

50376
1983
65

UNIVERSITÉ DES SCIENCES ET TECHNIQUES DE LILLE I

50376
1983
65

Année 1983

N° 330

THÈSE

présentée à l'Université de LILLE I
pour l'obtention du grade de

DOCTEUR INGÉNIEUR

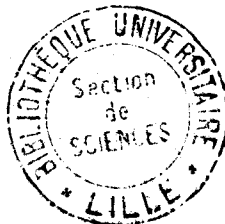
Option Biochimie

par

Denise BÉLAÏCHE-MAYEUR

LA DÉGRADATION D'EDMAN AUTOMATISÉE

PRINCIPE ET FONCTIONNEMENT DU SÉQUENCEUR
MODIFICATIONS TECHNIQUES DU SÉQUENCEUR BECKMAN 890 C



présentée le 10 Novembre 1983 devant la Commission d'examen

Président : Professeur J. MONTREUIL
Rapporteur : Docteur P. SAUTIERE
Examineurs : Professeur G. BISERTE
Professeur C. LOUCHEUX
Docteur J. VAN BEEUMEN
Docteur A. LOIEZ

DOYENS HONORAIRES DE L'ANCIENNE FACULTE DES SCIENCES

MM. R. DEFRETIN, H. LEFEBVRE, M. PARREAU.

PROFESSEURS HONORAIRES DES ANCIENNES FACULTES DE DROIT
ET SCIENCES ECONOMIQUES, DES SCIENCES ET DES LETTRES

MM. ARNOULT, Mme BEAUJEU, BONTE, BROCHARD, CHAPPELON, CHAUDRON, CORDONNIER,
CORSIN, DECUYPER, DEHEUVELS, DEHORS, DION, FAUVEL, FLEURY, P. GERMAIN, GLACET,
GONTIER, HEIM DE BALSAC, HOCQUETTE, KAMPE DE FERRET, KOURGANOFF, LAMOTTE,
LASSERRE, LELONG, Mme LELONG, LHOMME, LIEBAERT, MARTINOT-LAGARDE, MAZET,
MICHEL, PEREZ, ROIG, ROSEAU, ROUELLE, SAVARD, SCHILTZ, WATERLOT, WIEMAN,
ZAMANSKI.

PROFESSEUR EMERITE

M. A. LEBRUN.

ANCIENS PRESIDENTS DE L'UNIVERSITE DES SCIENCES ET
TECHNIQUES DE LILLE

MM. R. DEFRETIN, M. PARREAU, J. LOMBARD, M. MIGEON.

PRESIDENT DE L'UNIVERSITE DES SCIENCES ET TECHNIQUES
DE LILLE

M. J. CORTOIS.

PROFESSEURS - CLASSE EXCEPTIONNELLE

M. CONSTANT Eugène	Electronique
M. FOURET René	Physique du Solide
M. GABILLARD Robert	Electronique
M. MONTREUIL Jean	Biochimie
M. PARREAU Michel	Analyse
M. TRIDOT Gabriel	Chimie appliquée
M. VIVIER Emile	Biologie cellulaire
M. WERTHEIMER Raymond	Physique atomique et moléculaire

PROFESSEURS - 1ère CLASSE

M. BACCHUS Pierre	Astronomie
M. BEAUFILS Jean Pierre	Chimie physique
M. BIAYS Pierre	Géographie
M. BILLARD Jean	Physique du solide
M. BOILLY Bénoni	Biologie

M. BOUGHON Pierre	Algèbre
M. BOURIQUET Robert	Biologie végétale
M. BREZINSKI Claude	Analyse numérique
M. CELET Paul	Géologie générale
M. CHAMLEY Hervé	Géotechnique
M. COEURE Gérard	Analyse
M. CORDONNIER Vincent	Informatique
M. DEBOURSE Jean-Pierre	Gestion des entreprises
M. DYMENT Arthur	Mécanique
M. ESCAIG Bertrand	Physique du solide
M. FAURE Robert	Mécanique
M. FOCT Jacques	Métallurgie
M. GRANELLE Jean-Jacques	Sciences économiques
M. GRUSON Laurent	Algèbre
M. GUILLAUME Jean	Microbiologie
M. HECTOR Joseph	Géométrie
M. LABLACHE COMBIER Alain	Chimie organique
M. LACOSTE Louis	Biologie végétale
M. LAVEINE Jean-Pierre	Paléontologie
M. LEHMANN Daniel	Géométrie
Mme LENOBLE Jacqueline	Physique atomique et moléculaire
M. LHOMME Jean	Chimie organique biologique
M. LOMBARD Jacques	Sociologie
M. LOUCHEUX Claude	Chimie physique
M. LUCQUIN Michel	Chimie physique
M. MAILLET Pierre	Sciences économiques
M. PAQUET Jacques	Géologie générale
M. POUZET Pierre	Analyse numérique
M. PROUVOST Jean	Minéralogie
M. ROUSSEAU Jean-Paul	Physiologie animale
M. SALMER Georges	Electronique
M. SEGUIER Guy	Electrotechnique
M. STANKIEWICZ François	Sciences économiques
M. TILLIEU Jacques	Physique théorique
M. VIDAL Pierre	Automatique
M. ZEYTOUNIAN Radyadour	Mécanique

PROFESSEURS - 2ème classe

M. AL FAKIR Sabah	Algèbre
M. ALLAMANDO Etienne	Electronique et électrotechnique
M. ANCIAN Bernard	Spectrochimie
M. ANTOINE Philippe	Analyse
M. BART André	Biologie animale
Mme BATTIAU Yvonne	Géographie
M. BEGUIN Paul	Mécanique
M. BELLET Jean	Physique atomique et moléculaire
M. BERZIN Robert	Analyse
M. BKOUCHE Rudolphe	Algèbre
M. BODARD Marcel	Biologie végétale
M. BOIVIN Jean-Claude	Chimie minérale
M. BONNELLE Jean-Pierre	Catalyse
M. BOSCO Denis	Probabilités
M. BOUQUELET Stéphane	Biochimie structurale
M. BRASSELET Jean-Paul	Géométrie et topologie

M. BRIDOUX Michel	Chimie physique
M. BRUYELLE Pierre	Géographie
M. CAPURON Alfred	Biologie animale
M. CARREZ Christian	Informatique
M. CHAPOTON Alain	Electronique
M. COQUERY Jean-Marie	Psychophysiologie
Mme CORSIN Paule	Paléontologie
M. CORTOIS Jean	Physique nucléaire et corpusculaire
M. COUTURIER Daniel	Chimie organique
M. CRAMPON Norbert	Hydrogéologie et environnement
M. CROSNIER Yves	Electronique
Mlle DACHARRY Monique	Géographie
M. DAUCHET Max	Informatique
M. DEBRABANT Pierre	Géologie appliquée
M. DEGAUQUE Pierre	Electronique
M. DELORME Pierre	Physiologie animale
M. DE MASSON D'AUTUME Antoine	Sciences économiques
M. DEMUNTER Paul	Sociologie
M. DENEL JACQUES	Informatique
M. DE PARIS Jean-Claude	Analyse
M. DEPREZ Gilbert	Physique du solide et cristallographie
M. DERIEUX Jean-Claude	Microbiologie
Mlle DESSAUX Odile	Spectroscopie de la réactivité chimique
M. DEVRAINNE Pierre	Chimie minérale
M. DHAINAUT André	Biologie animale
Mme DHAINAUT Nicole	Biologie animale
M. DORMARD Serge	Sciences économiques
M. DOUKHAN Jean-Claude	Physique du solide
M. DUBOIS Henri	Spectroscopie hertzienne
M. DUBRULLE Alain	Spectroscopie hertzienne
M. DUBUS Jean-Paul	Spectrométrie des solides
M. DUPONT Christophe	Vie de la firme (I.P.A.)
Mme EVRARD Micheline	Chimie appliquée
M. FONTAINE Hubert	Dynamique des cristaux
M. FOUQUART Yves	Optique atmosphérique
M. FOURNET Bernard	Biochimie structurale
M. FRONTIER Serge	Ecologie numérique
M. GAMBLIN André	Géographie urbaine, industrielle et démographie
M. GLORIEUX Pierre	Physique moléculaire et rayonnements atmosphériques
M. GOBLOT Rémi	Algèbre
M. GOSSELIN Gabriel	Sociologie
M. GOUDMAND Pierre	Chimie Physique
M. GREMY Jean-Paul	Sociologie
M. GREVET Patrick	Sciences économiques
M. GUILBAULT Pierre	Physiologie animale
M. HENRY Jean-Pierre	Génie mécanique
M. HERMAN Maurice	Physique spatiale
M. HOUDART René	Physique atomique et moléculaire
M. JACOB Gérard	Informatique
M. JACOB Pierre	Probabilités et statistiques
M. JACQUILLAT Bertrand	Gestion
M. JEAN Raymond	Biologie des populations végétales
M. JOFFRE Patrick	Vie de la firme (I.P.A.)
M. JOURNAL Gérard	Spectroscopie hertzienne
M. KREMBEL Jean	Biochimie
M. LANGRAND Claude	Probabilités et statistiques

Mme LECLERCQ Ginette	Catalyse
M. LEFEVRE Christian	Pétrologie
Mlle LEGRAND Denise	Algèbre
Mlle LEGRAND Solange	Algèbre
Mme LEHMANN Josiane	Analyse
M. LEMAIRE Jean	Spectroscopie hertzienne
M. LE MAROIS Henri	Vie de la firme (I.P.A.)
M. LEROY Jean Marie	Chimie appliquée
M. LEROY Yves	Electronique, électrotechnique, automatique
M. LESENNE Jacques	Electrotechnique
M. LHENAFF René	Géographie
M. LOCQUENEUX Robert	Physique théorique
M. LOSFELD Joseph	Informatique
M. LOUAGE Francis	Electronique
M. MACKE Bruno	Physique moléculaire et rayonnements atmosphériques
M. MAHIEU Jean-Marie	Physique atomique et moléculaire
M. MAIZIERES Christian	Automatique
M. MESMACQUE Gérard	Génie mécanique
M. MESSELYN Jean	Physique atomique et moléculaire
M. MESSERLIN Patrick	Sciences économiques
M. MIGNOT Fulbert	Analyse numérique
M. MONTEL Marc	Physique du solide
M. MONTUELLE Bernard	Biologie et biochimie appliquées
Mme MOUNIER Yvonne	Physiologie des structures contractiles
Mme N'GUYEN VAN CHI Régine	Géographie
M. NICOLE Jacques	Chimie analytique
M. NOTELET Francis	Electronique, électrotechnique, automatique
M. PARSY Fernand	Mécanique
M. PASZKOWSKI Stéphan	Analyse numérique
Mlle PAUPARDIN Colette	Biologie physiologie végétales
M. PECQUE Marcel	Chimie organique
M. PERROT Pierre	Chimie appliquée
M. PERTUZON Emile	Physiologie animale
M. PETIT Francis	Chimie organique, minérale et analytique
M. PONSOLLE Louis	Chimie physique
M. PORCHET Maurice	Biologie animale
M. POVY Lucien	Automatique
M. RACZY Ladislas	Electronique
M. RAOULT Jean-François	Géologie structurale
M. RICHARD Alain	Biologie animale
M. RIETSCH François	Physique des polymères
M. ROGALSKI Marc	Analyse
M. ROY Jean-Claude	Psychophysiologie
M. SCHAMPS Joël	Spectroscopie moléculaire
Mme SCHWARZBACH Yvette	Géométrie
M. SIMON Michel	Sociologie
M. SLIWA Henri	Chimie organique
M. SOMME Jean	Géographie
Mlle SPIK Geneviève	Biochimie
M. STERBOUL François	Informatique
M. TAILLIEZ Roger	Génie alimentaire
M. THERY Pierre	Electronique, électrotechnique, automatique
M. TOULOTTE Jean-Marc	Automatique
M. TURREL Georges	Spectrochimie infrarouge et Raman
M. VANDORPE Bernard	Chimie minérale
M. VAST Pierre	Chimie inorganique
M. VERBERT André	Biochimie
M. VERNET Philippe	Génétique
M. VILETTE Michel	Résistance des matériaux

M. WALLART Francis	Spectrochimie Infrarouge et Raman
M. WARTEL Michel	Chimie inorganique
M. WATERLOT Michel	Géologie générale
M. WERNER Georges	Informatique fondamentale appliquée
M. WOSNIAK Michel	Hydrométallurgie
Mme ZINN Justin Nicole	Algèbre

TABLE DES MATIERES

	Pages
<u>INTRODUCTION</u>	1
<u>CHAPITRE I</u>	
<u>LA DEGRADATION D'EDMAN</u>	7
<u>I - 1 LES DIFFERENTES METHODES MANUELLES</u>	12
I - 1-1 Dégradation d'Edman couplée à la dansylation (27)	14
I - 1-2 La méthode d'Edman soustractive	14
I - 1-3 Méthode d'Edman modifiée utilisant le 4-NN-diméthylaminoazobenzène 4'-isothiocyanate (DABITC) (12).	17
<u>I - 2 LES METHODES AUTOMATISEES DE LA DEGRADATION D'EDMAN</u>	20
I - 2-1 Le séquenceur en phase liquide	20
I - 2-1-1 Description de l'appareil	22
I - 2-1-1-1 La coupelle de réaction	22
I - 2-1-1-2 Le système de vide	24
I - 2-1-1-3 Les réservoirs de réactifs et de solvants	25
I - 2-1-1-4 Les vannes de délivrance	26
I - 2-1-1-5 La vanne "élimination/collection"	26
I - 2-1-1-6 Le collecteur de fractions	28
I - 2-1-1-7 Le pupitre de commande	28

I - 2-1-2	Fonctionnement de l'appareil	31
I - 2-1-3	Les modifications	36
I - 2-1-3-1	Les modifications techniques	36
I - 2-1-3-2	Les modifications chimiques	39
I - 2-2	Le séquenceur en phase solide	49
I - 2-3	Le séquenceur en phase solide, liquide, gazeuse	50

CHAPITRE II

LES AMELIORATIONS APORTEES AU SEQUENCEUR BECKMAN 890 C ET 53

A LA METHODOLOGIE DE LA DEGRADATION D'EDMAN AUTOMATISEE

II - I	<u>MODIFICATIONS TECHNIQUES DU SEQUENCEUR BECKMAN 890C</u>	55
II - 1-1	Les vannes de délivrance	55
II - 1-2	Le système de vide	57
II - 1-3	Le piège à froid	57
II - 1-4	La coupelle de réaction	62
II - 1-5	La conversion automatique	65
II - 1-6	Autres modifications	71
II - 1-7	Détermination de la séquence N-terminale de la myoglobine de Cheval après modifi- cation du séquenceur Beckman 890C	72
II - 2	<u>TRAITEMENT DES REACTIFS ET DES SOLVANTS</u>	78
II - 3	<u>MICROSEQUENCE EN PHASE SOLIDE, LIQUIDE, GAZEUSE</u>	79

CHAPITRE III

LES METHODES D'IDENTIFICATION DES PTH-AA 85

III - 1	<u>LA CHROMATOGRAPHIE SUR COUCHE MINCE</u>	87
III - 2	<u>LA CHROMATOGRAPHIE EN PHASE GAZEUSE</u>	89

III - 3	<u>LA REGENERATION DES ACIDES AMINES A PARTIR DES PTH-AA</u>	91
III - 4	<u>LA SPECTROMETRIE DE MASSE</u>	92
III - 5	<u>LA CHROMATOGRAPHIE LIQUIDE A HAUTE PERFORMANCE (HPLC)</u>	93
III - 5-1	La chromatographie solide-liquide	93
III - 5-2	La chromatographie liquide-liquide	94
 <u>CHAPITRE IV</u>		
	<u>CONTRIBUTION A LA DETERMINATION DE LA STRUCTURE PRIMAIRE DES DIFFERENTES PROTEINES</u>	110
	<u>PUBLICATIONS 1 et 2</u>	112
IV - 1	<u>DETERMINATION DE LA STRUCTURE PRIMAIRE DE LA PROTEINE HRm DE <i>Rhizobium meliloti</i></u>	113
	<u>PUBLICATIONS 3, 4 et 5</u>	122
IV - 2	<u>DETERMINATION DE LA STRUCTURE PRIMAIRE D'HIS- NES H2A ISOLEES D'INVERTEBRES MARINS</u>	123
IV - 3	<u>DETERMINATION DE LA STRUCTURE PRIMAIRE DE PROTAMINES DE MAMMIFERES ET DE POISSONS</u>	134
IV - 3-1	Protamines de Mammifères	134
IV - 3-1-1	Structure primaire de la protamine de sperme de Bélier (<i>Ovis aries</i>)	134
IV - 3-1-2	Séquence N-terminale de la protamine de sperme d'étalon (<i>Equus equus</i>)	139
IV - 3-2	Protamines de Poissons. Structure pri- maire de la scylliorhinine 4 de sperme de Roussette (<i>Scylliorhinus caniculus</i>)	145
	<u>CONCLUSION</u>	152
	<u>BIBLIOGRAPHIE</u>	154

Ce travail a été réalisé, sous la direction scientifique du Docteur P.SAUTIERE, Maître de Recherche au CNRS, dans le service de Séquence automatisée, commun aux Unités INSERM 124 et 16 (co-Directeurs : Docteur P.SAUTIERE, Professeur P.DEGAND) et implanté dans le laboratoire des Protéines Nucléaires de l'Unité de Recherches Ultrastructurales et Biochimiques sur les Cellules Normales et Cancéreuses (Unité 124 de l'INSERM) et de l'Institut de Recherches sur le Cancer de Lille (Directeur : Professeur G.BISERTE).

Nous remercions vivement le Docteur G.BRIAND (responsable scientifique du service de Séquence automatisée) et O.HANNEDOUCHE, pour l'aide qu'ils ont apportée à la réalisation de ce mémoire.

ABREVIATIONS UTILISEES

ATZ-AA	: 2-anilino-5-thiazolinone d'acide aminé
DABITC	: 4-NN-diméthylaminoazobenzène 4'-isothiocyanate
DABTH-AA	: 4-NN-diméthylaminoazobenzène 4'-thiohydantoïne d'acide aminé
DABTZ-AA	: 4-NN-diméthylaminoazobenzène 4'-thiazolinone d'acide aminé
DMAA	: diméthylallylamine
DMBA	: diméthylbenzylamine
DTT	: dithiothréitol
G.C	: chromatographie en phase gazeuse
HFBA	: acide heptafluorobutyrique
HPLC	: chromatographie liquide à haute performance
n.d.	: non dosé
PITC	: phénylisothiocyanate
PTH-AA	: phénylthiohydantoïne d'acide aminé
TFA	: acide trifluoroacétique

Ce travail a fait l'objet des publications suivantes
présentées dans le chapitre IV.

- Publication n°1

LAINE B., BELAÏCHE D., SAUTIERE P. and BISERTE G. (1982)
Biochem.Biophys.Res.Comm., 106, 101-107.

- Publication n°2

LAINE B., BELAÏCHE D., KHANAKA H. and SAUTIERE P. (1983)
Eur.J.Biochem., 131, 325-331.

- Publication n°3

WOUTERS-TYROU D., MARTIN-PONTHIEU A., BRIAND G.,
SAUTIERE P. and BISERTE G. (1982)
Eur.J.Biochem., 124, 489-498.

- Publication n°4

MARTINAGE A., BELAÏCHE D., DUPRESSOIR T. and SAUTIERE P.
(1983)
Eur.J.Biochem., 130, 465-472.

- Publication n°5

KMIECIK D., COUPPEZ M., BELAÏCHE D. and SAUTIERE P. (1983)
Eur.J.Biochem., 135, 113-121.

INTRODUCTION

La détermination de la structure primaire, ou enchaînement des acides aminés d'une protéine est une étape nécessaire à la compréhension de son rôle biologique.

Les phénomènes du vivant se traduisent toujours en termes d'interactions moléculaires. Dès lors la biologie se prolonge par la biochimie et celle-ci explique au niveau moléculaire les phénomènes du monde vivant : elle les explique par l'étude chimique des constituants cellulaires, scrute leurs formes, analyse les interactions de leurs architectures. Les protéines, macromolécules biologiques n'échappent pas à cette investigation.

Dans la vie cellulaire, les protéines assurent les fonctions les plus diverses et les plus essentielles:

- Elles forment les structures des organites cellulaires (noyau, membranes, ribosomes, squelette cellulaire...)
- Elles réalisent les fonctions métaboliques (anabolisme, catabolisme, respiration cellulaire) et effectuent tous les transports membranaires (dans la cellule et de cellule à cellule).
- Elles contribuent à l'expression de l'information génétique en assurant le renouvellement des acides nucléiques, leurs lectures et toutes les étapes de la synthèse protéique.

Un rapide inventaire des activités des protéines fait pressentir qu'une grande part des questions qu'elles suscitent, recevra un début de réponse par l'étude de leur structure chimique. La connaissance de cette structure chimique apporte un complément d'information nécessaire à la pleine exploitation des données fournies par l'utilisation des calculs de prédiction de structure secondaire, et l'analyse des cristaux de protéines (lorsque l'on peut les obtenir) par la diffraction des rayons X.

Dans la période qui a suivi la seconde guerre mondiale, les techniques de détermination de la séquence des protéines se sont développées grâce à la mise au point du dosage des acides aminés de MOORE et STEIN (32), à l'utilisation des nouveaux supports de chromatographie d'échanges ioniques et de phase liquide sur papier (MOORE et coll.) (46).

SANGER (53) par la technique de la dinitrophénylation du résidu α -NH₂ d'une chaîne protéique ou d'un peptide réalise l'identification de l'acide aminé NH₂-terminal. Cette technique permet aussi de déterminer le nombre de chaînes polypeptidiques d'une molécule. Bien que longue et fastidieuse cette technique lui permettra de mener à bien la détermination de la séquence de l'insuline, travail récompensé par son premier prix Nobel en 1950. Doué d'une prodigieuse richesse inventive en même temps que d'un sens pratique hors du commun pour la mise en oeuvre de ses idées, SANGER recevra un second prix Nobel pour la détermination de la séquence des acides nucléiques en 1977.

Le principe de la dégradation récurrente des acides aminés d'une protéine à partir de l'extrémité NH_2 -terminale a été inventé par le suédois Pehr Victor EDMAN (1916-1977) en 1950 avec l'emploi du phénylisothiocyanate et la formation d'une anilino-thiazolinone "composé chimique clef" de la dégradation en trois étapes : couplage, cyclisation et conversion. Cette découverte géniale se prêtait à l'automatisation qui perfectionnait le rendement de chaque étape de la dégradation par la précision des opérations standardisées par la machine. Cela fut réalisé en 1967 avec la naissance du premier séquenceur Beckman en phase liquide conceptualisé par EDMAN et BEGG.

Depuis lors, le principe de cette dégradation n'a pas été fondamentalement changé. Cependant, des modifications visant à la fois à augmenter le nombre de résidus identifiables et à diminuer les quantités de protéine (ou de peptide) et de réactifs nécessaires pour la séquence ont été progressivement introduites. C'est ainsi que l'ensemble des opérations de la dégradation automatisée bénéficie de l'utilisation de réactifs d'un haut degré de pureté, permettant d'éviter les oxydations ou réactions chimiques annexes préjudiciables au rendement d'obtention de l'anilino-thiazolinone. L'utilisation de "mainteneurs" a permis une meilleure adhésion du film protéique sur la paroi de la coupelle.

Par ailleurs, la coupelle qui constitue l'élément essentiel du séquenceur, a été redessinée et son architecture adaptée à la mécanique des fluides réactifs.

La conversion automatique de l'anilino-thiazolinone en phénylthiohydantoïne contribue à limiter la dégradation de ce dérivé instable en réalisant sa transformation dès son extraction.

Une amélioration essentielle a également été apportée avec la quantification des PTH-AA qui permet le calcul du rendement de chaque étape et même la dégradation de deux peptides simultanément pourvu que leurs concentrations soient suffisamment différentes. C'est là que l'apport de la chromatographie en phase gazeuse et surtout de la chromatographie liquide haute performance s'est révélé décisif en permettant une identification certaine de tous les PTH-AA en même temps que leur dosage. Enfin, l'avènement des microprocesseurs a permis de réaliser une coordination et une reproductibilité parfaite des différentes étapes de l'automatisation.

Un autre aspect des possibilités d'amélioration de la séquence des protéines est illustré par l'emploi du marquage avec des isotopes radioactifs de la protéine lors de sa biosynthèse. Les possibilités de l'identification sont rendues plus sélectives et plus sensibles par le dosage des PTH-AA radioactives formées.

Dans un premier chapitre, notre travail décrira les bases de la réaction chimique d'EDMAN et les principes des séquenceurs utilisés à ce jour. Nous aborderons cette question en insistant sur les particularités de la chimie de la réaction

d'EDMAN, qui doivent retenir toute notre attention pour obtenir un rendement d'étape optimum de 96 à 99 p.100, facteur clef pour réaliser un nombre de cycles maximum.

Dans le chapitre II nous décrirons les modifications que nous avons apportées au séquenceur dans notre laboratoire, en particulier le système de la conversion automatique des thiazolinones d'acides aminés. Une étude critique de cette modification majeure nous permettra de faire ressortir les avantages et les inconvénients de cette opération.

Puis nous examinerons dans le chapitre III les méthodes d'identification des PTH-AA en montrant l'apport des techniques récentes qui permettent aussi le dosage des PTH-AA.

Le chapitre IV traitera des travaux de séquence que nous avons effectués sur des substrats protéiques variés et nous permettra de faire ressortir les adaptations nécessaires imposées par la nature chimique des protéines étudiées afin d'obtenir une efficacité optimum de l'appareil concrétisée par un nombre de cycles de dégradation le plus élevé possible.

CHAPITRE I

LA DEGRADATION D'EDMAN

PLAN DU CHAPITRE I - LA DEGRADATION D'EDMAN

I - 1 LES DIFFERENTES METHODES MANUELLES

- I - 1-1 Dégradation d'Edman couplée à la dansylation
- I - 1-2 La méthode d'Edman soustractive
- I - 1-3 Méthode d'Edman modifiée utilisant le 4-NN-diméthylaminoazobenzène 4'-isothiocyanate (DABITC)

I - 2 LES METHODES AUTOMATISEES DE LA DEGRADATION D'EDMAN

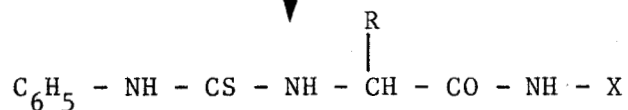
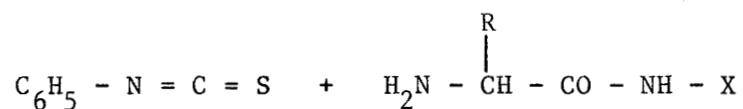
- I - 2-1 Le séquenceur en phase liquide
- I - 2-2 Le séquenceur en phase solide
- I - 2-3 Le séquenceur en phase "gazeuse"

récurrence d'EDMAN (18,19) décrite en 1950 et qui a subi depuis un certain nombre de modifications ou d'adaptations - mais le principe et les réactions chimiques restent fondamentalement les mêmes.

Cette méthode est dite récurrente car elle permet d'enlever les acides aminés un par un à partir de l'extrémité N-terminale sans détruire les autres liaisons peptidiques. Le schéma des différentes réactions chimiques qui interviennent au cours d'un cycle de dégradation est présenté dans la figure 1.

Cette méthode consiste à faire réagir, à 57°C, le phénylthiocyanate avec la fonction α -aminée libre de la chaîne peptidique (réaction de couplage). Le composé obtenu est le phénylthiocarbamylpeptide qui en milieu acide fort anhydre donne la chaîne peptidique résiduelle amputée d'un acide aminé et un composé instable : la 2-anilino-5-thiazolinone d'acide aminé (réaction de clivage). Cette 2-anilino-5-thiazolinone d'acide aminé est ensuite convertie à chaud en milieu acide dilué, en un dérivé plus stable : la phénylthiohydantoïne d'acide aminé (réaction de conversion). A la fin d'un cycle de dégradation, le peptide résiduel peut réagir avec le phénylthiocyanate tandis que la phénylthiohydantoïne de l'acide aminé est identifiée par différentes méthodes essentiellement chromatographiques qui seront exposées dans le troisième chapitre de cette thèse.

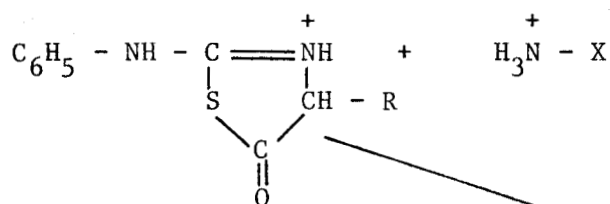
La méthode de dégradation d'EDMAN peut être employée soit manuellement, soit automatiquement.

