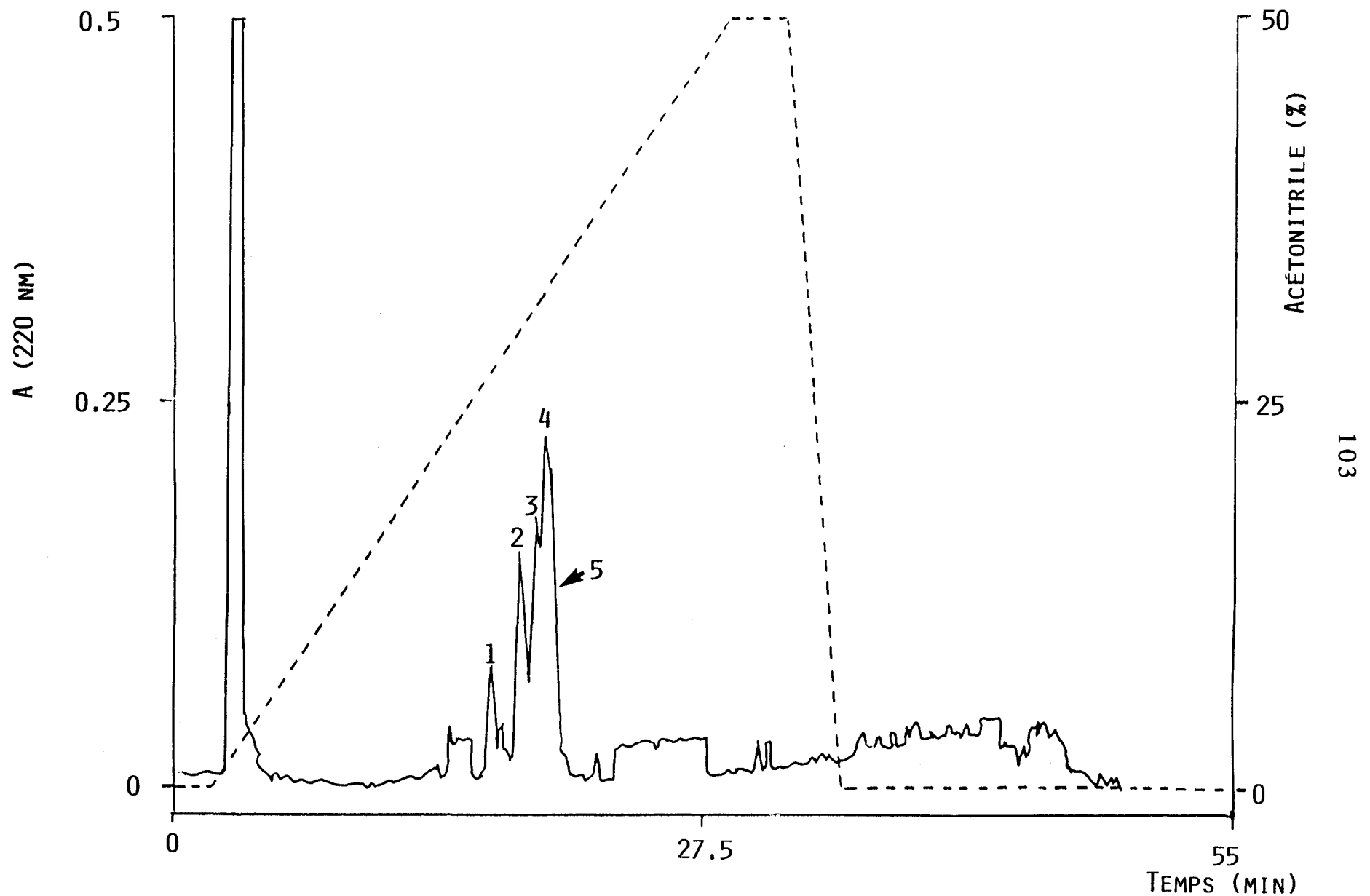


Figure29: Séparation en CLHP-phase inverse des peptides obtenus après clivage du peptide NH₂-terminal de la protéine intermédiaire TP2 de bœuf, par l'endoprotéinase Arg-C.

XII. Coupure du peptide N-terminal par l'endoprotéase Lys-C:

Le peptide N-terminal de la protéine réduite et alkylée par l'IE-TFA a été hydrolysé par l'endoprotéase Lys-C. En effet, cet enzyme coupe au niveau C-terminal de la lysine, mais également, par mimétisme, au niveau C-terminal de l'aminoéthylcystéine.

20 nmoles de peptide sont hydrolysées par l'endoprotéase Lys-C en tampon bicarbonate d'ammonium 0.1M à pH 8.5 à 37°C pendant 4h avec un rapport enzyme/substrat de 1/64 (p/p). Les peptides obtenus sont séparés en CLHP-phase inverse. Une composition en acides aminés a été effectuée sur une aliquote de chacune des fractions.

XIII. Coupure du peptide N-terminal par l'endoprotéase Asn-C:

20 nmoles de peptide réduit et alkylé par l'IE-TFA sont hydrolysées par l'endoprotéase Asn-C en milieu acétate d'ammonium 20mM, pH 5.58, DTT 10mM, EDTA 1mM, pendant 15h à 37°C avec un rapport enzyme/substrat de 2µ/20nmol. Les différents peptides obtenus sont séparés en CLHP-phase inverse. Une composition en acides aminés a été réalisée sur chaque fraction obtenue.

XIV. Conclusion:

Actuellement, nous disposons seulement de résultats partiels concernant la séquence de la protéine TP2. La séquence obtenue est surtout localisée du côté C-terminal de la protéine. Le tableau III montre les homologies de séquence entre la protéine TP2 de bœuf et celles déduites du c-DNA chez d'autres mammifères.

Nous attendons les résultats de séquence des peptides issus de l'hydrolyse du peptide N-terminal de TP2 par l'endoprotéase Lys-C et l'endoprotéase Asn-C, pour analyser les différentes fractions en spectrométrie de masse.

Nous voyons encore par ces exemples que la spectrométrie de masse est complémentaire de la méthode récurrente d'Edman, pour confirmer ou compléter la structure des peptides, plus particulièrement lorsqu'ils sont en mélange.

Souris	D	T	K	M	Q	S	L	-	-	¹⁰ P	T	T	H	P	H	P	H	S	S	²⁰ S	R	P	Q	S	H	T	S	N	Q	³⁰ C
Rat	G	G	H	T	R	R	G	R	A	P	T	T	H	P	H	P	H	S	S	S	R	P	Q	S	H	T	N	N	Q	C
Verrat	D	T	K	T	Q	S	L	P	N	A	H	T	Q	P	H	S	N	S	G	P	Q	S	H	A	C	N	Q	C	S	C
Taureau	D	T	K	T	Q	S	L	P	N	T	H	A	Q	P	H	S	N	S	R	P	Q	S	H	A	C	H	H	C	S	C
Bélier		T	K																											

Souris	N	Q	C	T	C	S	H	H	C	⁴⁰ R	S	C	S	Q	A	G	H	A	G	⁵⁰ S	S	S	S	P	S	P	G	P	P	⁶⁰ M
Rat	A	-	-	-	C	S	H	H	C	R	S	C	S	Q	A	G	H	P	S	S	S	S	S	P	S	P	G	P	P	T
Verrat	S	H	H	C	Q	N	C	S	Q	S	C	D	R	S	Q	S	C	S	R	S	R	S	S	Q	S	P	T	G	H	
Taureau	S	H	H	C	Q	S	R	S	R	S	R	-	-	-	-	S	C	-	R	S	R	S	S	S	R	R	P	R	S	H
Bélier																												S	H	

Souris	K	H	P	K	P	S	V	H	S	⁷⁰ R	H	S	P	A	R	P	S	H	R	⁸⁰ G	S	C	P	K	N	R	K	T	F	⁹⁰ E
Rat	K	H	P	K	T	P	M	H	S	R	Y	S	P	S	R	P	S	H	R	G	S	C	P	K	N	R	K	T	L	E
Verrat	R	S	L	P	G	H	Q	S	Q	S	L	S	P	S	P	S	P	R	H	R	K	R	A	M	H	S	H	R	C	P
Taureau	R	S	P	T	G	H	Q	G	R	A	R	-	P	Q	P	S	S	E	A	P	Q	T	H	H	A	L	P	P	V	S
Bélier	R	S	P	P	G		Q			S	-	P						R	H	R	Y	T	-	M	H	S	H	Q	G	P

Souris	G	K	V	S	K	R	K	A	V	¹⁰⁰ R	R	R	K	R	T	H	R	A	K	R	R	S	S	G	R	R	Y	K		¹²⁰
Rat	G	N	V	S	K	R	K	A	V	R	R	R	K	R	T	H	R	A	K	R	R	S	S	G	R	R	Y	K		
Verrat	S	R	P	G	T	R	S	C	S	H	S	K	K	R	K	N	V	E	G	K	A	N	K	R	K	G	I	K	R	S
Taureau	S	R	A	V	T	H	S	C	S	H	S	K	N	R	K	N	L	E	G	K	V	I	K	R	K	Q	V	K	R	S
Bélier	S	R	A	V	T	H	S	C	S	H	S	K	N	R	K	N	L	E	G	K	V	I	K	R	K	Q	V	K	R	S

Souris										¹³⁰																				
Rat																														
Verrat	Q	Q	V	Y	K	T	K	R	R	S	S	G	R	K	Y	N														
Taureau	K	Q	V	Y	K	R	K	R	Q	S	S	G	R	K	Y	N														
Bélier	K	Q	V	Y	K	R	K	R	Q	S	S	G	R	K	Y	N														

Tableau III

Comparaison des séquences en acides aminés de la protéine intermédiaire TP2.

Souris, rat, taureau, verrat: Séquences déduites à partir du gène. (177-180)

Bélier: Séquence partielle obtenue à partir de la protéine.

Chapitre 6

Utilisation de la
spectrométrie de masse
pour la détermination de
modifications chimiques et
post-traductionnelles

Etude de modifications chimiques et post- traductionnelles de protéines basiques

I. Introduction:

Après avoir montré l'importance de la spectrométrie de masse dans l'élucidation de la structure primaire des protéines basiques, nous avons étudié son rôle dans la détermination de modifications chimiques et post-traductionnelles qui affectent ces protéines. L'essentiel de nos résultats est présenté dans la publication 3:

Application of electrospray and fast atom bombardment mass spectrometry to the identification of chemical and post-translational modifications of proteins and peptides.

Kouach M., Belaïche D., Jaquinod M., Couppez M., Kmiecik D., Ricart G., Van Dorsselaer A., Sautière P. and Briand G. (1993)
Biol. Mass Spectrom. soumis pour publication

Application of Electrospray and Fast Atom Bombardment Mass Spectrometry to the Identification of Chemical and Post-translational Modifications of Proteins and Peptides*.

Mostafa Kouach¹, Denise Belaïche¹, Michel Jaquinod², Maurice Couppez¹, Daniel Kmiecik¹, Guy Ricart³, Alain Van Dorsselaer², Pierre Sautière¹ and Gilbert Briand^{1**}

¹ URA 409 CNRS, Université de Lille II, Institut de Recherches sur le Cancer, Place de Verdun, 59045 Lille Cedex, France

² Laboratoire de Spectrométrie de Masse Bio-Organique, Institut de Chimie, Université Louis Pasteur, 67008 Strasbourg Cedex, France

³ Laboratoire de Spectrométrie de Masse, Université des Sciences et Techniques de Lille-Flandres Artois, 59655 Villeneuve d'Ascq Cedex, France

**Correspondence should be addressed to :

Dr. Gilbert BRIAND

Present adress : Service Commun de Spectrométrie de masse, Faculté de Médecine

3 ème étage, Place de Verdun, 59045 Lille Cedex, France

Tel : (33) 20.62.68.99

FAX : (33) 20.53.55.77

* The authors dedicate this work to the memory of the late Professor Bernard Fournet.

ABBREVIATIONS

Alpha-chymotrypsin : (EC 3.4.21.1)

Endoproteinase Lys-C : (EC 3.4.21.50)

Trypsin TPCK : (EC 3.4.21.4)

Pepsin (EC 3.4.23.1)

SUMMARY

Mass spectrometry (MS) is a very powerful tool in the identification of chemical and post-translational modifications of proteins and peptides. Often these modifications cannot be determined by conventional techniques. This report describes the combined use of electrospray ionization mass spectrometry (ES-MS) and fast atom bombardment mass spectrometry (FAB-MS) to complete the primary structure of proteins and peptides. Examples illustrate how the MS is used to locate sites of phosphorylation, methylation, acetylation, verify the presence of a sulfate group and identify blocking groups and unexpected side reactions such as deamidation or alkylation.

INTRODUCTION

In conventional protein structure determination, sequence of a protein is carried out by sequencing all peptides obtained from a chemical or enzymatic cleavage of the protein. Automated Edman degradation is used for this purpose, but it cannot handle peptides with blocked N-terminus. Another problem encountered during sequencing is the inability to detect certain modified amino acids, either because the chromatographic behaviour of the corresponding phenylthiohydantoin (PTH) derivatives is not known, or the modifying group is lost during a chemical step of the Edman procedure, or the PTH derivative of the modified amino acid is not extracted with organic solvents. Mass spectrometry (MS) may be used to resolve these structural problems that are not accessible by classical techniques. In the past several years, new ionization technologies have been developed to allow high mass measurements : Fast Atom Bombardment (FAB)¹⁻³, ²⁵²Cf Plasma Desorption (PD)⁴, Matrix Assisted Laser Desorption (LD)⁵⁻⁸ and Electrospray Ionization (ES)⁹⁻¹¹. These methods show extremely high ionization efficiencies and, for FAB-MS and ES-MS the capacity for very precise molecular mass measurements.

During FAB-MS analysis, the analyte is dissolved in a viscous liquid matrix, such as glycerol, and acids such as TFA are often added to enhance protonation. Desorption-ionization is accomplished by bombarding the sample-matrix mixture with an energetic beam of atoms or ions.

In ES-MS, molecules are ionized at atmospheric pressure. This ionization method produces a coherent series of multiply charged molecular ions for each molecular species present in the sample, from highly charged droplets formed by electrostatic nebulization. The charge state of each of these peaks may be simply determined and from this information the molecular mass of each component is obtained. The ES-MS allows the measurement of relative mass in the range of 5-40 kDa with precision better than 0.01%. To date, samples with molecular mass of up to 200 kDa¹² have been successfully analysed.

This report describes the use of ES-MS and FAB-MS to identify post-translational modifications (sulfatation, acetylation, methylation, phosphorylation) or chemical modifications (cyclization, deamidation, alkylation) by accurately measuring the molecular mass of peptides and proteins.

EXPERIMENTAL

Materials

Endoproteinase Lys-C was from Boehringer. Pepsin was purchased from Worthington. Chymotrypsin was from Merck. N-(Iodoethyl)-trifluoroacetamide was from Pierce. Acetonitrile for reverse-phase HPLC was obtained from Rathburn. E.coli alkaline phosphatase was from Sigma. All solvents and reagents for gas-phase sequencing were from Applied Biosystems. All other reagents were of the highest purity available.

Samples

Peptide (1-13) of histone H5 from *Xenopus laevis* erythrocyte. Histone H5 was prepared as described¹³. The protein (250 nmol) was hydrolysed with pepsin at 37°C for 2h, using an enzyme-to-substrate ratio of 1:50 (w/w). The hydrolysate was freeze-dried prior to fractionation. The N-terminal peptide (1-32) was hydrolysed with trypsin at 37°C for 2h in 0.1 M ammonium bicarbonate pH 8.0 using an enzyme-to-substrate ratio of 1:50 (w/w). The hydrolysate was acidified and freeze-dried prior to fractionation.

Peptide (1-23) from calf thymus histone H4 was generated by cleavage at aspartyl residues and purified as described¹⁴. This peptide was subsequently hydrolysed with

endoproteinase Lys-C in 300 μ l of 0.1M ammonium bicarbonate, pH 8.5 at 37°C for 4h using an enzyme-to-substrate ratio of 1:50 (w/w). The digestion was stopped by lowering the pH with formic acid. The peptides are designated as K.

Synthetic peptides (26-33) cholecystokinin (CCK) and (26-33) cholecystokinin sulfated were obtained from UCB-Bioproducts S.A. (Braine l'Alleud, Belgium).

Carboxamidomethylated peptides of ram spermatidal protamine. After reduction and alkylation by iodoacetamide¹⁵, protamine (50 nmol), was digested with trypsin at 37°C for 4 h in 0.1 M ammonium bicarbonate, pH 8.5 using an enzyme-to-substrate ratio of 1:50 (w/w). The hydrolysate was acidified and freeze-dried prior to fractionation.

Phosphorylated peptides from ram spermatidal protamine. Ram spermatidal proteins was extracted as described¹⁶. After reduction and alkylation by N-(Iodoethyl)-trifluoroacetamide (IE-TFA)¹⁷, the proteins were then submitted to ion-exchange chromatography on a column (200 x 20mm) of carboxymethyl-cellulose CM-52 (Whatman) equilibrated with 0.3 M guanidinium chloride and eluted with a discontinued gradient of guanidinium chloride (respectively 0.52 M, 0.8 M and 1.5 M) in 0.05 M sodium acetate at pH 5.0. The protamine, eluted by 0.8 M GuCl, was dialysed and lyophilized.

Protamine (50 nmol), dissolved in 1ml of Tris/HCl 50 mM pH 8.0, MgCl₂ 1 mM, ZnCl₂ 0.5 mM, was submitted to alkaline phosphatase at 37°C for 3 h using an enzyme-to-substrate ratio of 1:6 (w/w). The reaction was stopped by the addition of 12 M HCl up to a final concentration of 0.25 M. Protamine was then precipitated by acetone at -20°C for 1 h. After centrifugation, the protein was dissolved in deionised water and lyophilized.

Hydrolysis of protamine with endoproteinase Lys-C was carried out in 300 μ l of 0.1 M ammonium bicarbonate, pH 8.5 at 37°C for 4 h using an enzyme-to-substrate ratio of 1:100 (w/w). The digestion was stopped by lowering the pH with formic acid and the hydrolysate was freeze-dried prior to fractionation.

40 nmol of protamine, dissolved in 300 μ l of 0.1 M ammonium acetate (pH 5.0) were digested with chymotrypsin at 30°C for 4 h using an enzyme-to-substrate ratio of 1:100 (w/w).

Purification of peptides

Peptides generated by enzymatic or chemical cleavages were separated by reverse-phase high performance liquid chromatography on a Merck superspher 100 RP18 endcapped column (244 x 4.0 mm, particle size 5 μ m). The elution was performed with a linear gradient from 0% to 50 % of acetonitrile in 0.1 % TFA, at a flow rate of 1ml/min. The absorbance was measured at 220 nm and 280 nm.

Mass Spectrometry

Electrospray mass spectrometry (ES-MS): ESMS was performed on a VG BIO-Q mass spectrometer¹⁸, an instrument which consists essentially of an electrostatic spray ion source operating at atmospheric pressure followed by a quadrupole mass analyser with a mass range of 4000. The voltage of the extracting cone (Vc) was adjusted to 150 Volts. The mass spectrometer was scanned from $m/z = 100$ to $m/z = 1000$ in 15s at unit resolution. Calibration was performed using charged ion from poly(ethylene glycol) 800 obtained from a separate introduction. For positive ESMS, the protein sample was first dissolved in water containing 5% acetic acid and an equivalent volume of methanol was added. For negative ESMS, the protein was dissolved in water/acetonitrile 50:50. The

concentration used was between 30-100 pmol/ μ l. This solution (2-10 μ l) was introduced into the ion source at a flow rate of 2 μ l/min.

Fast Atom Bombardment mass spectrometry (FAB-MS): A Concept II HH (Kratos Analytical Ltd, Urmston, Manchester U.K.) four-sector tandem mass spectrometer¹⁹, which consists of two double-focusing forward-geometry instruments joined back-to-back (E.B.E.B) was used to acquire spectra. It is equipped with a commercial Kratos FAB source, an Ion Tech B11 NF saddle-field fast atom gun (Ion Tech Ltd, Teddington, England) and a Kratos DS-90 data system. A beam of Xe atoms of approximately 7 keV impact energy and equivalent to 1 mA emission current was employed to ionize peptides dissolved in the matrix (glycerol/water/trifluoroacetic acid in ratio 10:88:2 in positive ion mode or triethanolamine/tetramethylurea in ratio 50:50 in negative ion mode). The FAB-produced ions were then accelerated through a potential of 8 kV and mass-selected by using MS-I (ESA-I and the magnetic sector) at a resolution of about 1500 (full width at 5% height). Cesium iodide was used as a standard compound for mass calibration of samples. The peptides were dissolved in deionised water at a concentration of 1 nmol/ μ l. 1 μ l of peptide solution was deposited on a stainless steel target, and 1 μ l of matrix was added. Spectra were recorded in negative ion mode and positive ion mode using peak format or multi-channel acquisition (MCA)²⁰.

Protein sequencing

Microsequencing was performed on a gas-phase sequencer, Applied Biosystems 470A, using the 03RPTH standard program. The phenylthiohydantoin derivatives of amino acids were identified on an on-line phenylthiohydantoin analyser, Applied Biosystems 120A, equipped with an CR4A Shimadzu integrator-calculator. The amount of each sample submitted to Edman degradation was approximately 1 nmol.

RESULTS

Peptide of the histone H5 of *Xenopus laevis* erythrocyte

The amino acid sequence of histone H5 from *Xenopus laevis* erythrocyte has been deduced from its DNA sequence¹³. But that approach gave no information regarding post-translational modifications.

Figure 1 shows the FAB mass spectrum of the N-terminal tryptic peptide with the sequence from Ala at the N-terminus to Lys at position 13 of histone H5 (sequence : A-E-N-S-A-A-T-P-A-A-K-P-K). The mass of the peptide was observed at $m/z = 1298.5$ (theoretical value = 1255.7). No free amino terminal group was found by Edman degradation. These results suggested that an acetyl group (molecular mass = 43 Da) was attached to the N-terminal of peptide. Moreover the supplementary unit is likely due to a deamidation of asparagine at position 3, during the preparation of the peptide. A doubly charged ion (not shown) was observed at 649.3 giving a single ion of 1297.7. The presence of this doubly charged ion may be due to the protonation of the two lysine residues by the TFA used in the matrix. The mass observed for the singly charged ion at 1298.5 and that observed for the doubly charged ion at 649.3 obviously show the partial deamidation of the asparagine residue.

Figure 2 shows the theoretical isotopic ion distributions of Asn-peptide (A), Asp-peptide (B), and the MCA spectra of the N-terminal tryptic peptide (C). Comparison of theoretical patterns of molecular region with MCA clearly indicate the presence of a mixture of asparagine and aspartic peptides.

Deamidation of Asn residue to Asp is a common nonenzymatic post-translational modification of proteins. This reaction proceeds via the formation of an intramolecular cyclic imide intermediate and particularly when asparagine is followed by Gly, Thr or Ser²¹ (Ser in the present peptide). This succinimide can reopen to either

alpha-Asp or beta-Asp. Tandem mass spectrometry with high energy CID may be used to differentiate between alpha-Asp and beta-Asp peptides²².

Peptide(1-23) from calf thymus histone H4

Among the post-translational modifications which may affect histone H4, acetylation is the most important due to its correlation with major biological events including replication, transcription, chromatin assembly and displacement of histones from chromatin²³. The four lysine residues at position 5, 8, 12 and 16 of the amino acid sequence of histone H4 can be reversibly acetylated^{14, 24}. Moreover, the N-terminal group of the histone H4 is always blocked by an acetyl group. On the other hand, the presence of epsilon-N-methyllysine and epsilon-N-dimethyllysine in this histone has been noted earlier²⁵⁻²⁷; methylation may be required for mitotic condensation^{28, 29}.

Figure 3 shows the elution diagram of the peptides derived from the cleavage of the monoacetylated peptide (1-23) with endoproteinase Lys-C.

The FAB-MS results are summarized in Table 1 which show perfect agreement between mass spectrometry values obtained and calculated mass, excepted for K5 and K6. For K5, the mass difference (729 - 687) corresponds to the mass of an additional N-acetyl group. From the enzyme specificity, we concluded that the peptide K5 results from the acetylation of lysine 12 which is a minor site of acetylation in histone H4 from calf thymus.