I Couplage

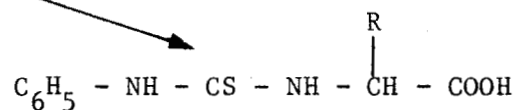
Phénylthiocarbamyl-(PTC) peptide

II Clivage

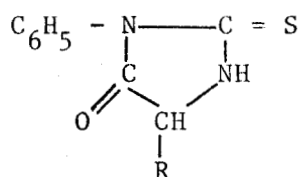
milieu anhydre



2-anilino-5-thiazolinone

III Conversion

milieu aqueux



3-phényl-2-thiohydantoïne

FIGURE 1 SCHEMA DES REACTIONS INTERVENANT AU COURS DES DIFFERENTES ETAPES DE LA DEGRADATION D'EDMAN.



I - 1 LES DIFFERENTES METHODES MANUELLES

Le facteur limitant de la méthode de dégradation d'EDMAN est l'oxygène. C'est un facteur particulièrement important dans le déroulement des opérations de dégradation manuelle. En effet, la présence d'oxygène provoque une désulfuration oxydative du groupe phénylthiocarbamyl-peptide et il en résulte à plus ou moins brève échéance un blocage de la réaction. Cette méthode manuelle est mise en oeuvre dans un tube conique sous atmosphère d'azote mais les risques de contact avec l'oxygène de l'air ne peuvent être totalement écartés au cours des manipulations. La réaction de couplage est réalisée pendant 40 minutes à 55°C en tampon diméthylallylamine (DMAA) 0.8M ajusté à pH 9.5 avec l'acide acétique. En fin de couplage, l'excès de phénylthiocyanate (PITC) et les dérivés secondaires formés (aniline, phénylthiourée, diphenylthiourée) sont extraits par le benzène. La réaction de clivage est ensuite effectuée à 55°C pendant 7 minutes par l'acide trifluoroacétique anhydre. La 2-anilino-5-thiazolinone d'acide aminé (ATZ-AA) formée est alors extraite puis convertie à 80°C pendant 10 minutes, en milieu acide chlorhydrique 1N, en un dérivé stable : la phénylthiohydantoïne d'acide aminé (PTH-AA). Différentes variantes de cette méthode de dégradation d'EDMAN présentées dans la figure 2 sont également utilisées.

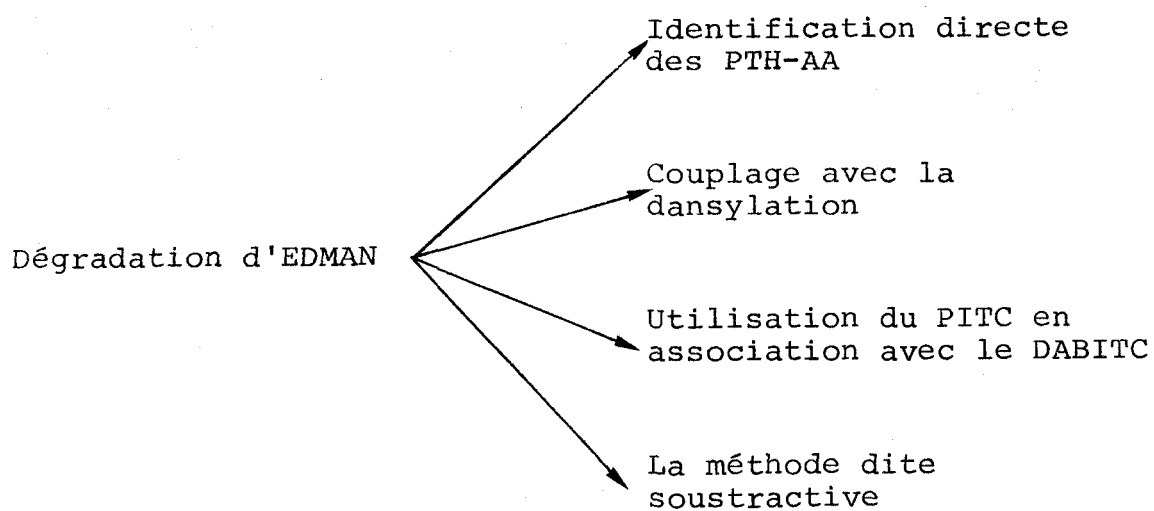


FIGURE 2 LES VARIANTES DE LA DEGRADATION D'EDMAN

I - 1-1 Dégradation d'EDMAN couplée à la dansylation (27)

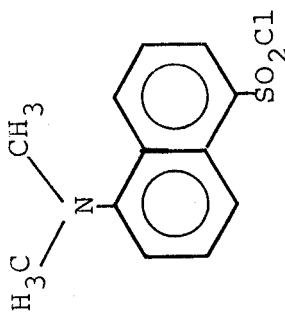
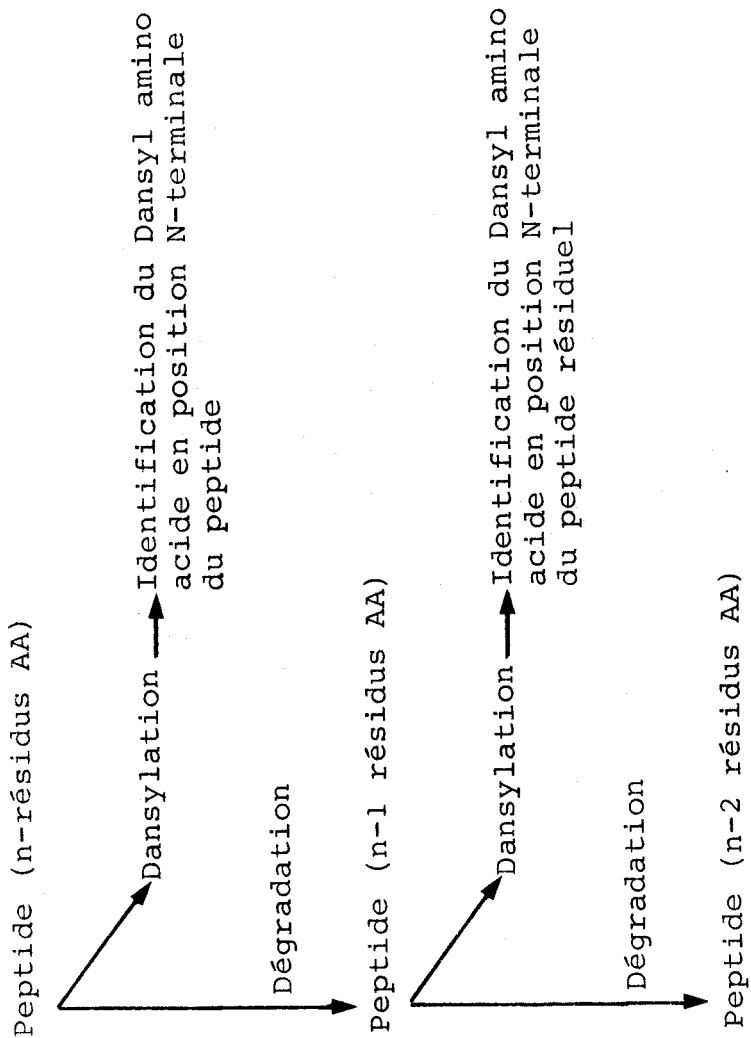
Cette méthode dont les différentes étapes sont schématisées dans la figure 3 fait réagir le chlorure de l'acide diméthyl amino-1-naphtalène-5-sulfonique (chlorure de dansyle) avec la fonction α -aminée libre de l'acide aminé situé en position N-terminale. Le dansyl peptide obtenu libère par hydrolyse acide un dansyl amino-acide qui est fluorescent en lumière ultraviolette.

Simultanément, la protéine est soumise à la dégradation d'EDMAN. A la fin de chaque cycle, une fraction aliquote de la protéine ou du peptide résiduel est prélevée pour la réaction avec le chlorure de dansyle.

L'intérêt de cette technique réside dans le fait qu'une très faible quantité de protéine ou de peptide (20 à 50 nanomoles) suffit pour déterminer la séquence.

I - 1-2 La méthode d'EDMAN soustractive

Dans cette méthode présentée dans la figure 4, le peptide est soumis à la dégradation d'EDMAN. A la fin de chaque cycle, une fraction aliquote du peptide résiduel est prélevée et est soumise à une hydrolyse totale acide.



Chlorure de dansyle

FIGURE 3 LA DEGRADATION D'EDMAN COUPLEE A LA DANSYLATION



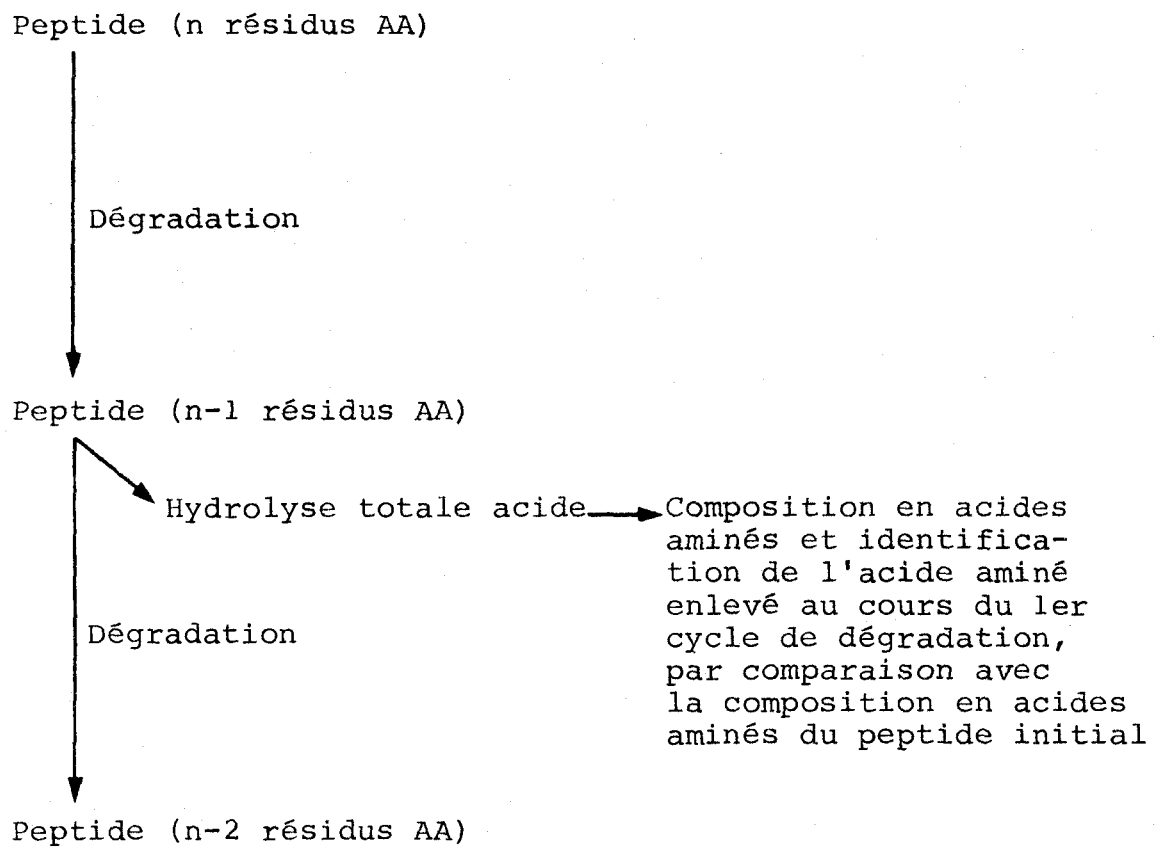


FIGURE 4 LA METHODE DE DEGRADATION SOUSTRACTIVE

On détermine alors la composition en acides aminés du peptide résiduel et par comparaison avec la composition en acides aminés du peptide initial, on en déduit l'acide aminé qui a été enlevé au cours de la dégradation.

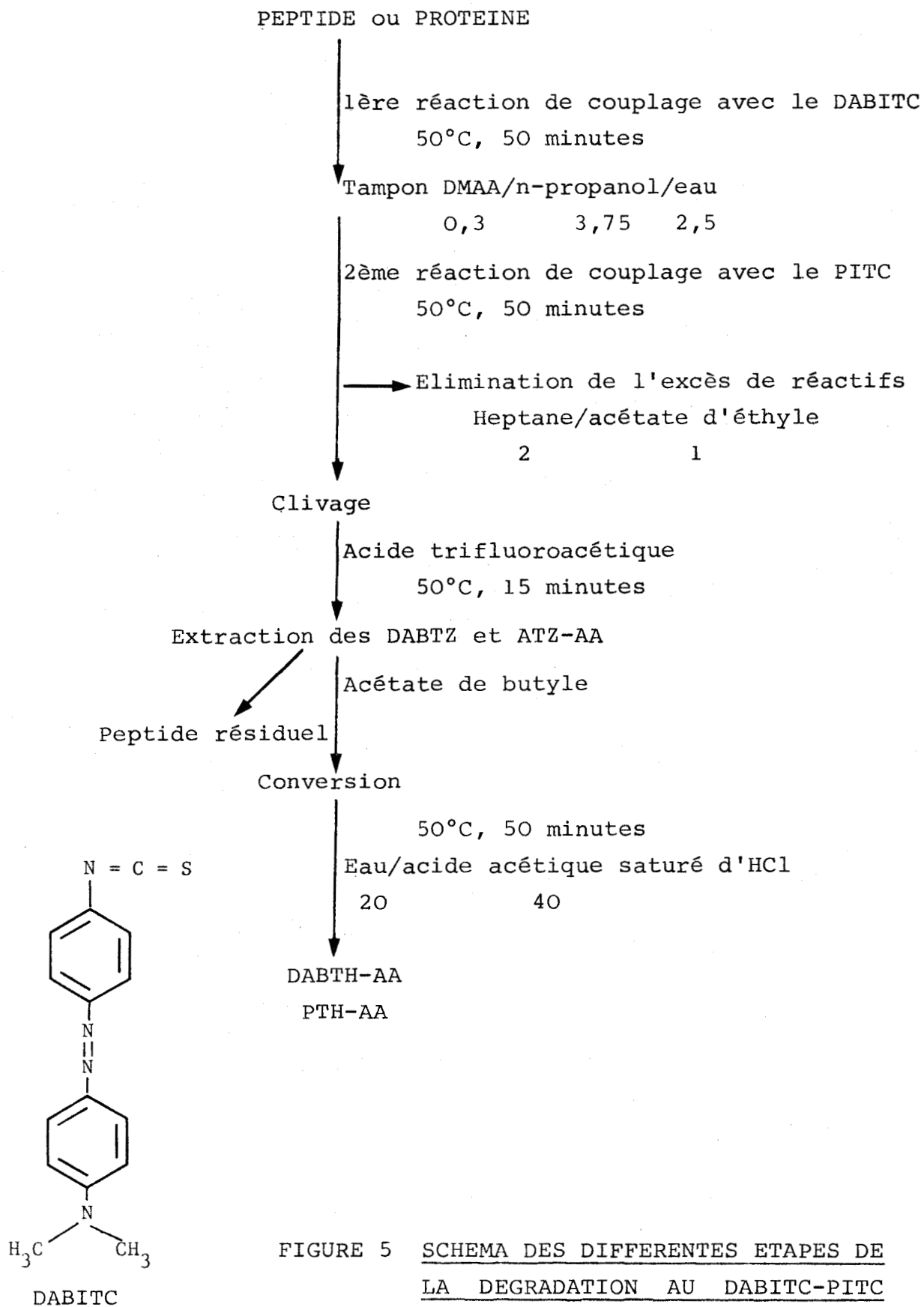
Cette méthode est comme la précédente assez sensible, mais son inconvénient majeur est sa durée. En effet, l'hydrolyse acide totale demande 24 heures et la détermination de la composition en acides aminés 3 heures.

I - 1-3 Méthode d'EDMAN modifiée utilisant le 4-NN-diméthyl-
aminoazobenzène 4'-isothiocyanate (DABITC) (12)

Les différentes étapes de la dégradation au DABITC sont présentées dans la figure 5.

Afin d'éviter l'emploi d'une température très élevée (75°C) nécessaire à un couplage de 100 p.100 du résidu N-terminal du peptide, mais néfaste à la stabilité du DABITC, cette méthode comporte deux couplages à 50°C. L'un est réalisé avec le DABITC tandis que l'autre se fait avec le PITC.

L'identification des DABTH-AA formés est effectuée soit en chromatographie bidimensionnelle sur couche mince sur plaque de polyamide de 2,5x2,5cm (12), soit en chromatographie liquide à haute performance sur colonne RP8 (Kanuer) à 40°C équilibrée en tampon acétate de sodium 8 mM, pH 5.0 avec un gradient linéaire de 60 à 85 p.100 de méthanol à 90 p.100 (13).



L'intérêt de cette méthode de dégradation est qu'elle augmente de 10 à 20 fois la sensibilité des méthodes manuelles. Elle présente toutefois un inconvénient non négligeable : les témoins DABTH-AA ne sont pas commercialisés, il faut de ce fait les synthétiser et l'on a alors des problèmes de stabilité des DABTH-AA en milieu acide.

Toutes ces méthodes manuelles de dégradation d'EDMAN présentent incontestablement des avantages qui sont :

- la faible quantité de protéine ou de peptide nécessaire.
- la possibilité de dégrader plusieurs peptides en même temps.
- enfin, c'est une technique irremplaçable pour déterminer la séquence en acides aminés des peptides courts.

Elles présentent aussi quelques inconvénients, qui cependant ne font pas oublier les avantages, et que l'on peut éliminer en partie :

- la difficulté de préserver le milieu réactionnel de tout contact avec l'oxygène de l'air.
- la longueur de la méthodologie. On peut en effet effectuer 2 à 3 cycles de dégradation au maximum par jour. On risque ainsi de voir se transformer la glutamine en acide pyroglutamique, ce qui a pour effet de bloquer la dégradation d'EDMAN.

Toutefois, le manque de rapidité, par rapport aux méthodes automatisées, est compensé par la possibilité de dégrader plusieurs peptides en même temps.

I - 2 LES METHODES AUTOMATISEES DE LA DEGRADATION D'EDMAN

La dégradation automatique d'EDMAN est effectuée soit en phase liquide, soit solide, soit gazeuse.

Les avantages de l'automatisation de la technique résident dans le contrôle rigoureux des conditions de la dégradation (la dégradation est effectuée constamment sous atmosphère inerte d'azote, les réactifs et les solvants utilisés sont eux aussi mis sous azote avant leur emploi). De plus, la technique est rapide (un cycle de dégradation dure en moyenne 75 minutes), et l'appareil travaille 24 heures sur 24.

Nous nous attarderons tout d'abord sur la méthode de dégradation réalisée en phase liquide.


I - 2-1 Le séquenceur en phase liquide

Cet appareil qui est schématisé dans la figure 6 est composé de compartiments pilotés par un pupitre de commande :

- la coupelle de réaction
- le système de vide
- le compartiment réactifs et solvants
- la distribution d'azote
- le collecteur de fractions.

L'appareil sur lequel nous avons effectué nos travaux est le séquenceur Beckman 890 C qui est présenté dans la figure 7, et qui est le seul séquenceur en phase liquide existant sur le marché.

FIGURE 6 SCHEMA SIMPLIFIE DU SEQUENCEUR BECKMAN 890 C

 = électro-vanne à 2 voies

 = électro-vanne à 3 voies

 = manomètre

T = régulateur

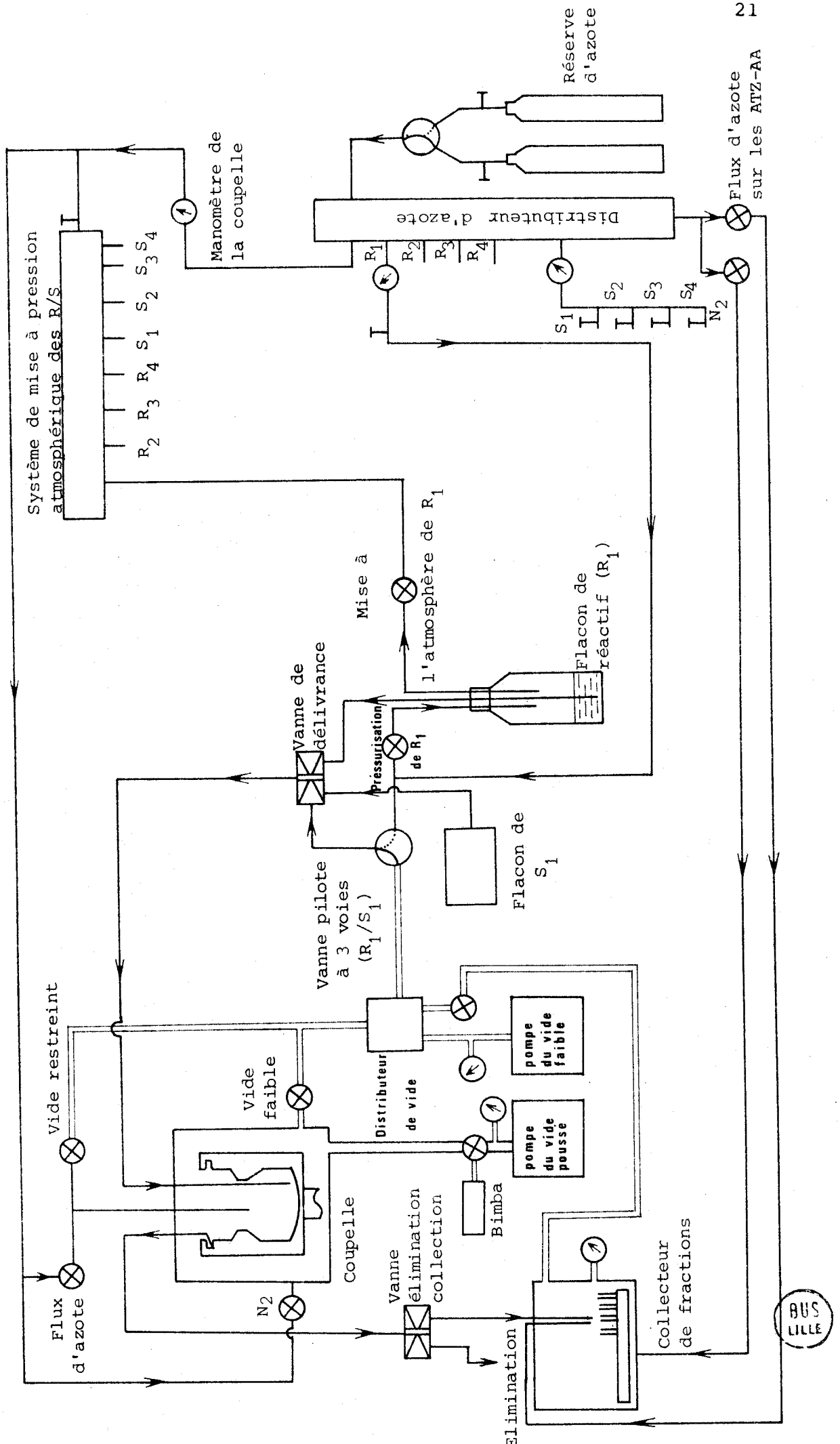
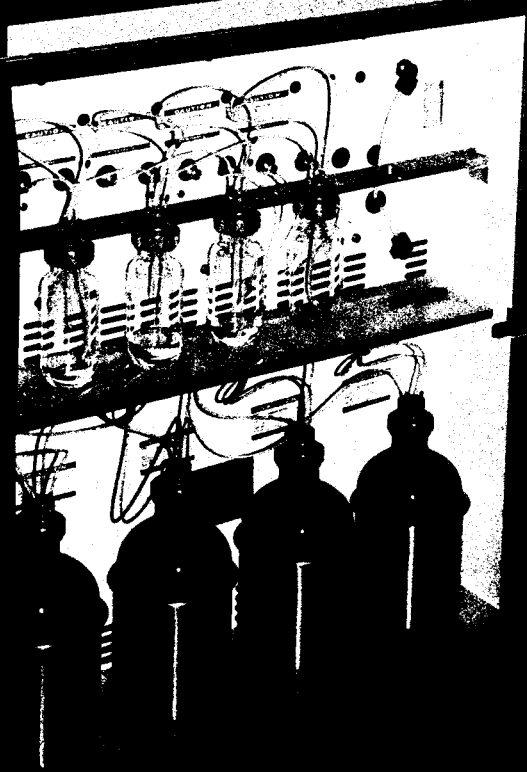


FIGURE 7 LE SEQUENCEUR BECKMAN 890 C





I - 2-1-1. Description de l'appareil

I - 2-1-1-1 La coupelle de réaction

La coupelle de réaction dans laquelle la protéine ou le peptide à dégrader est déposé sous forme d'un film mince et homogène constitue véritablement le "coeur" du séquenceur où vont se dérouler les différentes étapes d'un cycle de dégradation :

- Dépôt dans la coupelle de protéine ou de peptide à dégrader
- Couplage avec le phénylisothiocyanate (PITC)
- Elimination de l'excès de réactifs et des sous-produits de la dégradation
- Clivage avec l'acide heptafluorobutyrique (HFBA)
- Extraction de la 2-anilino-5-thiazolinone d'acide aminé (ATZ-AA)

La coupelle de réaction dont le schéma est présenté dans la figure 8 est un cylindre de verre épais, dont les parois internes sont polies optiquement, le plus petit défaut de surface pouvant provoquer une irrégularité du film protéique. A l'intérieur du cylindre de verre, on remarque un léger décrochement qui est appelé "undercut". Ce surplomb a pour but d'éviter, de manière mécanique, l'élimination de la protéine ou du peptide lors des lavages par le benzène.

Elle est animée d'un mouvement de rotation, et peut tourner soit à 1000 tours/minute, soit à 1500 tours/minute.

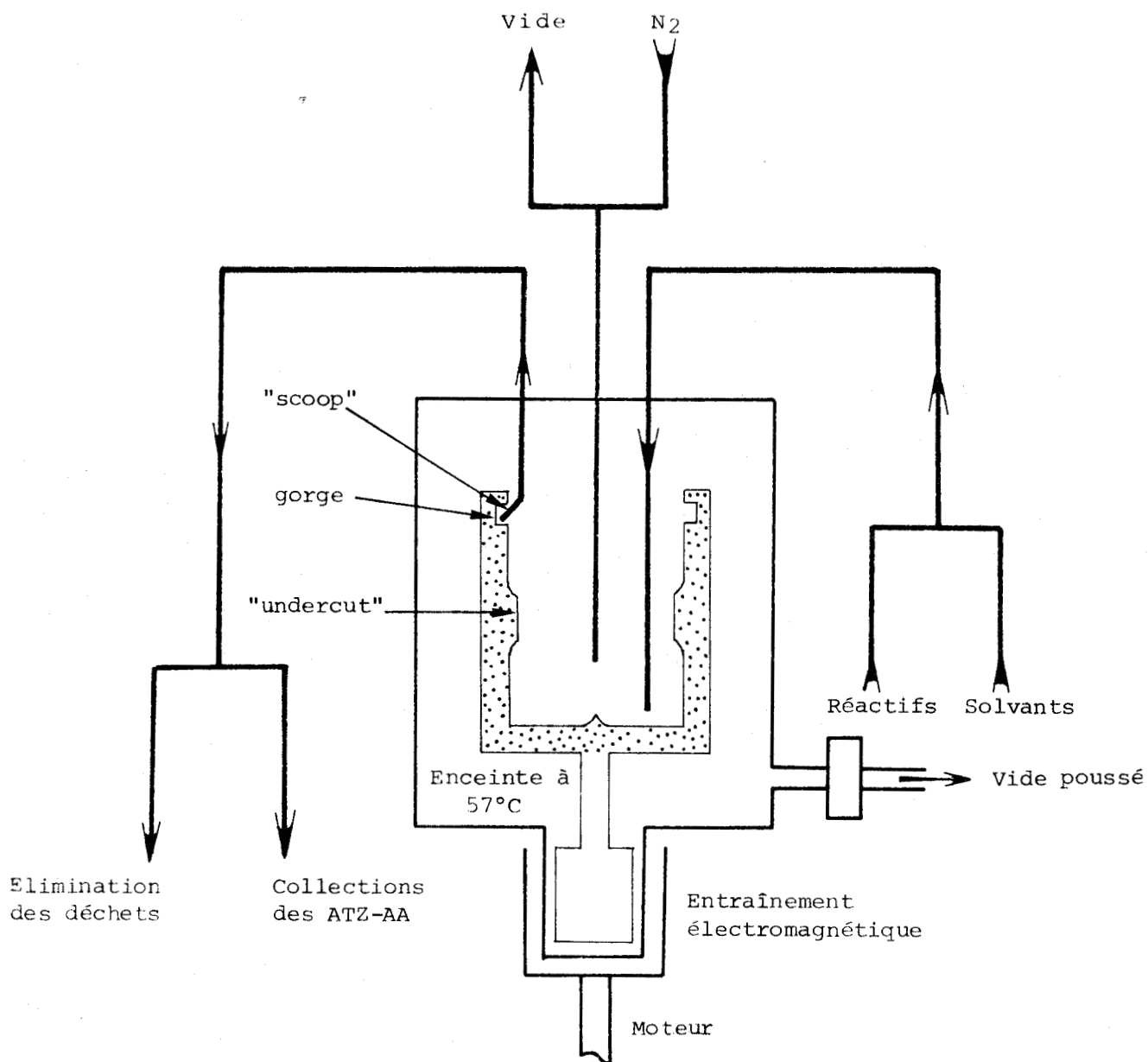


FIGURE 8 SCHEMA SIMPLIFIE DE LA COUPELLE DU SEQUENCEUR BECKMAN 890 C



La coupelle est située dans une enceinte chauffée à 57°C. Le chauffage et la régulation sont assurés par un thermocouple. On peut à l'intérieur de la coupelle réaliser trois vides différents. Le premier est un vide très faible de l'ordre de 5 Torr, le second est de l'ordre de 1 Torr, enfin le troisième est un vide très poussé inférieur à 100 millitorr. On envoie également dans la coupelle de l'azote soit pour casser le vide avant une délivrance de réactif ou de solvant, soit pour conserver une atmosphère inerte dans la coupelle, ou pour faire un pré-séchage après une délivrance de réactif. Cet azote de qualité "U" ($O_2 < 5$ ppm volume) arrive à l'intérieur de la coupelle par l'intermédiaire d'un diffuseur.