The mass difference of 115 Da between the two peptides of the fraction K6 corresponds to partial enucleation of an aspartyl residue after acetic acid hydrolysis of the protein. The mass difference of 70 Da between values obtained and calculated mass indicated that substituents with mass values of 43.0 Da for acetyl and 30.0 Da for dimethyl were attached to the epsilon-NH₂ group of lysine.

Figure 4 shows the FAB mass spectra of the fraction K6. Small peptides often yield fragments in their FAB mass spectra, but generally the series is incomplete. For fraction

K6, the spectrum contains fragment ion peaks of significant intensities to identify the epsilon-N-acetylated lysine and the epsilon-N-dimethyllysine. The mass difference observed between the Y8 and Y7 ions^{30,31} (m/z 1162 and m/z 992, respectively) is 170 Da. This is the expected value for the residue epsilon-N-acetyllysine. The mass difference between the C8 and C7 ions (m/z 978 and m/z 822, respectively) is 156 Da. This is the expected value for the residue epsilon-N-dimethyllysine.

We conclude that the structure of the post-translationally modified N-terminal (1-23) of monoacetylated histone H4 from calf thymus contains an acetylated lysine at position 16 (major site) or 12 (minor site) and a dimethylated lysine residue at position 20 (Fig.5)

Cholecystokinin sulfate octapeptide : D-Y(SO₃)-M-G-W-M-D-P(NH₂)

In FAB-MS, desulfatation is predominant in the positive ion mode (Fig. 6A) where the $(MH - SO_3)^+$ (m/z 1063) is much more abundant than the protonated molecular ion MH^+ . This is due to the propensity of the modified peptide to loss SO_3 and suggests that the peptide is weakly or not sulfated. This sulfotyrosine containing peptide also losses SO_4 (96 Da) from MH^+ . The mass of $(MH - SO_4)^+$ is observed at 1047 Da.

In the negative ion mode, the tyrosine sulfate group is stable (Fig. 6B). The mass difference of 80 Da between the two peptides (m/z 1143 - m/z 1063) corresponds to the presence of a sulfate group. When a molecule is thought to contain a sulfate group, it therefore is essential that the FAB mass spectra be acquired in the negative ion mode.

In contrast to the FAB-MS results, the sulfotyrosine containing peptide can be identified by electrospray mass spectrometry in the positive (Fig.7A) or in the negative (Fig.7B) ion mode. The sulfotyrosine containing peptide does not fragment in the negative ion mode. On the other hand, in the positive ion mode, the sulfate is partially destroyed, but the signal intensity of the sulfated peptide is significant enough to permit identification.

Cyclization in carboxamidomethylated peptide of ram spermatidal protamine

The tryptic peptide C-C-L-T-H-S-R (residues 5-11) from ram protamine has a theoretical mass of 933 Da, taking into account the carboxamidomethylation of the two cysteine residues. Its mass spectrum (Fig. 8A) is characterized by the presence of three peaks : a peak $(M+H)^+$ at 933 and a peak $(M_1+H)^+$ at 990 more abundant accompanied with an adduct ion $(M_1+Na)^+$. The difference of 57 Da is probably due to alkylation of the histidine residue. This peptide was also found, in a later fraction of HPLC. The mass spectrum (Fig. 8B) of this peptide is characterized by the presence of a strong signal m/z at 916. The weak signal at 973 shows that the histidine residue is weakly alkylated. The m/z 916 signal corresponds to the carboxymethylated peptide in which cyclization has occurred between the alpha-amino group of cysteine residue and the amide group of its S-alkyl derivative with a subsequent loss of mass equal to 17 Da, corresponding to the loss of an NH_3 molecule (Fig. 9).

The cyclization prevents straightforward use of Edman degradation for sequencing. This phenomenon is similar to the conversion of glutamine residue at the N-termini of peptides generated by enzymatic or chemical cleavages, to pyrrolidone carboxyl residue^{32,33}. The exposure of peptides to acidic conditions which promote the cyclization has to be minimized.

Phosphorylation site of ram spermatidal protamine

Phosphorylation is a major modification of nuclear proteins. Most works with nuclear proteins in this area have been concerned with phosphorylation of serine and threonine, but Henry and Hodge³⁴ have reported the presence of phosphotyrosine in nuclear matrix proteins.

In several mammals such as rat, ram and mouse, histones can be removed from chromatin during spermiogenesis and replaced by nuclear basic proteins specific of the spermatids. These spermatid-specific proteins also named intermediate proteins, are displaced at the end of spermiogenesis by the protamines, the only nuclear basic proteins present in mature sperm^{35,36}. Protamine isolated from testes contains usually mono-, di-phosphorylated forms and also a small proportion of tri-phosphorylated form while the protamine isolated from the epididymal spermatozoa mainly consist of monophosphoprotamine and unmodified protamine in similar proportions^{37,38}. The protamine of mature sperm is present in a nonphosphorylated form. The ram spermatidal protamine was shown to be monophosphorylated³⁹.

Figures 10 and 11 report the spectra of respectively reduced sperm protamine and protamine treated with IE-TFA. The ES-MS spectrum of reduced sperm protamine (Fig. 10) exhibits a series of multicharged ions with 9 to 14 charges. The measured mass ($M_r = 6711,4 \pm 1.3$ Da) is entirely consistent with the calculated mass of sperm protamine¹⁵. The ES-MS spectrum of protamine treated with IE-TFA (Fig.11) is more complex and shows five series of multicharged ions with 10 to 12 charges. The nonphosphorylated protamine (Serie A) shows a mass value at $m/z = 7294.7 \pm 7.3$ (theoretical value : 7013.5). The difference suggests that three of the aminoethylcysteine residues are present in the TFA blocked form (molecular mass of TFA = 96 Da). These data are not in agreement with previous results⁴⁰. The serie B ($M_r = 7375.4$) corresponds to a protein with an M_r increase of 80.5 Da compared to the serie A. This difference could be attributed to the presence of a phosphate group. The mass spectra also contains ions species (Series C, D, E) corresponding to the attachment to the analyte ions of moities with molecular mass 98 Da. These adducts are non covalently bound to the protonated protein molecular ions. Sulfate and/or phosphate impurities present in the peptide and protein samples produce small amounts of sulfuric acid and/or phosphoric acid⁴¹ that bind strongly to the accessible basic sites of the protein ions. The removal of these adducts from peptide

and protein can be done by collisional activation. Phosphatase alkaline treatment is realized for confirmation of phosphorylation.

Phosphorylated peptides were detected by comparing the Lys-C peptides in the HPLC chromatograms (Fig. 12) before (A) and after (B) treatment of the protamine with alkaline phosphatase. The peaks of the HPLC map (Fig. 12A) that disappear after alkaline phosphatase treatment were analysed by FAB-MS. Table 2 summarizes the sequence of each peptide, the expected mass and the observed mass. All the peptides contain the same sites of phosphorylation (Threonine and Tyrosine). Partial hydrolysis by endoproteinase Lys-C is consistent with aminoethylcysteine residues present in the TFA blocked form. The alkylation reaction of protein with N-(iodoethyl)-trifluoroacetamide is shown in figure 13.

In order to localize the phosphorylation site, the protamine was digested with chymotrypsin at pH 5.0 to exclude the tyrosine residue from the C-terminal peptide. The peptides were separated by reverse-phase HPLC (results not shown) and analysed in FAB-MS. Two peptides show masses values at 1006 and 1086 respectively (Fig. 14). These peptides correspond to the sequences TVVRCTRQ and PhosphoTVVRCTRQ respectively. The results obtained from the Lys-C peptides and this chymotryptic peptide indicate that the threonine residue 43 is indeed the site of phosphorylation of the ram spermatidal protamine.

DISCUSSION AND CONCLUSION

Mass spectrometry may be considered as one of the principal analytical tools for characterizing most of the post-translational and chemical modifications of peptides and proteins. Edman degradation of sulfated or phosphorylated peptides has limited utility due to weak extraction or acid lability of the corresponding modified amino acids. The method of Meyer *et al* is limited to the identification of phosphoserine in microsequence analysis after conversion to S-ethylcysteine residues⁴². Recently, Aebersold *et al* have described a

procedure based on solid-phase sequence analysis of phosphopeptides immobilized on carrier discs and identification of phosphotyrosine PTH by HPLC ⁴³.

One method available for the location of these residues in peptide sequences, at the subpicomole level, utilizes radiolabelling with ³²P in combination with Edman degradation⁴³. But these approaches are expensive and time consuming.

In contrast to these methods, phosphorylated and sulfated amino acid residues are readily detected by mass spectrometry. The presence of a phosphorylated or a sulfated residue increases the molecular mass of a peptide by 80 Da. We have used this technique to identify the phosphothreonine residue in ram spermatidal protamine and the sulfotyrosine residue in a cholecystokinin sulfate octapeptide. In addition, desulfatation is extensive during FAB-MS in the positive ion mode, whereas phosphorylated residues are stable under these conditions. Sulfated forms of peptide can be easily detected by FAB-MS in the negative ion mode or by ES-MS.

Post-translational modifications such as acetylation and methylation can also be determined by FAB-MS. These modifications are stable in positive ion mode. If a peptide, is modified at more than one site as is peptide (1-23) from calf thymus H4 histone, the presence of fragment ion peaks in the spectrum is invaluable in the peptide structure determination.

If the N-terminus of the protein is blocked preventing Edman degradation, the identification of the modification is based on precise measurement of the molecular mass of a little peptide obtained by chemical or enzymatic fragmentation of protein. Mass measurements of peptides were easily obtained with ES-MS or FAB-MS with accuracy and good sensitivity.

In FAB-MS analyses of peptides mixtures, a selective desorption and ionization of hydrophobic peptides occurs together with a suppression of hydrophilic components⁴⁴. These suppression effects observed in FAB-MS appear to be less serious in the electrospray mass spectrometry method. ES-MS not only confirms the mass determination, but gives a

critical purity evaluation.

In this work, we have identified a N-acetyl group at the N-terminal position of histone H5 from *Xenopus leavis* erythrocyte. The great precision of FAB-MS indicates a partial deamidation of an asparagine residue in this peptide. On the other hand the carboxamidomethylcysteine present at the N-terminus of a peptide may undergo cyclization with a subsequent loss of a NH_3 molecule, similarly to the cyclization of glutamine at the N-terminus of a peptide. Mass spectrometry analysis confirms this result with the shifting of peak of expected mass $(\text{M}+\text{H})^+$ towards $(\text{M}+\text{H} - 17)^+$. Moreover, under the conditions used for alkylating the thiol groups, histidine can be affected by the reagent. The spectrum is then characterized by the presence of a peak at $(\text{M}+\text{H}+57)^+$. Also, alkylation of protein with IE-TFA shows sometimes a peak at $(\text{M}+\text{H}+96)^+$. This peak is due to fixation of TFA in the alkylant group.

In conclusion, the analysis of protein and peptide sequence by mass spectrometry has received a dramatic impulse from the introduction of FAB-MS in the early 1980s³ and the complementary capacities of different techniques such as FAB-MS, PD-MS, LASER-MS and ES-MS demonstrates that mass spectrometry is now of significant aid to protein and peptide structure determination.

ACKNOWLEDGEMENTS

We thank A. Hemez for the separation in HPLC and O. Hannedouche-Moreau for sequence analysis.

REFERENCES

1. M. Barber, R.S. Bordoli, R.D. Sedgwick and A.N. Tyler, *J. Chem. Soc. Chem. Commun.* 7, 325 (1981)
2. M. Barber, R.S. Bordoli, R.D. Sedgwick and A.N. Tyler, *Nature* 293, 270 (1981)
3. M. Barber, R.S. Bordoli, G.J. Elliot, R.D. Sedgwick and A.N. Tyler, *Anal. Chem.* 54, 645A (1982)
4. R.D. MacFarlane and D.F. Torgenson, *Int. J. Mass Spectrom. Ion Phys.* 21, 81 (1976)
5. M. Karas, D. Bachmann, U. Bahr and F. Hillenkamp, *Int. J. Mass Spectrom. Ion Processes* 78, 53 (1987)
6. M. Karas and F. Hillenkamp, *Anal. Chem.* 60, 2299 (1988)
7. M. Karas, U. Bahr and F. Hillenkamp, *Int. J. Mass Spectrom. Ion Processes* 92, 231 (1989)
8. K. Tanaka, H. Waki, Y. Ido, S. Akita, Y. Yoshida and T. Yoshida, *Rapid Commun. Mass. Spectrom.* 2, 151 (1988)
9. J.B. Fenn, M. Mann, C.K. Meng, S.F. Wong and C.M. Whitehouse, *Mass Spectrom. Rev.* 9, 37 (1990)
10. R.D. Smith, J.A. Loo, C.G. Edmonds, C.J. Barinaga and H.R. Udseth, *Anal. Chem.* 62, 882 (1990)
11. T.R. Covey, R.F. Bonner, B.I. Shushan and J. Henion, *Rapid Commun. Mass Spectrom.* 2, 249 (1988)
12. R. Feng and Y. Konoshi, *Anal. Chem.* 64, 2090 (1992)
13. A.F.M. Moorman, P.A.J. De Boer, R. Charles and W.H. Lamers, *Differentiation* 35, 100 (1987)
14. M. Couppez, A. Martin-Ponthieu and P. Sautière, *J. Biol. Chem.* 262, 2854 (1987)
15. P. Sautière, D. Belaïche, A. Martinage and M. Loir, *Eur. J. Biochem.* 144, 121 (1984)

16. T. Dupressoir, P. Sautière, M. Lanneau and M. Loir, *Exp. Cell. Res.* 161, 63 (1985)
17. W. E. Schwartz, P.K. Smith and G.P. Royer, *Anal. Biochem.* 106, 43 (1980)
18. A. Van Dorsselaer, F. Bitsch, B. Green, S. Jarvis, P. Lepage, R. Bischoff, H.V.J. Kolbe and C. Roitsch, *Biomed. Environ. Mass Spectrom.* 19, 692 (1990)
19. D. Despeyroux, J. Bordas-Nagy and K.R. Jennings, *Rapid Commun. Mass Spectrom.* 5, 156 (1991)
20. R. L. Johnson and L.C.E. Taylor, *OMS Letters* 22, 807 (1990)
21. T. Geiger and S. Clarke, *J. Biol. Chem.* 262, 785 (1987)
22. S.A. Carr, M.E. Hemling, M.F. Bean and G.D. Roberts, *Anal. Chem.* 63, 2802 (1991)
23. A. Csordas, *Biochem. J.* 265, 23 (1990)
24. K.E. Van Holde, *Chromatin*, New York : Springer-Verlag (1989)
25. K. Murray, *Biochemistry* 3, 10 (1964)
26. W.C. Starbuck, C.M. Mauritzen, C.W. Taylor, J.S. Saroya and H. Busch, *J. Biol. Chem.* 243, 2038 (1968)
27. W.K. Paik, S. Kim, *Biochem. Biophys. Res. Commun.* 27, 479 (1967)
28. B.M. Honda, E.P.M. Candido and G.H. Dixon, *J. Biol. Chem.* 250, 8686 (1975b)
29. R. Camato and R.M. Tanguay, *Embo J.* 1, 1529 (1982)
30. P. Roepstorff and J. Fohlman, *Biomed. Mass Spectrom.* 11, 601 (1984)
31. K. Biemann, *Biomed. Environ. Mass Spectrom.* 16, 99 (1988)
32. J. Melville, *Biochem. J.* 29, 179 (1935)
33. D.G. Smyth, W.H. Stein and S. Moore, *J. Biol. Chem.* 237, 1845 (1962)
34. S.M. Henry and L.D. Hodge, *Eur. J. Biochem.* 133, 23 (1983)
35. R.B. Golberg, R. Geremia and W.R. Bruce, *Differentiation* 7, 167 (1977)

36. S.R. Grimes, M.L. Meistrich, R.D. Platz and L.S. Hnilica *Exp. Cell Res.* 110, 31 (1977)
37. Y. Marushige and K. Marushige, *J. Biol. Chem.* 230, 39 (1975)
38. Y. Marushige and K. Marushige, *Biochim. Biophys. Acta* 518, 440 (1978)
39. F. Chirat, Thèse de Doctorat de l'Université de Lille Flandres-Artois (1990)
40. W.E. Schwartz, P.K. Smith and G.P. Royer, *Anal. Biochem.* 106, 43 (1980)
41. S.K. Chowdhury, W. Katta, R.C. Beavis and B.T. Chait, *J. Am. Soc. Mass Spectrom.* 1, 382 (1990)
42. H.E. Meyer, E. Hoffmann-Posorske, H. Korte and L.M.G. Heilmeyer, *FEBS Lett.* 204, 61 (1986)
43. R. Aebersold, J.D. Watts, H.D. Morrison and E.J. Bures, *Anal. Biochem.* 199, 51 (1991)
44. S. Naylor, A.F. Findeis, B.W. Gibson and D.H. Williams, *J. Am. Chem. Soc.* 108, 6359 (1986)

LEGENDES TO FIGURES

Table 1. Sequences, calculated molecular weights and observed m/z values of endoproteinase Lys-C fragments of monoacetylated peptide (1-23) from calf thymus histone H4. Sequences of the fragments are numbered according to protein sequence.

Table 2. Sequences, calculated molecular weights and observed m/z values of endoproteinase Lys-C fragments of ram spermatidal protamine treated with IE-TFA. Sequences of the fragments are numbered according to their positions in the protamine sequence.

Figure 1 : Section of the FAB mass spectrum of the N-terminal peptide derived from the cleavage with trypsin of the N-terminal pepsin peptide (1-32) of histone H5 from *Xenopus laevis* erythrocyte.

Figure 2 : Theoretical isotopic ion distributions of Asn-tryptic peptide (A), Asp-tryptic peptide (B) and MCA FAB spectrum of the N-terminal tryptic peptide (C) of histone H5 from *Xenopus laevis* erythrocyte.

Figure 3 : Separation of endoproteinase Lys-C fragments of monoacetylated peptide (1-23) of calf thymus histone H4. The column was superspher 100 RP18. Elution was performed with a linear gradient from 0% to 50% of acetonitrile in 0.1% TFA, at a flow rate of 1ml/min. The solid bars indicate peptides submitted to mass analysis.

Figure 4 : FAB mass spectrum of the fraction K6 derived from the cleavage with endoproteinase Lys-C of peptide (1-23) from calf thymus histone H4. Fragments

correspond to N-terminal fragmentation ions of the C type and to C-terminal fragmentation of the Y type^{30,31}.

Figure 5 : Amino acid sequence of monoacetylated peptide (1-23) of calf thymus histone H4. K indicate fragments obtained by cleavage with endoproteinase Lys-C.

Figure 6 : Section of the FAB mass spectra of the cholecystokinin sulfate octapeptide (26-33) in the positive ion mode (A) and in the negative ion mode (B).

Figure 7 : Electrospray mass spectra of the cholecystokinin sulfate octapeptide (26-33) in the positive ion mode (A) and in the negative ion mode (B).

Figure 8 : Section of the FAB mass spectra of the alkylated tryptic peptide C-C-L-T-H-S-R of ram spermatidal protamine before cyclization (A) and after cyclization (B).

Figure 9 : Mechanism for the cyclisation of the alkylated tryptic peptide C-C-L-T-H-S-R of the ram spermatidal protamine.

Figure 10 : Electrospray mass spectrum of the ram reduced sperm protamine. One series of multiply charged ions with 9-14 charges is detected.

Figure 11 : Electrospray mass spectrum of ram sperm protamine treated with IE-TFA. Five series (A, B, C, D, and E) of multiply charged ions with 10-12 charges were detected. The series A corresponds to expected protamine with three of the aminoethylcysteinyl residues present in the TFA blocked form. The series B corresponds to the monophosphorylated protamine. The series C, D and E corresponds to the protamine with the attachment of

moities with molecular mass 98 Da.

Figure 12 : Separation of endoproteinase Lys-C fragments of ram sperm protamine treated with IE-TFA before (top) and after (bottom) treatment after alkaline phosphatase. The column was superspher 100 RP18. Elution was performed with a linear gradient from 0% to 50% of acetonitrile in 0.1% TFA at a flow rate of 1ml/min. The solid bars indicate peptides submitted to mass analysis.

Figure 13 : Alkylation reaction with N-(iodoethyl) trifluoroacetamide (IE-TFA) of a protein.

Figure 14 : Section of FAB mass spectra of two alkylated chymotryptic peptides with the sequence TVVRCTRQ. A: nonphosphorylated form, B: phosphorylated form of ram spermatidal protamine.

Structures of peptides from calf histone H4 analysed by mass spectrometry

Number	Amino acid sequence	Quasi molecular ion		Difference and attribution
		Calculated	Determined	
K1	¹ Ac-S- ⁵ G-R-G-K	546	546	-
K2	¹ Ac-S-G-R-G-K-G-G- ⁸ K	788	788	-
K3	⁶ G-G- ⁸ K	261	261	-
K4	⁹ G-L-G- ¹² K	374	374	-
K5	⁹ G-L-G-K-G-G-A- ¹⁶ K	687	729	42 acetyl
K6	¹³ G-G-A-K-R-H-R-K-V-L-R ²³	1277	1347	70 acetyl(42) and dimethyl(28)
	¹³ G-G-A-K-R-H-R-K-V-L-R-D ²⁴	1392	1462	70 acetyl(42) and dimethyl(28)

Table 1

Structures of peptides from ram protamine analysed by mass spectrometry

Fraction	Amino acid sequence	Quasi molecular ion		Difference and attribution
		Calculated	Determined	
1	⁴⁰ R-R-Y-T-V-V-R-C ⁴⁷	1095	1175	80(phosphate)
2	³⁹ C-R-R-Y-T-V-V-R-C ⁴⁷	1241	1417	80(phosphate) and 96(TFA)
3	⁴⁰ R-R-Y-T-V-V-R-C-T-R-Q ⁵⁰	1481	1657	80(phosphate) and 96(TFA)
4	³⁹ C-R-R-Y-T-V-V-R-C-T-R-Q ⁵⁰	1626	1898	80(phosphate) and 2×96(2 TFA)

Table 2

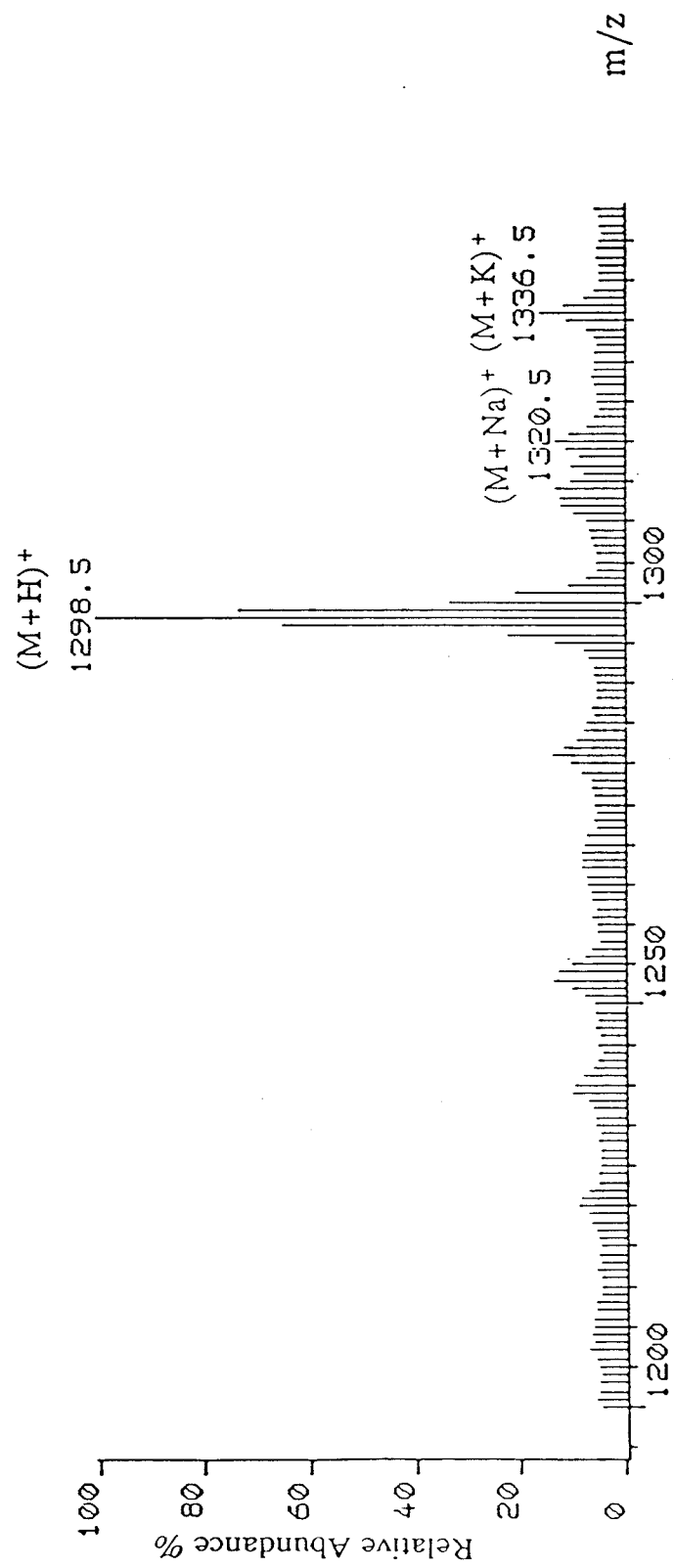


Figure 1

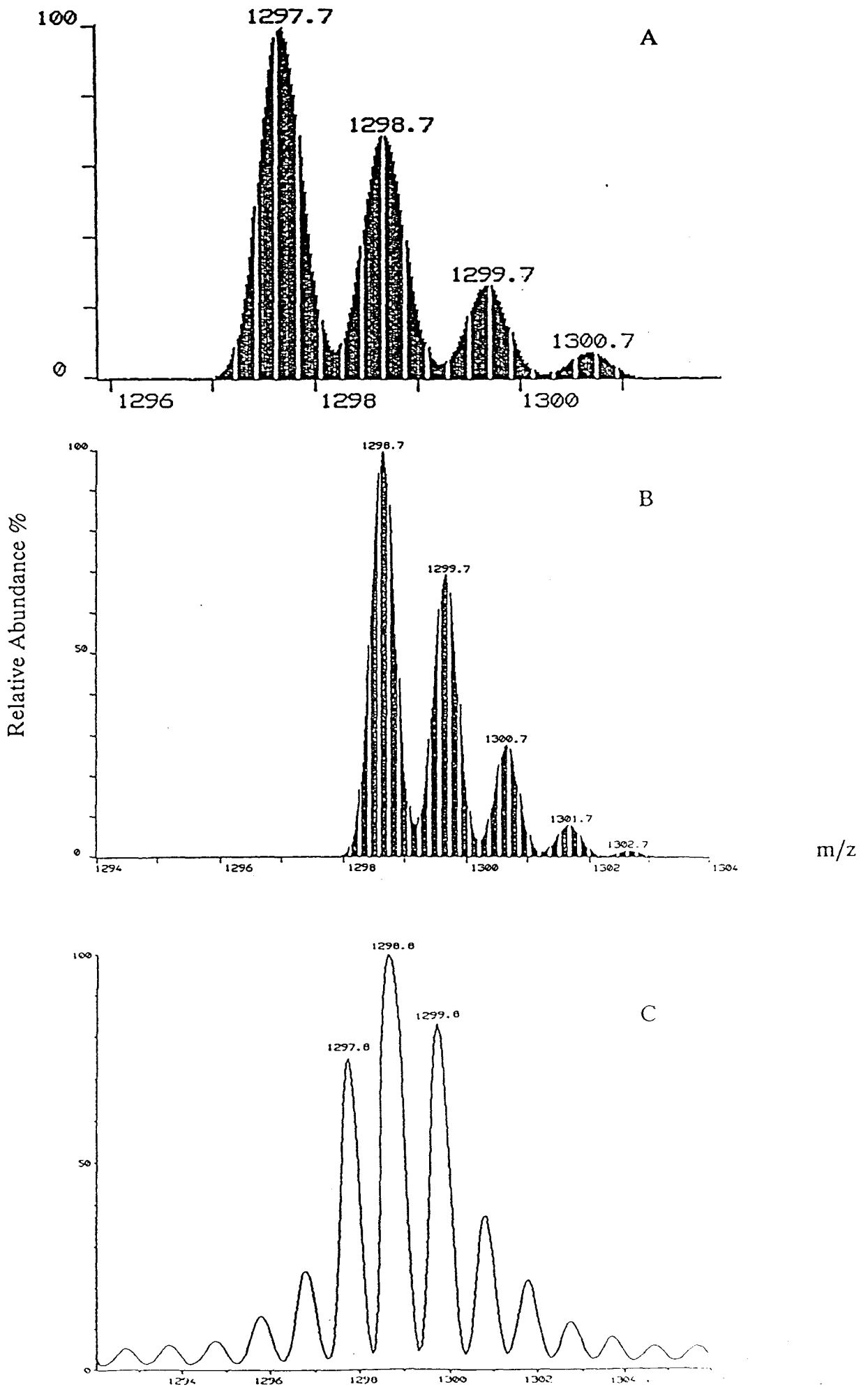


Figure 2

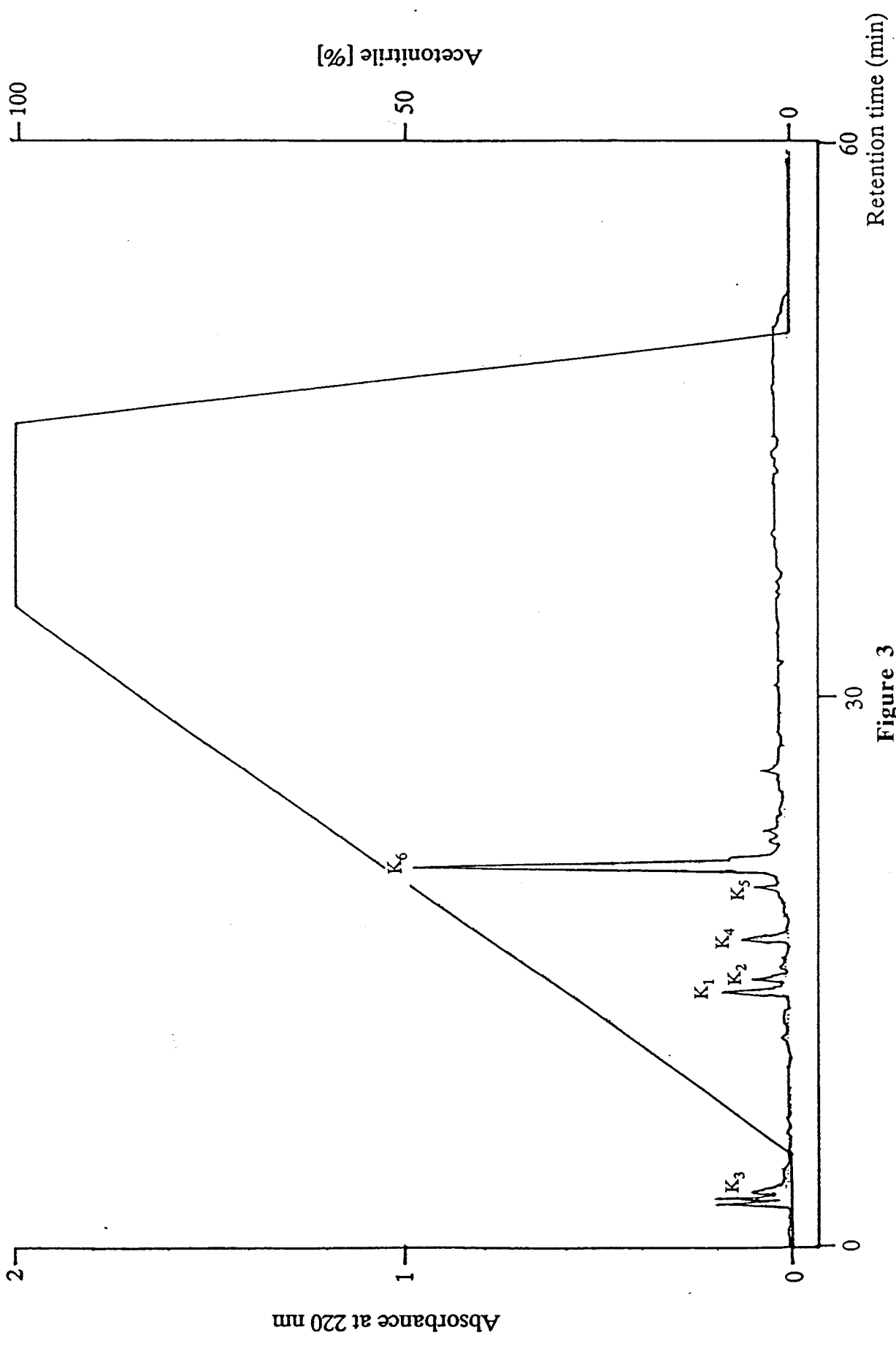
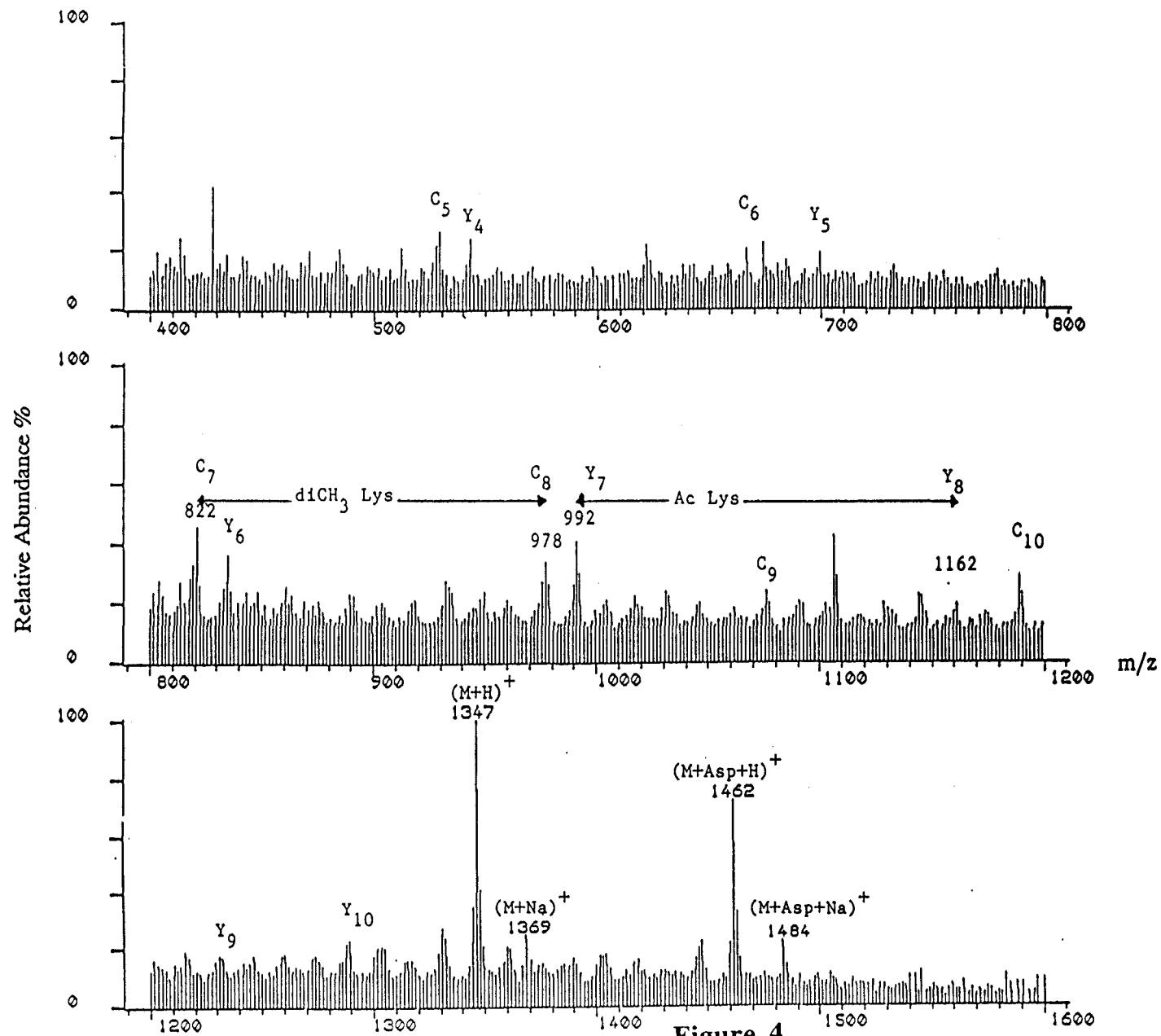


Figure 3



<-----K1-----> <-K3-> <--K4--> <-----K6----->

Ac-S-G-R-G-K-G-G-K-G-L-G-K-G-G-A-K-R-H-R-K-V-L-R-(D)

$\begin{array}{ccc}
| & | & | \\
\text{Ac} & \text{Ac} & \text{di}(\text{CH}_3)
\end{array}$

<-----K2-----> <-----K5----->

Figure 5

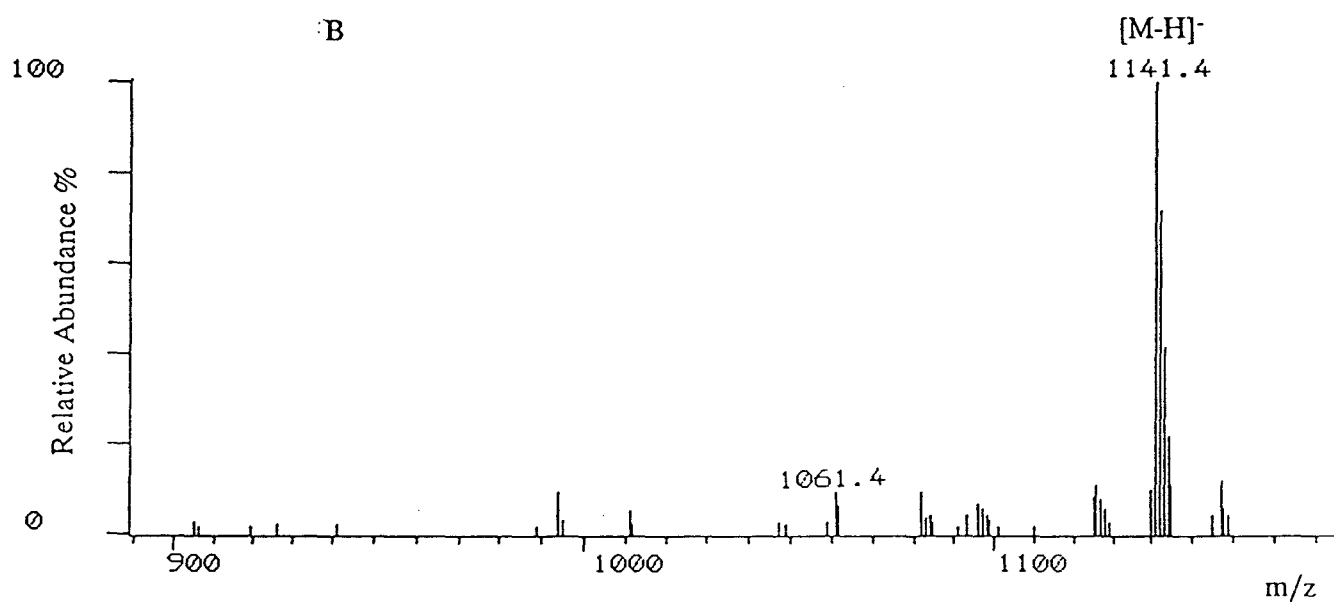
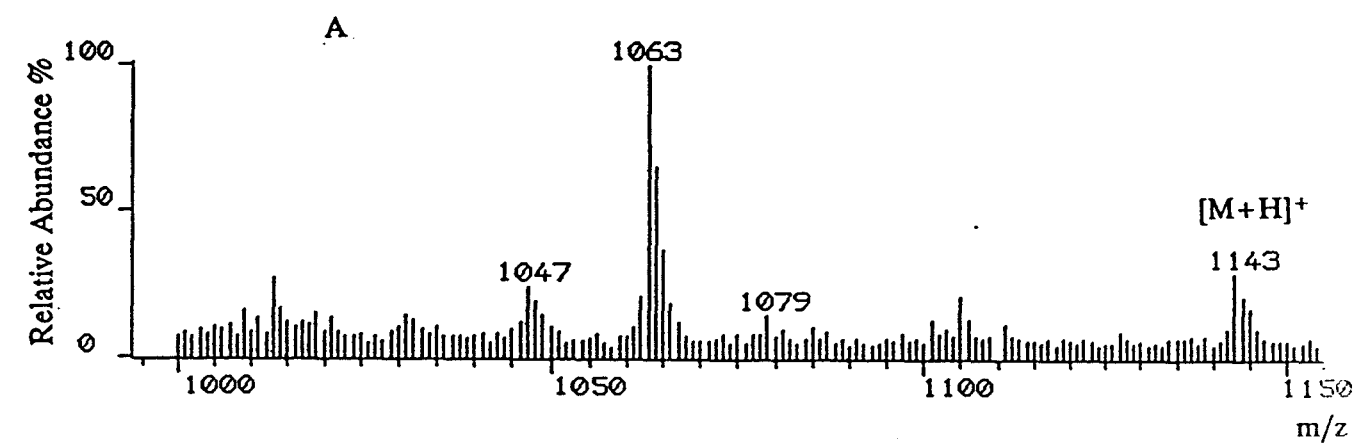


Figure 6

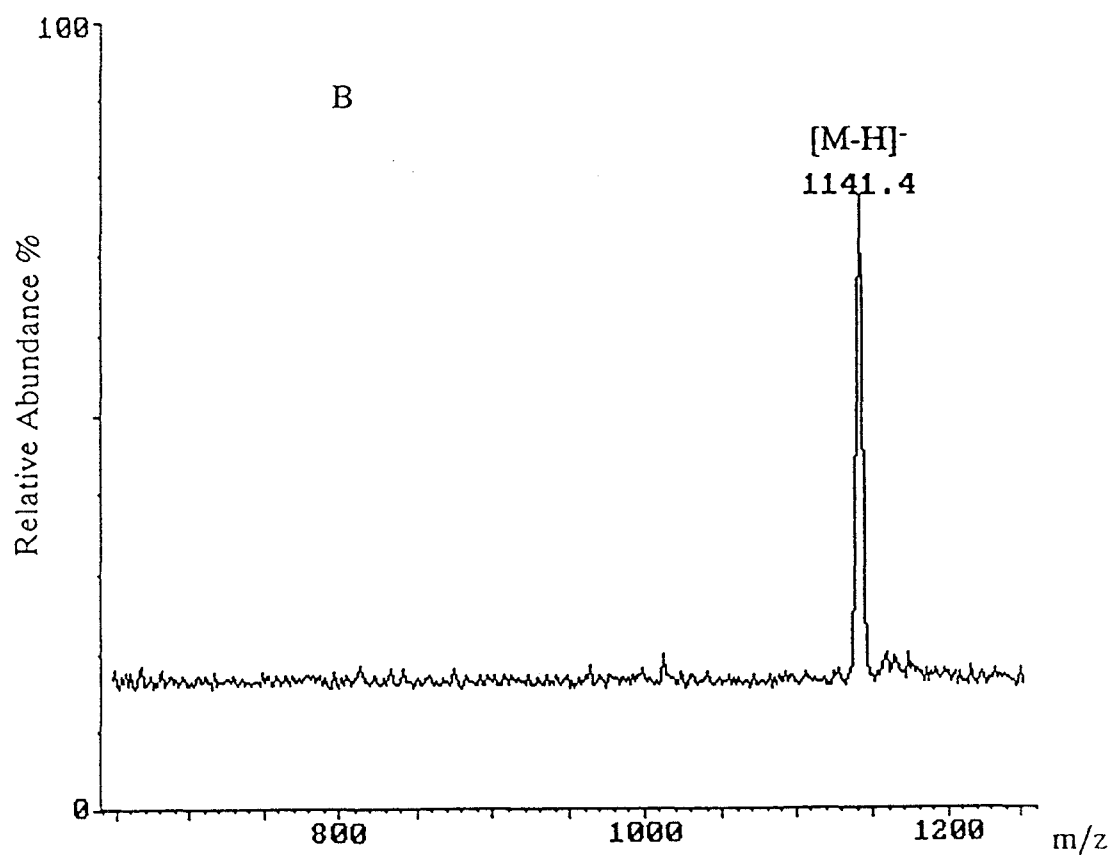
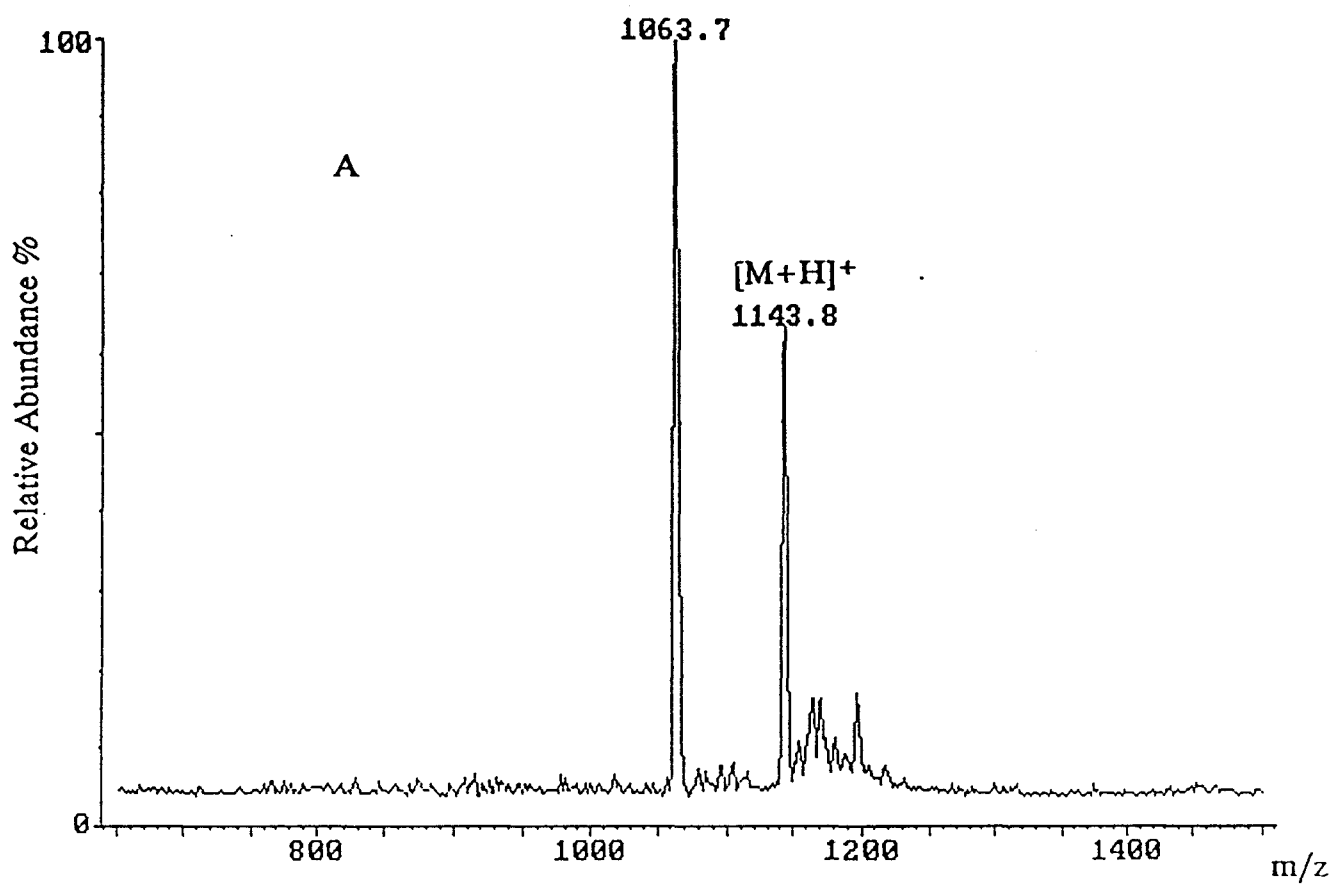