Les réactifs et les solvants nécessaires à la dégradation sont amenés dans la coupelle par 4 tubes fins en téflon. Le premier est commun à la délivrance du réactif 1 (PITC) et du solvant 1 (benzène), le second à la délivrance du réactif 2 et du solvant 2 et ainsi de suite. Les liquides délivrés sont ensuite aspirés par l'intermédiaire d'un tube en téflon appelé "scoop" qui s'ouvre près de la gorge de la coupelle.

I - 2-1-1-2 Le système de vide

Le vide est assuré par deux pompes Welch 1402 M à double étage. Le vide très faible (< 5 Torr) est réalisé par l'intermédiaire d'une vanne électromagnétique de type Skinner à deux voies ; celui de 1 Torr est commandé par une seconde vanne

électromagnétique du même type que la précédente ; enfin le vide poussé (< 100 millitorr) est réalisé par l'intermédiaire d'une vanne électropneumatique de type Bimba. Ces trois différentes vannes de vide travaillent successivement et à l'intérieur de la coupelle les trois entrées de vide sont séparées.

I - 2-1-1-3 Les réservoirs de réactifs et de solvants

Les huit bouteilles de réactifs et de solvants sont situées à l'avant de l'appareil. Les bouteilles de solvants sont en verre brun. Lors d'une délivrance d'un réactif ou d'un solvant, la bouteille désirée est d'abord mise à la pression atmosphérique pendant 10 secondes par l'intermédiaire d'une vanne électromagnétique à deux voies, puis sous pression d'azote pendant 30 secondes par l'intermédiaire d'un autre électro-aimant.

Ces bouteilles contiennent :

- Le phénylthiocyanate en solution à 5 p.100 dans le n-heptane
- Le tampon de couplage :
 - Tampon quadrol *N,N,N',N' tétra (2-hydroxypropyl)-éthylène diamine Quadrol 1M, 0.33M, 0.1M dans le mélange n-propanol/eau (3:4, v/v) ajusté à pH 9.0 avec de l'acide trifluoroacétique.*
 - Tampon DMAA (diméthylallylamine) *DMAA/n-propanol/eau (2:50:34, v/v/v) ajusté à pH 9.4 avec de l'acide acétique.*
 - Tampon DMBA (diméthylbenzylamine) *DMBA/n-propanol/eau (10:34:40, v/v/v) ajusté à pH 9.4 avec de l'acide acétique.*
- L'acide heptafluorobutyrique (HFBA)

- Le benzène
- L'acétate de butyle
- Le chlorobutane dans lequel nous ajoutons du dithioérythritol à 0.0005 p.100 (p/v) qui a pour but de préserver les ATZ-AA de la destruction.

I - 2-1-1-4 Les vannes de délivrance

Ces vannes sont réalisées entièrement en téflon. Un schéma de celles-ci est présenté dans la figure 9. Ces vannes sont au nombre de quatre et sont à deux voies c'est-à-dire que chacune d'entre elles est commune pour un réactif et le solvant correspondant à celui-ci. Elles ont deux entrées indépendantes et une sortie commune. De chaque côté du corps de vanne, une membrane et un pointeau en téflon commandé par l'intermédiaire d'une vanne électromagnétique à trois voies permet l'écoulement du réactif ou du solvant désiré vers la coupelle de réaction.

I - 2-1-1-5 La vanne "élimination/collection"

Comme les vannes de délivrance, elle est réalisée entièrement en téflon. Elle est commandée par des vannes électromagnétiques et permet soit d'envoyer les solvants de lavage (benzène, acétate d'éthyle) vers le flacon récepteur de déchets, soit d'envoyer la thiazolinone, extraite par le chlorobutane, dans le collecteur de fractions.

Un système de sécurité empêche l'ouverture simultanée du côté "élimination" et du côté "collection" de la vanne.

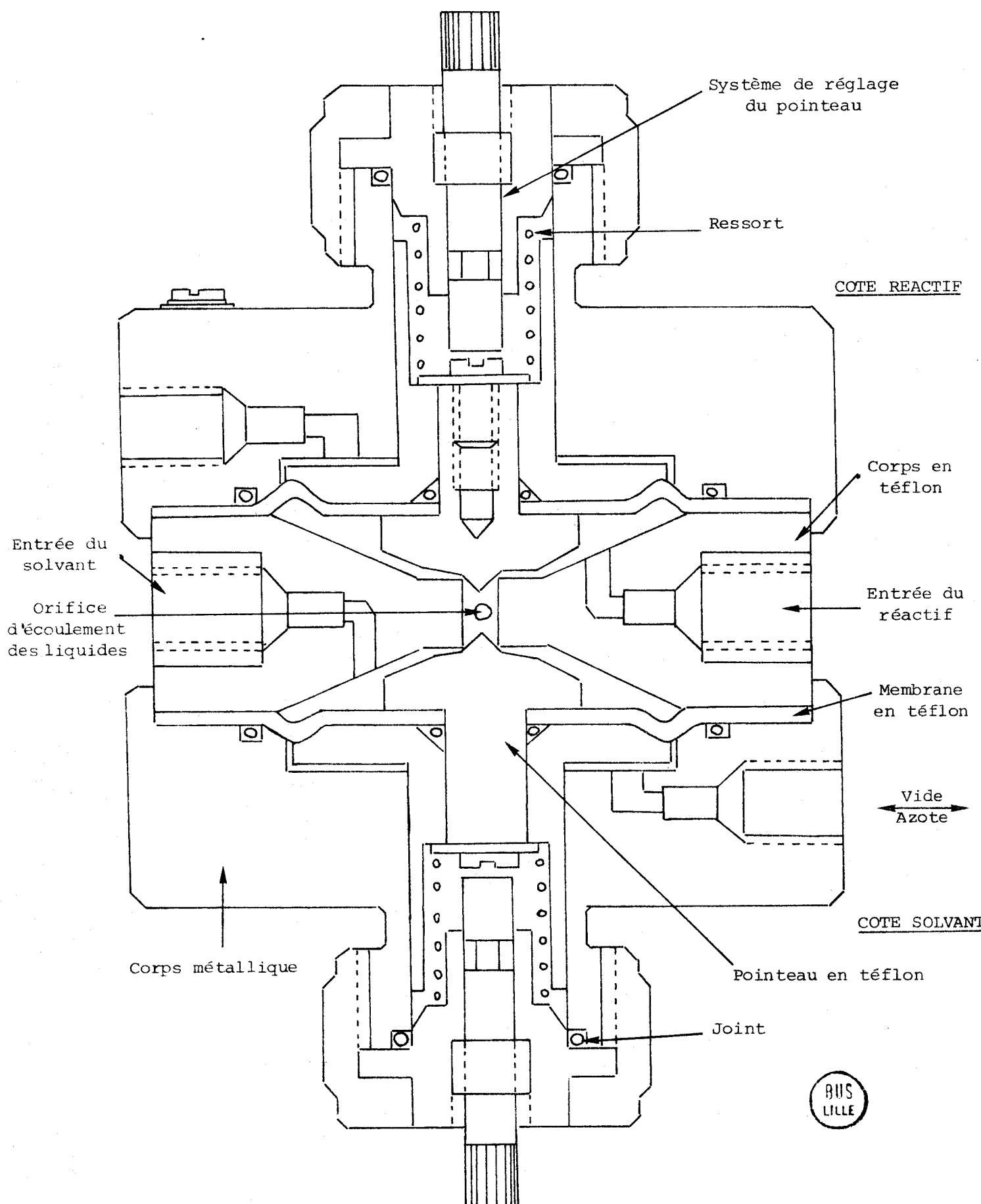


FIGURE 9 SCHEMA D'UNE VANNE DE DELIVRANCE BECKMAN

I - 2-1-1-6 Le collecteur de fractions

Le collecteur est situé dans une enceinte réfrigérée à + 8°C. Il contient 88 tubes. Il est facilement démontable et sa rotation est assurée automatiquement à la fin de chaque cycle de dégradation. Cette enceinte peut être placée sous vide partiel ou sous atmosphère d'azote.

I - 2-1-1-7 Le pupitre de commande

Ce pupitre commande par l'intermédiaire d'une bande perforée, toutes les opérations de la dégradation. La durée de chaque étape peut aller de 1 à 999 secondes. Le nombre d'étapes réalisables pour un cycle de dégradation est illimité. Le déroulement de chacune d'entre elles se fait automatiquement. De même, lorsqu'un cycle de dégradation est terminé, un second recommence automatiquement.

Le pupitre de commande comporte trois compteurs qui mémorisent :

- le nombre de cycles à effectuer
- le nombre de cycles déjà réalisés
- l'étape de dégradation à laquelle on se trouve.

Quand le séquenceur est en fonctionnement, des voyants lumineux visualisent les fonctions qui sont en cours. En outre, il est possible, grâce à un interrupteur à deux positions, de faire revenir la bande programme à une étape déterminée de la

séquence ceci afin d'effectuer soit un double couplage, soit un double clivage. De plus, tous les interrupteurs de commande de fonction de l'appareil sont à trois positions :

- automatique
- manuelle
- arrêt

Au cours d'un cycle, ils sont tous en position automatique mais la commande manuelle est prioritaire ce qui permet d'intervenir rapidement en cas de mauvaise lecture de la bande programme par exemple.

Le séquenceur comporte également un certain nombre de sécurités et d'impossibilités.

- Les sécurités

Elles jouent pour :

le manque d'azote : l'appareil s'arrête alors à l'étape où il se trouve, fait un très faible vide dans la coupelle et met le collecteur de fractions sous vide.

la coupure de courant : on retrouve l'appareil dans les mêmes conditions que précédemment s'il y a eu une coupure de courant supérieure à 1 minute.

la délivrance de réactif : il existe à ce niveau une sécurité réglable manuellement de 0 à 2 minutes. Si pour une raison quelconque (mauvaise lecture de la bande programme ou encore problème au niveau des compteurs de temps) le temps de délivrance d'un réactif est supérieur à la limite qu'on s'est imposée, l'appareil coupe la délivrance de ce réactif, reste à l'étape où il se trouve, fait un faible vide dans la coupelle et met également sous vide le collecteur de fractions.

l'arrêt du ventilateur-extracteur : si celui-ci, qui est situé à l'arrière du séquenceur, et qui assure l'évacuation vers l'extérieur de tous les effluents gazeux, s'arrête, l'alimentation électrique de l'appareil est disjonctée : ceci a pour effet d'arrêter les pompes, le chauffage et la rotation de la coupelle afin d'éviter toute étincelle susceptible de provoquer une explosion due à l'accumulation de vapeurs de solvants.

- Les impossibilités

Elles interviennent au niveau de :

la délivrance de réactif, de solvant, de l'azote : il est en effet impossible de demander une délivrance de réactif ou de solvant dans la coupelle si celle-ci est sous vide. De même, on ne peut pas non plus faire un faible vide dans la coupelle et y envoyer de l'azote en même temps, par contre, on peut faire un vide poussé et un balayage d'azote simultanément.

la collection : si à cause d'une erreur de perforation dans la bande programme, il est demandé à l'appareil d'envoyer l'ATZ-aminoacide à la fois vers le flacon récepteur de déchets et vers le collecteur de fractions, l'appareil choisira et enverra automatiquement celle-ci vers le collecteur de fractions.

I - 2-1-2 Fonctionnement de l'appareil

Nous venons de décrire l'appareil et ses "organes vitaux". Voyons maintenant les différentes opérations qui interviennent au cours de la dégradation automatisée d'une protéine ou d'un peptide.

La protéine (10 mg au maximum) est dissoute dans 500 μ l d'eau bidistillée ou d'acide formique à 10 p.100. La solution est alors introduite dans la coupelle en rotation. Le film liquide qui s'est déposé sur la paroi de la coupelle est séché. Le PITC en solution dans le n-heptane (5:100, v/v) est délivré dans la coupelle pendant 2 secondes afin de permettre le recouvrement du film protéique. On passe ensuite à la délivrance du tampon de couplage pendant 7 à 16 secondes selon le tampon choisi. La réaction de couplage s'effectue à 57°C pendant 20 minutes. Après séchage, l'excès de PITC et les dérivés secondaires formés (aniline, phénylthiourée, diphenylthiourée) sont éliminés par lavage au benzène et par des étapes de vide croissant (5 Torr, 1 Torr) suivies d'un balayage d'azote sous vide poussé (500 millitorr). Le benzène a pour but d'éliminer tous les constituants organiques tandis que le vide croissant (5 Torr, 1 Torr, 500 millitorr) suivi du balayage d'azote permet une évaporation partielle des constituants du tampon de couplage à base de DMAA (n-propanol, diméthylallylamine, eau). Un second lavage par l'acétate d'éthyle sera nécessaire si le tampon de couplage utilisé est à base de quadrol car celui-ci n'étant pas volatil, il ne sera pas éliminé lors des étapes de vide.

Après élimination de l'excès de PITC et des produits secondaires formés au cours de la réaction de couplage, le film est à nouveau séché par un vide croissant de 5 à 1 Torr suivi d'un balayage d'azote sous vide partiel (500 millitorr), afin d'éliminer le solvant résiduel présent dans la coupelle, et en vue de la préparation de l'étape suivante qui est le clivage.

Le dérivé phénylthiocarbamyl-protéine (ou peptide) est traité par l'acide heptafluorobutyrique anhydre. Le clivage et la cyclisation en 2-anilino-5-thiazolinone (ATZ) d'acide aminé s'effectue à 57°C pendant 7 minutes. Le film est ensuite séché en vue de l'extraction de la 2-anilino-5-thiazolinone d'acide aminé ainsi formée. A la fin d'un cycle, l'ATZ-AA est extraite par le chlorobutane. L'extrait "chlorobutane" est collecté dans la gorge de la coupelle à l'aide du "scoop" et recueilli dans le collecteur de fractions. L'ATZ-AA en solution dans le chlorobutane est recueillie dans un tube conique contenant 0.2 ml d'HCl 1N-éthylmercaptan 1 p.1000 (v/v). Le tube est placé dans un collecteur de fractions maintenu en atmosphère d'azote et à + 8°C. L'ATZ-AA collectée est ensuite convertie en un dérivé plus stable : la phénylthiohydantoïne d'acide aminé (PTH-AA).

Dans la coupelle, le film protéique est alors séché : la protéine (ou le peptide) amputée d'un résidu d'acide aminé est ainsi prête pour subir un nouveau cycle de dégradation.

La conversion de l'ATZ-AA se fait manuellement. Après évaporation de l'extrait "chlorobutane" sous azote, l'ATZ-AA est convertie à 80°C en PTH-AA par traitement avec l'HCl 1N - éthylmercaptan 1 p.1000 (v/v) pendant 10 minutes. Ce temps de conversion est un temps moyen. Il peut être réduit à 7 minutes si on suspecte la présence d'un dérivé d'acide aminé hydroxylé (sérine ou thréonine). En effet, la PTH-sérine se décompose lors de la conversion en PTH-déhydrosérine qui est instable. Quant à la PTH-thréonine elle donne un produit de β -élimination : la PTH-déhydrothréonine qui est stable mais dont le spectre présente un maximum d'absorption à 313 nm alors que celui des PTH-AA "normales" a un maximum d'absorption à 269 nm.

Par ailleurs, le temps de conversion est porté à 15 minutes afin d'avoir une conversion totale d'ATZ-valine, d'ATZ-leucine, d'ATZ-isoleucine, d'ATZ-phénylalanine ou d'ATZ-lysine en PTH-AA correspondantes.

Ces différentes opérations que nous venons de décrire et qui interviennent au cours de la dégradation automatisée d'une protéine ou d'un peptide ne sont pas illimitées. En effet, les performances de l'appareil, c'est-à-dire le nombre de cycles de dégradation possibles, sont directement liées aux problèmes de la chimie de la dégradation d'EDMAN.

Ces différents problèmes sont (64) :

- La présence d'oxygène qui produit une désulfuration oxydative du groupe phénylthiocarbamyl et bloque la formation de thiazolinone. Il est donc très important d'opérer en atmosphère inerte d'azote très pur, d'avoir des vannes étanches et

d'atteindre un bon vide, ces deux dernières conditions étant nécessaires aussi pour éviter le mélange de réactifs qui favoriserait la formation de sels.

- Les réactions incomplètes au niveau des étapes de couplage ou de clivage, qui peuvent être dues à certains acides aminés à chaîne latérale encombrante ou encore à la proline dont l'anilino-thiazolinone ne se clive que partiellement, et qui donneront de ce fait des "overlaps" aux cycles suivants.

- L'augmentation du bruit de fond due à la présence d'eau dans l'acide heptafluorobutyrique (agent de clivage de la réaction), qui provoque des réactions secondaires d'hydrolyse de la chaîne polypeptidique. Ce réactif doit donc être parfaitement anhydre.

- La formation d'aniline due à l'eau qui contamine le PITC, ainsi que la formation de dérivés thiourée et dithiourée.

- La pureté des réactifs et des solvants qui influence également le rendement de la dégradation. Ceux-ci doivent être dépourvus de toutes traces d'aldéhydes et de peroxydes.

- La séparation de la protéine (ou du peptide) des sous-produits et des réactifs résiduels au cours de la dégradation.

En plus de tous ces problèmes, il existe des difficultés qui sont spécifiques à la protéine (ou au peptide) et qui provoquent également une diminution des rendements. Ce sont :

- Les régions comportant des séquences répétitives avec en l'occurrence des résidus encombrants (ex: lysine, arginine) qui influent fortement sur le rendement de la dégradation. Les séquences Pro-Pro ou Pro-Ile ou encore Pro-Val diminuent elles aussi très fortement ce rendement (7).

- Les régions dans lesquelles un résidu d'arginine précède une séquence Pro-Ile ou Pro-Pro : le rendement de la dégradation est alors encore plus médiocre.

- Les régions contenant beaucoup de résidus basiques et qui donnent un taux élevé d'"overlaps".

- Les groupes de résidus acides et en particulier la séquence Glu-Glu-Asp affectent le rendement de la dégradation. Cependant dans ce dernier cas, une répétition du couplage permet de surmonter cette difficulté.

Malgré ces différents problèmes que nous venons de décrire, la dégradation d'EDMAN automatisée reste techniquement supérieure à la dégradation manuelle, d'autant plus que depuis quelques années les possibilités du séquenceur ont été élargies à la fois grâce à des modifications techniques de l'appareil et grâce à des modifications d'ordre chimique.

I - 2-1-3 Les modifications

I - 2-1-3-1 Les modifications techniques

Ces modifications ont été pour la plupart d'entre elles effectuées par WITTMANN-LIEBOLD et son groupe (65). En effet, le séquenceur Beckman 890 C vendu actuellement présente quelques points faibles au niveau :

- des vannes de délivrance
- du système de vide et du piège à froid
- de la coupelle
- de la conversion qui se fait manuellement

Les vannes de délivrance : ces vannes qui sont vendues très chères par la Société Beckman, ont un volume mort très important. De plus, elles amènent les réactifs et les solvants dans la coupelle par quatre lignes différentes ce qui peut représenter une source de fuites possibles. A l'intérieur de ces vannes, les cônes de téflon situés de part et d'autre du corps de téflon de la vanne peuvent subir des déformations. En outre, le vide à l'intérieur des vannes n'est pas excellent, et elles doivent donc être entretenues régulièrement pour conserver une bonne efficacité.

Tout ceci a pour conséquence que :

- l'intérieur des vannes n'est pas totalement vide après le vide restreint

- les vannes ne sont pas étanches très longtemps
- les volumes de solvants de lavage sont inutilement élevés à cause de la construction même de ces vannes
- la vidange des lignes de délivrance par le vide après réaction perturbe le film protéique : celui-ci a en effet tendance à monter le long de la paroi de la coupelle à cause de l'addition soit de tampon de couplage, soit d'acide heptafluorobutyrique provenant de la vidange des lignes par le vide restreint.

Le système de vide et le piège à froid : Dans le système de vide du séquenceur Beckman 890 C, il existe trois sorties séparées pour les différents vides effectués à l'intérieur de la coupelle, et qui sont opérationnelles l'une après l'autre ce qui provoque de la condensation dans les lignes de vide et diminue les performances de l'appareil. En outre, les différents vides sont réalisés grâce à deux pompes. Nous avons également remarqué que la vanne pneumatique de type Bimba ainsi que la coupelle ne sont pas parfaitement étanches. Les fuites sont causées par un trop grand nombre de connections et par le manque d'étanchéité des joints en viton.

Quant au piège à froid commercialisé par Beckman, il n'équipait pas notre séquenceur à l'origine. Depuis son installation, nous avons observé une amélioration du vide sur notre appareil, mais il n'est quand même pas parfait. En effet, les vapeurs de réactifs et de solvants qui se condensent dans le piège, sont recueillies dans un ballon placé à la température

ambiante. Dans ces conditions, les solvants s'évaporent et les vapeurs ainsi formées vont contaminer l'huile de pompe.

Tout ceci conduit à :

- une diminution de l'efficacité des pompes qui ont une durée de vie raccourcie
- une condensation de liquide et une formation de dépôts solides dans les lignes de vide
- une destruction ou une contamination des ATZ-AA.

La coupelle : le changement brutal de la vitesse de rotation de celle-ci, qui est effectué au cours des différentes étapes de la dégradation, fragmente le film protéique. De plus, la coupelle du séquenceur Beckman 890 C comporte beaucoup trop de sorties (lignes de délivrance, "scoop", thermistance) qui sont sources de problèmes (perte de vide, pénétration d'oxygène dans la coupelle).

Différentes améliorations ont donc été effectuées sur le séquenceur Beckman 890 C notamment au niveau :

- des vannes de délivrance
- du système de vide et du piège à froid
- de la coupelle
- de la conversion.

Elles feront l'objet du second chapitre.

I - 2-1-3-2 Les modifications chimiques

Les performances du séquenceur peuvent également être améliorées grâce à des modifications d'ordre chimique :

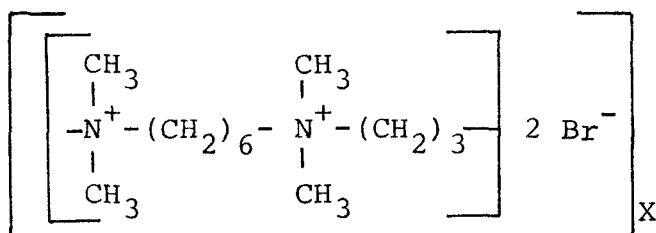
- Utilisation de mainteneurs
- Greffe de radicaux hydrophiles
- Utilisation de tampons de couplage autres que le quadrol 1M
- Emploi du PITC marqué
- Purification des réactifs et des solvants

Utilisation de mainteneurs . Les peptides ou les protéines relativement courts ou en trop faible quantité, sont très facilement extraits par les solvants de lavage au cours de la dégradation automatique. Différents auteurs ont donc eu l'idée d'utiliser un mainteneur afin de stabiliser mécaniquement la protéine (ou le peptide) dans la coupelle. Cependant, ce mainteneur doit être inerte vis-à-vis de la dégradation d'EDMAN. Différents mainteneurs ont été utilisés. Il s'agit d'abord :

- de protéines bloquées naturellement à leur extrémité N-terminale : la parvalbumine (51), l'apocytochrome C (49),
- de protéines bloquées artificiellement : l'apomyoglobine succinylée (38), le lysozyme modifié (24),
- de polymères artificiels : polyornithine succinylée (55), copolymère norleucine-arginine (48), polyornithine sulfonylée (60).

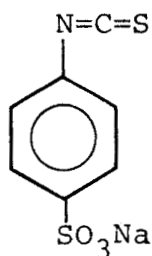
Cependant, les protéines bloquées ne sont plus employées car on a remarqué qu'au cours de la dégradation elles subissent des coupures dans leur chaîne peptidique. Ces coupures sont dues à l'exposition de la protéine aux clivages répétés en milieu acide heptafluorobutyrique. De ce fait, après un certain nombre de cycles de dégradation, on voit apparaître un "bruit de fond" important qui provient de la dégradation du mainteneur.

Le mainteneur qui est le plus employé actuellement est le polybrène (24,55,57). C'est un ammonium quaternaire de formule :

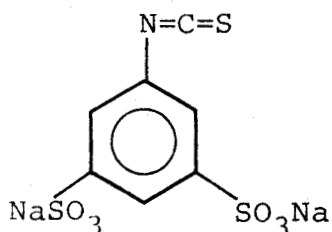


Son mécanisme d'action n'est pas très bien connu mais la quantité minimale à utiliser est de 2 à 3 mg (37). L'inconvénient de ce mainteneur est qu'il provoque des artefacts qui gênent la détection des PTH-AA en chromatographie liquide à haute performance. HUNKAPILLER et HOOD (33) ont remédié à ce problème en dégradant dans un premier temps un dipeptide Gly-Gly en présence de ce mainteneur. On effectue trois cycles de dégradation au cours desquels les artefacts, dûs au polybrène réagissent avec le dipeptide et sont éliminés. Une fois cette opération terminée, la protéine (ou le peptide) est déposée dans la coupelle puis dégradée.

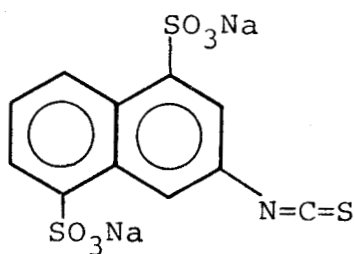
Greffe de radicaux hydrophiles . Certains chercheurs ont remarqué que les peptides riches en lysine posaient un problème au cours de la dégradation. En effet, les peptides contenant beaucoup de lysine, sont plus extractibles par le benzène et l'acétate d'éthyle du fait de la formation du groupement phénylthiocarbamylé sur la fonction ϵ -NH₂. De ce fait, le peptide a tendance à être extrait au cours des étapes de lavage, le rendement de la dégradation devient alors faible et on remonte beaucoup moins loin dans la séquence peptidique. Certains auteurs ont remédié à ce problème par greffe de radicaux hydrophiles. Les réactifs les plus utilisés sont ceux décrits par BRAUNITZER (10,11). Ce sont des dérivés isothiocyanate. Leur formule est présentée dans la figure 10. Selon BRAUNITZER, les dérivés sulfonés du naphthalène isothiocyanate donnent un bon film ce qui est une condition essentielle pour réaliser une bonne dégradation. Cette modification chimique présente cependant un inconvénient : les résidus de lysine ainsi modifiés ne sont plus solubles dans les solvants utilisés habituellement pour l'extraction des ATZ-AA, et donc ne sont pas identifiables directement. On peut pourtant remédier à ce problème en utilisant moins de réactif afin de n'avoir qu'une modification partielle. Dans ce cas, les résidus de lysine restent identifiables. Enfin, le 4-sulfo-PITC possède un rendement de cyclisation dans l'acide heptafluorobutyrique qui est assez bas. Il en résulte une coupure incomplète du premier résidu de la protéine et de ce fait la production d'un "overlap" qui gêne l'identification des résidus suivants.



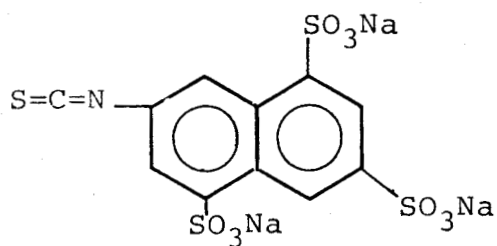
Réactif I : 1 isothiocyano-4 benzène sulfonate



Réactif II : 5 isothiocyano-1-3-benzène disulfonate



Réactif III : 3 isothiocyano-1,5-naphtalène disulfonate



Réactif IV : 7 isothiocyano-1,3,5-naphtalène trisulfonate

FIGURE 10 LES DERIVES ISOTHIOCYANATE DE BRAUNITZER (10,11)



DWULET et coll. (17) utilisent pour leur part le 3-sulfo-PITC dont le rendement de cyclisation est légèrement supérieur (5 p.100) à celui du 4-sulfo-PITC. Quant à FOSTER et coll. (22), ils greffent un composé hydrophile sur le groupe α -carboxylique du résidu C-terminal. Il s'agit de l'acide 2-amino-naphtalène-1,5-disulfonique qui est fixé par l'intermédiaire d'une carbodiimide. Si on effectue la réaction à pH 4.0, il n'est pas nécessaire de bloquer au préalable les groupements α -aminés libres de la chaîne peptidique dont le pK est environ de 9.5. Le groupement carboxylique des chaînes latérales des acides dicarboxyliques réagit de la même façon que celui du résidu C-terminal du peptide. Les acides dicarboxyliques sont identifiés après hydrolyse totale acide de la phase aqueuse. L'avantage de l'addition de ce composé naphtalénique est qu'il augmente la quantité du film protéique.