Figure 7

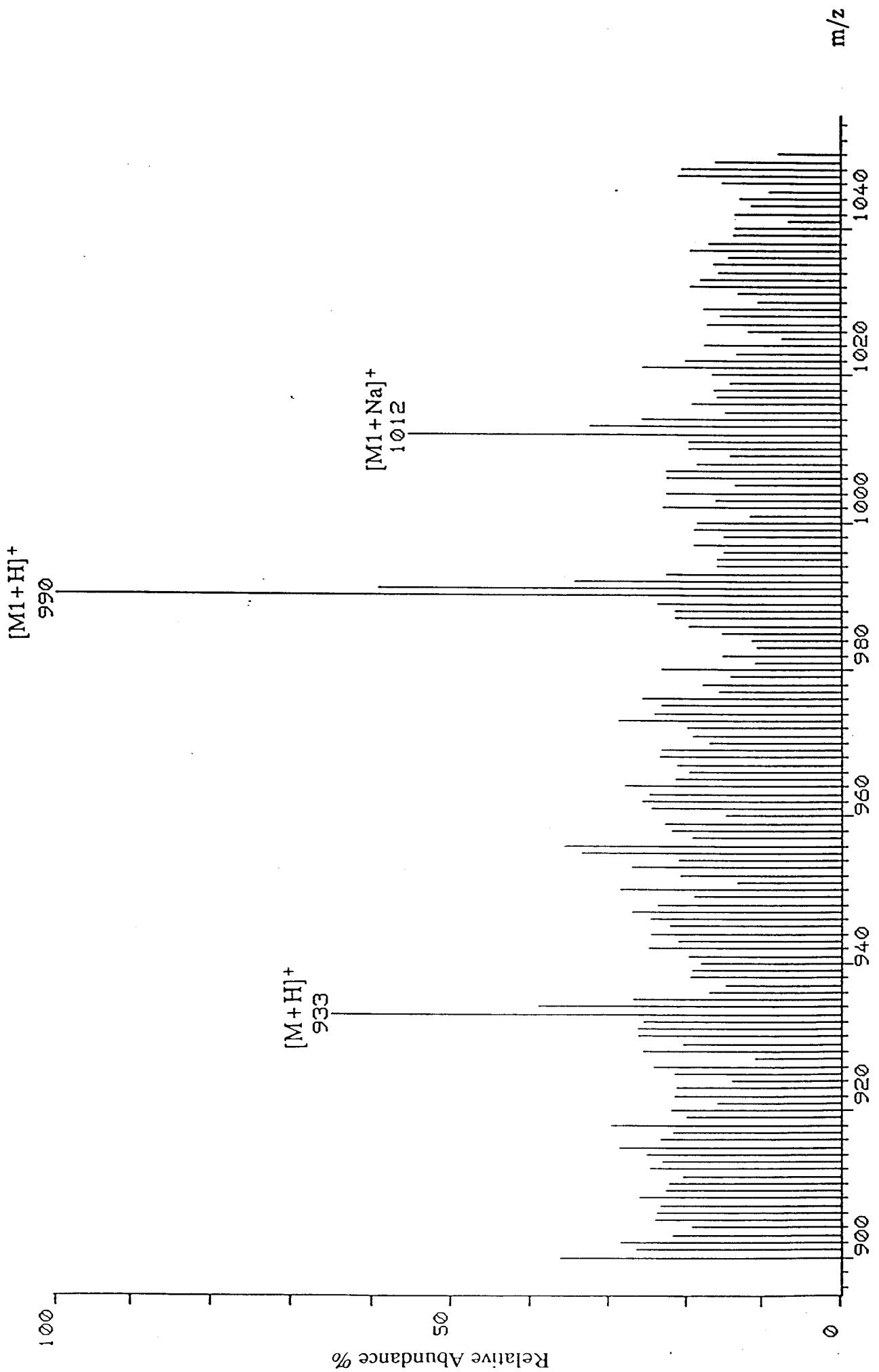


Figure 8A

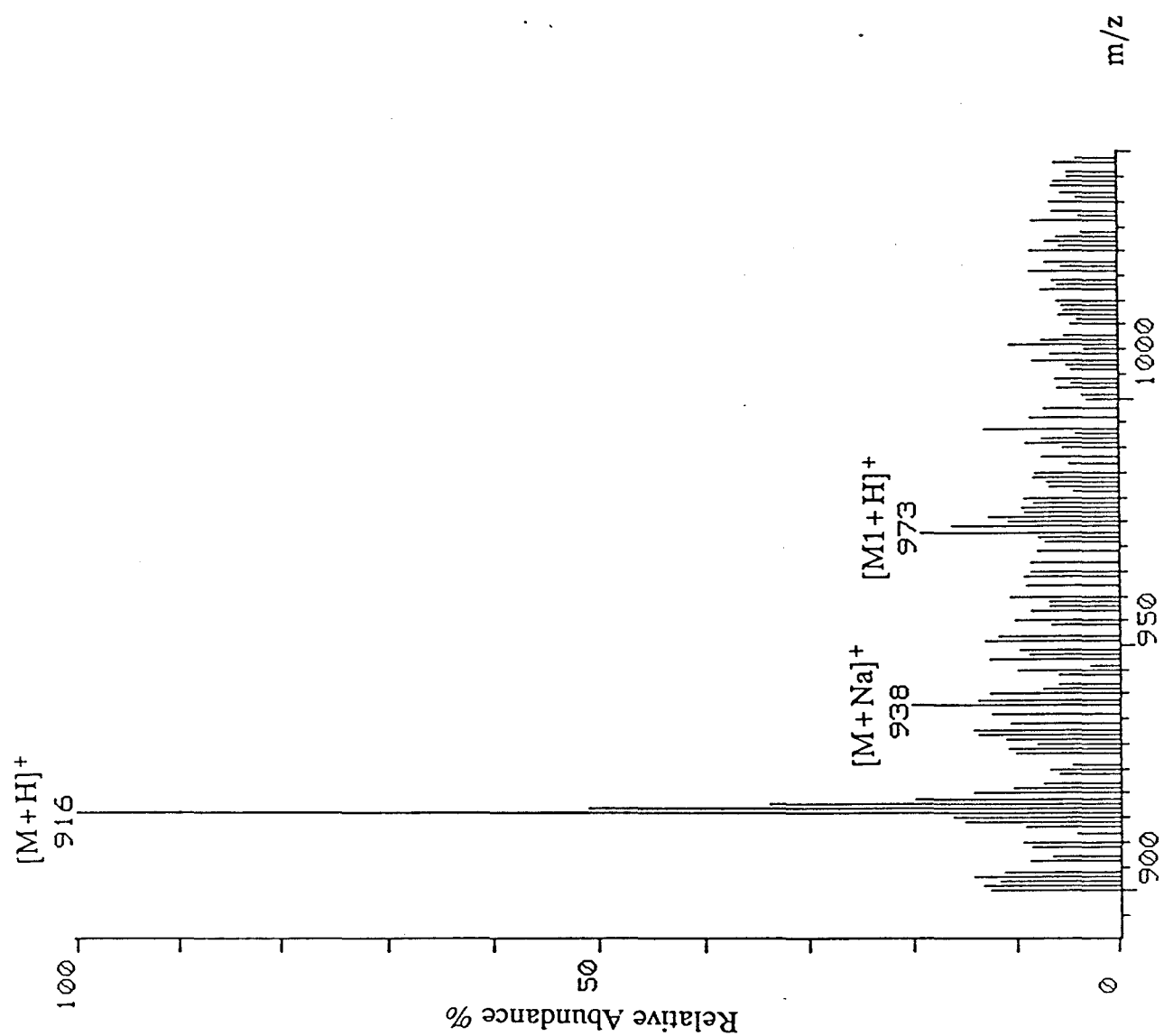


Figure 8B

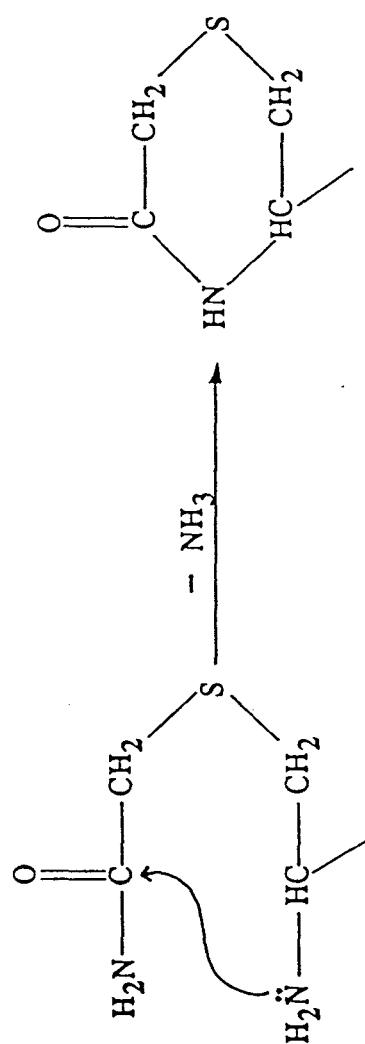


Figure 9

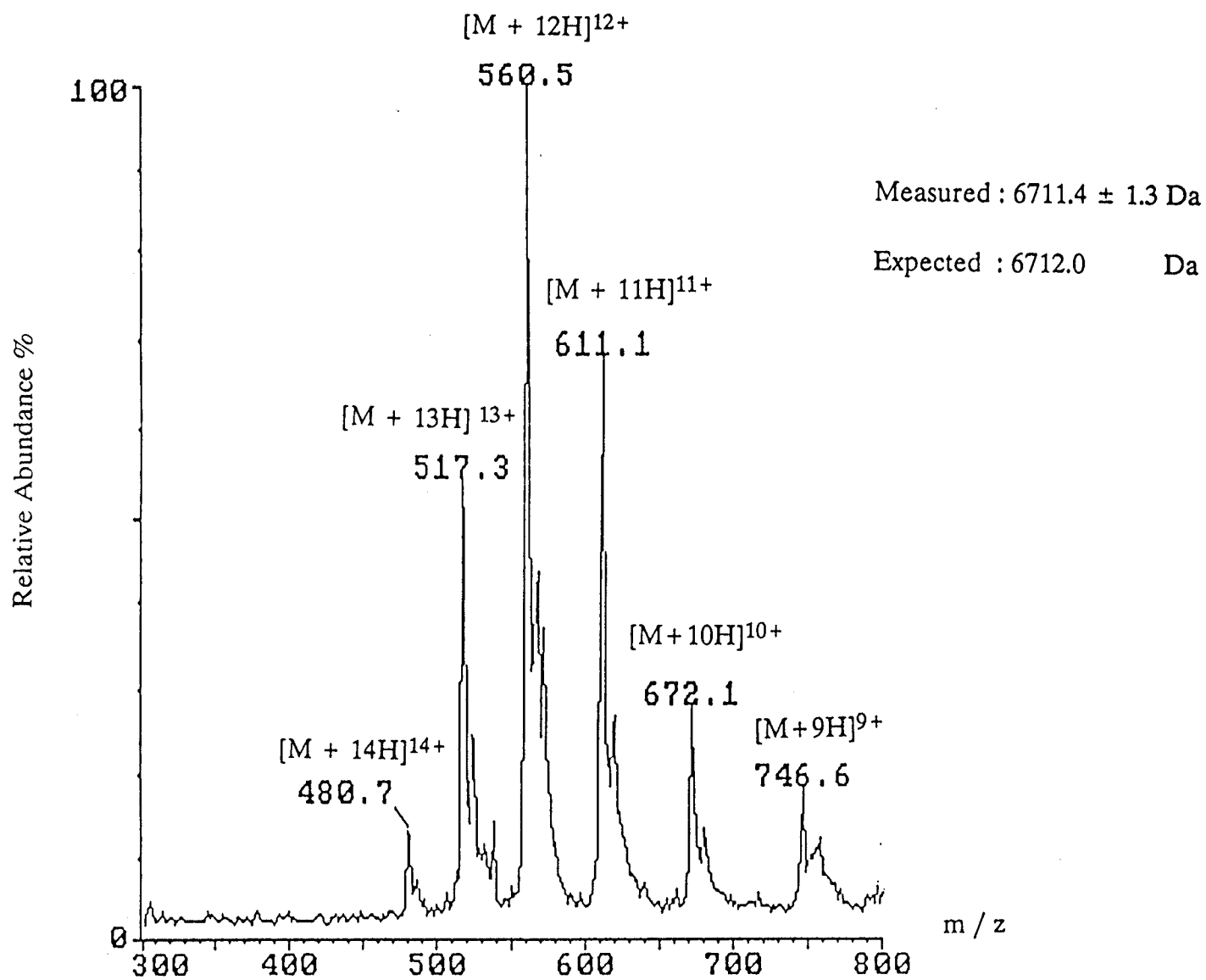


Figure 10

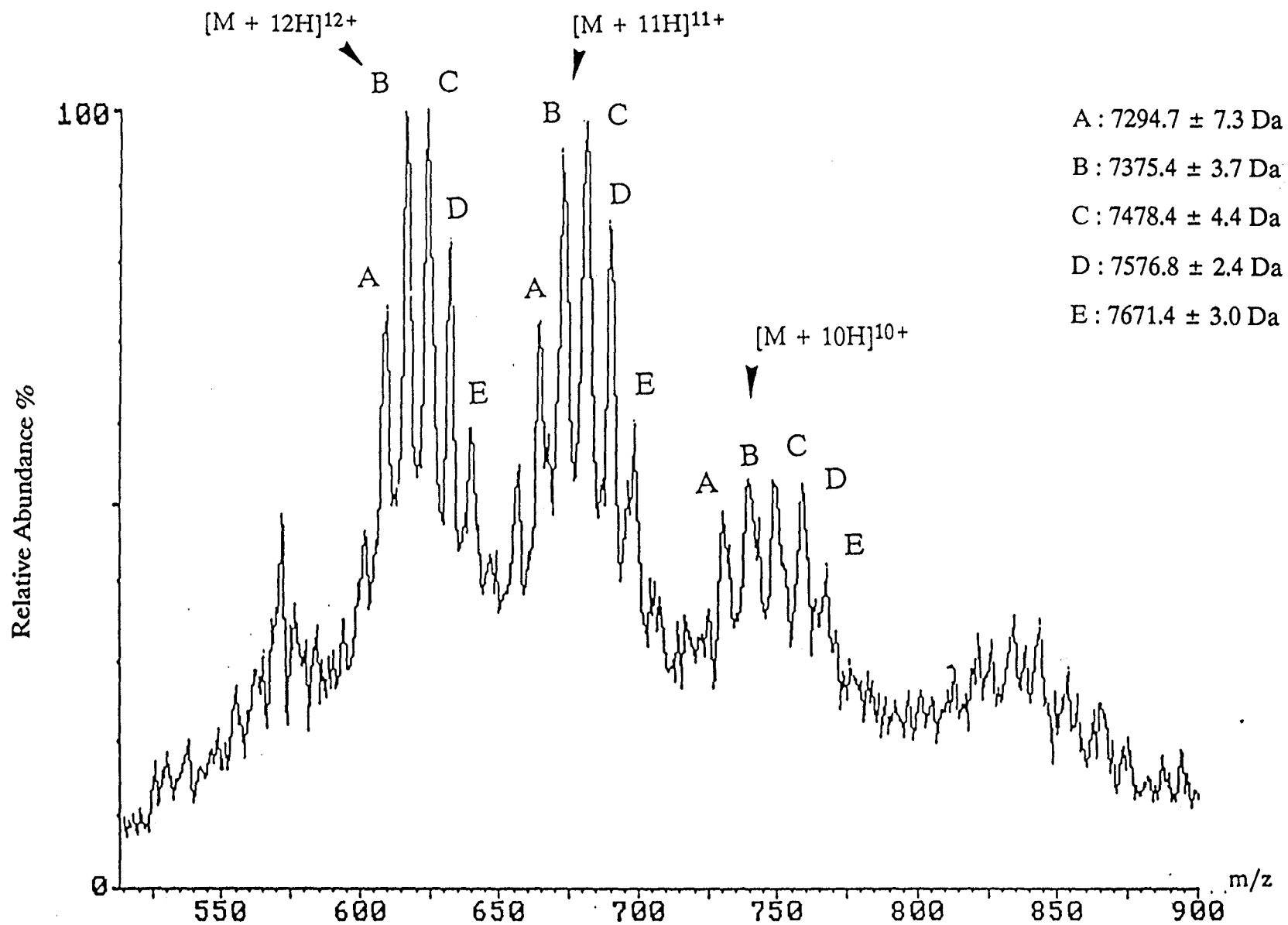


Figure 11

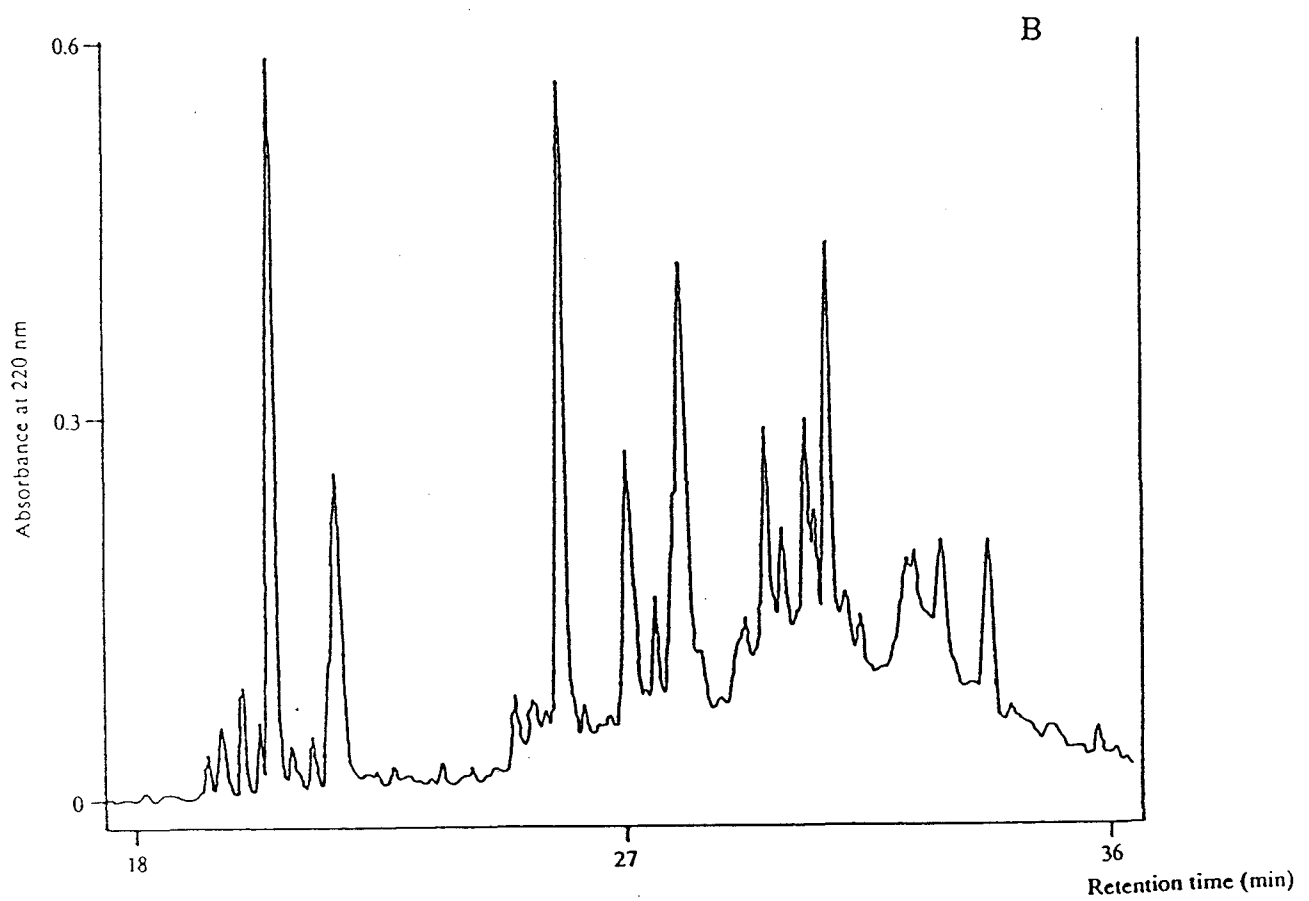
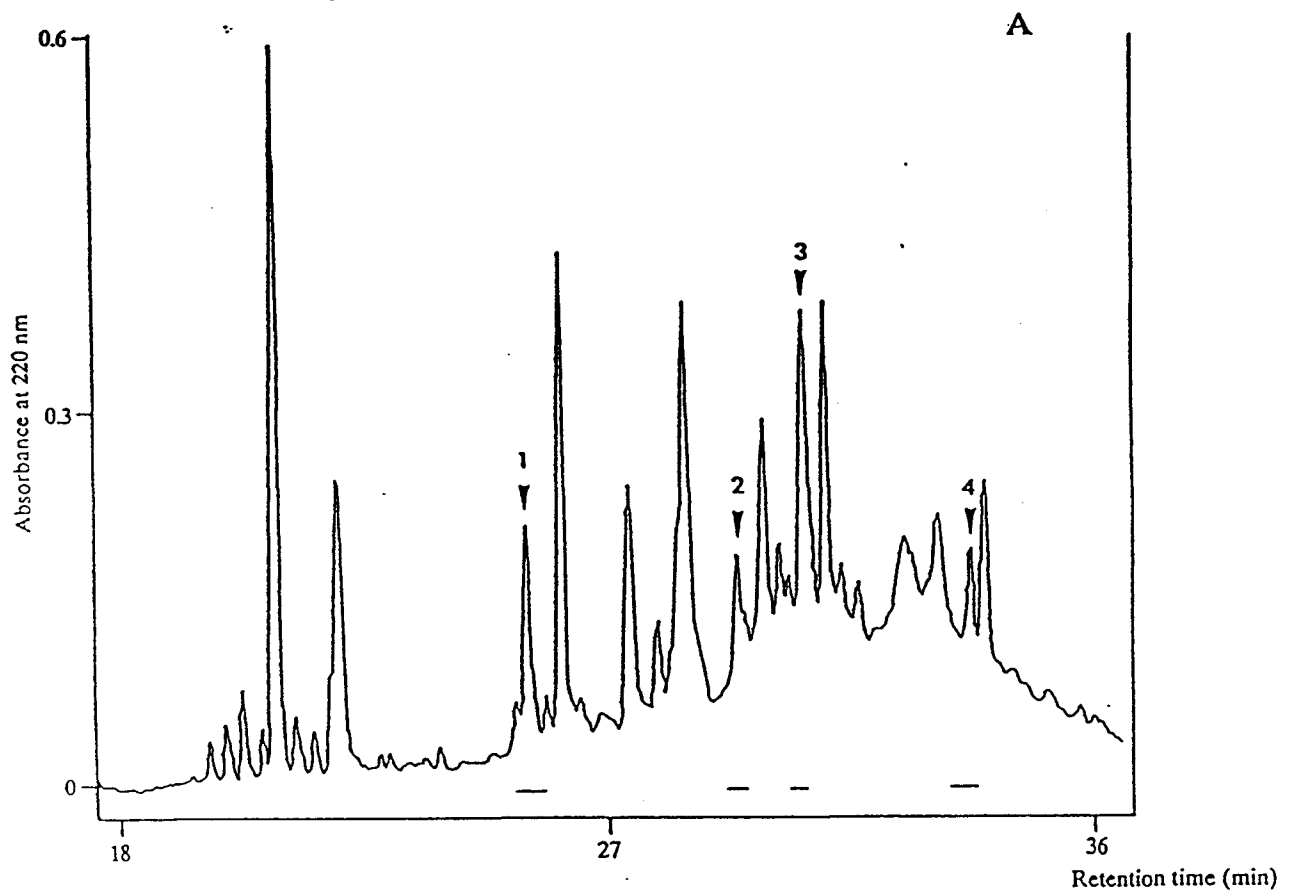


Figure 12

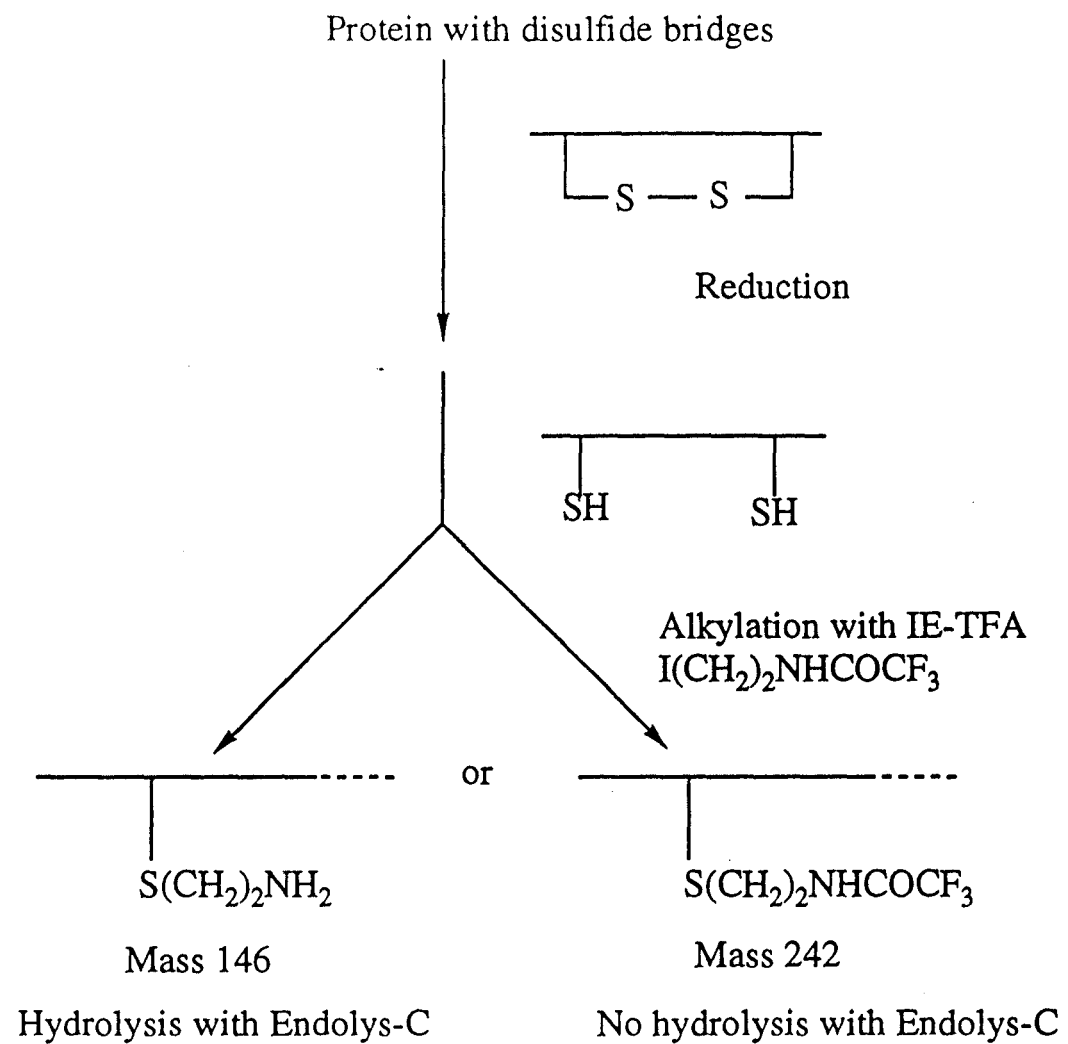


Figure 13

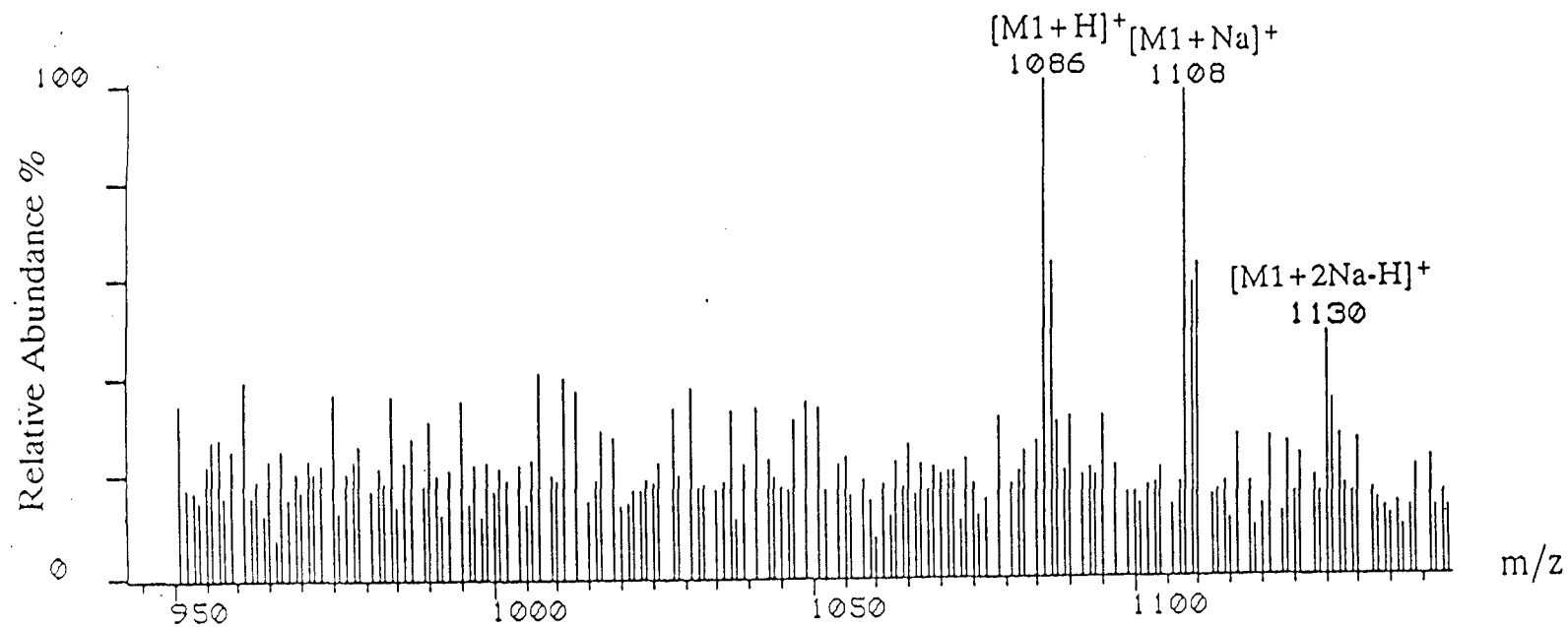
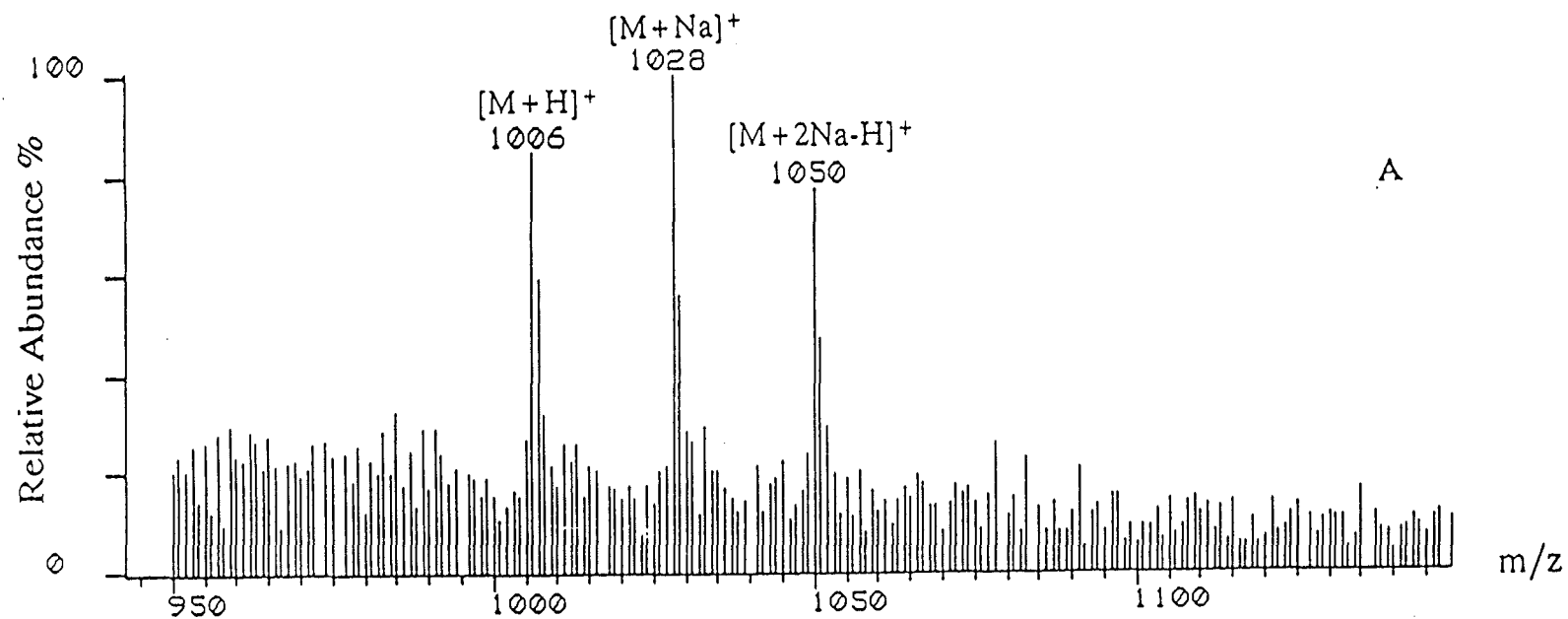


Figure 14

III. Discussion:

III.1. Peptide N-terminal de l'histone H5 de *Xenopus laevis*

Ce travail nous a permis de déterminer la nature de la modification en NH₂-terminal de l'histone H5 de *Xenopus laevis*. Il s'agit d'un groupement acétyl de masse 43 Da.

Le massif isotopique a été déterminant pour montrer la désamidation de l'asparagine. En effet, la FAB, par sa précision, permet de distinguer une fonction amide d'une fonction carboxylique (1uma de différence). La comparaison du profil isotopique du spectre obtenu avec les profils simulés par ordinateur montre clairement qu'il s'agit d'un mélange de deux peptides l'un contenant un résidu d'acide aspartique, l'autre un résidu d'asparagine. La masse de ce peptide a été également vérifiée en FAB mode négatif dans le glycérol/eau/TEA 3/2/1.

III.2. Peptide de synthèse de la cholécystokinine:

L'étude du comportement du groupement sulfate a été réalisée sur un peptide de synthèse de la cholécystokinine (CCK 26-33). Les résultats obtenus en mode négatif, en FAB-MS et en ES-MS montrent la stabilité du sulfate. En mode positif, il est préférable d'utiliser l'ES-MS.

La précision de l'analyse a permis, là encore, de distinguer la fonction carboxylique de la fonction amide. Contrairement au peptide de l'histone H5, le spectre montre que le peptide 26-33 de la cholécystokinine est à 100% sous la forme amide. Nous avons étudié ce peptide dans d'autres matrices. Les meilleurs résultats ont été obtenus dans le glycérol/eau/TFA 10:88:2, glycérol/eau 10:90 ou encore dans le glycérol/eau/triéthylamine 3:2:1 en mode positif et dans le glycérol/eau/triéthylamine 3:2:1 ou triéthylamine/tétraméthylurée 50:50 en mode négatif.

III.3. Peptide carboxamidométhylé de la protamine de béliér:

Par la spectrométrie de masse, nous avons mis en évidence la cyclisation d'une carboxamidométhylcystéine en position N-terminale d'un peptide de la protamine de béliér. Un tel peptide soumis à la dégradation d'Edman ne donne aucun résultat. Nous avons observé le même phénomène

dans un peptide de la protamine mP2 de souris (188) (Fig.30) de séquence C-R-R-H-H-H. La masse $(M+H)^+$ attendue est de 902.4 Da. Le spectre montre deux pics : un majoritaire à 885.4 uma correspondant au peptide sous la forme cyclique de la carboxamidométhylcystéine (902 Da) après départ de NH_3 (17 Da). Le pic à 903.3 uma correspond au peptide avec la cystéine sous la forme carboxyméthylcystéine. Le spectre a été réalisé sur 0.5 nmole de peptide dans la matrice glycérol/eau/TFA 10:88:2. Cette modification peut être comparée à la cyclisation de la glutamine en acide pyroglutamique avec perte de NH_3 (189).

III.4. Phosphorylation de la protamine de béliér:

Sachant que la protamine de béliér est phosphorylée dans le testicule et déphosphorylée dans le spermatozoïde, notre travail était de vérifier en spectrométrie de masse le degré de phosphorylation et de localiser le ou les sites de phosphorylation. Nous avons réalisé ce travail sur la protamine de béliér que nous avons obtenue lors de la purification des protéines PCA insolubles (TP2, TP3 et protamine) (cf. Fig19, page 86). La protamine a été réduite et alkylée par le N-(iodoéthyl)-trifluoroacétamide. Le but de cette alkylation est de pouvoir hydrolyser la protéine TP2 et la protamine par l'endoprotéase Lys-C (hydrolyse par mimétisme au niveau des résidus d'aminoéthylcystéine). En effet, la protamine de béliér est dépourvue de résidus de lysine alors qu'elle contient 7 résidus de cystéine. Sa structure primaire a été déterminée dans notre laboratoire (190). Le spectre de masse que nous avons obtenu en électrospray (Fig.10, publication 3) confirme cette séquence. En effet, la masse attendue de la protamine sous la forme réduite est de 6712 Da; la masse observée est de 6711.4 +/- 1.3 Da. Cette masse confirme également l'existence de la protamine sous la forme déphosphorylée dans le sperme.

L'analyse en ES-MS de la protamine réduite et alkylée par l'IE-TFA, isolée du testicule de béliér, a donné un spectre avec 5 familles de pics chargés de 10 à 12 fois (Fig.11, publication 3). Les masses attendues pour la protamine réduite et alkylée sous la forme non- et mono-phosphorylée, sont respectivement 7013 Da et 7093 Da. Les masses observées sont supérieures d'environ 288 uma par rapport aux masses attendues. L'hypothèse de l'alkylation avec l'IE-TFA sans départ de un ou plusieurs groupements TFA (de masse = 96 Da) a été confirmée par l'analyse des peptides, issus d'hydrolyses enzymatiques, en FAB. La différence de masse de 80 uma entre

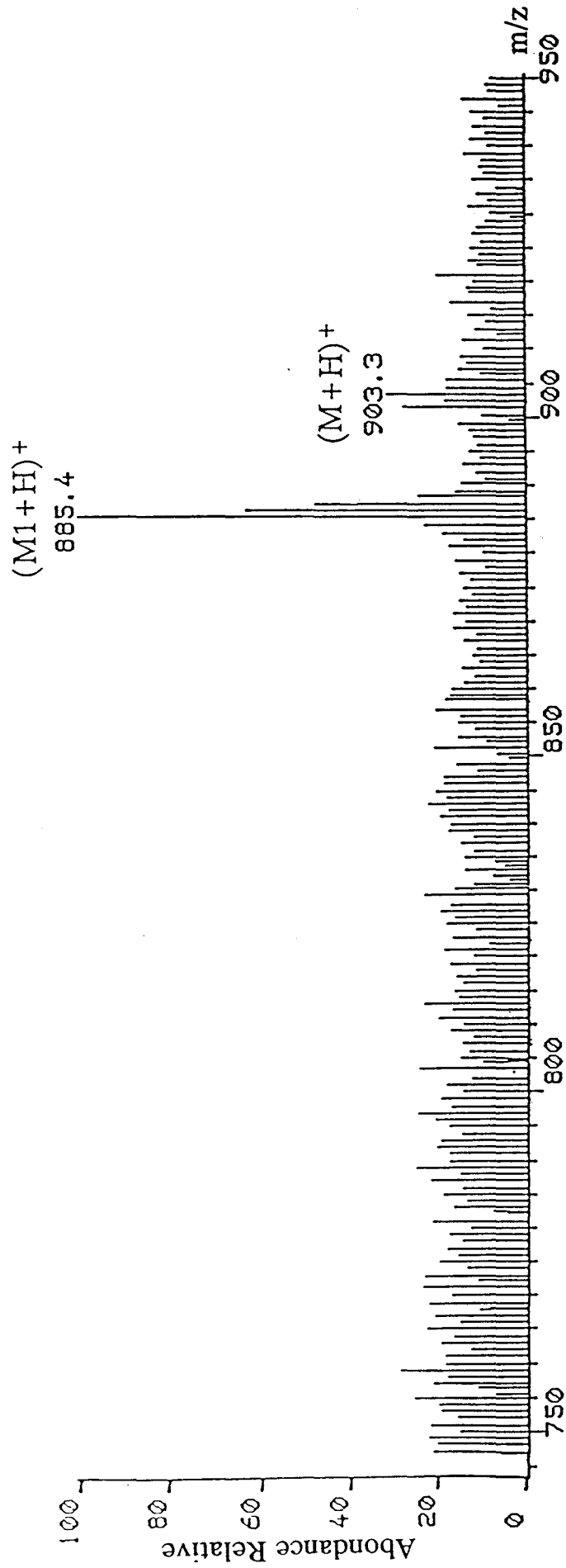


Figure30: Spectre FAB-MS du peptide CRRHHH de la protamine mP2 de souris.

la première et la deuxième série de pics, confirme l'existence d'une forme monophosphorylée et d'une forme non phosphorylée. Les autres séries de pics sont attribuées à la protamine additionnée d'adduits non covalents de phosphate et/ou de sulfate de masse 98 Da.

La protamine, avant et après traitement par la phosphatase alcaline, a été soumise à une hydrolyse par l'endoprotéase Lys-C. La présence d'un groupement phosphate a été confirmée par la disparition de certains pics (numérotés de 1 à 4) dans le chromatogramme de la protamine après traitement par la phosphatase (Fig.12, publication3). Ces 4 peptides provenant de l'hydrolyse limitée par la présence ou non de groupement TFA sur les résidus de cystéine, contiennent deux sites probables de phosphorylation (Thréonine et Tyrosine) de la protéine. Les autres fractions (Fig.12, publication3) ont été analysées en FAB. Les masses observées correspondent soit aux masses attendues, soit à des masses de peptides contenant un ou plusieurs résidus de cystéine sous la forme TFA-aminoéthylcystéine (Tableau IV).

Dans le but de séparer, dans deux peptides différents, les sites potentiels de phosphorylation, nous avons hydrolysé la protamine par l' α -chymotrypsine à pH 5.0. L'analyse des différentes fractions séparées en CLHP en FAB-MS nous a permis de vérifier les masses des peptides attendus. La masse observée du peptide contenant la tyrosine correspond à la masse attendue sans phosphorylation. Quant au peptide contenant la thréonine (43 à 50), la masse 1086 uma observée correspond à la masse calculée avec le groupement phosphate. Ce résultat montre clairement que la thréonine en position 43 est le site de phosphorylation de la protamine de bélier.

L'ensemble de ce travail nous permet d'affirmer que:

- la protamine de bélier est phosphorylée au niveau de la thréonine 43,
- dans le testicule, la protamine existe sous les deux formes, non et mono phosphorylée avec un rapport de 2/1 environ, alors que dans le sperme, la protamine est non phosphorylée,

-l'alkylation par l'IE-TFA est efficace pour réaliser des hydrolyses par l'endoprotéase Lys-C. Mais le groupement TFA reste parfois fixé sur le groupement alkylant contrairement à ce qui a été publié par Schwartz et *al.* (ref. 40, publication3). Dans ce cas, l'hydrolyse par l'endoprotéase Lys-C n'a pas lieu au niveau de ce résidu.

Peptides attendus après hydrolyse par l'endoprotéase Lys-C de la protamine isolée du testicule de bœlier (dans l'ordre de la séquence)	Masses (M+H) ⁺ calculées	Masses (M+H) ⁺ observées	Commentaires sur les masses inattendues
1 5 ARYRC(C)	711 / 857	711 / 953	Présence du groupement TFA sur la cystéine en position 5
6 14 (C)LTHSRSRC	1002 / 1148	1002 / 1148 1244	Présence du groupement TFA sur la cystéine en position 6
15 22 RRRRRRRC	1258	1258	
23 38 RRRRRRFGRRRRRVC(C)	2341 / 2487	-	
39 47 (C)RRYTVVRC	1095 / 1241	1095 / 1241/ 1175 / 1337 1417	1175: Présence d'un groupement phosphate dans le peptide 40-47 1337: Peptide 39-47 avec un groupement TFA sur la cystéine 39 1417: Présence d'un groupement phosphate dans le peptide observé à 1337
48 50 TRQ	404	-	
Peptides non prévus après hydrolyse par l'endolys-C:			
40 50 RRYVVRCTRQ	1481	1577 / 1657	1577: Présence d'un groupement TFA sur la cystéine 47 1657: Présence d'un groupement phosphate dans le peptide observé à 1577
'39 50 CRRYTVVVRCTRQ	1626	1818 / 1898	1818: Présence d'un groupement TFA sur les cystéines 39 et 47 1898: Présence d'un groupement phosphate dans le peptide observé à 1818

Tableau IV

Masses observées en FAB sur les peptides obtenus après hydrolyse par l'endoprotéase Lys-C de la protamine isolée du testicule de bœlier. Commentaires sur les masses inattendues.

Contrairement à la sulfatation, la phosphorylation est stable en ES-MS et en FAB en mode positif. Nous avons également montré cette stabilité sur la protéine de transition T isolée du testicule de calmar. L'analyse de cette protéine en ES-MS a montré 5 familles de pics multichargés de 11 à 15 charges (Fig.31). La masse calculée de cette protéine est de 10 790 Da. Les pics de la série A donnent une masse de 10951.2 +/- 2.7 Da. La différence de 160 Da entre la masse observée et la masse calculée correspond à deux groupements phosphate. Les séries B-E montrent que la protéine se trouve également sous des formes tri à hexaphosphorylées.

III.5. Acétylation et méthylation de l'histone H4 de thymus de veau:

Les modifications post-traductionnelles qui affectent le plus souvent les histones sont l'acétylation, la méthylation et la phosphorylation. Nous avons montré que l'histone H5 de *Xenopus laevis* était acétylée en position N-terminale. Nous avons étudié la méthylation et l'acétylation au niveau de l'histone H4 de thymus de veau. Ces modifications n'avaient jamais été étudiées par spectrométrie de masse au niveau de cette histone. Après hydrolyse de l'histone H4 monoacétylée par l'acide acétique 0.25M pendant 16h à 105°C, les peptides ont été séparés en CLHP-phase inverse (Fig.32). Nous avons utilisé deux critères pour distinguer le peptide 1-23, élué dans la fraction 2, des autres peptides: l'absorbance à 275 nm et l'indice de Bull et Breese. En effet, ce peptide est le seul qui ne contient pas de tyrosine et son indice de Bull et Breese à +400 indique un caractère très hydrophile.

Les peptides obtenus par hydrolyse du peptide 1-23 par l'endoprotéase Lys-C ont été séparés en CLHP-phase inverse et analysés en FAB-MS. Les résultats confirment l'acétylation de cette histone en NH₂-terminal. Ils montrent par ailleurs qu'il existe 2 sites de monoacétylation (acétylation de la chaîne latérale de la lysine), un site mineur (lysine 12) et un site majeur (lysine 16). Quant à la lysine en position 20, elle est sous la forme de diméthyllysine.

Nous avons également analysé en FAB l'hydrolysats total du peptide 1-23. Les figures 33 et 34 montrent les spectres obtenus respectivement dans le glycérol/eau/TFA et glycérol/eau/TEA. L'absence des pics de masses 729 et 788 uma correspondant aux peptides K5 et K2 est due au phénomène de suppression. En effet, leurs indices de Bull et Breese respectivement de +390 et de +567 sont supérieurs aux indices des autres peptides présents. On note

également l'absence des pics de cationisation $(M+Na)^+$ quand la matrice utilisée est le glycérol/eau/TEA. Les sels sont piégés par celle-ci.