Enfin, GARRICK et SLOAN (26) proposent de brancher un peptide par son résidu en position C-terminale, à un polypeptide tel que le cytochrome C par l'intermédiaire d'une carbodiimide.

Utilisation de tampons de couplage autres que le quadrol 1M .

Une autre façon de rendre la dégradation plus performante est d'utiliser un tampon différent de celui préconisé par EDMAN. En effet, le tampon d'EDMAN qui est le quadrol (N,N,N',N', tétra(2-hydroxypropyl)-éthylène diamine) 1M a un excellent pouvoir tampon mais comme il n'est pas volatil, il faut beaucoup de benzène puis d'acétate d'éthyle pour l'extraire. Ces

lavages prolongés risquent d'extraire les peptides et à un degré moindre, les protéines. De ce fait, BRAUER et coll. (8) utilisent le quadrol 0.1M qui demande moins de solvant de lavage. Cependant, le quadrol n'étant soluble que dans l'acétate d'éthyle, il est nécessaire de maintenir le lavage avec ce solvant.

BEGG et MORGAN (2) emploient pour leur part un tampon: le THEED (tétrahydroxyéthyl/éthylène diamine) analogue au quadrol et qui contient moins d'aldéhydes que celui-ci.

HERMODSON et coll. (30) recommandent l'utilisation de la diméthylbenzylamine (DMBA) 0.8M. Ce tampon est soluble dans le benzène ce qui évite l'utilisation de l'acétate d'éthyle et de plus il est beaucoup moins coûteux que le quadrol 1M. Son utilisation sera donc très intéressante pour les protéines et les peptides hydrophobes. Ce tampon a toutefois pour inconvénient de salir fortement toute l'enceinte de la coupelle.

NIALL et coll. (47) ainsi que THOMSEN et coll. (58) utilisent pour leur part la diméthylallylamine (DMAA) 0.8M. Ce tampon est très volatil. Il est donc éliminé par une simple évaporation sous vide ce qui diminue la quantité de benzène nécessaire au lavage du film. Cependant, les auteurs ont remarqué une dégradation anormale de l'histidine en présence de diméthylallylamine. En effet, à la fin de l'étape de couplage, le pH du milieu réactionnel peut descendre jusqu'à 6 pendant l'évaporation du tampon, ce qui est suffisant

pour obtenir un clivage partiel du phénylthiocarbamyl-histidine peptide (PTC-His). Il en résulte un "overlap" au cycle suivant qui s'explique par le fait que le groupe α -aminé du résidu placé immédiatement après l'histidine réagit avec le PITC. De plus, la PTH-histidine sera mal quantifiée car une partie de la PTC-histidine est extraite à la fin du couplage, par les solvants de lavage. Cette réaction de clivage partiel se fera d'autant plus facilement si le résidu d'histidine est suivi d'un résidu de proline ou de glycine.

Enfin, BRANDT et VON HOLT (6) lors de la détermination de la séquence de certaines histones ainsi que BRAUNITZER et SCHRANK (9) ont utilisé la diméthylaminopropylène (DMAP). Ceux-ci trouvent que les rendements obtenus en dégradation sont meilleurs avec ce tampon qui est absolument dépourvu d'aldéhydes, les peptides séquencés dans ces conditions sont moins extractibles et les diagrammes d'identification des PTH-AA sont plus facilement interprétables.

Emploi du PITC marqué. L'emploi de réactifs marqués, permet également d'atteindre une très bonne sensibilité qui est de l'ordre de la femtomole (10^{-15} mole). Le réactif le plus utilisé est le PITC marqué au soufre 35 (^{35}S -PITC) qui a une activité spécifique beaucoup plus élevée que celle du ^{14}C -PITC ou du ^3H -PITC ce qui permet une analyse qualitative et quantitative des PTH radioactives. On effectue donc un premier couplage avec un mélange de ^{35}S -PITC et de PITC non marqué,

puis on fait un deuxième couplage avec du PITC froid afin d'éviter une asynchronie de la dégradation. Cependant, l'utilisation du ^{35}S -PITC reste très limitée car c'est un produit fort cher qui de surcroît n'est pas très pur d'où la nécessité de le purifier. A ces inconvénients viennent s'ajouter les problèmes liés à la contamination de tout le séquenceur.

Purification des réactifs et des solvants. La purification des réactifs et des solvants permet également d'augmenter les performances du séquenceur. WITTMANN-LIEBOLD (65) préconise les procédés de purification suivants :

- PITC et HFBA purifiés par passage sur colonne contenant de l'alumine acide et de la silice.

- Quadrol 0.1M purifié par passage sur colonne d'alumine neutre et de silice.

- Benzène et acétate d'éthyle distillés sous azote puis purifiés sur colonne d'alumine neutre et de silice.

- Polybrène purifié par passage sur colonne de Dowex 1x4 puis sur colonne d'alumine acide.

Grâce à toutes ces purifications, les PTH-AA obtenues sont très pures et les contaminants les plus importants que l'on remarque lors de l'identification en chromatographie liquide haute performance, notamment au niveau de l'alanine, disparaissent.

D'autres procédés de purification ont été décrits par FRANK (25) : ils concernent les solvants et le quadrol.

Les solvants (benzène/acétate d'éthyle) sont conservés sur de l'alumine neutre à raison de 20 g/litre. Quant au quadrol, il est conservé sur aminoéthylcellulose, pour fixer les aldéhydes contaminants habituels du quadrol. L'aminéthylcellulose est régénérée toutes les dix séquences. Afin d'éviter que des particules d'alumine et d'aminéthylcellulose ne bouchent les lignes de délivrance, FRANK place un verre fritté de porosité 2 à l'intérieur des flacons de réactifs et de solvants.

L'accumulation d'"overlaps" au fur et à mesure de la dégradation, les coupures qui se produisent dans la chaîne polypeptidique au cours des clivages répétés par l'acide heptafluorobutyrique ainsi que les impuretés contenues dans les réactifs et les solvants, donnent des pics parasites lors de l'identification des PTH-AA en chromatographie liquide à haute performance. Ce "bruit de fond" qui s'accumule tout au long de la dégradation représente un facteur limitant pour la détermination de la séquence. BHOWN et coll. (5) ont donc introduit un "nettoyeur" de "bruit de fond".

Le "nettoyeur" employé est la fluorescamine. Les auteurs l'utilisent pour masquer le "bruit de fond" lors de la détermination de la séquence de protéines riches en proline. En effet, les résidus de proline sont très difficilement clivés et de ce fait, créent beaucoup d'"overlaps". Le traitement à la fluorescamine est très facilement réalisable.

Dans un premier temps, les auteurs déterminent la séquence de la protéine afin de localiser les résidus de proline. La protéine est ensuite passée une seconde fois au séquenceur. Avant chaque résidu de proline, la coupelle est mise à température ambiante, puis le programme est démarré normalement mais sans délivrance de PITC. Immédiatement après la délivrance du tampon de couplage, les auteurs stoppent le programme de dégradation. La coupelle est alors ouverte pour ajouter la fluorescamine (3 à 4 mg/ml). Le cours de la dégradation est alors repris normalement mais l'extrait "chlorobutane" est envoyé au flacon récepteur de déchets. Un cycle normal de dégradation est alors effectué avec un double clivage. Ce traitement à la fluorescamine est ensuite répété avant chaque résidu de proline. Il a pour effet de bloquer les nouveaux N-terminaux formés au cours de la dégradation (N-terminaux dus aux coupures dans la chaîne polypeptidique, ou encore "overlaps" provenant des acides aminés précédents), il n'interfère pas dans la chimie d'EDMAN puisque le groupement imine de la proline ne réagit pas avec la fluorescamine et permet une identification normale des PTH-AA en chromatographie liquide à haute performance (HPLC).

La méthode de dégradation automatisée décrite par EDMAN et BEGG, et dont nous venons de présenter les principes, la mise en oeuvre et les perfectionnements, s'effectue sur un film de protéine ou de peptide collé sur une paroi de verre : c'est la méthode en phase liquide.

La dégradation d'EDMAN automatisée peut également s'effectuer en phase solide selon un procédé mis au point par LAURSEN (42,43) ou en phase "gazeuse" selon un procédé mis au point par HEWICK et coll. (31).

I - 2-2 Le séquenceur en phase solide

Dans le séquenceur en phase solide, les peptides sont attachés par leur extrémité C-terminale à un support insoluble. La dégradation d'EDMAN se fait ensuite classiquement. L'excès de réactif est éliminé et l'ATZ-AA extraite par lavage du support avec les solvants appropriés.

L'avantage de cette méthode réside dans le fait que les lavages et les extractions sont pratiqués sans perte de peptide celui-ci étant lié de manière covalente à la résine. Cette méthode pourrait paraître idéale, mais le point crucial est celui de la fixation du peptide au support. En effet, la bonne marche de la dégradation dépend essentiellement du rendement de fixation du peptide au support.

Les différents supports utilisés sont :

- les résines à base de polystyrène (aminopolystyrène, triéthylènetétramine polystyrène)

- les billes de verre sur lesquelles sont greffés des groupes aminopropyles ou β -aminoéthylaminopropyles.

La fixation du peptide au support restant délicate, la méthode de dégradation en phase solide reste beaucoup moins employée que celle en phase liquide. De plus, cette méthode se trouve également limitée par la nature des protéines. Ainsi, des protéines basiques comme les histones, ou les HMG-protéines, ne peuvent être séquencées par cette technique.

I - 2-3 Le séquenceur en phase solide, liquide, gazeuse

Dans le séquenceur en phase "gazeuse", la protéine (ou le peptide) en solution aqueuse est déposée sur un petit disque en fibre de verre monté dans la cellule de réaction. Cette cellule est en verre ultrasonique et elle est miniature (25 mm de diamètre x 25 mm de long). L'arrivée des réactifs et des solvants se fait par un tube capillaire de 0.5 mm de diamètre. Les vannes de délivrance et le système de conversion sont également miniaturisés. Le tampon de couplage employé est la triméthylamine- H_2O sous forme vapeur, et le réactif de clivage est l'acide trifluoroacétique vapeur au lieu de l'acide heptafluorobutyrique liquide. Les autres réactifs et solvants utilisés sont identiques à ceux employés en phase liquide.

Les caractéristiques essentielles du programme utilisé sur le séquenceur en phase "gazeuse" sont les suivantes :

- Délivrance d'une quantité de PITC à 15 p.100 dans le n-heptane juste suffisante pour mouiller complètement le disque en fibre de verre (20 μ l environ).

- L'heptane est ensuite évaporé par un courant d'argon à l'intérieur de la chambre de réaction.

- Le couplage est effectué grâce à un écoulement faible de triéthylamine- H_2O vapeur à travers la chambre de réaction.

- Le clivage par le TFA se fait de la même façon.

- Les ATZ-AA sont ensuite extraites du disque de fibre de verre par de nombreuses aliquotes de chlorobutane et sont converties automatiquement.

Les avantages de cette méthode par rapport aux deux précédentes sont les suivants :

- Immobilisation de l'échantillon pendant le couplage et le clivage : l'échantillon est enrobé dans une matrice de polybrène à l'intérieur de laquelle la combinaison des réactifs liquides et la diffusion des réactifs gazeux mènent à bien la dégradation d'EDMAN.

- L'échantillon n'est jamais exposé aux liquides dans lesquels il est soluble.

- La miniaturisation des différents éléments qui composent le séquenceur permet de séquencer de très faibles quantités de protéine ou de peptide (10 nanomoles à 5 picomoles pour la myoglobine), d'avoir une consommation très faible de réactifs et de solvants, ainsi qu'une maintenance assez faible.

Cette méthode semble donc être parfaite pour la dégradation des protéines et des peptides et le séquenceur en phase "gazeuse" qui vient d'être récemment commercialisé pourrait dans un avenir proche supplanter les séquenceurs en phase liquide et en phase solide.

CHAPITRE II

LES AMELIORATIONS APORTEES AU SEQUENCEUR BECKMAN 890C ET
A LA METHODOLOGIE DE LA DEGRADATION D'EDMAN AUTOMATISEE

PLAN DU CHAPITRE II

LES AMELIORATIONS APORTEES AU SEQUENCEUR BECKMAN 890C ET
A LA METHODOLOGIE DE LA DEGRADATION D'EDMAN AUTOMATISEEII - 1 MODIFICATIONS TECHNIQUES DU SEQUENCEUR BECKMAN 890C

II - 1-1 Les vannes de délivrance

II - 1-2 Le système de vide

II - 1-3 Le piège à froid

II - 1-4 La coupelle de réaction

II - 1-5 La conversion automatique

II - 1-6 Autres modifications

II - 1-7 Détermination de la séquence N-terminale de
la myoglobine de Cheval après modification
du séquenceur Beckman 890C.II - 2 TRAITEMENT DES REACTIFS ET DES SOLVANTSII - 3 MICROSEQUENCE EN PHASE "GAZEUSE"

II - 1 MODIFICATIONS TECHNIQUES DU SEQUENCEUR BECKMAN 890C

II - 1-1 Les vannes de délivrance

Des nouvelles vannes commercialisées par la Société W.K.G. (WITTMANN-LIEBOLD, KHOLS, GRAFFUNDER) à Berlin équipent déjà plusieurs séquenceurs. Elles ont l'avantage d'avoir un volume mort très restreint. Elles sont entièrement en téflon et sont montées en série. A l'intérieur de celles-ci, le liquide chemine en zig-zag. L'ouverture et la fermeture des entrées et des sorties individuelles se fait par une membrane en téflon. Leur ouverture est commandée par le vide tandis que la fermeture se fait par l'envoi d'azote sous pression de 5 bars. Enfin, les lignes sont totalement vidées après chaque délivrance de solvant ou de réactif grâce à un courant d'azote qui pousse le liquide.

Nous avons remplacé les grosses vannes de délivrance équipant le séquenceur Beckman 890C par deux blocs de vannes W.K.G. comportant chacun neuf unités. Un schéma simplifié de ces vannes est présenté dans la figure 11.

Cette modification s'est traduite par une économie très appréciable (de l'ordre de 50 p.100) pour les solvants de lavage, benzène ou acétate d'éthyle, utilisés après le couplage.

La diminution du volume du solvant nécessaire pour l'élimination de l'excès de réactif de couplage et des produits

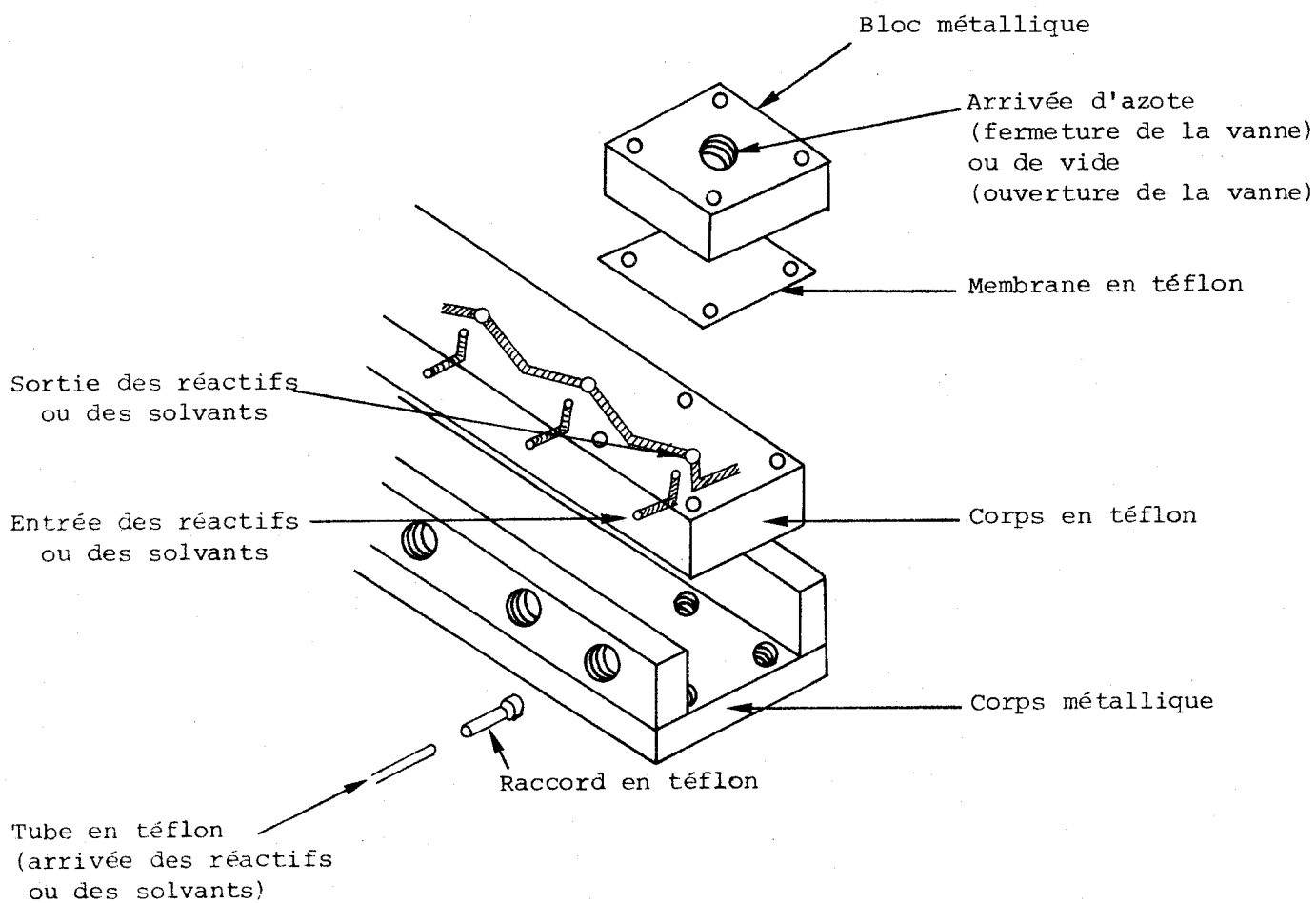


FIGURE 11 SCHEMA DU SYSTEME DE VANNE DE DELIVRANCE WKG (62)

secondaires de la réaction a un effet bénéfique sur la qualité et l'adhérence du film protéique sur la paroi de la coupelle. La dégradation automatisée de la protéine s'en trouve ainsi améliorée et peut être effectuée sur un plus grand nombre de cycles.

II - 1-2 Le système de vide

Il est possible de résoudre le problème des pertes de vide en remplaçant la vanne pneumatique Bimba par une grosse vanne électromagnétique comme l'ont fait BHOWN et coll. (3). Une solution plus radicale a été choisie par WITTMANN-LIEBOLD qui a complètement reconstruit le système de vide (65). Son système est réalisé entièrement en aluminium avec une seule entrée directe qui est celle du vide poussé. Deux vannes solénoïdes (Leybold-Heraeus) branchées en parallèle commandent le vide restreint et le vide poussé qui sont réalisés à l'aide d'une pompe unique de type Alcatel 2033 C.

II - 1-3 Le piège à froid

Le montage d'un piège à froid sur la ligne de vide, présente les avantages suivants :

- Condensation des vapeurs de solvant et élimination des retours de vapeurs d'huile dans la coupelle.

- Protection de la pompe à vide et absence de condensation des sous-produits de la dégradation d'Edman dans les lignes de vide, ce qui diminue considérablement la contamination des ATZ-AA par les sous-produits.
- Réduction des pertes par oxydation.

Nous avons installé sur notre séquenceur, un piège à froid Beckman dont le schéma est présenté dans la figure 12. Ce piège se compose de deux parties.

La première partie est constituée d'un vase de type "Dewar" en verre qui comporte des chicanes afin de mieux condenser les vapeurs de solvant extraites de la coupelle de réaction. Ce vase est emprisonné à l'intérieur d'un bâti métallique et il est isolé par de la mousse de polyuréthane. A l'extérieur du système réfrigéré est branché un ballon amovible dans lequel sont recueillis les produits de condensation. Le vase Dewar est rempli d'alcool absolu (600 ml) à l'intérieur duquel est immergé un "doigt" métallique qui est relié à la seconde partie de l'appareillage c'est-à-dire le groupe de froid.

Ce groupe de froid est constitué d'un cryostat de la marque NESLAB qui peut descendre jusqu'à -80°C , -100°C . Ce système a deux étages. Le premier étage de froid est un circuit fermé de pentane et le second un circuit fermé de fréon. Le froid est transmis au "doigt" métallique par l'intermédiaire d'un flexible qui est très fragile lorsqu'il est à -80°C . Ce piège à froid doit être décongelé toutes les semaines car on s'aperçoit qu'en fin de dégradation, le vide à l'intérieur de la coupelle devient moins bon à cause du colmatage du vase

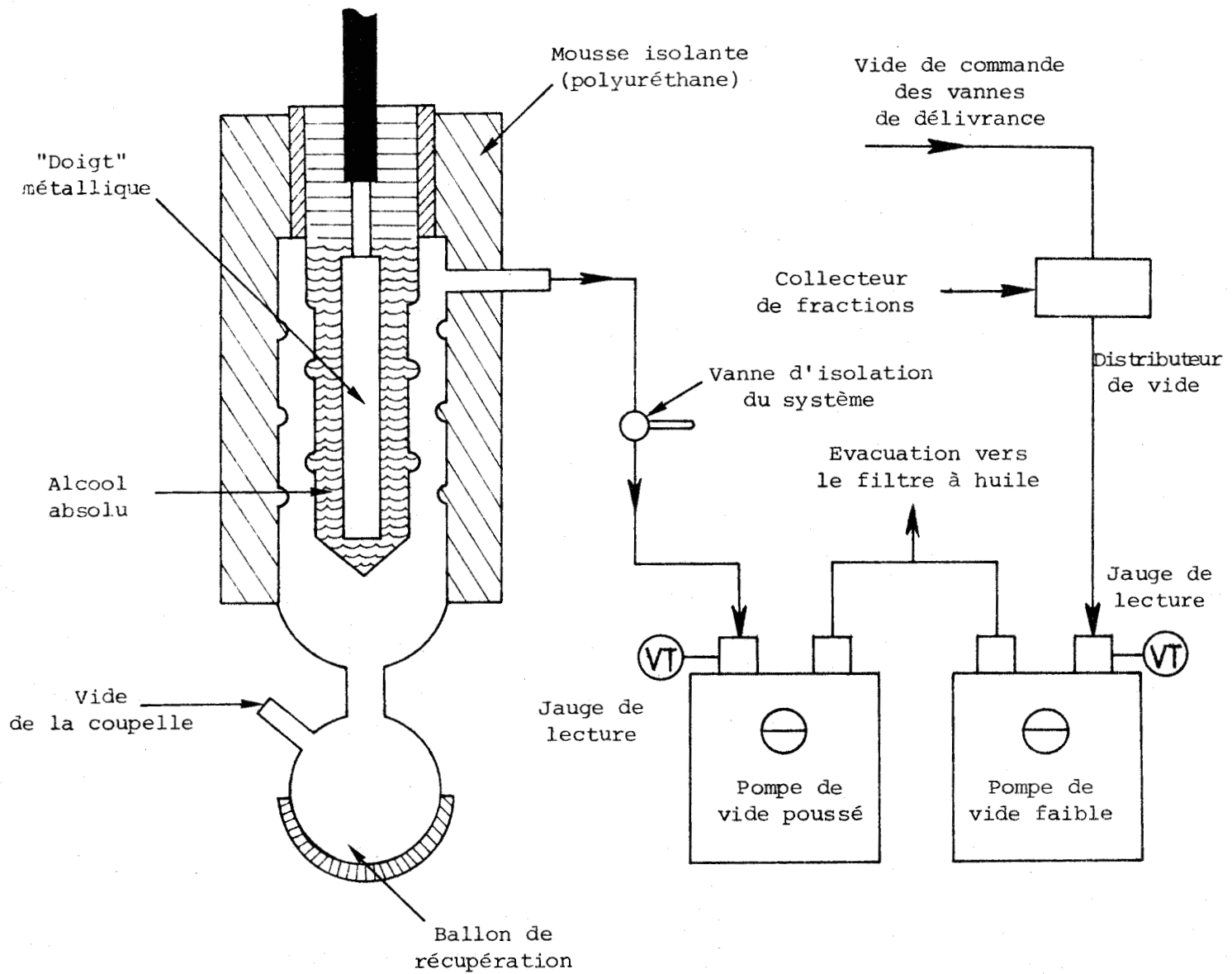


FIGURE 12 SCHEMA SIMPLIFIE DU PIEGE A FROID BECKMAN

Dewar. Cette opération de décongélation représente un inconvénient car pour avoir de nouveau un piège efficace, on est obligé de couper la réfrigération pendant 24 heures ce qui représente une perte d'utilisation du temps machine.

Par ailleurs, le ballon où sont collectés les produits de condensation doit être vidé tous les 25 cycles de dégradation environ. En outre, comme il se trouve à l'extérieur du système réfrigéré, les produits de condensation ont tendance à s'évaporer et les vapeurs ainsi formées sont entraînées dans la pompe, ce qui est un inconvénient.

Dans un avenir proche, nous espérons remplacer ce piège par un autre système qui permettrait de recueillir les vapeurs condensées à l'intérieur d'un ballon qui serait réfrigéré afin d'éviter la réévaporation des solvants et d'avoir un piège très efficace. Cependant, en dépit de cet inconvénient, l'installation du piège à froid Beckman sur notre séquenceur a apporté une amélioration sensible de la dégradation d'EDMAN automatisée.

En effet, afin d'estimer l'efficacité de ce piège, nous avons soumis la myoglobine de coeur de Cheval à la dégradation d'EDMAN, avant et après la mise en place du piège à froid.

Une même quantité de protéine (200 nanomoles) a été séquencée avec un programme quadrol 0.1M. La comparaison des rendements des différentes PTH-AA déterminées a été effectuée sur les 20 premiers résidus de la protéine (tableau I).

Cette comparaison nous montre que le rendement des PTH-AA est bien meilleur après la mise en place du piège à froid; nous

TABLEAU I

COMPARAISON DES RENDEMENTS OBTENUS LORS DE LA DETERMINATION DE LA SEQUENCE DE LA MYOGLOBINE DE COEUR DE CHEVAL
AVANT ET APRES INSTALLATION DU PIEGE A FROID

PTH-AA déterminée	Gly	Leu	Ser	Asp	Gly	Glu	Trp	Gln	Gln	Val	Leu	Asn	Val	Trp	Gly	Lys	Val	Glu	Ala	Asp
Position dans la séquence	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Rendement obtenu avant installation (en pourcentage)	15,7	41,7	6,6	40,5	11,5	22	30,4	16,5	12	11,8	24,4	20,1	15,1	11,5	6	15,1	13	n.d	11,2	11,8
Rendement obtenu après installation (en pourcentage)	26,9	82,8	12,1	47,3	19,1	44,4	37,5	34,7	37,3	31,6	42,4	28,1	23,7	22,3	11,2	28,1	17,2	18,9	22,1	20,9



avons notamment une quantité de PTH-sérine qui est doublée, ce qui est très intéressant car la PTH-sérine pose souvent des problèmes d'identification. La PTH-sérine en effet n'est pas stable (80 p.100 de PTH-sérine sont détruites) et donne un produit de β -élimination (figure 13). De plus, elle donne en HPLC des pics de décomposition qui lui sont caractéristiques.

Des expériences ultérieures nous ont montré que la PTH-thréonine, dont 65 p.100 se décompose en déhydrothréonine (figure 14) était elle aussi beaucoup mieux préservée.

II - 1-4 La coupelle de réaction

Différents types de coupelle ont été proposés. Tout d'abord, FAIRWELL et BREWER (20) ont adapté sur leur séquenceur une coupelle qui tourne à 1800 et 3600 tours/minute c'est-à-dire deux fois plus vite que la coupelle d'origine. Celle-ci a un diamètre supérieur à la coupelle Beckman et son "undercut" est situé juste sous la gorge. Du fait de la rotation plus rapide, la surface du film augmente de 75 p.100, les échanges avec les réactifs sont meilleurs et à partir de 600 nmoles d'apomyoglobine de sperme de Baleine, les auteurs ont pu déterminer sans ambiguïté les 85 premiers résidus de la protéine.

INGLIS et coll. (35) ont pour leur part remplacé la coupelle traditionnelle en verre par une coupelle en métal.

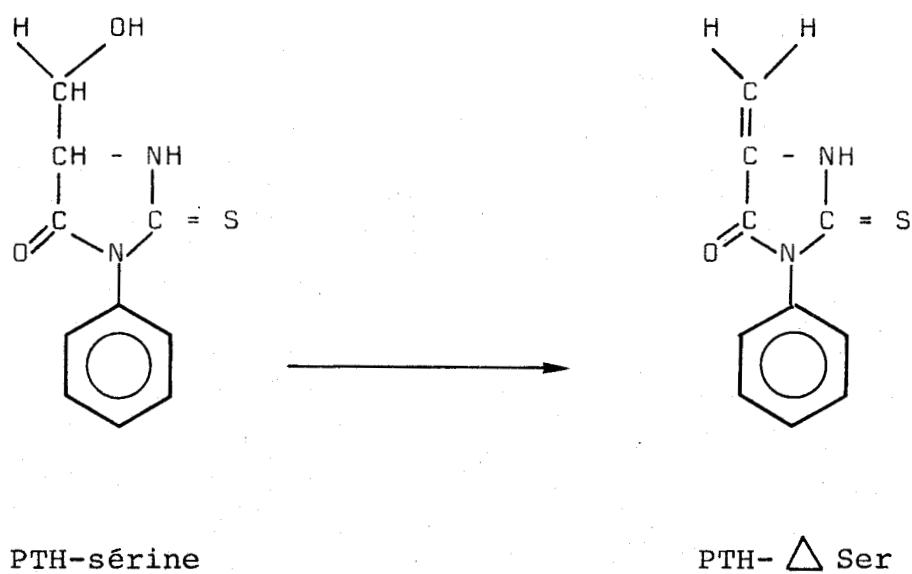


FIGURE 13 DECOMPOSITION DE LA PTH-SERINE

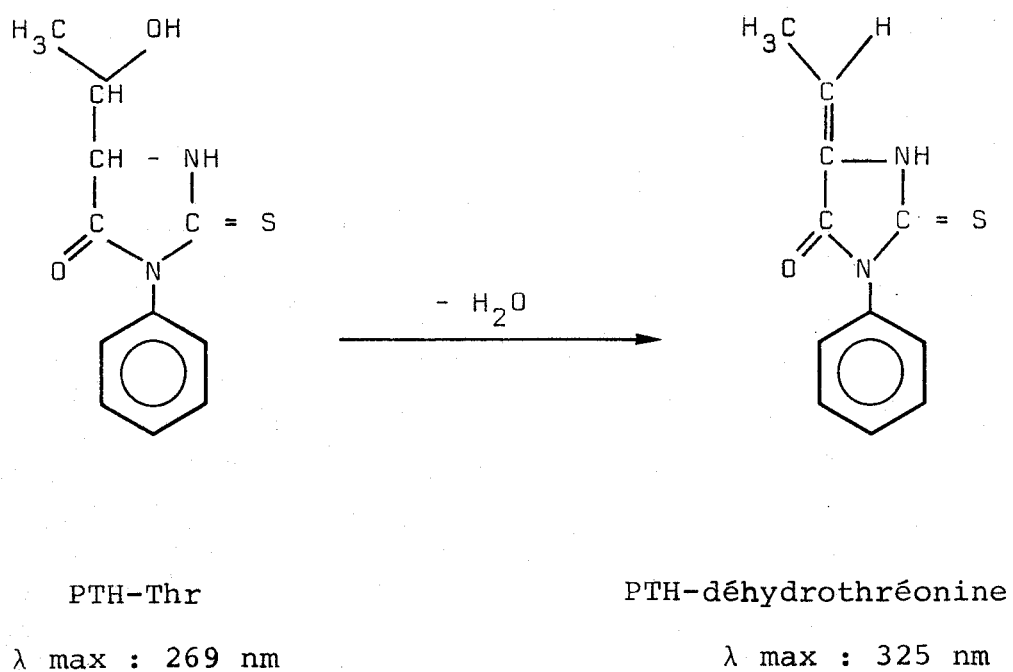


FIGURE 14 DECOMPOSITION DE LA PTH-THREONINE EN DEHYDROTHREONINE

Celle-ci est réalisée en aluminium. Deux lignes différentes arrivent dans cette coupelle. Dans un premier temps, le solvant ou le réactif arrive dans la coupelle puis les auteurs ferment l'arrivée de liquide. Les échanges entre le film et les réactifs s'effectuent. Dans un second temps, la rotation de la coupelle est arrêtée, le liquide descend dans le fond de celle-ci et monte dans la deuxième ligne grâce à une poussée d'azote. Les auteurs ont réussi grâce à ce nouveau système, à diminuer leur consommation de solvant par un facteur dix. Enfin, WITTMANN-LIEBOLD (65) a équipé son séquenceur d'une coupelle en verre à double paroi. La température est maintenue constamment à 55°C à l'intérieur de celle-ci grâce à une circulation d'eau. Cette coupelle comporte 3 sorties :

- une pour l'unique ligne de délivrance
- une pour l'azote
- une pour le scoop qui est en or. Le positionnement de ce scoop dans la gorge de la coupelle a également été rendu plus aisé.

Enfin, le système de chauffage Beckman du compartiment coupelle est utilisé comme chauffage additionnel.

II - 1-5 La conversion automatique

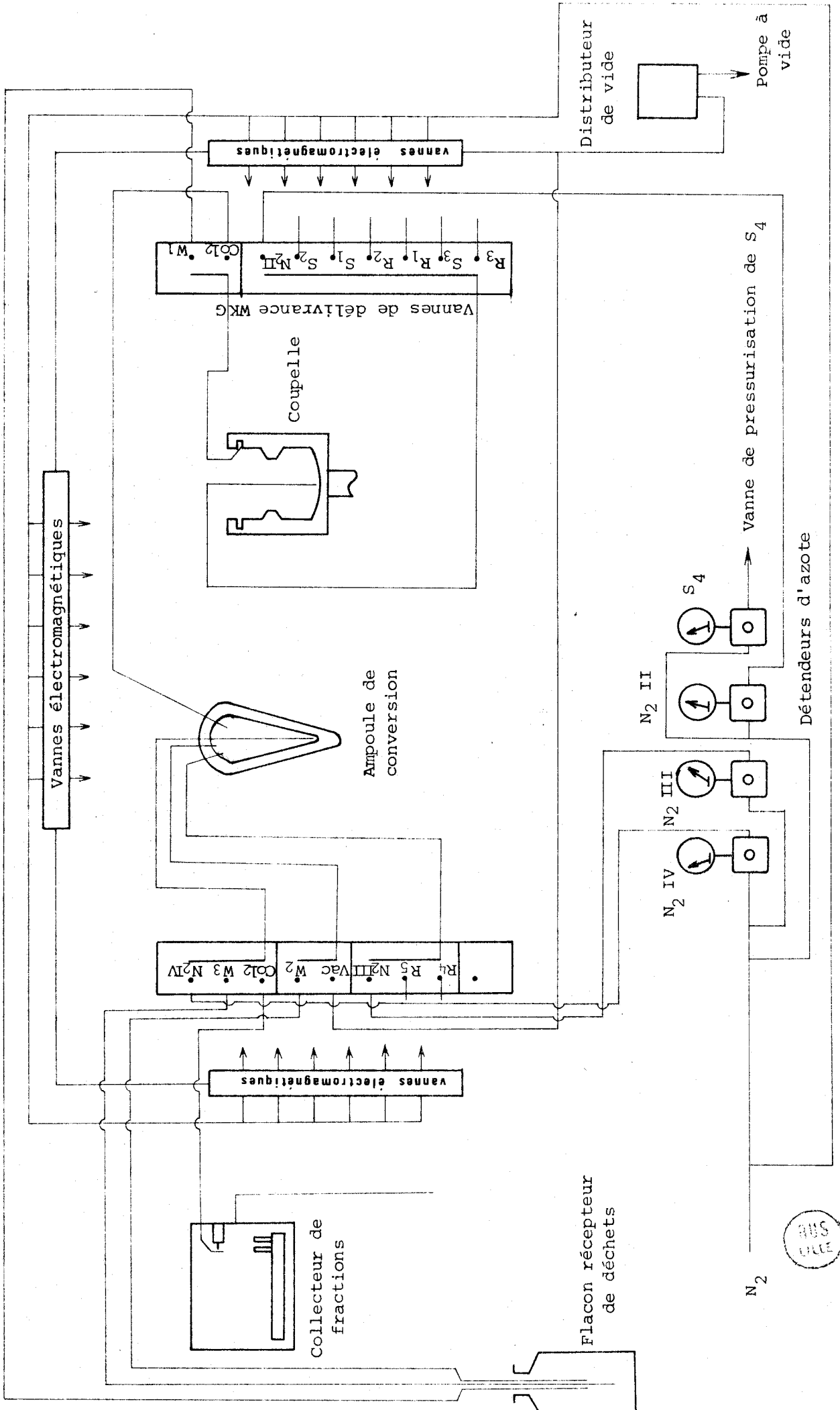
Outre les modifications des vannes de délivrance et du système de vide, WITTMANN-LIEBOLD et coll.(63) ont également proposé un système de conversion automatique pour l'équipement du séquenceur Beckman 890C. En effet, jusqu'à présent,

la conversion devait être réalisée manuellement. Or, Les ATZ-AA sont très sensibles à l'oxydation, ainsi qu'à l'exposition aux acides forts et aux acides dilués, conditions dans lesquelles elles se trouvent au moment du clivage par l'acide heptafluorobutyrique anhydre, et de la conversion par l'acide chlorhydrique 1N. Ceci conduit soit à une désamidation, soit à une destruction des ATZ-AA. Il est donc nécessaire que la conversion se fasse immédiatement après le clivage. Le système de conversion automatique proposé par WITTMANN-LIEBOLD et coll. est piloté par les fonctions 36 à 42 qui étaient inoccupées sur le programmeur Beckman. La conversion se fait dans une ampoule thermostatée à 55°C. Le réactif de conversion est l'acide trifluoroacétique à 20 p.100 dans l'eau et les PTH-AA formées sont extraites par un mélange 1,2 dichloroéthane/méthanol (7:3,v/v). Ce réactif et ce solvant sont délivrés par une vanne en téflon du même type que celles décrites précédemment. Grâce à cette conversion automatique, les auteurs ont constaté qu'il n'y avait plus de désamidation des PTH-asparagine et PTH-glutamine et que le rendement des autres PTH-acides aminés était accru.

Le système de conversion automatique que nous avons mis en place sur notre séquenceur est dérivé de celui décrit par WITTMANN-LIEBOLD et coll.(63). Son installation a nécessité la mise en place des pompes à vide à l'extérieur du séquenceur car le système de conversion est placé dans le compartiment initialement prévu pour les pompes.

Ce système de conversion est présenté dans la figure 15.

FIGURE 15 LE SYSTEME DE CONVERSION AUTOMATIQUE



La conversion se déroule dans une ampoule en verre thermostatée à 55°C. Cette ampoule comporte deux entrées et une sortie placées en périphérie et une sortie centrale:

- une entrée pour la thiazolinone
- une entrée commune au réactif de conversion (TFA à 20 p.100 dans l'eau distillée), au solvant d'extraction des PTH-AA (dichloroéthane/méthanol, 7:3,v/v), et à l'azote (ligne N₂III)
- une sortie commune pour le vide ou l'évacuation vers le flacon récepteur de déchets (ligne W₂)

La sortie centrale plonge jusqu'au fond de l'ampoule et est commune à l'extraction de la PTH-AA formée vers le collecteur de fractions, à l'azote (ligne N₂IV) et à l'évacuation vers le flacon récepteur de déchets (ligne W₃).

Une représentation schématique de cette ampoule est présentée dans la figure 16. Les étapes de conversion et de dégradation ont lieu simultanément (figure 17). A la fin d'un cycle de dégradation, la thiazolinone extraite par le chlorobutane, est envoyée dans l'ampoule de conversion. Elle est alors séchée complètement par un courant d'azote (lignes N₂IV et N₂III).

Le réactif de conversion est alors envoyé pendant 9 secondes, puis la conversion s'effectue sous bullage d'azote pendant 30 minutes. A la fin de ce temps, le TFA est presque évaporé entièrement. Le séchage de l'ampoule est alors poursuivi sous vide.

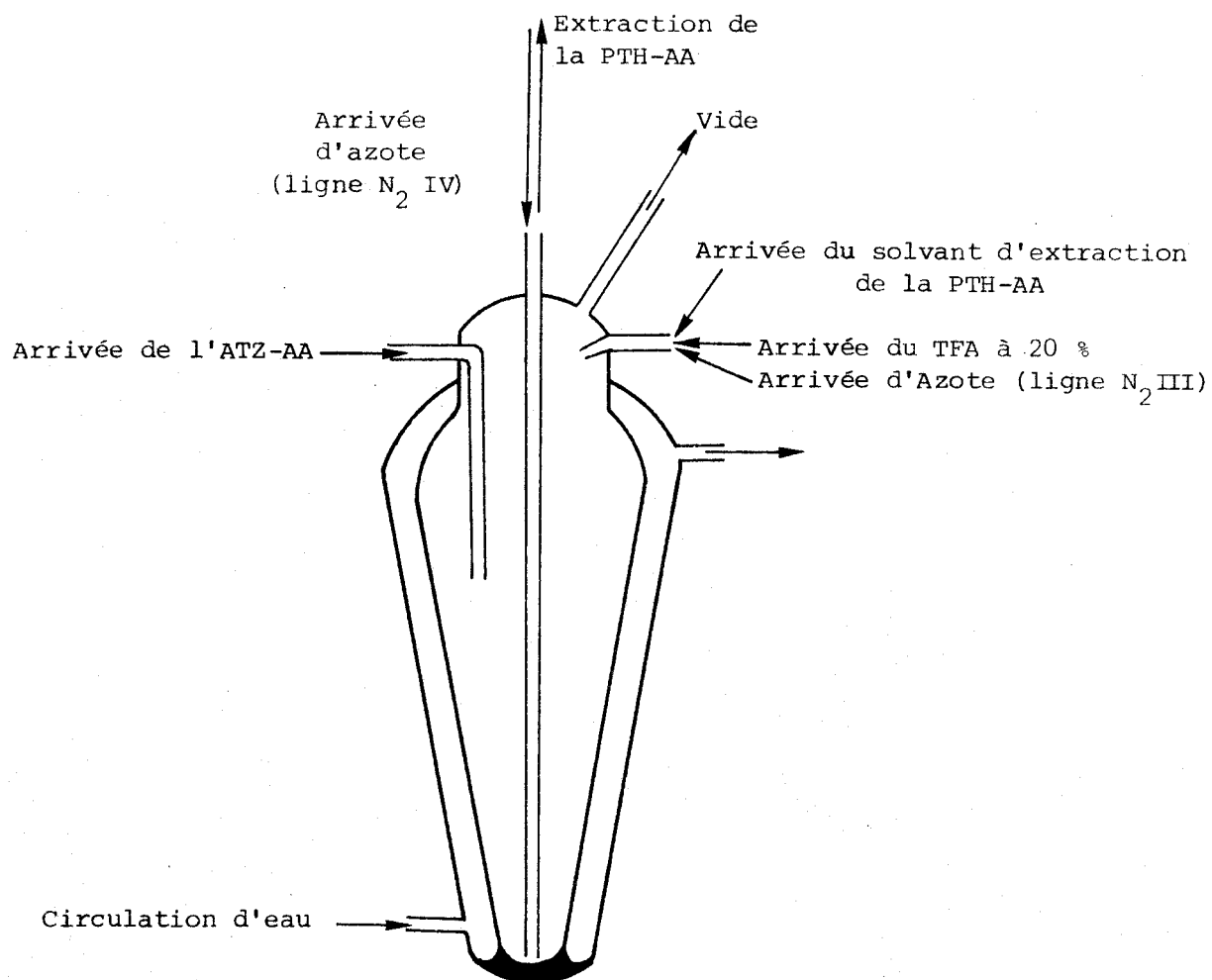


FIGURE 16 SCHEMA SIMPLIFIE DE L'AMPOULE DE CONVERSION AUTOMATIQUE (63)

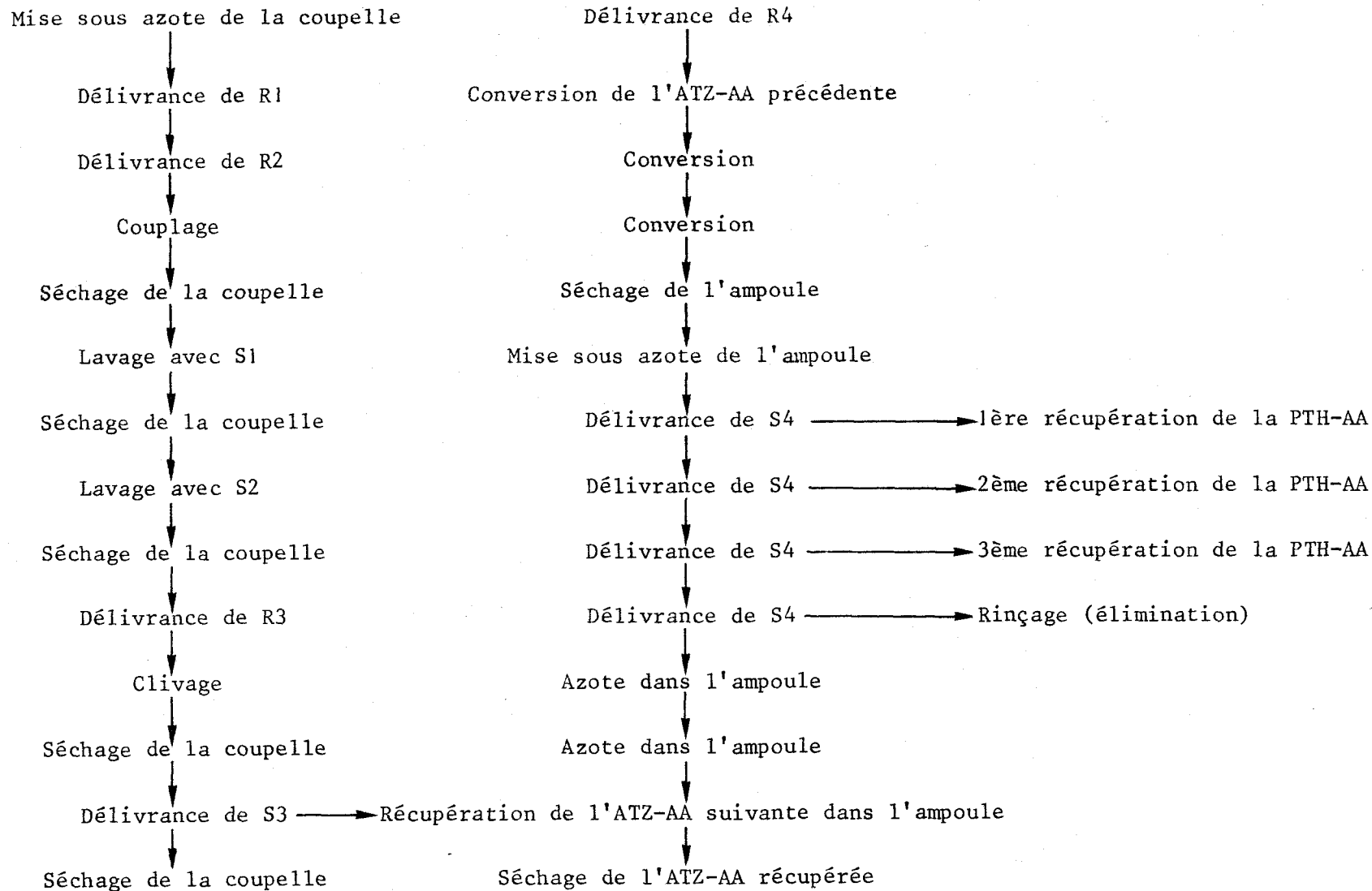


FIGURE 17 SCHEMA DE COORDINATION DES ETAPES DE DEGRADATION ET DE CONVERSION



Une fois l'ampoule parfaitement sèche, on la remet à la pression atmosphérique par un jet d'azote avant d'envoyer le solvant d'extraction. Le solvant d'extraction est projeté sur les parois de l'ampoule grâce à l'azote (ligne N₂III). Après un léger bullage d'azote, la PTH-AA est envoyée vers le collecteur de fractions par une poussée d'azote (ligne N₂III). Cette opération de récupération de la PTH-AA est effectuée quatre fois. Les trois premières extractions sont envoyées dans le collecteur de fractions, et la dernière, qui sert à nettoyer l'ampoule de réaction, est dirigée vers le flacon récepteur de déchets. Le réactif de conversion et les solvants d'extraction sont délivrés par des vannes en téflon à faible volume "mort" (vannes W.K.G.) dont une description a été faite précédemment. La commande du système de conversion automatique est assurée par les sept fonctions (36 à 42) qui restaient libres sur le programmeur, et une alimentation électrique* indépendante de celle du séquenceur Beckman.

II - 1-6 Autres modifications

Il s'agit d'une part de la pose d'un piège à oxygène réalisée par BHOWN et coll.(4). Ce piège (Oxyclear) est placé sur la ligne d'alimentation d'azote. Il a pour but d'éliminer les traces d'oxygène et d'eau contenues dans l'azote et qui

* Cette alimentation électrique a été installée avec l'aide de Gérard LECOQ (Ingénieur INSERM, Unité 124 INSERM) que nous remercions vivement.

diminuent le rendement de la dégradation. De ce fait, le piège est placé le plus près possible de la coupelle de réaction afin que le gaz qui arrive dans celle-ci ne se recontamine pas en cours de route.

D'autre part, HUNKAPILLER et HOOD (33) remplacent l'azote par un gaz plus lourd et d'une très grande pureté : *l'argon*.

II - 1-7 Détermination de la séquence N-terminale de la myo-
globine de Cheval après modification du séquenceur
Beckman 890C

La séquence N-terminale de la myoglobine de coeur de Cheval (200 nanomoles) a été déterminée après transformation du séquenceur Beckman, avec un programme quadrol 1M.

La comparaison des rendements des différentes PTH-AA identifiés avant et après transformation du séquenceur est présentée dans le tableau II. On peut remarquer que les rendements individuels des différentes PTH-AA sont assez nettement supérieurs après les modifications techniques apportées au séquenceur. En particulier, on n'observe aucune désamidation de la PTH-glutamine et de la PTH-asparagine, alors qu'avant la transformation de l'appareil, nous avons environ 50 p.100 de désamidation de la PTH-glutamine et 25 p.100 de désamidation de la PTH-asparagine. Cependant, un problème s'est posé au niveau de l'identification en HPLC de la PTH-sérine et de la PTH-thréonine.

TABLEAU II

COMPARAISON DES RENDEMENTS OBTENUS LORS DE LA DETERMINATION DES 15 PREMIERS RESIDUS DE LA MYOGLOBINE DE COEUR DE CHEVAL
AVANT ET APRES TRANSFORMATION DU SEQUENCEUR BECKMAN

PTH-AA déterminée	Gly	Leu	Ser	Asp	Gly	Glu	Trp	Gln	Gln	Val	Leu	Asn	Val	Trp	Gly
Position dans la séquence	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Rendements obtenus avant transformation (en pourcentage)	48	59,7	18,6	23	34,8	16,2	22,4	8,9 5,4 (Glu)	10,8 4,5 (Glu)	32,6	30,7	13,2 3,1 (Asp)	26,6	15,9	15,8
Rendements obtenus après transformation (en pourcentage)	25	65,8	n.d.	54,4	22,6	63,8	43,2	43,2	44,2	33,4	34,4	33,4	31,8	20,2	23,6



En effet, nous avons observé une destruction pratiquement totale de ces deux dérivés d'acides aminés, destruction due à l'utilisation de la conversion automatique.

La destruction de la PTH-sérine a pu être partiellement évitée en ajoutant à tous nos solvants ainsi qu'aux réactifs de clivage et de conversion, du dithiothréitol (DTT) à raison de 10 mg/litre (54). De ce fait, la PTH-sérine est identifiée à 254 nm sous la forme de deux pics de "PTH-sérine-DTT" comme le montre la figure 18. Cette PTH-AA modifiée n'étant pas présente dans notre standard, sa quantification ne peut être effectuée et ne figure donc pas dans notre tableau de comparaison.

L'effet de la conversion automatique sur la PTH-thréonine résulte en une transformation totale de la PTH-thréonine en PTH-déhydrothréonine dont l'identification n'est possible qu'à 313 nm (figure 19).

On peut enfin remarquer dans le tableau II que le rendement de la PTH-glycine est beaucoup plus faible après transformation du séquenceur. Ceci est dû à une modification partielle de la PTH-glycine, liée à la fois à l'addition de dithiothréitol aux solvants et aux réactifs de clivage et de conversion, et à la conversion automatique. La PTH-glycine ainsi modifiée est identifiée sous la forme d'un pic supplémentaire, non quantifié, élué entre PTH-tyrosine et PTH-glutamique méthylester comme le montre la figure 20.