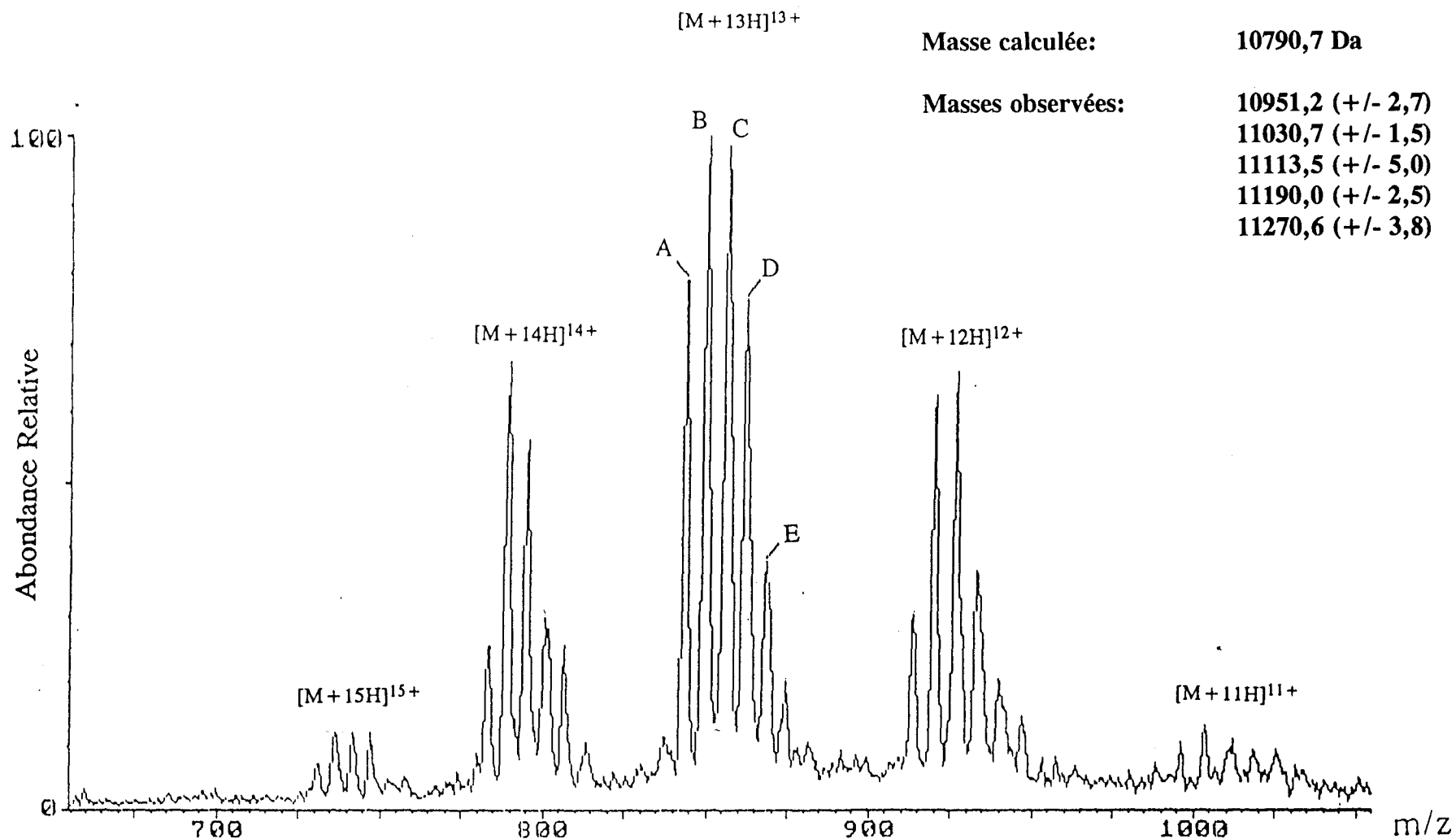


Figure31: Spectre ES-MS de la protéine de transition T isolée du testicule de calmar.

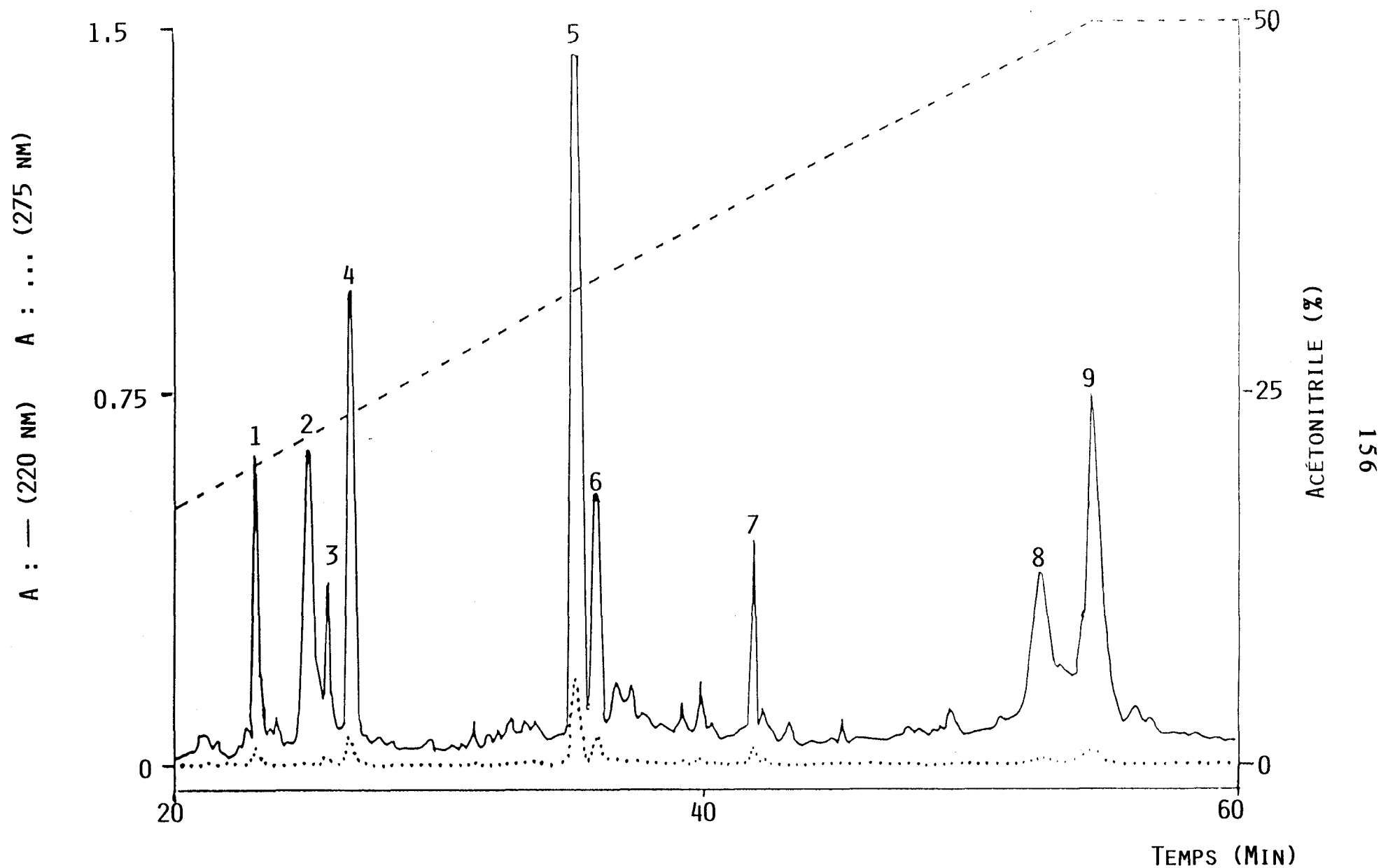


Figure32: Séparation en CLHP-phase inverse des peptides de l'histone H4 monoacétylée de thymus de veau après hydrolyse par l'acide acétique 0.25M.

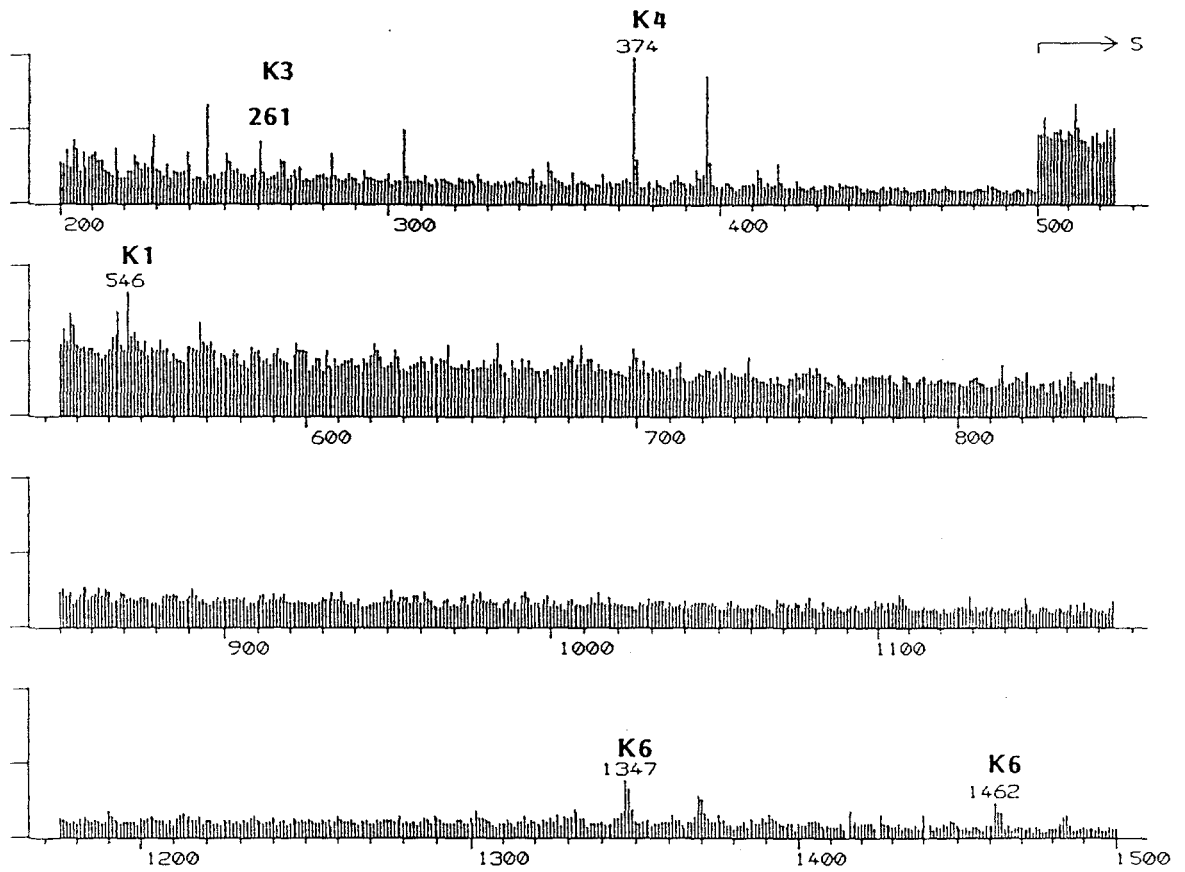


Figure 33: "FAB-mapping" de l'hydrolysate Endolys-C du peptide 1-23 de l'histone H4 monoacétylée de thymus de veau, dans le glycérol/eau/TFA 10:88:2

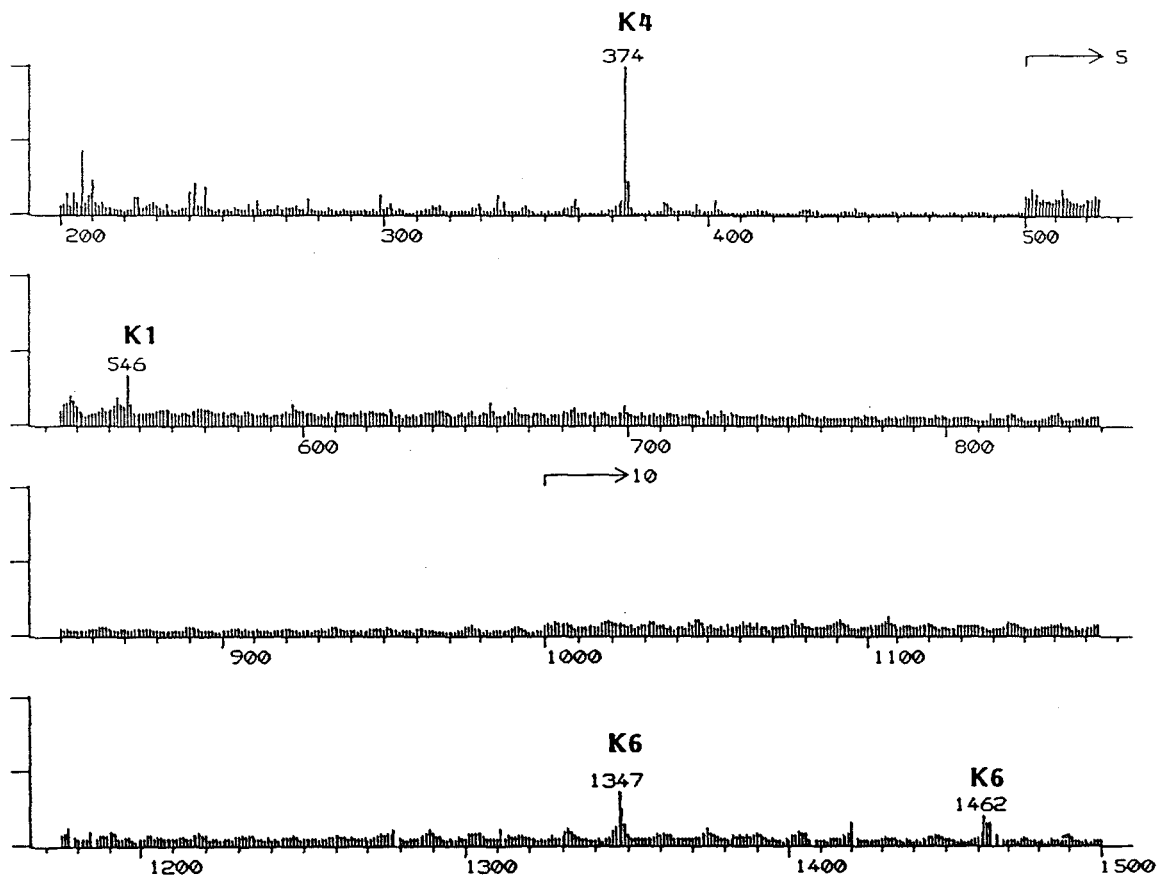


Figure 34: "FAB-mapping" de l'hydrolysate Endolys-C du peptide 1-23 de l'histone H4 monoacétylée de thymus de veau, dans le glycérol/eau/triéthylamine 3/2/1.

IV. Conclusion:

Cette publication montre l'intérêt de la spectrométrie de masse dans l'étude de modifications chimiques telles que l'alkylation, la désamidation ou la cyclisation, et de modifications post-traductionnelles telles que l'acétylation, la méthylation ou la phosphorylation. Les modifications post-traductionnelles qui affectent souvent les protéines basiques de la spermatogenèse (histones, protéines intermédiaires ou protamines) peuvent être actuellement étudiées aisément grâce à la spectrométrie de masse, moyen sûr et rapide, malgré le caractère très hydrophile de ces protéines. La présence de diméthylproline en position NH_2 -terminale des histones H2B d'étoile de mer et de siponcle a également été mise en évidence par la technique FAB (191-192). Toutefois, l'adaptation de cette technique à ce type de protéines ou de peptides peut être améliorée par la recherche de matrices adéquates, par la purification de ces molécules pour l'élimination de sels. Le développement de la CF-FAB dans le but de gagner en sensibilité et de diminuer le phénomène de suppression serait d'un intérêt capital.

Conclusion et Perspectives

Conclusion:

L'étude structurale des protéines basiques de la spermatogenèse est une étape importante et nécessaire pour aborder la relation structure-activité. Pour cela, il est indispensable de connaître la structure exacte de la protéine étudiée ainsi que les modifications post-traductionnelles qui l'affectent.

Nous avons évoqué, dans l'introduction sur la spermatogenèse, l'implication de ces protéines dans la condensation progressive de la chromatine au cours de la spermatogenèse. De nombreux travaux sur l'étude structurale de ces protéines ont été réalisés dans notre laboratoire depuis l'orientation de l'équipe vers cette thématique. Au cours de ces études, différents problèmes ont été rencontrés. En effet, les techniques chromatographiques de séparation sont limitées par le caractère très hydrophile de ces protéines et par l'existence parfois de variants de structure très voisine. La composition en acides aminés ne permet pas de déterminer le nombre exact de résidus d'arginine en proportion trop importante par rapport aux autres résidus (par exemple la protamine Sp de seiche contient 70% de résidus d'arginine). La faible diversité en acides aminés limite les coupures chimiques et enzymatiques. Enfin, la méthode récurrente d'Edman est limitée par les séquences répétitives de groupements arginyl, par les modifications post-traductionnelles, voir même impossible lorsque la fonction NH_2 -terminale de la protéine est bloquée.

Nous avons tenté de pallier ces difficultés en faisant appel à la spectrométrie de masse. L'application de cette technique à l'étude des protéines nucléaires basiques de la spermatogenèse concernait essentiellement deux modes d'ionisation: le bombardement par particules neutres ou FAB et l'électrospray ou ES-MS.

La technique FAB nous a permis de déterminer le groupement acétyl bloquant le NH_2 -terminal de l'histone H5 mais également de positionner les résidus acétylés et méthylés du peptide 1-23 de l'histone H4.

Dans le cas des protéines intermédiaires, chez le bélier, les techniques FAB et ES-MS ont mis en évidence l'existence de variants structuraux de la protéine TP1 et apporté des éléments de structure de la protéine TP2.

Ces techniques ont été déterminantes dans l'élucidation de la structure complète de la protamine Z3 de roussette et dans la localisation du site de phosphorylation de la protamine isolée du testicule de bélier.

Nous avons également, au cours de notre thèse, analysé dans le cadre de collaborations, d'autres molécules telles que des protéines recombinantes (193), des peptides de synthèse, des acylcarnitines ou encore un médicament: le piroxicam. Les résultats obtenus n'ont pas été présentés dans ce mémoire limité aux protéines basiques de la spermatogenèse.

L'ensemble de ces résultats nous a convaincu que la spectrométrie de masse, par ces différents modes d'ionisation (FAB, PD-MS, Laser et ES-MS) est souvent une méthode complémentaire indispensable des techniques classiques de détermination de structure. Nous avons également montré par ce travail la complémentarité des différentes techniques de spectrométrie de masse entre elles. La technique FAB, limitée aux masses inférieures à 10 000 Da, reste une technique très utile pour sa précision mais également par les informations de structure apportées par la fragmentation.

L'ES-MS s'impose de plus en plus comme méthode de vérification de structure et de pureté de différentes molécules, tant dans le domaine des protéines naturelles ou recombinantes que dans celui des peptides de synthèse ou des petites molécules. La précision de ses mesures devrait atteindre 0.001% grâce à l'utilisation d'appareils à secteurs magnétique et électrostatique.

Récemment disponible, la technique de désorption laser, qui supplante actuellement la technique CF-PD-MS, est la troisième solution dans le cas de structure particulière comme les glycoprotéines très riches en glycannes. Facile d'utilisation, elle permettra de résoudre très rapidement la plupart des problèmes structuraux laissant les plus complexes aux techniques FAB et ES-MS.

Perspectives:

Nous espérons, comme perspective à court terme, élucider la structure complète de la protéine intermédiaire TP2 de bélier. Nous étudierons, également, par spectrométrie de masse, les différentes formes acétylées de l'histone H4 isolée du testicule de bélier. La localisation des sites d'acétylation, pour chaque forme, nous permettra d'établir une relation entre la méthylation et l'acétylation de cette histone.

Enfin, dans le cadre du service commun de spectrométrie de masse de l'Université de Lille II récemment créé et équipé d'un Ionspray, et en collaboration avec le service commun de spectrométrie de masse de l'Université de Lille I équipé du FAB-MS, nous voulons contribuer à

l'application de ces techniques aux problèmes structuraux de différents groupes de recherche.

BIBLIOGRAPHIE

- 1) Das B.C., Gero S.D. and Lederer E. (1967)
N-methylation of N-acyl oligopeptides.
Biochem. Biophys. Res. Commun. 29, 211-215
- 2) Biemann K. (1978)
Mass spectrometric sequencing of peptides and proteins.
Pure and Appl. Chem. 50, 149-158
- 3) Barber M., Bordoli R.S., Sedgwick R.D. and Tyler A.N. (1981)
Fast atom bombardment of solids (F.A.B.): A new ion source for mass spectrometry.
J. Chem. Soc. Chem. Commun. 325-327
- 4) Aberth W., Straub K.M. and Burlingame A.L. (1982)
Secondary ion mass spectrometry with cesium ion primary beam and liquid target matrix for analysis of bioorganic compounds.
Anal. Chem. 54, 2029-2034
- 5) McFarlane R.D. (1981)
²⁵²Californium plasma desorption mass spectrometry.
Biomed. Mass Spectrom. 8, 449-453
- 6) Karas M. and Hillenkamp F. (1988)
Laser desorption ionization of proteins with molecular masses exceeding 10000 daltons.
Anal. Chem. 60, 2299-2301
- 7) Whitehouse C.M., Dreyer R.N., Yamashita M. and Fenn J.B. (1985)
Electrospray interface for liquid chromatographs and mass spectrometers.
Anal. Chem. 57, 675-679
- 8) Bruins A.P., Covey T.R. and Henion J.D. (1987)
Ion spray interface for combined liquid chromatography/atmospheric pressure ionization mass spectrometry.
Anal. Chem. 59, 2642-2646
- 9) Yergey J., Heller D., Hansen G., Cotter R.J. and Fenselau C. (1983)

Isotopic distributions in mass spectra of large molecules.
Anal. Chem. 55, 353-356

- 10) Buko A.M., Phillips L.R. and Fraser B.A. (1983)
Peptide studies using a fast atom bombardment high field mass spectrometer and data system. 1- Sample introduction, data acquisition and mass calibration.
Biomed. Mass Spectrom. 10, 324-332
- 11) Surnam D.J. and Vickerman J.C. (1981)
Fast atom bombardment quadrupole mass spectrometry.
J. Chem. Soc. Chem. Commun. 324-325
- 12) Picard C. and Couderc F. (1988)
Utilisation d'un appareil de spectrométrie de masse quadripolaire pour l'analyse en FAB-mapping de protéines de synthèse.
Technologies de purification des protéines - Toulouse 419-424
- 13) Celma C. and Giralt E. (1990)
Optimization of the experimental procedures in fast atom bombardment mass spectrometry of peptides with a quadrupole mass spectrometer.
Biomed. Environ. Mass Spectrom. 19, 235-239
- 14) Picard C., Couderc F. and Maftouh M. (1988)
Characterization of two metabolic products of GRF (1-44) NH₂ in human serum by a combination of HPLC, amino-acid analysis and FAB-mapping mass spectrometry.
Technologies de purification des protéines - Toulouse 425-430
- 15) Gaskell S.J. (1992)
Hybrid tandem mass spectrometers in biological research.
Biol. Mass Spectrom. 21, 413-419
- 16) Buko A.M., Phillips L.R. and Fraser B.A. (1983)
Peptide studies using a fast atom bombardment high field mass spectrometer and data system. 2- Characteristics of positive ionization spectra of peptides, m/z 858 to m/z 5729.