Tampon A : Acétate de sodium 10 mM pH 4,27/Méthanol
(95:5) (v/v)
Tampon B : Méthanol/acétate de sodium 10 mM pH 4,27
(95:5) (v/v)
Sensibilité : 0,1 AUFS
Débit : 2,2 ml/min

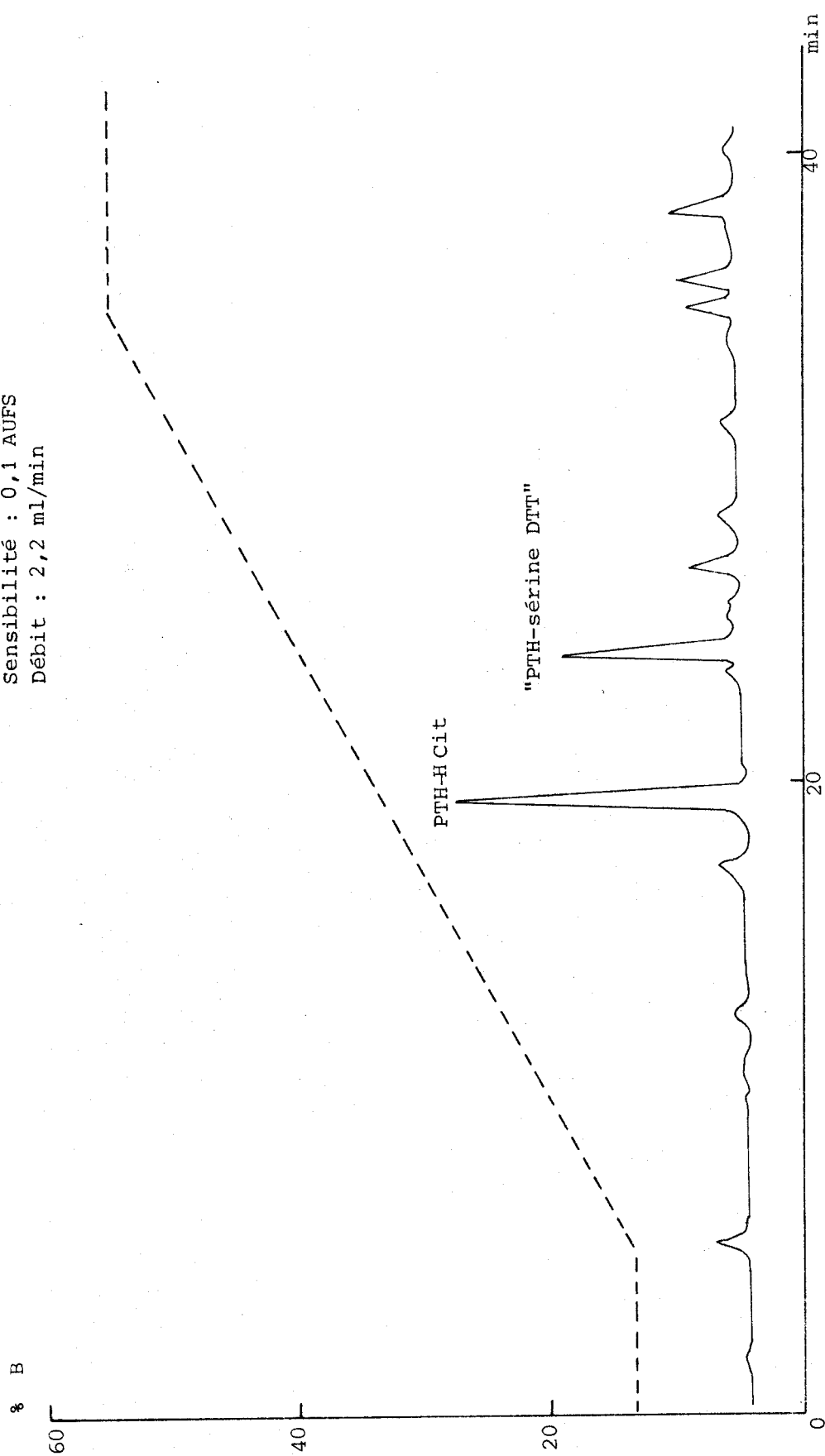


FIGURE 18 IDENTIFICATION DE LA "PTH-SÉRINE DTT" A 254 nm



Tampon A : Acétate de sodium 10 mM pH 4,27/
Méthanol (95:5) (v/v)

Tampon B : Méthanol/acétate de sodium 10 mM
pH 4,27 (95:5) (v/v)

Sensibilité : 0,1 AUFS

Débit : 2,2 ml/min

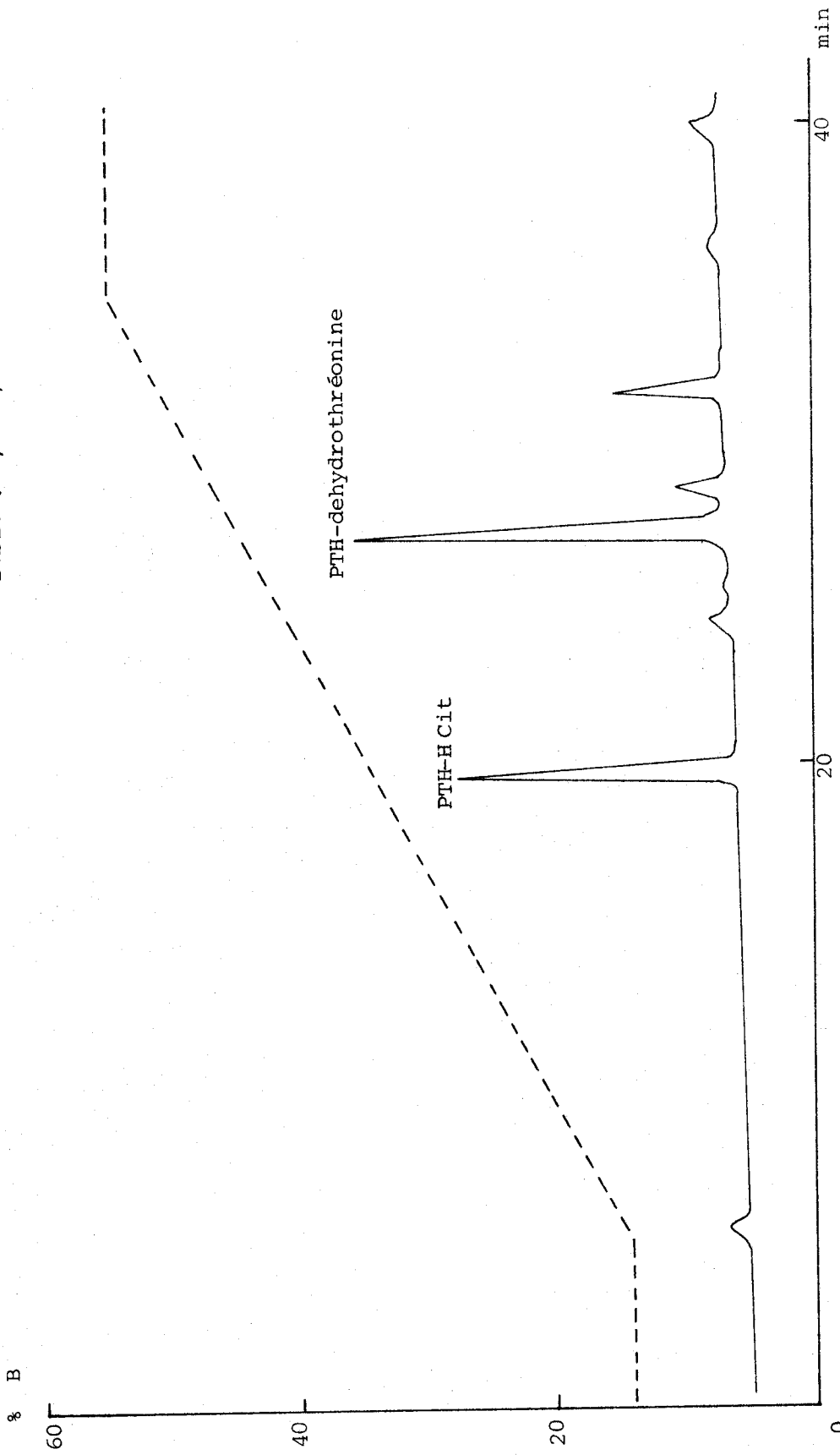


FIGURE 19 IDENTIFICATION DE LA PTH-DEHYDROTHREONINE A 313 nm





PTH-Gly modifiée

Tampon A : Acétate de sodium 10 mM pH 4,27/
Méthanol (95:5) (v/v)

Tampon B : Méthanol/acétate de sodium 10 mM
pH 4,27 (95:5) (v/v)

Sensibilité : 0,1 AUFS

Débit : 2,2 ml/min

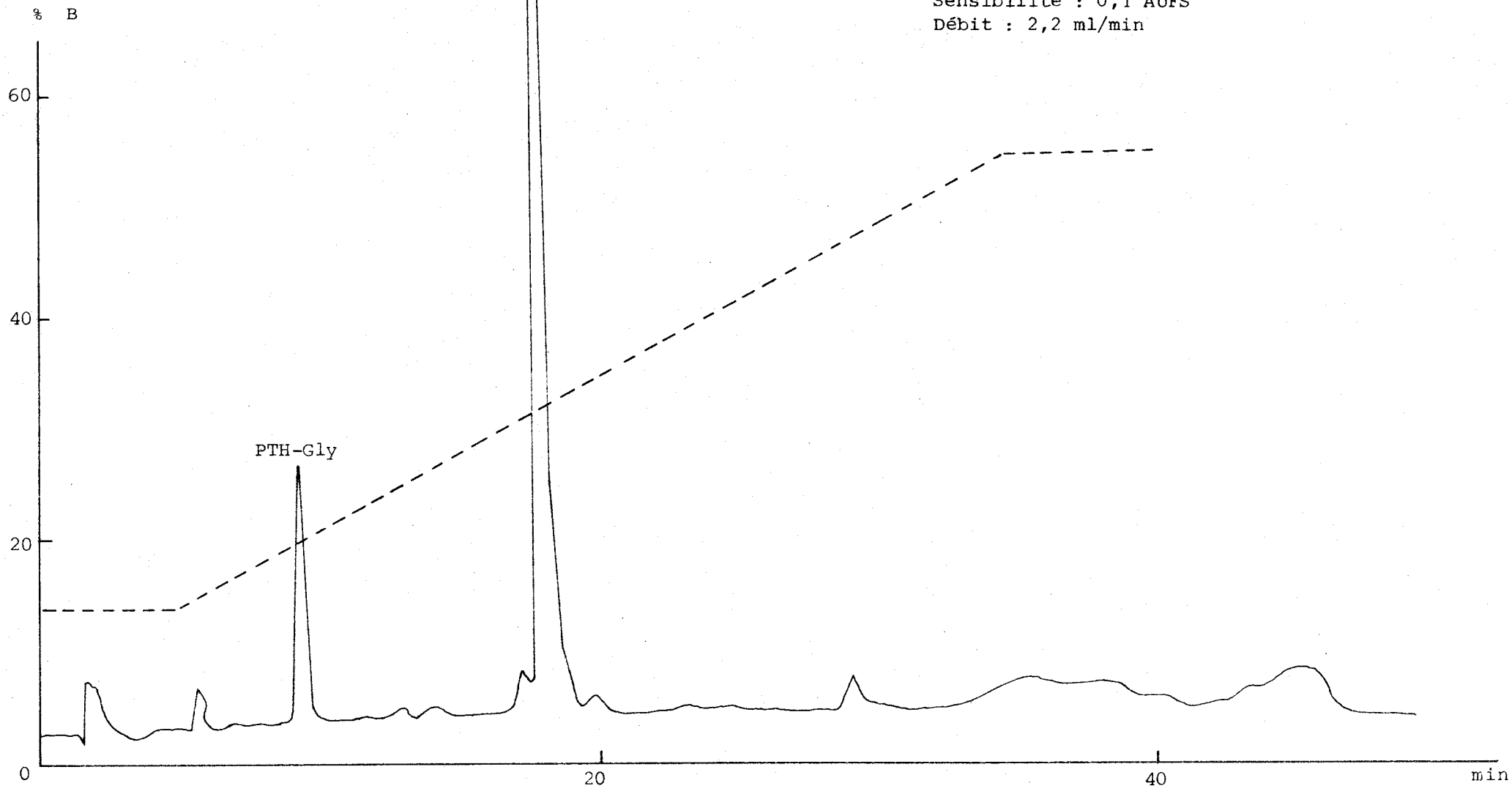


FIGURE 20

IDENTIFICATION DE LA PTH-GLYCINE SUR COLONNE μ BONDAPAK C₁₈

II - 2 TRAITEMENT DES REACTIFS ET DES SOLVANTS

Ce traitement est réalisé selon la méthode de FRANK (25). La purification du quadrol est réalisée par absorption des impuretés sur aminoéthyl-cellulose. 1,5 g d' aminoéthyl-cellulose Serva sont lavés deux fois avec 25 ml d'une solution 2M de trifluoroacétate de sodium dans l'acide trifluoroacétique 0.1M. L' aminoéthyl-cellulose est ensuite chauffée pendant 15 minutes à 50°C dans 25 ml de solution de lavage, puis elle est filtrée sur un verre fritté de porosité 1. Le filtrat obtenu est brun. La procédure est répétée trois fois pour que l' aminoéthyl-cellulose devienne blanche, puis on lave à l'eau distillée jusqu'à pH neutre. L' aminoéthyl-cellulose est alors séchée par aspiration à la trompe à eau puis introduite dans 100 ml de tampon de couplage (quadrol 0.33M dans un mélange n-propanol/eau, 3:4, v/v, ajusté à pH 9). L'ensemble est alors mis en place dans le séquenceur, la ligne de délivrance du tampon de couplage étant emprisonnée dans un filtre de porosité 1. Quant aux solvants (benzène et acétate d'éthyle), ils sont purifiés par conservation sur alumine Woelm W 200 neutre super I (20 g d'alumine/litre de solvant). Les lignes de délivrance des solvants sont emprisonnées dans des verres frittés de porosité 2. Une comparaison de la séquence d'un hexadécapeptide (résidus 69-84) de l'histone H4 de thymus de Veau, obtenu par coupure de la protéine* au niveau des résidus d'acide aspartique, a été

* Ce peptide nous a été aimablement donné par Maurice COUPPEZ.

effectuée avant et après traitement du quadrol et des solvants de lavage. Une quantité de peptide voisine de 50 nanomoles a été séquencée en programme quadrol en présence de polybrène. La comparaison des rendements des différentes PTH-AA déterminées est présentée dans le tableau III.

On peut remarquer que les rendements obtenus avec le programme quadrol traité sont supérieurs à ceux obtenus avec le quadrol non traité, notamment à partir du 6ème cycle où le rendement est deux fois plus important alors qu'au départ les rendements sont voisins. De plus, en utilisant le quadrol traité, nous avons déterminé la séquence complète du peptide alors que précédemment nous n'avions pu déterminer que les 13 premiers résidus.

II - 3 MICROSEQUENCE EN PHASE SOLIDE, LIQUIDE, GAZEUSE

Cette technique d'application récente et dont le principe a été exposé dans le chapitre I, a été expérimentée grâce à l'aimable collaboration du Docteur J. Van BEEUMEN de l'Université de Gand.

Pour cette expérimentation nous avons utilisé six nanomoles d'un hexadécapeptide (peptide 69-84) de l'histone H4 de thymus de Veau obtenu par coupure de la protéine au niveau des résidus d'acide aspartique. L'appareil utilisé est fabriqué aux Etats-Unis par la firme Applied Biosystem et commercialisé en

TABLEAU III

COMPARAISON DES RENDEMENTS OBTENUS LORS DE LA DETERMINATION DE LA SEQUENCE DU PEPTIDE DE L'HISTONE H4 DE THYMUS DE VEAU
 (RESIDUS 69 à 84) EFFECTUES EN PROGRAMME QUADROL 0.33M AVANT ET APRES TRAITEMENT DES REACTIFS ET DES SOLVANTS

PTH-AA déterminée	Ala	Val	Thr	Tyr	Thr	Glu	His	Ala	Lys	Arg	Lys	Thr	Val	Thr	Ala	Met
Position dans la séquence	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Rendements obtenus avant traitement (en pourcentage)	85,7	64,1	44,3	55,5	32,5	20,5	n.d	24,1	20,4	13,7	14,6	8,04	9,6			
Rendements obtenus après traitement (en pourcentage)	88,5	69	50,9	61,9	40,8	44,2	n.d	52,7	37,7	27,4	26,1	22,7	35,6	15,9	12,7	2,7



France par la Société SEMSA. La méthodologie employée pour la séquence en phase "gazeuse" a été transposée de la méthodologie employée pour la séquence en phase liquide.

Le peptide a été soumis à la dégradation d'EDMAN automatisée en présence de polybrène avec un programme dont le tampon de couplage est une solution aqueuse de triméthylamine à 25 p.100. Avant l'introduction du peptide dans le séquenceur, trois cycles de dégradation sur le dipeptide glycine-glycine ont été réalisés dans le but de nettoyer le mainteneur.

Les PTH-AA obtenues après conversion automatique ont été identifiées sur colonne CN avec une double détection réalisée simultanément à 269 et 313 nm. Les quantités de PTH-AA injectées étaient comprises entre 0,3 et 2 nanomoles calculées par rapport à la quantité de peptide soumise à la dégradation.

La quantification de ces PTH-AA a été effectuée par rapport à un mélange standard de PTH-AA de 62 picomoles à une sensibilité de 0.01 AUFS.

Les rendements que nous avons obtenu sont présentés dans le tableau IV.

Le rendement répétitif calculé à partir des résidus d'alanine en position 1 et 8 est de 87 p.100. Bien que ce rendement ne soit pas très élevé, le peptide a pu être séquencé entièrement et les diagrammes d'identification des PTH-AA en HPLC ne présentaient aucune ambiguïté.

Cette technique de dégradation automatique en phase "gazeuse" requiert environ 10 fois moins de peptide que la dégradation en phase liquide. Nous rappelons en effet qu'au cours d'un

TABLEAU IV

RENDEMENTS OBTENUS LORS DE LA DEGRADATION D'EDMAN AUTOMATISEE DU
PEPTIDE 69-84 (6 NANOMOLES) DE L'HISTONE H4 DE THYMUS DE VEAU
PAR LA TECHNIQUE DE MICROSEQUENCE EN PHASE "GAZEUSE"

N° Résidu	Résidu	Rendements en nanomoles		
		R-1(1)	R(2)	R+1(3)
1	Ala		3,2	
2	Val		3,7	
3	Thr		1,2	
4	Tyr		2,4	
5	Thr		0,9	
6	Glu		1,3	
7	His		n.d(4)	n.d
8	Ala		1,2	0,4
9	Lys		1,9	0,3
10	Arg		1,4	0,2
11	Lys	0,3	0,8	0,2
12	Thr		0,2	0,05
13	Val		0,2	0,08
14	Thr	0,05	0,1	
15	Ala	0,03	0,1	0,04
16	Met		n.d	

(2) R = rendement en nanomoles de la PTH-AA obtenue au cours du cycle de dégradation.

(1) R-1 = rendement de cette PTH-AA au cycle précédent.

(3) R+1 = rendement de cette PTH-AA au cycle suivant.

(4) n.d = identifiée mais non dosée.



essai précédent (paragraphe II - 2, p.78), la séquence complète de l'hexadécapeptide avait été déterminée par dégradation d'EDMAN automatisée en phase liquide à partir de 50 nanomoles de peptide. De plus, les quantités de réactifs et de solvants nécessaires, sont de beaucoup inférieures à celles utilisées en phase liquide. Le tableau V nous montre une comparaison de consommation de réactifs et de solvants pour un cycle de dégradation, effectuée par HEWICK et coll.(31).

Cette microméthode de détermination de la structure primaire des protéines (ou des peptides) semble donc très intéressante et supplantera, sans doute, dans un avenir proche, les méthodes en phase liquide et solide qui demandent beaucoup plus de matériel de départ.

TABLEAU V

COMPARAISON DES VOLUMES DE REACTIFS ET DE SOLVANTS CONSOMMES POUR UN CYCLE DE DEGRADATION D'EDMAN

AVEC LE SEQUENCEUR EN PHASE LIQUIDE ET LE SEQUENCEUR EN PHASE "GAZEUSE" (31)

Volume en ml de réactifs et de solvants utilisés			
SEQUENCEUR EN PHASE LIQUIDE		SEQUENCEUR EN PHASE GAZEUSE	
R1 : PITC à 5 p.100 dans le n-heptane	0.4	PITC à 15 p.100 dans le n-heptane	0.05
R2 : Tampon quadrol 1M pH 9.0	0.7	Tampon triméthylamine à 25 p.100 dans l'eau	5 ml/min. de débit d'argon
R3 : HFBA, 0.01 p.100 de DTT	0.5	TFA, 0.01 p.100 de DTT	5 ml/min. de débit d'argon
R4 : TFA à 25 p.100 dans l'eau, 0.01 p.100 de DTT	0.3	TFA à 25 p.100 dans l'eau, 0.01 p.100 de DTT	0.05
S1 : Benzène	7	Benzène	1.1
S2 : Acétate d'éthyle, 0.05 p.100 d'acide acétique, 0.002 p.100 de DTT	16	Acétate d'éthyle, 0.05 p.100 d'acide acétique, 0.002 p.100 de DTT	2.7
S3 : Chlorobutane, 0.001 p.100 de DTT	7	Chlorobutane, 0.001 p.100 de DTT	1.2
S4 : Acétonitrile, 0.001 p.100 de DTT	4	Acétonitrile, 0.001 p.100 de DTT	0.3



CHAPITRE III

LES MÉTHODES D'IDENTIFICATION DES PTH-AA

PLAN DU CHAPITRE III - LES METHODES D'IDENTIFICATION DES PTH-AA

III - 1 LA CHROMATOGRAPHIE SUR COUCHE MINCE

III - 2 LA CHROMATOGRAPHIE EN PHASE GAZEUSE

III - 3 LA REGENERATION DES ACIDES AMINES A PARTIR DES PTH-AA

III - 4 LA SPECTROMETRIE DE MASSE

III - 5 LA CHROMATOGRAPHIE LIQUIDE A HAUTE PERFORMANCE

III - 5-1 La chromatographie solide-liquide

III - 5-2 La chromatographie liquide-liquide

III - 1 LA CHROMATOGRAPHIE SUR COUCHE MINCE

La chromatographie sur couche mince fut une méthode d'identification très populaire jusqu'à l'avènement de la chromatographie liquide à haute performance.

Les supports utilisés étaient le gel de silice ou le polyamide.

COHEN-SOLAL et BERNARD (14) ont décrit une méthode d'identification sur plaques de gel de silice, de 6,3cm x 6,3cm, amidonnées. Une première identification est réalisée par chromatographie ascendante jusqu'aux 2/3 de la hauteur de la plaque, avec le solvant I (chloroforme-méthanol, 1:1, v/v), suivie d'une migration jusqu'en haut de la plaque avec le solvant II (chloroforme pur). La révélation des PTH-AA s'effectue à l'aide d'un réactif iode-azide de sodium : dans ce cas les PTH-AA apparaissent en blanc sur fond ocre, ou d'un réactif à la ninhydrine-collidine (52) et dans ce cas les PTH-AA apparaissent sous forme de taches diversement colorées.

Les PTH-AA insuffisamment ou non séparées dans ce système-solvant, peuvent être séparées dans d'autres systèmes-solvants.

KULBE (39) utilise pour sa part, une méthode de chromatographie bidimensionnelle sur plaque de polyamide de 5cm x 5cm. La première dimension est réalisée dans un mélange toluène/n-pentane/acide acétique (60:30:16, v/v), et la seconde dans l'acide acétique à 25 p.100. La détection des PTH-AA se fait en lumière ultraviolette à 254 nm.

Ces techniques d'identification en couche mince ont l'avantage d'être très rapides, peu coûteuses, mais elles restent limitées par l'absence de dosage des PTH-AA détectées.

III - 2 LA CHROMATOGRAPHIE EN PHASE GAZEUSE

Cette technique est appliquée pour la séparation de produits volatilisables. Les PTH-AA n'ont pas toutes le même pouvoir de volatilité. On peut les classer en trois catégories :

- Les PTH-AA ayant une bonne volatilité :

alanine, glycine, valine, proline, isoleucine, leucine, méthionine, phénylalanine et tyrosine.

- Les PTH-AA ayant une volatilité moyenne :

tryptophane, asparagine, glutamine, histidine, thréonine, sérine, S-carboxyméthylcystéine, lysine.

- Les PTH-AA à volatilité nulle :

acide aspartique, acide glutamique, acide cystéique et arginine.

Pour identifier les PTH-AA des deux dernières catégories, il est nécessaire de les transformer en dérivés volatils par couplage avec le N.O.Bis triméthylsilylacétamide (BSA). Cependant, malgré cette transformation les PTH-lysine et arginine sont très difficilement détectables.

Les supports et les phases stationnaires les plus utilisés pour cette chromatographie en phase gazeuse, sont ceux décrits par PISANO et BRONZERT (50). La meilleure séparation est obtenue sur chromosorb WHP 100/120 mesh imprégné de chlorophénylsilicone (SP 400 à 10 p.100). Les colonnes sont en verre car les PTH-AA sont partiellement détruites à chaud au contact des colonnes métalliques.

Cependant, bien que cette méthode soit relativement rapide et assez sensible, elle présente quelques inconvénients :

- Transformation incomplète des PTH-AA en dérivés volatils.
- Décomposition thermique de PTH-thréonine, PTH-sérine, PTH-tryptophane et PTH-tyrosine.
- Désamidation partielle de PTH-glutamine et PTH-asparagine.

De ce fait, la quantification des PTH-AA ne peut être réellement faite que sur les PTH-AA volatiles classées dans la première catégorie.

III - 3 LA REGENERATION DES ACIDES AMINES A PARTIR DES PTH-AA

La régénération des acides aminés à partir des PTH-AA se fait généralement par hydrolyse acide soit par HCl 5.6N à 130-150°C pendant 16 à 24 heures (59), soit par HI 5N à 140°C pendant 24 heures (56).

Cette régénération par hydrolyse acide transforme les PTH-sérine et PTH-S-carboxyméthylcystéine en alanine, la PTH-thréonine en acide α -aminobutyrique, et les PTH-tryptophane et méthionine sont détruites. AFRICA et CARPENTER (1) préfèrent utiliser l'hydrolyse alcaline par la soude 0.1N à 120°C pendant 12 heures. Cette méthode de régénération transforme la PTH-thréonine en glycine et la PTH-arginine en ornithine. La PTH-tryptophane est partiellement détruite. Quant aux PTH-sérine et PTH-acide cystéique, elles sont totalement détruites.

On remarque donc que cette méthode n'est pas idéale pour identifier les PTH-AA. De plus, bien que cette méthode soit quantifiable, elle n'est jamais utilisée comme technique principale d'une part à cause de son manque de rapidité et d'autre part à cause du fait que la régénération ne se fait jamais à 100 p.100 et que les rendements obtenus varient d'une méthode d'hydrolyse à l'autre.

III - 4 LA SPECTROMÉTRIE DE MASSE

La spectrométrie de masse est souvent utilisée pour déterminer les phénylthiohydantoïnes d'acides aminés particuliers. Elle est employée soit couplée à la chromatographie en phase gazeuse, soit seule.

Dans le cas de la spectrométrie de masse couplée à la chromatographie en phase gazeuse, les échantillons sont introduits dans le chromatographe après avoir été modifiés chimiquement de manière à être volatils. Le diagramme des PTH-AA apparaît sur l'enregistreur et lorsqu'on est au sommet d'un pic, une simple pression sur une touche du tiroir de commande du spectromètre de masse, déclenche le balayage des masses et l'inscription du spectre de masse sur un enregistreur. Cette technique, qui ne peut être utilisée que si les PTH-AA sont modifiées de manière à être volatils, est trop longue.

Certains auteurs (61) préfèrent aujourd'hui utiliser directement le spectromètre de masse. Ils pratiquent pour cela un bombardement ionique rapide. Dans ce cas, il n'est plus nécessaire de modifier chimiquement les échantillons au préalable.

Cependant, comme la spectrométrie de masse nécessite un investissement très élevé, elle n'est utilisée que pour résoudre des problèmes structuraux bien précis, qui ne peuvent être résolus par des méthodes chimiques ou des méthodes physiques (spectroscopie de RMN, spectroscopie infrarouge).

III - 5 LA CHROMATOGRAPHIE LIQUIDE A HAUTE PERFORMANCE (HPLC)

Il y a deux types de chromatographie liquide à haute performance qui sont adaptés à la séparation des PTH-AA.

Ce sont : - la chromatographie solide-liquide
- la chromatographie liquide-liquide

III - 5-1 La chromatographie solide-liquide

Dans ce type de chromatographie, la phase stationnaire est constituée de silice ou encore d'alumine. Deux types de colonnes de silice commercialisées sous les noms de Spherisorb S 5W (Phase Separations Ltd), et de Lichrosorb Si 60 5 μm (Merck) sont généralement utilisés pour séparer les PTH-AA (23).

La colonne de Spherisorb S 5W est éluée par un mélange dichlorométhane contenant 4 p.100 d'éthanol, 4 p.100 de diméthylsulfoxyde et 0.35 p.100 d'acide formique à 20 p.100.

Elle sépare les PTH-AA polaires qui sont : PTH-acide aspartique, PTH-asparagine, PTH-glycine, PTH-acide glutamique, PTH-glutamine, PTH-sérine, PTH-thréonine et PTH-tyrosine.

La colonne de Lichrosorb Si 60 est éluée par un mélange de dichlorométhane contenant 0.8 p.100 d'isopropanol et 175 ppm de diméthyle sulfoxyde.

Elle sépare les PTH-AA apolaires qui sont : PTH-phénylalanine, PTH-leucine, PTH-isoleucine, PTH-valine, PTH-proline, PTH-alanine et PTH-tryptophane.

Les résultats obtenus de cette façon sont bons mais la méthode est longue car il est nécessaire d'avoir systématiquement recours à deux chromatographies pour séparer tous les PTH-AA.

III - 5-2 La chromatographie liquide-liquide

Cette méthode est beaucoup plus employée que la précédente. Les colonnes sont constituées d'un support de silice sur lequel sont greffés des groupements de polarité différente. Ils sont soit peu polaires : groupe octadécyle- $(\text{CH}_2)_{17}\text{-CH}_3$ et groupe phényle, soit de polarité intermédiaire : groupe cyano -CN .

La phase mobile utilisée étant une phase polaire, on réalise de cette façon une chromatographie de phase inverse. Cette phase mobile est souvent constituée d'acétate d'ammonium ou de sodium, et de méthanol ou d'acétonitrile. Certains auteurs préconisent l'utilisation de mélanges ternaires avec une association de deux solvants organiques :

- acétonitrile-méthanol
- acétonitrile-tétrahydrofurane.

Cependant, vu le coût élevé et la toxicité de l'acétonitrile, beaucoup d'entre eux utilisent le méthanol.

De nombreuses méthodes de séparation des PTH-AA par chromatographie liquide-liquide ont été décrites dans la littérature.

ZIMERMAN et coll. (66) séparent les PTH-AA sur une colonne Zorbax ODS (type C₁₈) à 50°C avec un gradient linéaire d'acétate de sodium (0.01N) et d'acétonitrile.

BHOWN et coll. (4) préfèrent employer une colonne Altex Ultrasphère ODS (type C₁₈) à température ambiante avec un gradient linéaire d'acétate de sodium 0.04M pH 3.72 et de méthanol.

D'autres auteurs comme FOHLMAN et coll. (21) séparent les PTH-AA sur colonne micro-Bondapak C₁₈ (Waters) à 37°C avec un gradient concave d'acétate d'ammonium pH 5.1 et d'éthanol.

Citons encore la méthode proposée par HENDERSON et coll. (28) qui utilisent une colonne micro-Bondapak phényle (Waters) avec un gradient linéaire d'acide propionique à 0.85 p.100 pH 3.80 et de méthanol à 90 p.100 ; ainsi que les travaux de JOHNSON et coll. (36) qui emploient une colonne micro-Bondapak cyano (Dupont Zorbax) à 31°C avec un gradient linéaire d'acétate de sodium 0.024M pH 5.40 et de méthanol/acétonitrile (17:3, v/v).

Nous citerons enfin la technique utilisée par HUNKAPILLER et HOOD (34) qui séparent les PTH-AA sur une colonne cyano (HPLC analytical grade, IBM Instrument) avec un gradient linéaire d'acétate de sodium 0.015M à pH 5.8 et de méthanol/acétonitrile (1:4, v/v).

Cependant, aucune de ces méthodes ne permet de séparer toutes les PTH-AA. C'est pourquoi, nous avons été amenée à utiliser en routine deux types de supports : une colonne de micro-Bondapak C₁₈ (Waters) et une colonne de micro-Bondapak phényle (Waters). Ainsi, les PTH-AA qui ne sont pas séparées sur le premier support, sont parfaitement résolues sur le second.

Dans un premier temps, les PTH-AA sont donc identifiées sur la colonne micro-Bondapak C₁₈ (Waters) en milieu acétate de sodium 10 mM pH 4.27 avec un gradient linéaire de 10 à 48 p.100 de méthanol, dérivé de celui décrit par HERMANN et coll. (29). Cette séparation qui est présentée dans la figure 21 est réalisée à température ambiante en 25 minutes. La PTH-homocitrulline est utilisée comme témoin interne pour la quantification. La PTH-valine et la PTH-méthionine, de même que la PTH-phénylalanine et la PTH-isoleucine ne sont pas séparées dans ce système.

Dans un second temps, les PTH-AA sont séparées sur une colonne micro-Bondapak phényle (Waters). La séparation que nous obtenons est présentée dans la figure 22. Elle est réalisée en 42 minutes en milieu acétate de sodium 10 mM pH 4.27 avec un gradient linéaire de 10 à 55 p.100 de méthanol. Grâce à cette colonne, les PTH-méthionine et PTH-valine ainsi que les PTH-phénylalanine et PTH-isoleucine sont parfaitement séparées alors qu'elles ne le sont pas sur la colonne de C₁₈.

Tampon A : Acétate de sodium 10 mM
 pH 4,27/Méthanol (95:5) (v/v)
 Tampon B : Méthanol/acétate de sodium
 10 mM pH 4,27 (95:5) (v/v)
 Quantité injectée : 10 nanomoles
 Débit : 2,2 ml/min

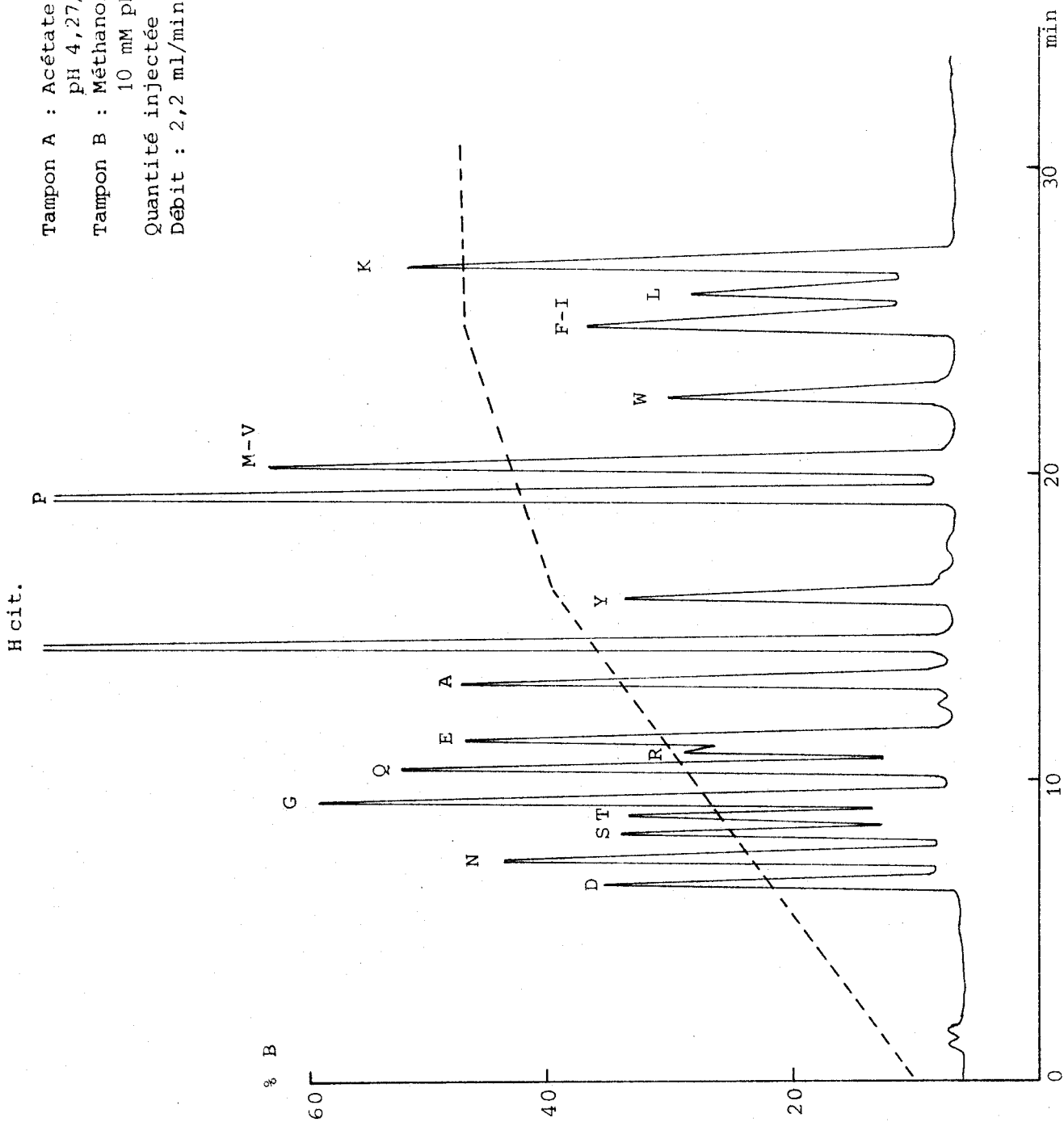


FIGURE 21 SEPARATION DES PTH-AA SUR COLONNE μBONDAPAK C₁₈



Tampon A : Acétate de sodium 10 mM pH 4,27/
méthanol (95:5) (v/v)

Tampon B : Méthanol/acétate de sodium 10 mM
pH 4,27 (95:5) (v/v)

Quantité injectée : 10 nanomoles
Débit : 2,2 ml/min

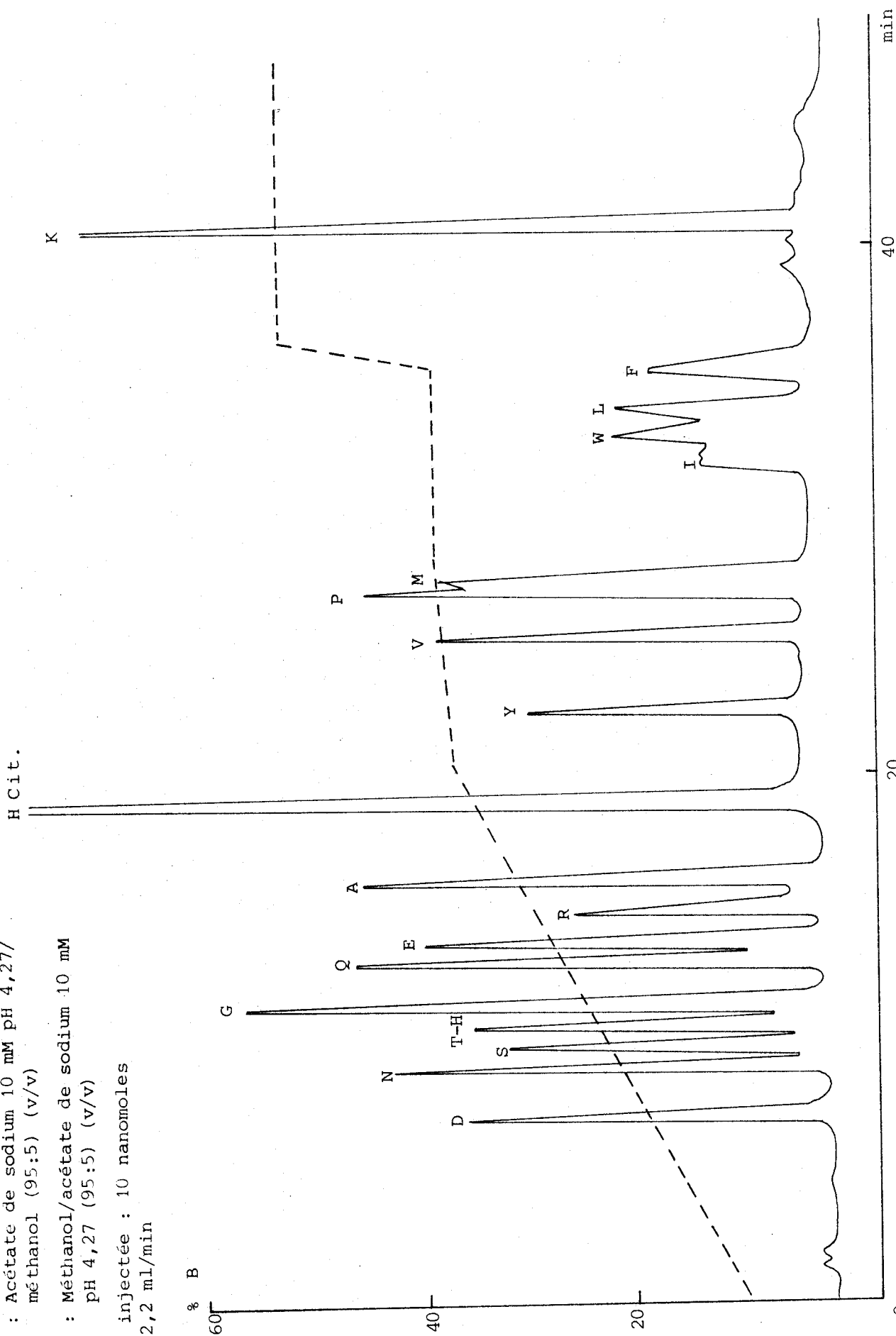


FIGURE 22 SEPARATION DES PTH-AA SUR COLONNE μ BONDAPAK PHENYLE

Pour l'identification de la PTH-histidine, nous avons été amenée à utiliser un troisième type de support. Il s'agit de la colonne micro-Bondapak cyano (Waters). La séparation que nous obtenons est présentée dans la figure 23. Elle est effectuée en 25 minutes, en milieu acétate de sodium 3 mM pH 5.20 avec un gradient linéaire de 2 à 65 p.100 de méthanol. Comme nous le remarquerons, cette colonne qui donne une très bonne résolution de la PTH-histidine, permet également de séparer les PTH-méthionine et PTH-valine ainsi que les PTH-phénylalanine et PTH-isoleucine.

Cependant, lorsque le tampon de couplage utilisé pour la dégradation contient de la diméthylallylamine le recours à la chromatographie en phase gazeuse est nécessaire pour identifier de façon certaine la PTH-valine car la DMAA donne un artefact qui, en HPLC, est élué avec la PTH-valine.

Au cours des différents essais de séparation des PTH-AA, nous avons été amenée à utiliser une colonne Altex Ultrasphère ODS commercialisée par la Société Beckman. Un diagramme de séparation des PTH-AA obtenu sur cette colonne est présenté dans la figure 24. L'analyse est effectuée en 60 minutes en milieu acétate de sodium 10 mM, pH 4.27 avec un gradient linéaire de 14 à 50 p.100 de méthanol. Comme nous pouvons le remarquer, cette colonne donne une très bonne résolution pour les PTH-AA et notamment pour PTH-méthionine



Tampon A : Acétate de sodium 3 mM
pH 5,2

Tampon B : Méthanol/acétate de sodium
3 mM pH 5,2 (85:15) (v/v)

Quantité injectée : 10 nanomoles
Débit : 1,8 ml/min

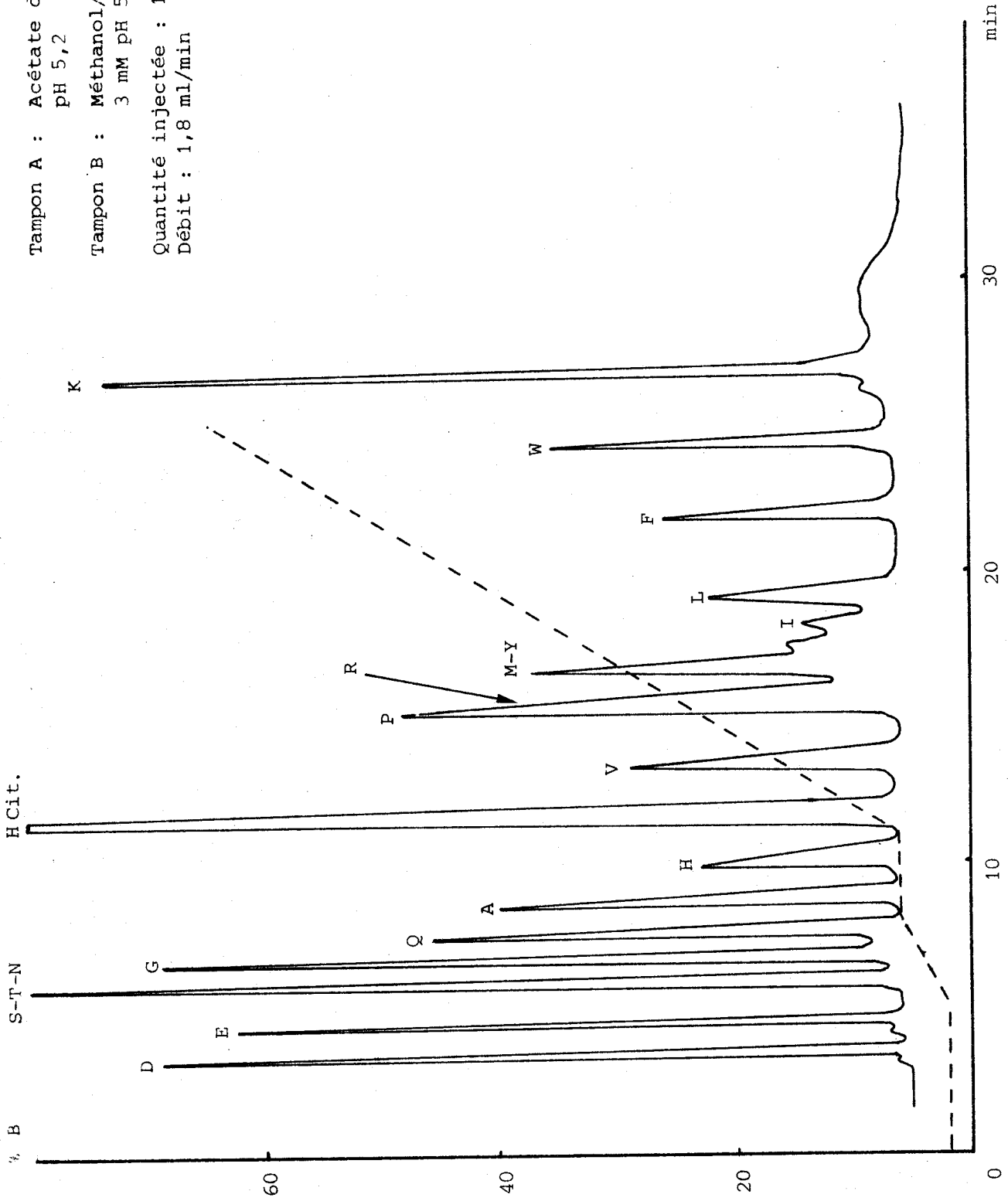


FIGURE 23 SEPARATION DES PTH-AA SUR COLONNE μBONDAPAK CN



Tampon A : Acétate de sodium 10 mM pH 4,27/méthanol (95:5) (v/v)
Tampon B : Méthanol/acétate de sodium 10 mM pH 4,27 (95:5) (v/v)
Quantité injectée : 2,5 nanomoles
Débit : 1 ml/min

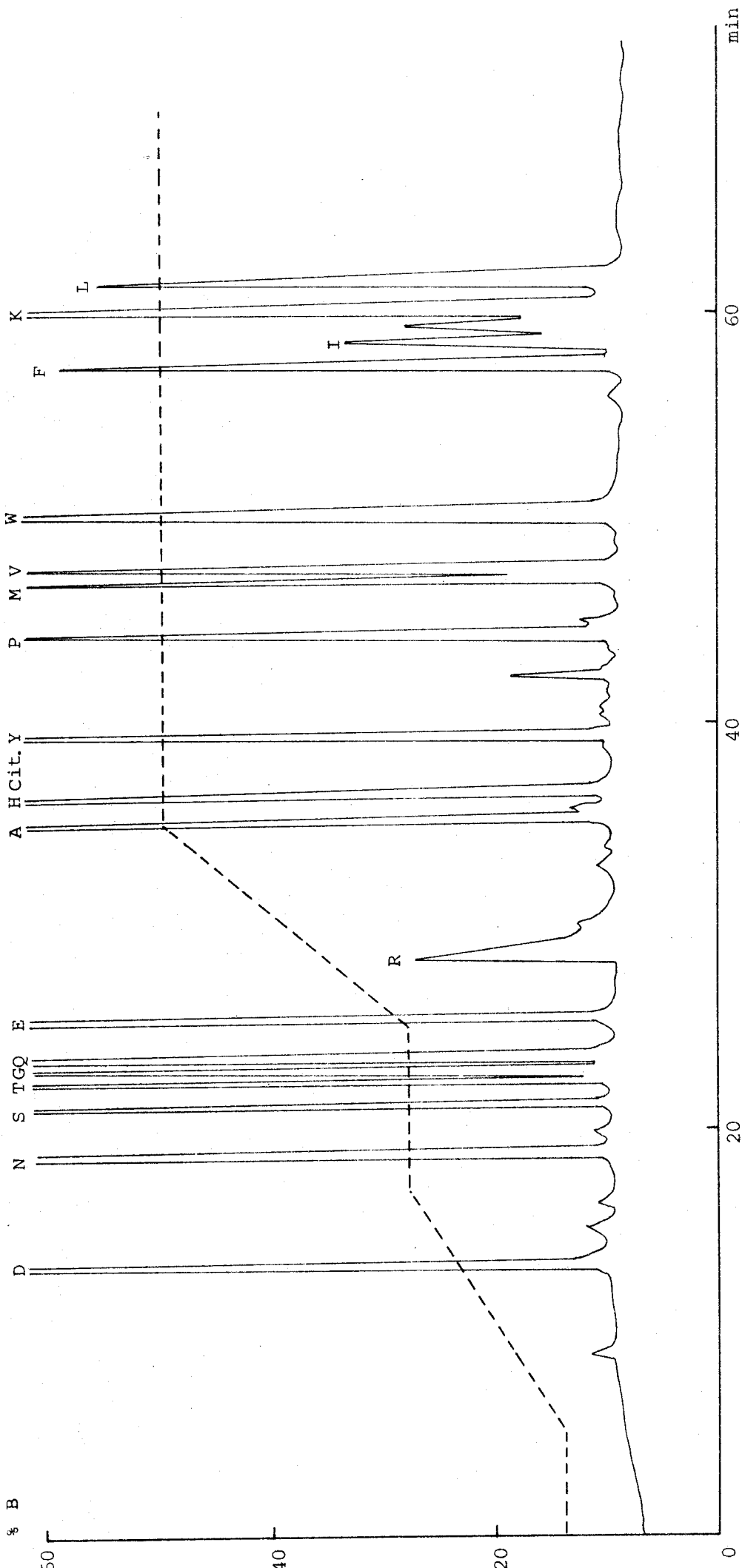


FIGURE 24 SEPARATION DES PTH-AA SUR COLONNE ALTEX ULTRASPHERE ODS

et PTH-valine ainsi que pour PTH-phénylalanine et PTH-isoleucine (PTH-isoleucine étant sous ses deux formes PTH-alloisoleucine et PTH-isoleucine) mais, les PTH-arginine et PTH-histidine sont très mal résolues.

Ces PTH-arginine et PTH-histidine ont en plus un comportement assez particulier et ceci sur n'importe quelle colonne. Elles sont en effet très sensibles au pH et à la molarité des tampons. Il est possible par exemple, en modifiant un de ces paramètres de les éluer très rapidement ou très tardivement. De plus, le vieillissement des colonnes entraîne des modifications sur le temps de rétention de ces PTH-AA. En effet, au cours du temps, les radicaux C₁₈, cyano et phényl greffés sur la silice sont partiellement libérés, et de ce fait, au phénomène de séparation par hydrophobicité, se superpose un phénomène d'échange d'ions avec le support de silice. On pallie cependant à cet inconvénient en augmentant ou en diminuant la molarité du tampon.

Au cours de toutes les déterminations de séquence que nous avons effectuées grâce au séquenceur, nous avons également été confrontée aux problèmes de l'identification des phénylthiohydantoïnes d'acides aminés modifiées, ces acides aminés pouvant être modifiés par des réactifs chimiques. Les résidus de cystéine par exemple sont alkylés après réduction

ou transformés en résidus d'acide cystéique. Cette modification a pour but de cliver les ponts disulfures et de dérouler la chaîne polypeptidique afin de faciliter la détermination de la séquence.

Les principales PTH des dérivés de la cystéine qui sont présentées dans la figure 25 sont obtenues :

- par oxydation performique qui modifie la cystéine en acide cystéique mais également la méthionine en méthionine sulfone.
- par l'acide iodoacétique qui donne la S-carboxyméthylcystéine.
- par l'iodoacétamide qui transforme la cystéine en S-carboxyamidocystéine.
- par l'éthylène imine : on obtient alors l'aminoéthylcystéine.
- par le méthylparanitrobenzène sulfonate qui donne la méthylcystéine.
- par la 4-vinyl-pyridine qui donne la pyridyléthylcystéine.

Sur cette figure 25, nous avons également fait apparaître les PTH des dérivés méthylesters des acides aminés dicarboxyliques. Ceux-ci sont obtenus au cours de la conversion des ATZ-AA si le milieu de conversion est l'HCl/méthanol, ou encore au cours du temps si les PTH-AA ont été reprises par le méthanol après une conversion à l'HCl/eau.

FIGURE 25

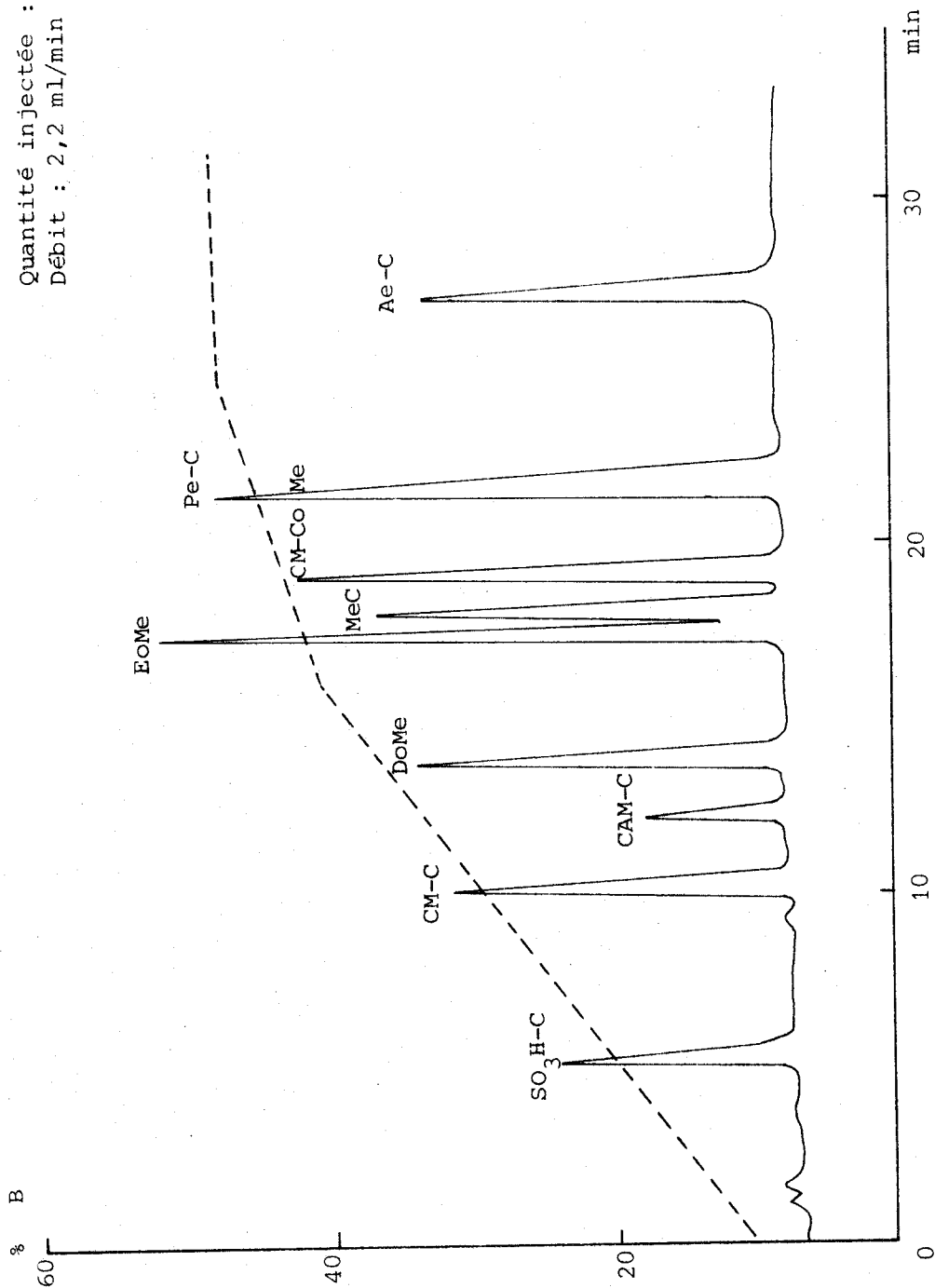
SEPARATION DES DIFFERENTES PTH-AA DES DERIVES
DE LA CYSTEINE ET DES PTH-AA DES DERIVES METHYL-
ESTER DES ACIDES DICARBOXYLIQUES SUR COLONNE
 μ BONDAPAK C₁₈

SO₃-C : PTH acide cystéique
CM-C : PTH S-carboxyméthyl cystéine
CAM-C : PTH S-carboxyamidométhyl cystéine
DoMe : PTH aspartique méthylester
EoMe : PTH glutamique méthylester
MeC : PTH méthyl cystéine
CM-CoMe : PTH carboxyméthyl cystéine méthylester
Pe-C : PTH pyridylethyl cystéine
Ae-C : PTH aminoéthyl cystéine

Tampon A : Acétate de sodium 10 mM pH 4,27/
méthanol (95:5) (v/v)

Tampon B : Méthanol/acétate de sodium 10 mM
pH 4,27 (95:5) (v/v)

Quantité injectée : 10 nanomoles
Débit : 2,2 ml/min



Ces PTH des dérivés d'acides dicarboxyliques sont :

- PTH-aspartique méthylester
- PTH-glutamique méthylester
- PTH-carboxyméthylcystéine méthylester

Une autre modification que nous avons également rencontrée est celle de la lysine transformée en ϵ -N-succinyllysine. Cette modification irréversible est effectuée dans le but de bloquer le résidu N-terminal ou de limiter les coupures par la trypsine dans la chaîne polypeptidique, aux seules liaisons arginyles. Cette modification s'effectue par traitement de la protéine ou du peptide avec l'anhydride succinique. La PTH- ϵ -N-succinyllysine, dont l'identification en HPLC sur colonne micro-Bondapak C₁₈ (Waters) est présentée dans la figure 26, apparaît sous la forme de trois pics caractéristiques.

Nous avons également été amenée à séparer un certain nombre de PTH-AA modifiées résultant de la modification post-synthétique des protéines étudiées.

C'est le cas notamment de certains résidus de lysine que l'on rencontre en particulier dans les histones H3 et H4 sous forme de ϵ -N-méthyllysine, ϵ -N-diméthyllysine et ϵ -N-acétyllysine.

Un diagramme de séparation des PTH-AA correspondantes est présenté dans la figure 27.