Biomed. Mass Spectrom. 10, 408-429

- 17) Buko A.M., Phillips L.R. and Fraser B.A. (1983)
Peptide studies using a fast atom bombardment high field mass spectrometer and data system. 3- Negative ionization: mass calibration, data acquisition and structural characterization.
Biomed. Mass Spectrom. 10, 387-393
- 18) Falick A.M., Wang G.H. and Walls F.C. (1986)
Ion source for liquid matrix secondary ionization mass spectrometry.
Anal. Chem. 58, 1308-1311
- 19) Gower J.L. (1985)
Matrix compounds for fast atom bombardment mass spectrometry.
Biomed. Mass Spectrom. 12, 191-196
- 20) Fenselau C. and Cotter R.J. (1987)
Chemical aspects of fast atom bombardment.
Chem. Rev. 87, 501-512
- 21) Cook K.D., Todd P.J. and Friar D.H. (1989)
Physical properties of matrices used for fast atom bombardment.
Biomed. Environ. Mass Spectrom. 18, 492-497
- 22) Sunner J., Morales A. and Kebarle P. (1988)
Mechanism of formation of FAB spectra.
Int. J. Mass Spectrom. Ion Phys. 86, 169-186
- 23) Martin S.A., Costello C.E. and Biemann K. (1982)
Optimization of experimental procedure for fast atom bombardment mass spectrometry.
Anal. Chem. 54, 2362-2368
- 24) Naylor S., Findeis A.F., Gibson B.W. and Williams D.H. (1986)
An approach toward the complete FAB analysis of enzymic digests of peptides and proteins.
J. Am. Chem. Soc. 108, 6359-6363

- 25) Caprioli R.M., Moore W.T. and Fan T. (1987)
Improved detection of "suppressed" peptides in enzymic digests analysed by FAB mass spectrometry.
Rapid Commun. Mass Spectrom. 1, 15-18

- 26) Naylor S., Moneti G. and Guyan S. (1988)
Hydrophobic effects in the fast atom bombardment mass spectra of proteins and large peptides.
Biomed. Environ. Mass Spectrom. 17, 393-397

- 27) Kouach M. (1990)
Application de la spectrométrie de masse à l'étude structurale des peptides et des protéines.
DEA Sciences de la Vie et de la Santé, Universités de Lille I et de Lille II

- 28) Chirat F., Martinage A., Briand G., Kouach M., Van Dorsselaer A., Loir M. and Sautière P. (1991)
Nuclear transition protein 1 from ram elongating spermatids. Mass spectrometric characterization, primary structure and phosphorylation sites of two variants.
Eur. J. Biochem. 198, 13-20

- 29) Falick A.M. and Maltby D.A. (1989)
Derivatization of hydrophilic peptides for liquid secondary ion mass spectrometry at the picomole level.
Anal. Biochem. 182, 165-169

- 30) Schindler P., Bitsch F., Klarskov K., Roepstorff P., Briand G., Wouters-Tyrou D., Sautière P. and Van Dorsselaer A. (1991)
Cuttlefish sperm protamines. Part 2. Mass spectrometry of protamines and related peptides.
Eur. J. Biochem. 195, 621-629

- 31) Bull H.B. and Breese K. (1974)
Surface tension of amino acid solutions: a hydrophobicity scale of the amino acid residues.
Arch. Biochem. Biophys. 161, 665-670

- 32) Caprioli R.M., Da Gue B., Fan T. and Moore W.T. (1987)
Microbore HPLC/mass spectrometry for the analysis of peptide mixtures using a continuous flow interface.
Biochem. Biophys. Res. Commun. 146, 291-299
- 33) Caprioli R.M. and Fan T. (1986)
High sensitivity mass spectrometric determination of peptides: direct analysis of aqueous solutions.
Biochem. Biophys. Res. Commun. 141, 1058-1065
- 34) Van Breemen R.B., Martin L.B. and Le J.C. (1991)
Continuous flow fast atom bombardment mass spectrometry of oligonucleotides.
J. Am. Soc. Mass Spectrom. 2, 157-163
- 35) Caprioli R.M. (1988)
Analysis of biochemical reactions with molecular specificity using fast atom bombardment mass spectrometry.
Biochemistry 27, 513-521
- 36) Cole S.M., Macrae P.V., Merson J.R., Pullen F.S. and Rance D.J. (1991)
Microbore liquid chromatography coupled to a flow fast atom bombardment probe for the on-line detection of the Tyr-Pro cleavage of a nonapeptide by recombinant HIV-1 protease.
J. Chromatogr. Biomed. Appl. 562, 67-72
- 37) Davis M.T. and Lee T.D. (1992)
Analysis of peptide mixtures by capillary high performance liquid chromatography: a practical guide to small-scale separations.
Protein Science 1, 935-944
- 38) Deterding L.J., Parker C.E., Perkins J.R., Moseley M.A., Jorgenson J.W. and Tomer K.B. (1991)
Capillary liquid chromatography-mass spectrometry and capillary zone electrophoresis-mass spectrometry for the determination of peptides and proteins.
J. Chromatogr. 554, 329-338

- 39) Frenz J., Bourell J. and Hancock W.S. (1990)
High-performance displacement chromatography-mass spectrometry of tryptic peptides of recombinant human growth hormone.
J. Chromatogr. 512, 299-314
- 40) Hemling M.E., Roberts G.D., Johnson W., Carr S.A. and Covey T.R. (1990)
Analysis of proteins and glycoproteins at the picomole level by on-line coupling of microbore high-performance liquid chromatography with flow fast atom bombardment and electrospray mass spectrometry: a comparative evaluation.
Biomed. Environ. Mass Spectrom. 19, 677-691
- 41) Silberring J. and Nyberg F. (1991)
Analysis of tyrosine and methionine-containing neuropeptides by fast atom bombardment mass spectrometry.
J. Chromatogr. Biomed. Appl. 562, 459-467
- 42) Smith J.B., Thévenon-Emeric G., Smith D.L. and Green B. (1991)
Elucidation of the primary structures of proteins by mass spectrometry.
Anal. Biochem. 193, 118-124
- 43) Van Breemen R.B. and Davis R.G. (1992)
Rates of peptide proteolysis measured using liquid chromatography and continuous-flow fast atom bombardment mass spectrometry.
Anal. Chem. 64, 2233-2237
- 44) Coutant J.E., Chen T.M. and Ackermann B.L. (1990)
Interfacing microbore and capillary liquid chromatography to continuous-flow fast atom bombardment mass spectrometry for the analysis of glycopeptides.
J. Chromatogr. Biomed. Appl. 529, 265-275
- 45) Ito Y., Takeuchi T., Ishii D. and Goto M. (1985)
Direct coupling of micro high-performance liquid chromatography with fast atom bombardment mass spectrometry.
J. Chromatogr. 346, 161-166

- 46) Deterding L.J., Moseley M.A., Tomer K.B. and Jorgenson J.W. (1989)
Coaxial continuous flow fast atom bombardment in conjunction with
tandem mass spectrometry for the analysis of biomolecules.
Anal. Chem. 61, 2504-2511
- 47) Kassel D.B., Musselman B.D. and Smith J.A. (1991)
Primary structure determination of peptides and enzymatically
digested proteins using capillary liquid chromatography/mass
spectrometry and rapid linked-scan techniques.
Anal. Chem. 63, 1091-1097
- 48) Kenny P.T.M. and Orlando R. (1992)
Tandem mass spectrometric analysis of peptides at the femtomole
level.
Anal. Chem. 64, 957-960
- 49) Moseley M.A., Deterding L.J., De Wit J.S.M., Tomer K.B., Kennedy R.T.,
Bragg N. and Jorgenson J.W. (1989)
Optimization of a coaxial continuous flow fast atom bombardment
interface between capillary liquid chromatography and magnetic
sector mass spectrometry for the analysis of biomolecules.
Anal. Chem. 61, 1577-1584
- 50) Moseley M.A., Deterding L.J., Tomer K.B. and Jorgenson J.W. (1991)
Nanoscale packed-capillary liquid chromatography coupled with mass
spectrometry using a coaxial continuous-flow fast atom bombardment
interface.
Anal. Chem. 63, 1467-1473
- 51) Tomer K.B., Perkins J.R., Parker C.E. and Deterding L.J. (1991)
Coaxial continuous flow fast atom bombardment for higher-molecular-
weight peptides: comparison with static fast atom bombardment and
electrospray ionization.
Biol. Mass Spectrom. 20, 783-788
- 52) Barber M. and Green B. (1987)
The analysis of small proteins in the molecular weight range 10-24

kDa by magnetic sector mass spectrometry.

Rapid Commun. Mass Spectrom. 1, 80-83

- 53) Kouach M., Belaïche D., Jaquinod M., Couppez M., Kmiecik D., Ricart G., Van Dorsselaer A., Sautière P. and Briand G. (1993)
Application of electrospray and fast atom bombardment mass spectrometry to the identification of chemical and post-translational modifications of proteins and peptides.
Biol. Mass Spectrom. soumis pour publication
- 54) Clench M.R., Garner G.V., Gordon D.B. and Barber M. (1985)
Surface effects in FAB mapping of proteins and peptides.
Biomed. Mass Spectrom. 12, 355-357
- 55) Gorman J.J. (1992)
Mapping post-translational modifications of viral proteins by mass spectrometry.
Trends in Anal. Chem. 11, 96-105
- 56) Gorman J.J., Corino G.L. and Shiell B.J. (1990)
Role of mass spectrometry in mapping strain variation and post-translational modifications of viral proteins.
Biomed. Environ. Mass Spectrom. 19, 646-654
- 57) Pramanik B.N., Tsabopoulos A., Labdon J.E., Trotta P.P. and Nagabhushan T.L. (1991)
Structural analysis of biologically active peptides and recombinant proteins and their modified counterparts by mass spectrometry.
J. Chromatogr. 562, 377-389
- 58) Pucci P., Carestia C., Fioretti G., Mastrobuoni A.M. and Pagano L. (1985)
Protein fingerprint by fast atom bombardment mass spectrometry: characterization of normal and variant human haemoglobins.
Biochem. Biophys. Res. Commun. 130, 84-90
- 59) Tsugita A., Takamoto K., Kamo M. and Iwadata H. (1992)
C-terminal sequencing of protein. A novel partial acid hydrolysis and analysis by mass spectrometry.

- Eur. J. Biochem. 206, 691-696
- 60) Carr S.A., Hemling M.E., Bean M.F. and Roberts G.D. (1991)
Integration of mass spectrometry in analytical biotechnology.
Anal. Chem. 63, 2802-2824
- 61) Biemann K. (1992)
Mass spectrometry of peptides and proteins.
Annu. Rev. Biochem. 61, 977-1010
- 62) Burlingame A.L., Baillie T.A. and Russell D.H. (1992)
Mass Spectrometry
Anal. Chem. 64, 467R-502R
- 63) Torgenson D.F., Skowronski R.P. and McFarlane R.D. (1974)
New approach to the mass spectroscopy of non-volatile compounds.
Biochem. Biophys. Res. Commun. 60, 616-621
- 64) McFarlane R.D. and Torgenson D.F. (1976)
Californium-252 Plasma Desorption Mass Spectroscopy.
Science 191, 920-925
- 65) Widdiyasekera S., Hakansson P. and Sundqvist B. (1988)
Initial velocity distributions of secondary ions in electronic sputtering
of biomolecules.
Nucl. Instr. and Meth. Phys. Res. B33, 836-839
- 66) Mann M., Rahbek-Nielsen H. and Roepstorff P. (1990)
Practical aspects of calibration and effect of non-protein compounds
on spectrum quality in protein analysis by PDMS.
In: "Ion formation from organic solids" 5th IFOS 47-54
- 67) Demirev P., Fenyö D., Hakansson P. and Sundqvist B.U.R. (1991)
Mechanism of formation of alkali metal cation adducts in plasma
desorption mass spectrometry of biomolecules.
Org. Mass Spectrom. 26, 471-475
- 68) Bouchonnet S., Gonnord M.F., Hoppilliard Y. and Mauriac C. (1992)

Fragmentations of dipeptides in plasma desorption mass spectrometry: new rules for the sequence elucidation of oligopeptide. Biol. Mass Spectrom. 21, 576-584

- 69) Talbo G. and Roepstorff P. (1990)
Sequence determination of N-acetylated-N,O-permethyated peptides by plasma desorption mass spectrometry.
Biomed. Environ. Mass Spectrom. 19, 589-596
- 70) Sundqvist B. and McFarlane R.D. (1985)
 ^{252}Cf -Plasma desorption mass spectrometry.
Mass Spectrom. Rev. 4, 421-460
- 71) Jonsson G.P., Hedin A.B., Hakansson P.L., Sundqvist B.U.R., Säve G.S., Nielsen P.F., Roepstorff P., Johansson K.E., Kamensky I. and Lindberg M.S.L. (1986)
Plasma desorption mass spectrometry of peptides and proteins absorbed on nitrocellulose.
Anal. Chem. 58, 1084-1087
- 72) Kamensky I. and Craig A.G. (1987)
 ^{252}Cf -Plasma desorption mass spectrometry: recent advances and applications.
Anal. Instrum. 16, 71-91
- 73) Hoffmann J.A., Chance R.E. and Johnson M.G. (1990)
Purification and analysis of the major components of chum salmon protamine contained in insulin formulations using high-performance liquid chromatography.
Prot. Express. Purif. 1, 127-133
- 74) Jardine I., Scanlan G.F. and Tsarbopoulos A. (1988)
Plasma desorption mass spectrometry of peptides adsorbed on nitrocellulose from a glutathione matrix.
Anal. Chem. 60, 1086-1088
- 75) Cotter (1988)
Plasma desorption mass spectrometry: coming of age.

Anal. Chem. 60, 781A-793A

- 76) Sundqvist B., Kamensky I., Hakansson P., Kjellberg J., Salehpour M., Widdiyasekera S., Fohlman J., Peterson P.A. and Roepstorff P. (1984)
Californium-252 plasma desorption time of flight mass spectroscopy of proteins.
Biomed. Mass Spectrom. 11, 242-256
- 77) Nielsen P.F., Klarskov K., Højrup P. and Roepstorff P. (1988)
Optimization of sample preparation for plasma desorption mass spectrometry of peptides and proteins using a nitrocellulose matrix.
Biomed. Environ. Mass Spectrom. 17, 355-362
- 78) Roepstorff P. (1989)
Plasma desorption mass spectrometry of peptides and proteins.
Acc. Chem. Res. 22, 421-427
- 79) Klarskov K., Breddam K. and Roepstorff P. (1989)
C-terminal Sequence Determination of peptides Degraded with carboxypeptidases of different specificities and Analysed by 252-Cf Plasma Desorption Mass Spectrometry.
Anal. Biochem. 180, 28-37
- 80) Tsarbopoulos A., Gross M., Kumar R. and Jardine I. (1989)
Rapid identification of Calbindin-D28K cyanogen bromide peptide fragments by plasma desorption mass spectrometry.
Biomed. Environ. Mass Spectrom. 18, 387-393
- 81) Tsarbopoulos A., Becker G.W., Occolowitz J.L. and Jardine I. (1988)
Peptide and protein mapping by ²⁵²Cf-plasma desorption mass spectrometry.
Anal. Biochem. 177, 113-123
- 82) Chen L., Cotter R.J. and Stults J.T. (1989)
Plasma desorption mapping of the tryptic digest of 23-kDa recombinant human growth hormone.
Anal. Biochem. 183, 190-194

- 83) Jensen O.N., Højrup P. and Roepstorff P. (1991)
Plasma desorption mass spectrometry as a tool in characterization of abnormal proteins. Application to variant human haemoglobins.
Anal. Biochem. 199, 175-183
- 84) Nielsen P.F. and Roepstorff P. (1989)
Suppression effects in peptide mapping by plasma desorption mass spectrometry.
Biomed. Environ. Mass Spectrom. 18, 131-137
- 85) Schmitter J.M. (1991)
Performances and limits of plasma desorption mass spectrometry in the primary structure determination of proteins.
J. Chromatogr. 557, 359-368
- 86) Aduru S. and Chait B.T. (1991)
Californium-252 plasma desorption mass spectrometry of oligosaccharides and glycoconjugates: control of ionization and fragmentation.
Anal. Chem. 63, 1621-1625
- 87) Townsend R.R., Alai M., Hardy M.R. and Fenselau C.C. (1988)
Assessment of glycosylation-site heterogeneity using plasma desorption mass spectrometry.
Anal. Biochem. 171, 180-191
- 88) Jardine I., Scanlan G., McNeil M. and Brennan P.J. (1989)
Plasma desorption mass spectrometric analysis of mycobacterial glycolipids.
Anal. Chem. 61, 416-422
- 89) Chowdhury S.K. and Chait B.T. (1989)
A mass spectrometric technique for detecting and identifying by-products in the synthesis of peptides.
Anal. Biochem. 180, 387-395

- 90) Jensen O.N., Roepstorff P., Rozynov B., Horanyi M., Szelenyi J., Hollan S.R., Aseeva E.A. and Spivak V.A. (1991)
Plasma desorption mass spectrometry of haemoglobin tryptic peptides for the characterization of a Hungarian alpha-chain variant.
Biol. Mass Spectrom. 20, 579-584
- 91) Tsarbopoulos A. (1989)
Plasma desorption mass spectrometry of natural and recombinant peptides and proteins.
Peptides Res. 2, 258-266
- 92) Khan A.H., Shaikh B., Allen E.H. and Sokoloski E.A. (1988)
Californium-252 plasma desorption mass spectrometry of aminoglycoside antibiotics.
Biomed. Environ. Mass Spectrom. 1, 329-335
- 93) Viari A., Ballini J.P., Vigny P., Shire D. and Dousset P. (1986)
²⁵²Cf-plasma desorption mass spectrometry of unprotected tri- and tetra-deoxyribonucleotides.
In: "Mass Spectrometry in the anal. of large molecules", 199-205
Ed: McNeal C.J.
- 94) Roepstorff P., Nielsen P.F., Klarskov K. and Højrup P. (1988)
Applications of plasma desorption mass spectrometry in peptide and protein chemistry.
Biomed. Environ. Mass Spectrom. 16, 9-18
- 95) Karas M., Bachmann D., Bahr U. and Hillenkamp F. (1987)
Matrix-assisted ultraviolet laser desorption of non-volatile compounds
Int. J. Mass Spectrom. Ion Proc. 78, 53-68
- 96) Hillenkamp F., Karas M., Beavis R.C. and Chait B.T. (1991)
Matrix-assisted laser desorption/ionization mass spectrometry of biopolymers.
Anal. Chem. 63, 1193A-1203A
- 97) Beavis R.C. and Chait B.T. (1989)

High accuracy molecular mass determination of proteins using matrix-assisted laser desorption mass spectrometry.

Anal. Chem. 62, 1836-1840

- 98) Eckerskorn C., Strupat K., Karas M., Hillenkamp F. and Lottspeich F. (1992)
Mass spectrometric analysis of blotted proteins after gel electrophoretic separation by matrix-assisted laser desorption/ionization.
Electrophoresis 13, 664-665
- 99) Karas M., Ingendoh A., Bahr U. and Hillenkamp F. (1989)
Ultraviolet-laser desorption/ionization mass spectrometry of femtomolar amounts of large proteins.
Biomed. Environ. Mass Spectrom. 18, 841-843
- 100) Spengler B. and Cotter R.J. (1990)
Ultraviolet laser desorption/ionization mass spectrometry of proteins above 100 000 Daltons by pulsed ion extraction time-of-flight analysis.
Anal. Chem. 62, 793-796
- 101) Mock K.K., Davey M. and Cottrell J.S. (1991)
The analysis of underivatised oligosaccharides by matrix-assisted laser desorption mass spectrometry.
Biochem. Biophys. Res. Commun. 77, 644-651
- 102) Finnigan Mat (1991)
Characterization of glycoprotein carbohydrates using Lasermat.
Anal. Biochem. Lasermat Appl. Data Sheet N°2
- 103) Jacobson K.B., Arlinghaus H.F., Buchanan M.V., Chen C.H., Glish G.L., Hettich R.L. and McLuckey S.A. (1991)
Application of mass spectrometry to DNA sequencing.
Gata 8, 223-229
- 104) Finnigan Mat (1991)
Quality control of peptide synthesis using LASERMAT to check for purity and structural integrity.

Anal. Biochem. Lasermat Appl. Data Sheet N°3

- 105) Steiner V., Börnsen K.O., Schär M., Gassmann E., Hoffstetter-Kuhn S., Rink H and Mutter M. (1992)
Analysis of synthetic peptides using matrix-assisted laser desorption ionization mass spectrometry.
Peptides Res. 5, 25-29
- 106) Chait B.T. and Kent S.B.H. (1992)
Weighing naked proteins: practical, high-accuracy mass measurement of peptides and proteins.
Science 257, 1885-1894
- 107) Beavis R.C. and Chait B.T. (1990)
Rapid, sensitive analysis of protein mixtures by mass spectrometry.
Proc. Natl. Acad. Sci. USA 87, 6873-6877
- 108) Chiarelli M.P. and Gross M.L. (1989)
Amino acid and tripeptide mixture analysis by laser desorption Fourier transform mass spectrometry.
Anal. Chem. 61, 1895-1900
- 109) Castoro J.A., Chiu R.W., Monning C.A. and Wilkins C.L. (1992)
Matrix-assisted laser desorption/ionization of capillary electrophoresis effluents by Fourier transform mass spectrometry.
J. Am. Chem. Soc. 114, 7571-7572
- 110) Keough T., Takigiku R., Lacey M.P. and Purdon M. (1992)
Matrix-assisted laser desorption mass spectrometry of proteins isolated by capillary zone electrophoresis.
Anal. Chem. 64, 1594-1600
- 111) Köster C., Castoro J.A. and Wilkins C.L. (1992)
High-resolution matrix-assisted laser desorption/ionization of biomolecules by Fourier transform mass spectrometry.
J. Am. Chem. Soc. 114, 7572-7574

- 112) Dole M., Mach L.L., Hines R.L., Mobley R.C., Fergusson L.D. and Alice M.B. (1968)
Molecular beams of macroions.
J. Chem. Phys. 49, 2240-2249
- 113) Fenn J.B., Mann M., Meng C.K., Wong S.F. and Whitehouse C.M. (1989)
Electrospray ionization for mass spectrometry of large biomolecules.
Science 246, 64-71
- 114) Covey T.R., Bonner R.F., Shushan B.I. and Henion J.D. (1988)
Determination of protein, oligonucleotide and peptide molecular weights by ion spray mass spectrometry.
Rapid Commun. Mass Spectrom. 2, 249-256
- 115) Fenn J.B., Mann M., Meng C.K., Wong S.F. and Whitehouse C.M. (1990)
Electrospray ionization-principles and practice.
Mass Spectrom. Rev. 9, 37-70
- 116) Iribarne J.V. and Thomson B.A. (1976)
On the evaporation of small ions from charged droplets.
J. Chem. Phys. 64, 2287-2294
- 117) Thomson B.A. and Iribarne J.V. (1979)
Field induced ion evaporation from liquid surfaces at atmospheric pressure.
J. Chem. Phys. 71, 4451-4463
- 118) Roellgen F.W., Bramer-Weger E. and Buetfering L. (1987)
Field ion emission from liquid solutions: ion evaporation against electrohydrodynamic disintegration.
J. Phys. Col. (Paris) C6, 48, 253-258
- 119) Smith R.D., Loo J.A., Edmonds C.G., Barinaga C.J. and Udseth H.R. (1990)
New developments in biochemical mass spectrometry: electrospray ionization.
Anal. Chem. 62, 882-899

- 120) Ikonomou M.G., Blades A.T. and Kebarle P. (1991)
Electrospray-ion spray: a comparison of mechanisms and performance.
Anal. Chem. 63, 1702-1708
- 121) Hiraoka K. and Kudaka I. (1992)
Formation of multiply charged ions of the oligopeptide Arg-Arg-Arg by electrospray ionization.
Anal. Chem. 64, 75-81
- 122) Chowdhury S.K., Katta V., Beavis R.C. and Chait B.T. (1990)
Origin and removal of adducts (molecular mass=98u) attached to peptide and protein ions in electrospray ionization mass spectra.
J. Am. Soc. Mass Spectrom. 1, 382-388
- 123) Voyksner R.D. (1992)
Electrospray LC/MS- Can it be used to determine lower molecular weight molecules ?
Nature 356, 86-87
- 124) Loo J.A., Edmonds C.G., Udseth H.R. and Smith R.D. (1990)
Effect of reducing disulfide-containing proteins on electrospray ionization mass spectra.
Anal. Chem. 62, 693-698
- 125) Bojensen G. (1987)
The order of proton affinities of the 20 common L- α -amino acids
J. Am. Chem. Soc. 109, 5557-5558
- 126) Kouach M., Jaquinod M., Belaïche D., Sautière P., Van Dorsselaer A., Chevaillier P. and Briand (1993)
A corrected primary structure for dog-fish *Scylliorhinus caniculus* protamine Z3.
Biochim. Biophys. Acta 1162, 99-104
- 127) Loo J.A., Loo R.R.O., Light K.J., Edmonds C.G. and Smith R.D. (1992)
Multiply charged negative ions by electrospray ionization of polypeptides and proteins.

- Anal. Chem. 64, 81-88
- 128) Mann M., Meng C.K. and Fenn J.B. (1989)
Interpreting mass spectra of multiply charged ions.
Anal. Chem. 61, 1702-1708
- 129) Braun T. and Zsindely S. (1992)
Electrospray ionization-mass spectrometry: a quantitative snapshot on
an instrumental breakthrough.
Trends in Anal. Chem. 11, 307-309
- 130) Mutigner R., Gagnon J., Van Dorsselaer A. and Job D. (1992)
Stabilization of sea urchin flagellar microtubules by histone H1.
Nature 360, 33-39
- 131) Reymann J.M., Rondeau J.M., Barth P., Jaquinod M., Van Dorsselaer A.
and Biellmann J.F. (1992)
Purification and electrospray mass spectrometry of aldose reductase
from pig lens.
Biochim. Biophys. Acta 1122, 1-5
- 132) Van Dorsselaer A., Bitsch F., Green B., Jarvis S., Lepage P., Bischoff R.,
Kolbe H.V.J. and Roitsch C. (1990)
Application of electrospray mass spectrometry to the characterization
of recombinant proteins up to 44 kDa.
Biomed. Environ. Mass Spectrom. 19, 692-704
- 133) Feng R. and Konishi Y. (1992)
Analysis of antibodies and other large glycoproteins in the mass range
of 150 000-200 000 Da by electrospray ionization mass spectrometry.
Anal. Chem. 64, 2090-2095
- 134) Briand J.P., Van Dorsselaer A., Raboy B. and Muller S. (1989)
Total chemical synthesis of ubiquitin using BOP reagent: biochemical
and immunochemical properties of the purified synthetic product.
Peptides Res. 2, 381-388

- 135) Bischoff R., Lepage P., Jaquinod M., Cauet G., Akerklein M., Clesse D., Laporte M., Bayol A., Van Dorsselaer A. and Roitsch C. (1993)
Sequence-specific deamidation: isolation and biochemical characterization of succinimide intermediates of recombinant Hirudin.
Biochemistry 32, 725-734

- 136) Dass C., Kusmierz J.J., Desiderio D.M., Jarvis S.A. and Green B.N. (1991)
Electrospray mass spectrometry for the analysis of opioid peptides and for the quantification of endogenous methionine enkephalin and β -endorphin.
J. Am. Soc. Mass Spectrom. 2, 149-156

- 137) De Caterina M., Esposito P., Grimaldi E., Di Maro G., Scopacasa F., Ferranti P., Parlapiano A., Malorni A., Pucci P. and Marino G. (1992)
Characterization of hemoglobin Lepore variants by advanced mass-spectrometric procedures.
Clin. Chem. 38, 1444-1448

- 138) Oliver R.W.A. and Green B.N. (1991)
The application of electrospray mass spectrometry to the characterisation of abnormal or variant haemoglobins.
Trends in Anal. Chem. 10, 85-91

- 139) Wada Y. (1992)
Mass spectrometry in the integrated strategy for the structural analysis of protein variants.
Biol. Mass Spectrom. 21, 617-624

- 140) Promé D., Blooquit Y., Ponthus C., Promé J.C. and Rosa J. (1991)
Structure of the human adult hemoglobin minor fraction A_{1b} by electrospray and secondary ion mass spectrometry. Pyruvic acid as amino-terminal blocking group.
J. Biol. Chem. 266, 13050-13054

- 141) Chowdhury S.K., Katta V. and Chait B.T. (1990)

Electrospray ionization mass spectrometric peptide mapping: a rapid, sensitive technique for protein structure analysis.

Biochem. Biophys. Res. Commun. 167, 686-692

- 142) Fauda H., Nocerini M., Schneider R. and Gedutis C. (1991)
Quantitative analysis by high-performance liquid chromatography atmospheric pressure chemical ionization mass spectrometry: the determination of the renin inhibitor CP-80, 794 in human serum.
J. Am. Soc. Mass Spectrom. 2, 164-167

- 143) Ikonomou M.G., Blades A. and Kebarle P. (1990)
Investigations of the electrospray interface for liquid chromatography/mass spectrometry.
Anal. Chem. 62, 957-967

- 144) Garcia F. and Henion J.D. (1992)
Fast capillary electrophoresis-ion spray mass spectrometric determination of sulfonylureas.
J. Chromatogr. 606, 237-247

- 145) Garcia F. and Henion J.D. (1992)
Gel-filled capillary electrophoresis/mass spectrometry using a liquid junction-ion spray interface.
Anal. Chem. 64, 985-990

- 146) Tsuji K., Baczynskyj L. and Bronson G.E. (1992)
Capillary electrophoresis-electrospray mass spectrometry for the analysis of recombinant bovine and porcine somatotropins.
Anal. Chem. 64, 1864-1870

- 147) Mass Spectrometry
Eds. McDowell C.A. and McGraw-Hill, Book Comp. (1963)

- 148) Cotter R.J. (1989)
Time-of-flight mass spectrometry: an increase role in the life sciences.
Biomed. Environ. Mass Spectrom. 18, 513-532

- 149) Evans V. and Brayer G.D. (1988)

- Horse heart metmyoglobin. A 2.8 Å resolution three-dimensional structure determination.
J. Biol. Chem. 263, 4263-4268
- 150) Larry N.L., Wright L.G. and Cooks R.G. (1985)
New scan modes accessed with a hybrid mass spectrometer.
Anal. Chem. 57, 2918-2924
- 151) Heerma W., Kammerling J.P., Slotboom A.J., Van Scharrenburg G.J.M., Green B.N. and Lewis I.A.S. (1983)
Fast atom bombardment and collisional activation mass spectrometry of a heptapeptide.
Biomed. Mass Spectrom. 10, 13-16
- 152) Beynon J.H., Cooks R.G., Amy J.W., Baitinger W.E. and Ridley T.Y. (1973)
Design and performance of a mass-analysed ion kinetic energy (MIKE) spectrometer.
Anal. Chem. 45, 1023A-1031A
- 153) Siu K.W.M., Gardner G.J. and Berman S.S. (1988)
Multiply charged ions in ionspray tandem mass spectrometry.
Org. Mass Spectrom. 24, 931-942
- 154) Roepstorff P. and Fohlman J. (1984)
Proposal for a common nomenclature for sequence ions in mass spectra of peptides.
Biomed. Mass Spectrom. 11, 601
- 155) Biemann K. and Martin S.A. (1987)
Mass spectrometric determination of the amino acid sequence of peptides and proteins.
Mass Spectrom. Rev. 6, 1-76
- 156) Fohlman J., Peterson P.A., Roepstorff P., Hörup P., Kamensky I., Säwe G., Häkansson P. and Sundqvist B. (1985)
²⁵²Californium plasma desorption and fast atom bombardment mass spectrometry for analysis of small peptides.

- Biomed. Mass Spectrom. 12, 380-387
- 157) Johnson R.S., Martin S.A. and Biemann K. (1988)
Collision-induced fragmentation of $(M+H)^+$ ions of peptides. Side chain specific ions.
Int. J. Mass Spectrom. Ion Proc. 86, 137-154
- 158) Stults J.T. and Watson J.T. (1987)
Identification of a new type of fragment ion in the collisional activation spectra of peptides allows leucine/isoleucine differentiation.
Biomed. Environ. Mass Spectrom. 14, 583-586
- 159) Bean M.F. and Carr S.A. (1992)
Characterization of disulfide bond position in proteins and sequence analysis of cystine-bridged peptides by tandem mass spectrometry.
Anal. Biochem. 201, 216-226
- 160) Anderegg R.J., Carr S.A., Huang I.Y., Hiipakka R.A., Chang C. and Liano S. (1988)
Correction of the cDNA-derived protein sequence of prostatic spermine binding protein: pivotal role of tandem mass spectrometry in sequence analysis.
Biochemistry 27, 4214-4221
- 161) Gjerset R., Ibarra F., Saragosti S. and Eisen H. (1981)
Distribution of IP25 in chromatin and its possible involvement in chromatin condensation.
Biochem. Biophys. Res. Commun. 99, 349-357
- 162) Gjerset R., Gorka C., Hasthorpe S., Lawrence J.J. and Eisen H. (1982)
Developmental and hormonal regulation of protein H1⁰ in rodents.
Proc. Natl. Acad. Sci. USA 79, 2333-2337
- 163) Loir M. and Lanneau M. (1978)
Transformation of ram spermatid chromatin.
Exp. Cell Res. 115, 231-243

- 164) Dupressoir T., Sautière P., Lanneau M. and Loir M. (1985)
 Isolement and characterization of the ram spermatidal protein P1, 3
 and T.
 Exp. Cell Res. 161, 63-74

- 165) Le Gac-Jegou F. (1978)
 Acétylation et phosphorylation des nucléoprotéines basiques pendant
 la spermiogenèse chez le bélier.
 DEA Physiologie de la Reproduction, Université Paris VI

- 166) Dixon G.H., Candido E.P.M., Honda B.M., Louie A.J., McLeod A.R. and
 Sung M.T. (1975)
 The biological roles of post-synthetic modifications of basic nuclear
 proteins.
 In: "The structure and function of chromatin" Ciba Found. Symp.
 Elsevier, Amsterdam 28, 229-258

- 167) Christensen M. E. and Dixon G.H. (1982)
 Hyperacetylation of histone H4 correlates with the terminal
 transcriptionally inactive stages of spermatogenesis in rainbow trout.
 Develop. Biol. 93, 404-415

- 168) Grimes S.R., Chae C. and Irvin J.L. (1975)
 Acetylation of histones of rat testis.
 Arch. Biochem. Biophys. 168, 425-435

- 169) Grimes S.R. and Henderson N. (1983)
 Acetylation of histones during spermatogenesis in the rat.
 Arch. Biochem. Biophys. 221, 108-116

- 170) Bode J., Henco K. and Wingender E. (1980)
 Modulation of the nucleosome structure by histone acetylation.
 Eur. J. Biochem. 110, 143-152

- 171) Marushige Y. and Marushige K. (1975)
 Transformation of sperm histone during formation and maturation of
 rat spermatozoa.

- J. Biol. Chem. 250, 39-45
- 172) Marushige Y. and Marushige K. (1978)
Phosphorylation of sperm histone during spermiogenesis in mammals.
Biochim. Biophys. Acta 518, 440-449
 - 173) Willmitzer L., Bode J. and Wagner K.G. (1977a)
Phosphorylated protamines. I. Binding stoichiometry and thermal stability of complexes with DNA.
Nucl. Acid. Res. 4, 149-162
 - 174) Willmitzer L., Bode J. and Wagner K.G. (1977b)
Phosphorylated protamines. II. Circular dichroism of complexes with DNA, dependency on ionic strength.
Nucl. Acid. Res. 4, 163-176
 - 175) Ingles C.J. and Dixon G.H (1967)
Phosphorylation of protamine during spermatogenesis in trout testis.
Proc. Natl. Acad. Sci. USA 58, 1011-1018
 - 176) Wouters-Tyrou D., Martin-Ponthieu A., Richard A. and Sautière P. (1988)
Characterization of the nuclear basic proteins specific of spermiogenesis in cuttlefish, *Sepia officinalis*.
Biochim. Biophys. Acta 953, 86-94
 - 177) Kleene K.C. and Flynn J.F. (1987)
Characterization of a cDNA clone encoding a basic protein, TP2 , involved in chromatin condensation during spermiogenesis in the mouse.
J. Biol. Chem. 262, 17272-17277
 - 178) Luerssen H., Maier W.M., Hoyer-Fender S. and Engel W. (1989)
The nucleotide sequence of rat transition protein 2 (TP2) cDNA.
Nucl. Acids Res. 17, 3585
 - 179) Reinhart N., Kremling H., Luerssen H., Adham I.M. and Engel W. (1991)

- Characterization of a gene encoding a basic protein of the spermatid nucleus, TNP2, and its close linkage to the protamine genes in the bull. Biol. Chem. Hoppe-Seyler 372, 431-436
- 180) Keime S., Kumm S., Luerssen H. and Engel W. (1992)
The nucleotide sequence of boar transition protein-2 (TNP2) cDNA and haploid expression of the gene during spermatogenesis.
Animal Gen. 23, 373-378
- 181) Meyer H.E., Hoffmann-Posorske E., Korte H. and Heilmeyer L.M.G. (1986)
Sequence analysis of phosphoserine-containing peptides. Modification for picomolar sensitivity.
FEBS Lett. 204, 61-66
- 182) Chauvière M., Laine B., Sautière P. and Chevaillier P. (1983)
Purification and characterization of two basic spermatid-specific proteins isolated from the dog-fish *Scylliorhinus caniculus*.
FEBS Lett. 152, 231-235
- 183) Gusse M., Sautière P., Chauvière M. and Chevaillier P. (1983)
Extraction, purification and characterization of the sperm protamines of the dog-fish *Scylliorhinus caniculus*.
Biochim. Biophys. Acta 748, 93-98
- 184) Sautière P., Briand G., Gusse M. and Chevaillier P. (1981)
Primary structure of a protamine isolated from the sperm nuclei of the dog-fish *Scylliorhinus caniculus*.
Eur. J. Biochem. 119, 251-255
- 185) Kistler W.S., Noyes C., Hsu R. and Heinrikson R.L. (1975b)
The amino acid sequence of the testis-specific basic protein that is associated with spermatogenesis.
J. Biol. Chem. 250, 1847-1853
- 186) Panyim S. and Chalkley R (1969)
High resolution acrylamide gel electrophoresis of histones.
Arch. Biochem. Biophys. 130, 337-346

- 187) Chirat F. (1990)
Les protéines nucléaires basiques de la spermiogenèse du bœlier
Ovis aries. Structure primaire, phosphorylation.
Thèse de Doctorat de l'Université des Sciences et Techniques de
Lille Flandres-Artois.
- 188) Chauvière M., Martinage A., Debarle M., Sautière P. and Chevaillier P.
(1992)
Molecular characterization of six intermediate proteins in the
processing of mouse protamine P2 precursor.
Eur. J. Biochem. 204, 759-765
- 189) Baldwin M.A., Falick A.M., Gibson B.W., Prisiner S.B., Stahl N. and
Burlingame A.L. (1990)
Tandem mass spectrometry of peptides with N-terminal glutamine.
Studies on a prion protein peptide.
J. Am. Soc. Mass Spectrom. 1, 258-264
- 190) Sautière P., Belaïche D., Martinage A. and Loir M. (1984)
Primary structure of the ram (*Ovis aries*) protamine.
Eur. J. Biochem. 144, 121-125
- 191) Martinage A., Briand G., Van Dorsselaer A., Turner C.H., and Sautière P.
(1985)
Primary structure of Histone H2B from gonads of Starfish *Asterias
rubens*. Identification of a N-dimethylproline residue at the amino-
terminal.
Eur. J. Biochem. 147, 351-359
- 192) Kmiecik D., Belaïche D., Sautière P., Loucheux-Lefebvre M.H. and
Kerckaert J.P. (1991)
Complete Sequence of *Sipunculus nudus* erythrocyte histone H2B and
its gene. Identification of an N,N-dimethylproline at the amino-
terminus.
Eur. J. Biochem. 198, 275-283

Human Bronchial mucus proteinase inhibitor and its amino- and carboxy- terminal domains. Electrospray mass spectrometry characterization and comparison with the recombinant protein.
J. Biochem. soumis pour publication

- 194) Yip T.T. and Hutchens T.W. (1992)
Mapping and sequence-specific identification of phosphopeptides in unfractionated protein digest mixtures by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry.
FEBS Lett. 308, 149-153

- 195) Caprioli R.M., Whaley B., Mock K.K. and Cottrell J.S. (1991)
Sequence-ordered peptide mapping by time-course analysis of protease digests using laser desorption mass spectrometry.
Tech. Prot. Chem. II, 497-510

- 196) Hutchens T.W., Allen M.H., Li C.M. and Yip T.T. (1992)
Occupancy of a C₂-C₂ type "zinc-finger" protein domain by copper. Direct observation by electrospray ionization mass spectrometry.
FEBS Lett. 309, 170-174

- 197) Surovoy A., Waidelich D. and Jung G. (1992)
Nucleocapsid protein of HIV-1 and its Zn²⁺ complex formation analysis with electrospray mass spectrometry.
FEBS Lett. 311, 259-262

- 198) Schroeder W., Covey T. and Hucho F. (1990)
Identification of phosphopeptides by mass spectrometry.
FEBS Lett. 273, 31-35

- 199) Thorneley R.N.F., Abell C., Ashby G.A., Drummond M.H., Eady R.R., Huff S., Macdonald C.J. and Shneier A. (1992)
Posttranslational modification of *Klebsiella pneumoniae* flavodoxin by covalent attachment of coenzyme A, shown by ³¹P NMR and electrospray mass spectrometry, prevents electron transfer from the *nifJ* protein to nitrogenase. A possible new regulatory mechanism for biological nitrogen fixation.

Biochemistry 31, 1216-1224

- 200) Medzihradszky K.F., Gillece-Castro B.L., Settineri C.A., Townsend R.R., Masiarz F.R. and Burlingame A.L. (1990)
Structure determination of O-linked glycopeptides by tandem mass spectrometry.
Biomed. Environ. Mass Spectrom. 19, 777-781

- 201) Promé J.C. (1991)
Caractérisation de modifications post-traductionnelles de protéines par spectrométrie de masse .
Cahiers IMABIO, Spectrométrie de masse 1, 7-16

- 202) Huang E.C. and Henion J.D. (1990)
LC/MS and LC/MS/MS determination of protein tryptic digests.
J. Am. Soc. Mass Spectrom. 1, 158-165

- 203) Nakamura T., Nagaki H., Ohki Y. and Kinoshita T. (1990)
Differentiation of leucine and isoleucine residues in peptides by consecutive reaction mass spectrometry.
Anal. Chem. 62, 311-313

- 204) Williams D., Bradley C., Bojesen G., Santikarn S and Taylor L.C.E. (1981)
Fast atom bombardment mass spectrometry: a powerfull technique for the study of polar molecules.
J. Am. Chem. Soc. 103, 5700-5704

- 205) Williams D.H., Findeis F., Naylor S. and Gibson B.W. (1987)
Aspects of the production of FAB and SIMS mass spectra.
J. Am. Chem. Soc. 109, 1980-1986

- 206) Schnölzer M., Jones A., Alewood P.F. and Kent S.B.H. (1992)
Ion-spray tandem mass spectrometry in peptide synthesis: structural characterization of minor by-products in the synthesis of ACP (65-74).
Anal. Biochem. 204, 335-343

- 207) Loo J.A., Edmonds C.G. and Smith R.D. (1990)

Primary sequence information from intact proteins by electrospray ionization tandem mass spectrometry.

Science 248, 201-204

Annexes

Masses et Compositions des 20 acides aminés usuels

acide aminé	composition	masse monoisotopique	masse chimique	I _{BB}
Ala (A)	C ₃ H ₅ NO	71.03711	71.0788	- 200
Arg (R)	C ₆ H ₁₂ N ₄ O	156.10111	156.1875	- 120
Asn (N)	C ₄ H ₆ N ₂ O ₂	114.04293	114.1038	+ 80
Asp (D)	C ₄ H ₅ NO ₃	115.02694	115.0886	- 200
Cys (C)	C ₃ H ₅ NOS	103.00919	103.1388	- 450
Gln (Q)	C ₅ H ₈ N ₂ O ₂	128.05858	128.1307	+ 160
Glu (E)	C ₅ H ₇ NO ₃	129.04259	129.1155	- 300
Gly (G)	C ₂ H ₃ NO	57.02146	57.0519	0
His (H)	C ₆ H ₇ N ₃ O	137.05891	137.1411	- 120
Ile (I)	C ₆ H ₁₁ NO	113.08406	113.1594	- 2260
Leu (L)	C ₆ H ₁₁ NO	113.08406	113.1594	- 2460
Lys (K)	C ₆ H ₁₂ N ₂ O	128.09496	128.1741	- 350
Met (M)	C ₅ H ₉ NOS	131.04049	131.1926	- 1470
Phe (F)	C ₉ H ₉ NO	147.06841	147.1766	- 2330
Pro (P)	C ₅ H ₇ NO	97.05276	97.1167	- 980
Ser (S)	C ₃ H ₅ NO ₂	87.03203	87.0782	- 390
Thr (T)	C ₄ H ₇ NO ₂	101.04768	101.1051	- 520
Trp (W)	C ₁₁ H ₁₀ N ₂ O	186.07931	186.2132	- 2010
Tyr (Y)	C ₉ H ₉ NO ₂	163.06333	163.1760	- 2240
Val (V)	C ₅ H ₉ NO	99.06841	99.1326	- 1560

Les masses monoisotopiques sont calculées avec les masses des isotopes les plus abondants des éléments:

C = 12.0000000, H = 1.0078250, N = 14.0030740, O = 15.9949146,
F = 18.9984033, P = 30.9737634, S = 31.9720718

Les masses chimiques sont calculées avec les poids des éléments:

C = 12.011, H = 1.00794, N = 14.00674, O = 15.9994, F = 18.99840
P = 30.97376, S = 32.066

**Mass changes due to some post-translational
modifications of peptides and proteins**

Modification	Monoisotopic Mass Change	Average Mass Change
Homoserine formed from Met by CNBr treatment	-29.9928	-30.0935
Pyroglutamic acid formed from Gln	-17.0265	-17.0306
Disulphide bond formation	- 2.0157	- 2.0159
C-terminal amide formed from Gly	- 0.9840	- 0.9847
Deamidation of Asn and Gln	- 0.9840	- 0.9847
Methylation	14.0157	14.0269
Hydroxylation	15.9949	15.9994
Oxidation of Met	15.9949	15.9994
Proteolysis of a single peptide bond	18.0106	18.0153
Formylation	27.9949	28.0104
Acetylation	42.0106	42.0373
Carboxylation of Asp and Glu	43.9898	44.0098
Phosphorylation	79.9663	79.9799
Sulphation	79.9568	80.0642
Cysteinylation	119.0041	119.1442
Pentoses (Ara, Rib, Xyl)	132.0423	132.1161
Deoxyhexoses (Fuc, Rha)	146.0579	146.1430
Hexosamines (GalN, GlcN)	161.0688	161.1577
Hexoses (Fru, Gal, Glc, Man)	162.0528	162.1424
Lipoic acid (amide bond to lysine)	188.0330	188.3147
N-acetylhexosamines (GalNAc, GlcNAc)	203.0794	203.1950
Farnesylation	204.1878	204.3556
Myristoylation	210.1984	210.3598
Biotinylation (amide bond to lysine)	226.0776	226.2994
Pyridoxal phosphate (Schiff Base formed to lysine)	231.0297	231.1449
Palmitoylation	238.2297	238.4136
Stearoylation	266.2610	266.4674
Geranylgeranylation	272.2504	272.4741
N-acetylneuraminic acid (Sialic acid, NeuAc, NANA, SA)	291.0954	291.2579
Glutathionylation	305.0682	305.3117
N-glycolylneuraminic acid (NeuGc)	307.0903	307.2573
5'-Adenosylation	329.0525	329.2091
4'-Phosphopantetheine	339.0780	339.3294
ADP-ribosylation (from NAD)	541.0611	541.3052

The Masses of Some Terminal Groups

N-Terminal Groups	Composition	Monoisotopic Mass	Average Mass
Hydrogen	H	1.00782	1.0079
N-Formyl	HCO	29.00274	29.0183
N-Acetyl	CH ₃ CO	42.01839	43.0452
C-Terminal Groups			
Free acid	OH	17.00274	17.0073
Amide	NH ₂	16.01872	16.0226