Tampon A : Acétate de sodium 10 mM pH 4,27/
méthanol (95:5) (v/v)

Tampon B : Méthanol/acétate de sodium 10 mM
pH 4,27 (95:5) (v/v)

Débit : 2,2 ml/min



FIGURE 26 IDENTIFICATION DE LA PTH-SUCCINYLLYSINE SUR COLONNE μ BONDAPAK C18

Tampon A : Acétate de sodium 10 mM pH 4,27/méthanol
(95:5) (v/v)

Tampon B : Méthanol/acétate de sodium 10 mM pH 4,27
(95:5) (v/v)

Débit : 2,2 ml/min

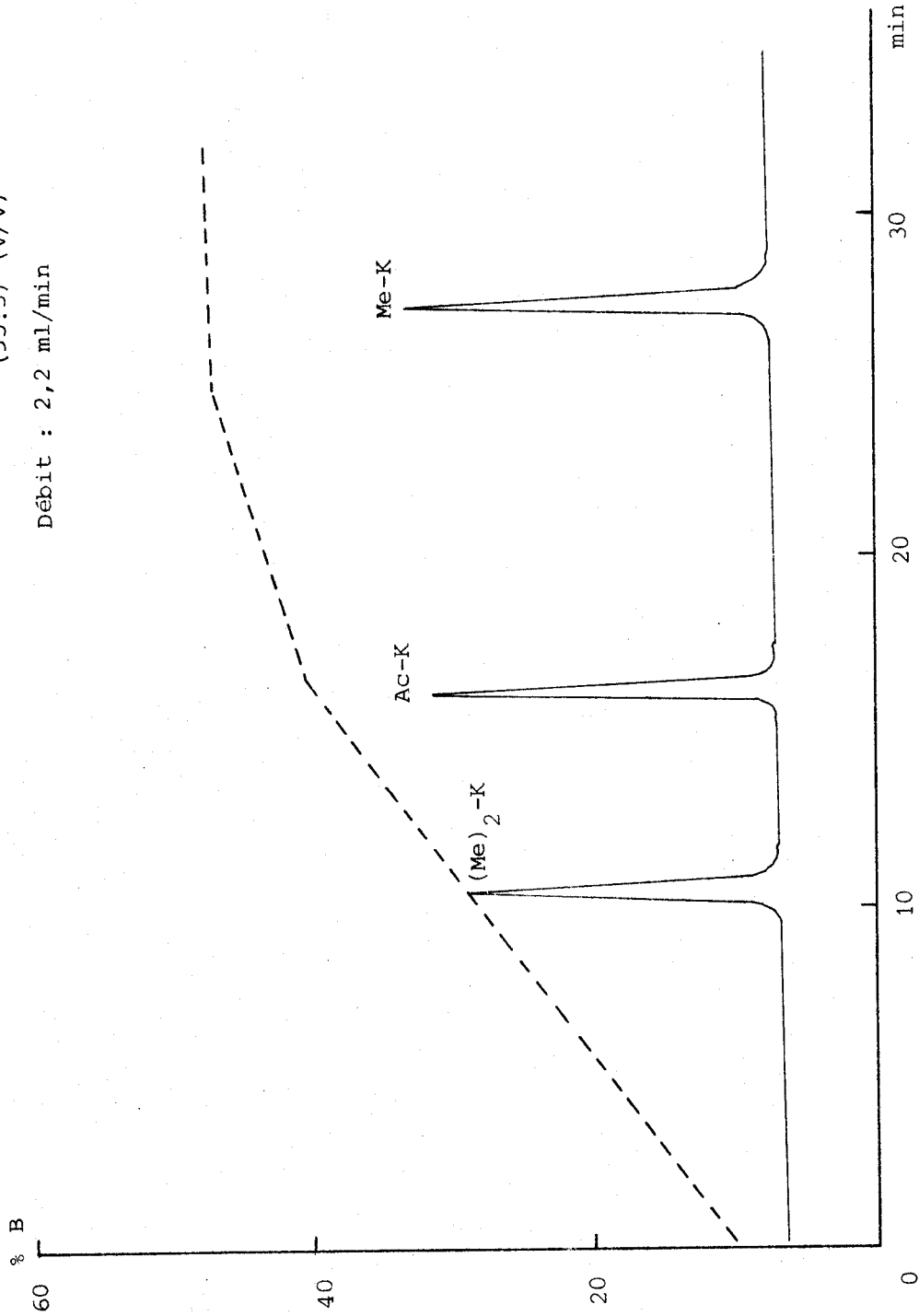


FIGURE 27 SEPARATION DES PTH-AA DES DERIVES METHYLES ET ACETYLES DE LA
LYSINE SUR COLONNE μBONDAPAK C18

Les résidus de sérine, thréonine et tyrosine peuvent quant à eux être phosphorylés sur leur groupe -OH. Une séparation des PTH-phosphosérine, PTH-phosphothréonine et PTH-phosphotyrosine est présentée dans la figure 28.

Enfin, nous avons également rencontré un résidu de tyrosine modifié en tyrosine sulfate lors de la détermination de la séquence de la protamine Z3 de Roussette.

L'identification de la PTH-AA correspondante n'a pu être réalisée que grâce à la spectrométrie de masse et à l'analyse en spectroscopie infrarouge.

Tampon A : Acétate de sodium 10 mM pH 4,27/méthanol
(95:5) (v/v)

Tampon B : Méthanol/acétate de sodium 10 mM pH 4,27
(95:5) (v/v)

Débit : 2,2 ml/min

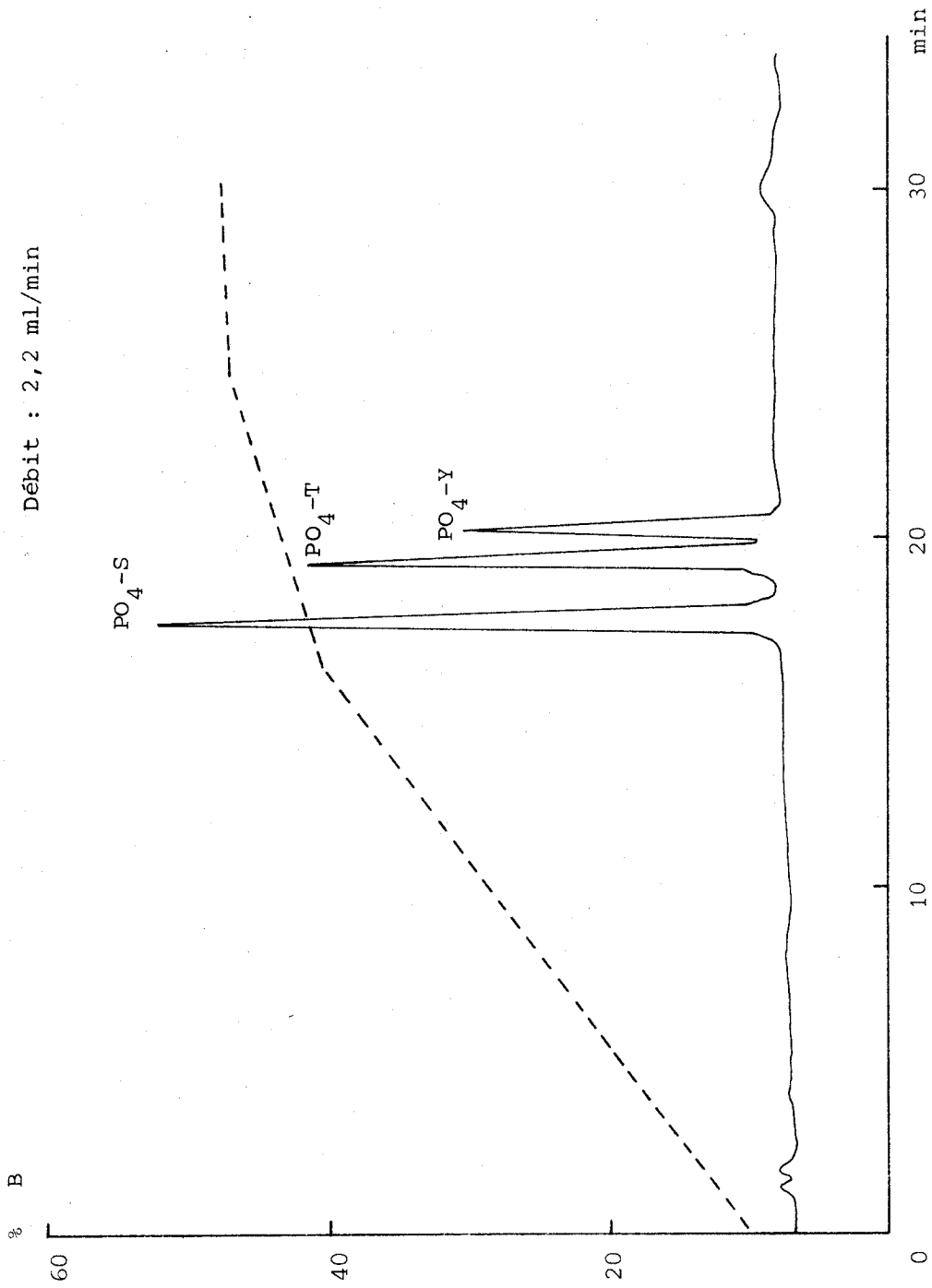


FIGURE 28 IDENTIFICATION DES PTH-AA DES DERIVES PHOSPHORYLES DE LA SERINE, THREONINE ET TYROSINE SUR COLONNE μ BONDAPAK C18

CHAPITRE IV

CONTRIBUTION A LA DETERMINATION DE LA STRUCTURE PRIMAIRE
DE DIFFERENTES PROTEINES

PLAN DU CHAPITRE IV

CONTRIBUTION A LA DETERMINATION DE LA STRUCTURE PRIMAIRE DE
DIFFERENTES PROTEINESIV - 1 DETERMINATION DE LA STRUCTURE PRIMAIRE DE LAPROTEINE HR_m DE *Rhizobium meliloti*IV - 2 DETERMINATION DE LA STRUCTURE PRIMAIRE D'HISTONES H2AISOLEES D'INVERTEBRES MARINSIV - 3 DETERMINATION DE LA STRUCTURE PRIMAIRE DE PROTAMINESDE MAMMIFERES ET DE POISSONS

IV - 3-1 Protamines de Mammifères

IV - 3-2 Protamines de Poissons. Structure primaire
de la Scylliorhinine 4 de sperme de Roussette
(*Scylliorhinus caniculus*)

Les travaux concernant la détermination de la structure primaire de la protéine HRm de *Rhizobium meliloti* ont fait l'objet des publications suivantes :

Publication 1

Characterization and structural study of the DNA-binding protein HRm from *Rhizobium meliloti*.

LAINE B., BELAÏCHE D., SAUTIERE P. and BISERTE G.

Biochem.Biophys.Res.Commun. (1982) 106, 101-107.

Publication 2

Primary structure of the DNA-binding protein HRm from *Rhizobium meliloti*.

LAINE B., BELAÏCHE D., KHANAKA H. and SAUTIERE P.

Eur.J.Biochem. (1983) 131, 325-331.

Primary Structure of the DNA-Binding Protein HRm from *Rhizobium meliloti*

Bernard LAINE, Denise BÉLAÏCHE, Hussain KHANAKA, and Pierre SAUTIÈRE

Unité 124 de l'Institut National de la Santé et de la Recherche Médicale, Lille; and
Laboratoire de Microbiologie Université des Sciences et Techniques de Lille, Villeneuve d'Ascq

(Received October 5, 1982) – EJB 6072

The amino acid sequence of protein HRm, a DNA-binding HU-type protein of 90 residues (M_r 9303), isolated from *Rhizobium meliloti*, has been established from automated sequence analysis of the protein and from structural data provided by peptides derived from cleavage of the protein at arginine and aspartic acid residues.

The comparison of the primary structure of protein HRm with that of other HU-type proteins shows that two short sequences, of 7 and 6 residues respectively, located in the median part of the molecule, appear highly conserved and may be important in the function of the protein.

Several authors have suggested that low-molecular-weight basic DNA-binding proteins isolated from *Escherichia coli* [1–4] could play a role in the condensation of the chromosomal DNA which is packaged in a compact structure [5] and has been visualized by electron microscopy to be condensed in a repeating beaded structure [6]. The most extensively studied of these proteins, the protein HU [1], is a hetero-complex constituted of two polypeptide chains termed HU-1 and HU-2, of 90 residues each [7–9]. Closely related proteins have been isolated from different species widely remote in the bacterial kingdom including Eubacteria (*Salmonella typhimurium* [10], *Pseudomonas aeruginosa* [11], *Bacillus subtilis* [12] and *Bacillus globigii* [13]), Cyanobacteria (*Anabaena* and *Synechocystis* [14]) and an Archaeobacteria (*Thermoplasma acidophilum* [15]). In support of the hypothesis of a role of the HU-type proteins in the condensation of DNA are the findings obtained essentially with the protein HU from *E. coli* and the protein HTa from *T. acidophilum*. The proteins HU and HTa have been shown to be associated with the chromosomal DNA [2, 10, 15] and to stabilize DNA against thermal denaturation [16, 17]. Protein HU can form nucleosome-like structures with circular double-stranded DNA which is condensed by a ratio of 2.4 relative to naked DNA [18]. Protein HTa condenses linear double-stranded DNA into globular particles [19]. The proteins HU-1 and HU-2 [8, 9] and the protein HTa [20] have been completely sequenced. The partial sequences of the HU-type proteins from *P. aeruginosa* [11] and from the Cyanobacterium *Synechocystis* [21] have also been reported.

A DNA-binding HU-type protein has been recently characterized in *Rhizobium meliloti* [22]. This protein termed HRm is a single polypeptide chain and exhibits lower affinity for double-stranded DNA than does *E. coli* protein HU.

In this paper we present the complete amino acid sequence of protein HRm established from automated sequence analysis of the protein and from structural data provided by peptides

derived from cleavage of the protein at arginine and aspartic acid residues.

MATERIALS AND METHODS

Materials

Rhizobium meliloti, strain 2011 Str 3 (streptomycin-resistant mutant) was provided by J. Denarie (Institut National de la Recherche Agronomique, Versailles, France) and was grown as described in [23]. The protein HRm was isolated and purified as in [22].

Carboxypeptidase A (treated with PhMeSO₂F) was obtained from Worthington. The protease from submaxillary gland of mouse (endoproteinase Arg-C) was from Boehringer.

Trifluoroacetic acid, dimethyl allylamine, propan-1-ol and dithioerythritol were sequanal grade from Pierce. Benzene and chloro-1-butane for sequential analysis were purchased from S.D.S. (Peypin, France). Phenylisothiocyanate, heptane and heptafluorobutyric acid were sequanal grade from Beckman. Polybrene was purchased from Aldrich and the dipeptide Gly-Gly from Sigma. Acetonitrile for HPLC was from Carlo Erba. All other reagents were of the highest purity available.

Structural Studies

Carboxy-Terminal Analysis. Protein HRm (80 nmol) dissolved in 0.8 ml of 0.2 M *N*-methyl-morpholine acetate, pH 8.0 was digested at 37°C with carboxypeptidase A (15 µg, 0.7 unit). At 10 min and 30 min, aliquots were taken off, acidified with glacial acetic acid and freeze-dried. The released amino acids were analysed on the amino acid analyser.

Cleavage at Arginyl Residues. Protein HRm (650 nmol) was digested with the protease from submaxillary gland of mouse in 0.1 M sodium bicarbonate (2 ml) pH 8.0 at 37°C for 4 h, using an enzyme-to-substrate ratio of 1:50 (w/w). The digestion was stopped by lowering the pH at 3.0 with glacial acetic acid.

Cleavage at Aspartyl Residues. Protein HRm (650 nmol) was hydrolysed with 1 ml of 0.25 M acetic acid in an evacuated sealed tube for 15 h at 105°C.

Abbreviations. HRm, DNA-binding HU-type protein isolated from *Rhizobium meliloti*; PhMeSO₂F, phenylmethylsulfonyl fluoride.

Enzymes. Carboxypeptidase A (EC 3.4.17.1); endoproteinase Arg-C (EC 3.4.–.).

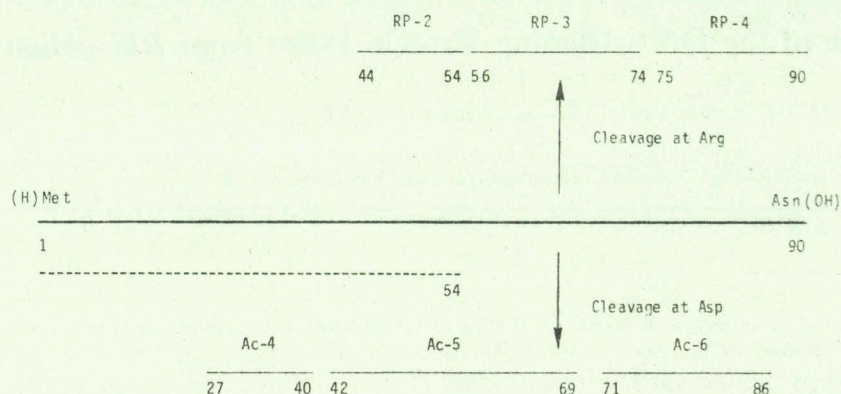


Fig. 1. Strategy used for the determination of the amino acid sequence of protein HRm. (---) Automated Edman degradation of protein HRm

Table 1. Automated Edman degradation of protein HRm

In this table and in Tables 3–6, HPLC = identification and quantification of phenylthiohydantoin of amino acids by high-pressure liquid chromatography [22,25]; GC = identification by gas chromatography [26]; – = not quantified

Cycle number	Residue number	Residue	Yield	Identification
			nmol	
1	1	Met	114	HPLC, GC
2	2	Asn	222	HPLC
3	3	Lys	264	HPLC
4	4	Asn	220	HPLC
5	5	Glu	195	HPLC
6	6	Leu	251	HPLC
7	7	Val	–	HPLC, GC
8	8	Ala	244	HPLC
9	9	Ala	268	HPLC
10	10	Val	–	HPLC, GC
11	11	Ala	210	HPLC
12	12	Asp	–	HPLC
13	13	Lys	114	HPLC
14	14	Ala	174	HPLC
15	15	Gly	143	HPLC
16	16	Leu	123	HPLC
17	17	Ser	27	HPLC
18	18	Lys	86	HPLC
19	19	Ala	127	HPLC
20	20	Asp	84	HPLC
21	21	Ala	118	HPLC
22	22	Ser	22	HPLC
23	23	Ser	26	HPLC
24	24	Ala	90	HPLC
25	25	Val	–	HPLC, GC
26	26	Asp	47	HPLC
27	27	Ala	74	HPLC
28	28	Val	–	HPLC, GC
29	29	Phe	49	HPLC, GC
30	30	Glu	20	HPLC, GC
31	31	Thr	12	HPLC
32	32	Ile	45	HPLC, GC
33	33	Gln	15	HPLC
34	34	Gly	37	HPLC, GC
35	35	Glu	15	HPLC, GC
36	36	Leu	18	HPLC, GC
37	37	Lys	29	HPLC
38	38	Asn	17	HPLC, GC
39	39	Gly	31	HPLC
40	40	Gly	38	HPLC
41	41	Asp	19	HPLC, GC

Table 1 (continued)

Cycle number	Residue number	Residue	Yield	Identification
42	42	Ile	19	HPLC, GC
43	43	Arg	22	HPLC, GC
44	44	Leu	8	HPLC, GC
45	45	Val	–	HPLC, GC
46	46	Gly	28	HPLC, GC
47	47	Phe	17	HPLC, GC
48	48	Gly	27	HPLC, GC
49	49	Asn	13	HPLC, GC
50	50	Phe	8	HPLC
51	51	Ser	1	HPLC
52	52	Val	–	HPLC, GC
53	53	Ser	1	HPLC
54	54	Arg	–	HPLC

Fractionation of Peptides. The enzymatic digest and the acid hydrolysate were freeze-dried and submitted to fractionation by reverse-phase high-pressure liquid chromatography. The separation of peptides was performed with a Waters Associates ALC/GPC 244 apparatus equipped with a U 6 K sample injector and a Cecil 212 variable-wavelength ultraviolet monitor. The samples were dissolved in 0.05% (v/v) trifluoroacetic acid in distilled water, centrifuged at $4000 \times g$ for 5 min and loaded on a C_{18} μ Bondapak column (9 μ m, 0.39×30 cm) (Waters Associates) equilibrated with 10% (v/v) acetonitrile in 0.05% trifluoroacetic acid. Peptides were eluted with a linear gradient of acetonitrile in 0.05% trifluoroacetic acid as indicated in Fig. 2 and 3. All runs were carried out at room temperature at a constant flow rate of 2 ml/min. The solvents were degassed for 20 min by sonication and for 1 min under vacuum with stirring.

Sequence Determination. Automated Edman degradation of the protein HRm and of large peptides was performed as indicated in [22]. Manual Edman degradation of peptides was carried out as in [24]. The phenylthiohydantoin derivatives of amino acids were identified by high-pressure liquid chromatography as indicated in [22].

Nomenclature of Peptides. Peptides obtained by cleavage of the protein at arginyl residues with the protease from submaxillary gland of mouse are designated by RP-. Peptides derived from cleavage at aspartyl residues are designated by Ac-. The peptides are numbered according to their position in the sequence of the protein.

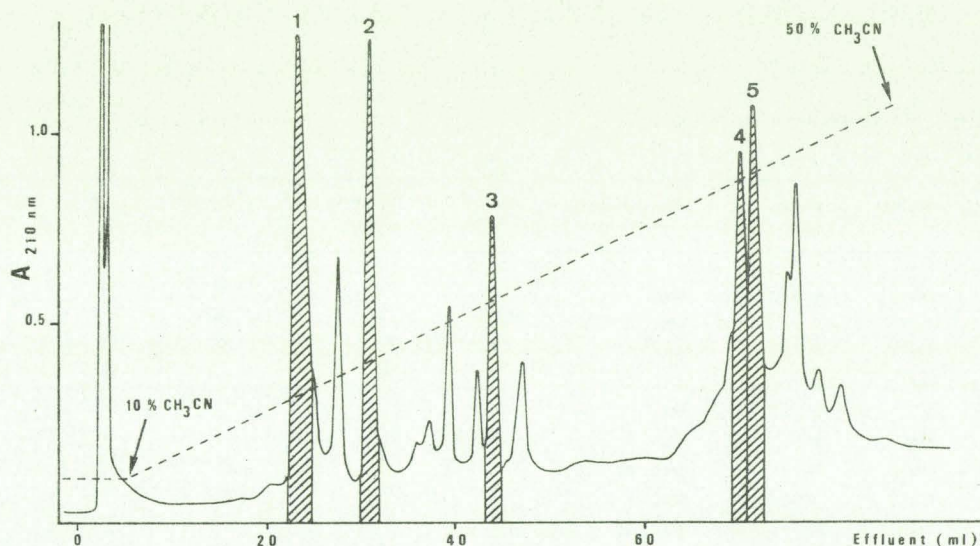


Fig. 2. Elution pattern of peptides generated by cleavage at arginyl residues with the protease from submaxillary gland of mouse and fractionated by reverse-phase HPLC. The digest (130 nmol) was loaded on a column of C_{18} microBondapak (0.39×30 cm) equilibrated with 10% (v/v) acetonitrile in 0.05% trifluoroacetic acid. The column was eluted with a linear gradient (---) from 10% to 50% acetonitrile in 0.05% trifluoroacetic acid. Fractions of 0.8 ml were collected at a flow rate of 2 ml/min. The elution of peptides was monitored at 210 nm. The fractions used for sequence determination were numbered by order of elution from 1 to 5

Table 2. Amino acid composition of peptides used for the sequence determination

Values are expressed as number of residues per mole of peptide. Values in parentheses are the nearest integers. No corrections were made for degradative losses of threonine and serine. The last column shows the amino acid composition of protein HRm determined in [22]

Amino acids	RP-1 residues 1 to 43	RP-2 residues 44 to 54	RP-3 residues 56 to 74	RP-4 residues 75 to 90	Ac-4 residues 27 to 40	Ac-5 residues 42 to 69	Ac-6 residues 71 to 86	Protein HRm
Aspartic acid	7.08 (7)	1.25 (1)	2.03 (2)	3.05 (3)	1.14 (1)	1.83 (2)	1.29 (1)	13
Threonine	0.90 (1)		0.99 (1)	0.99 (1)	0.81 (1)	0.88 (1)	0.95 (1)	3
Serine	2.30 (3)	1.42 (2)	1.38 (2)			3.14 (4)		7
Glutamic acid	4.47 (4)		2.11 (2)		2.67 (3)	2.13 (2)		6
Proline			1.96 (2)	1.03 (1)		0.81 (1)	1.73 (2)	3
Glycine	3.66 (14)	1.83 (2)	1.82 (2)	1.77 (2)	2.67 (3)	4.05 (4)	1.89 (2)	10
Alanine	7.91 (8)		2.91 (3)	1.89 (2)	1.32 (1)	2.21 (2)	2.34 (2)	13
Valine	4.06 (4)	2.06 (2)	1.02 (1)	2.12 (2)	1.06 (1)	2.70 (3)	1.30 (1)	9
Methionine	0.94 (1)							1
Isoleucine	2.10 (2)		1.03 (1)		0.90 (1)	1.05 (1)	0.89 (1)	3
Leucine	3.00 (3)	1.13 (1)		1.13 (1)	1.10 (1)	1.40 (1)	1.05 (1)	5
Phenylalanine	1.07 (1)	2.01 (2)		1.10 (1)	0.99 (1)	1.87 (2)	0.97 (1)	4
Lysine	3.94 (4)		1.12 (1)	2.92 (3)	1.18 (1)	1.49 (1)	2.75 (3)	8
Arginine	1.28 (1)	1.12 (1)	2.29 (2)			4.15 (4)	0.85 (1)	5
Number of residues	43	11	19	16	14	28	16	90
No. fraction	Fig. 2				Fig. 3			
	4-5	3	1	2	3	4	1	

RESULTS

The strategy used for the determination of the amino acid sequence of the protein HRm is presented in Fig. 1. The sequence of the first 54 amino acid residues was determined by automated Edman degradation of the protein. The sequence of the carboxy-terminal half of the molecule was established from structural data provided by large peptides generated by highly specific cleavages of the protein. The protein HRm was successively cleaved at arginyl and aspartyl residues and

only the peptides which allowed us to determine the sequence are presented.

Automated Edman Degradation of the Protein HRm

Automated Edman degradation of the protein HRm was performed in two different runs, with samples of 260 nmol and 420 nmol, respectively. The repetitive yield calculated from alanine residues at positions 8 and 24 was 94% in both ex-

Table 3. Automated Edman degradation of peptide RP-2

Cycle number	Residue number	Residue	Yield nmol	Identification
1	44	Leu	65	HPLC
2	45	Val	-	HPLC, GC
3	46	Gly	20	HPLC
4	47	Phe	52	HPLC, GC
5	48	Gly	17	HPLC
6	49	Asn	36	HPLC
7	50	Phe	33	HPLC, GC
8	51	Ser	7	HPLC
9	52	Val	-	HPLC, GC
10	53	Ser	2	HPLC
11	54	Arg	3	HPLC

Table 4. Automated Edman degradation of peptide RP-3

Cycle number	Residue number	Residue	Yield nmol	Identification
1	56	Glu	138	HPLC
2	57	Ala	86	HPLC
3	58	Ser	20	HPLC
4	59	Lys	86	HPLC
5	60	Gly	36	HPLC
6	61	Arg	-	HPLC
7	62	Asn	26	HPLC
8	63	Pro	36	HPLC, GC
9	64	Ser	-	HPLC, GC
10	65	Thr	7	HPLC
11	66	Gly	16	HPLC
12	67	Ala	14	HPLC
13	68	Glu	2	HPLC, GC
14	69	Val	-	HPLC, GC
15	70	Asp	12	HPLC, GC
16	71	Ile	2	HPLC, GC
17	72	Pro	10	HPLC
18	73	Ala	7	HPLC

Table 5. Automated Edman degradation of peptide RP-4

Cycle number	Residue number	Residue	Yield nmol	Identification
1	75	Asn	71	HPLC, GC
2	76	Val	-	HPLC, GC
3	77	Pro	63	HPLC
4	78	Lys	66	HPLC, GC
5	79	Phe	66	HPLC, GC
6	80	Thr	25	HPLC
7	81	Ala	40	HPLC
8	82	Gly	-	HPLC
9	83	Lys	34	HPLC
10	84	Gly	-	HPLC
11	85	Leu	24	HPLC
12	86	Lys	22	HPLC, GC
13	87	Asp	3	HPLC
14	88	Ala	13	HPLC, GC
15	89	Val	-	HPLC, GC
16	90	Asn	6	HPLC

Table 6. Automated Edman degradation of peptide RP-5

Cycle number	Residue number	Residue	Yield nmol	Identification
1	42	Ile	137	HPLC
2	43	Arg	59	HPLC
3	44	Leu	67	HPLC, GC
4	45	Val	121	HPLC, GC
5	46	Gly	59	HPLC, GC
6	47	Phe	51	HPLC, GC
7	48	Gly	27	HPLC
8	49	Asn	17	HPLC, GC
9	50	Phe	23	HPLC, GC
10	51	Ser	5	HPLC, GC
11	52	Val	45	HPLC, GC
12	53	Ser	4	HPLC
13	54	Arg	5	HPLC
14	55	Arg	8	HPLC
15	56	Glu	8	HPLC
16	57	Ala	4	HPLC
17	58	Ser	-	HPLC
18	59	Lys	-	HPLC
19	60	Gly	1	HPLC
20	61	Arg	-	HPLC
21	62	Asn	-	HPLC
22	63	Pro	-	HPLC

periments. Positive identifications were achieved through step 54 (Table 1) [22].

Cleavage at Arginyl Residues

The peptides derived from cleavage of the protein HRM at arginyl residues by digestion with the protease from submaxillary gland of mouse were fractionated by reverse-phase high-pressure liquid chromatography (Fig. 2). The five hatched

fractions correspond to pure peptides. Their amino acid compositions are presented in Table 2. The peptide RP-1 was found in both fractions 4 and 5. This is likely due to a partial deamidation of asparagine and glutamine residues which makes the peptide more hydrophilic. The peptides RP-2, RP-3 and RP-4 were obtained in fractions 3, 1 and 2 respectively. The peptide RP-2 (residues 44–54), RP-3 (residues 56–74) and RP-4 (residues 75–90) were completely sequenced by automated Edman degradation (Table 3, 4 and 5 and Fig. 4).

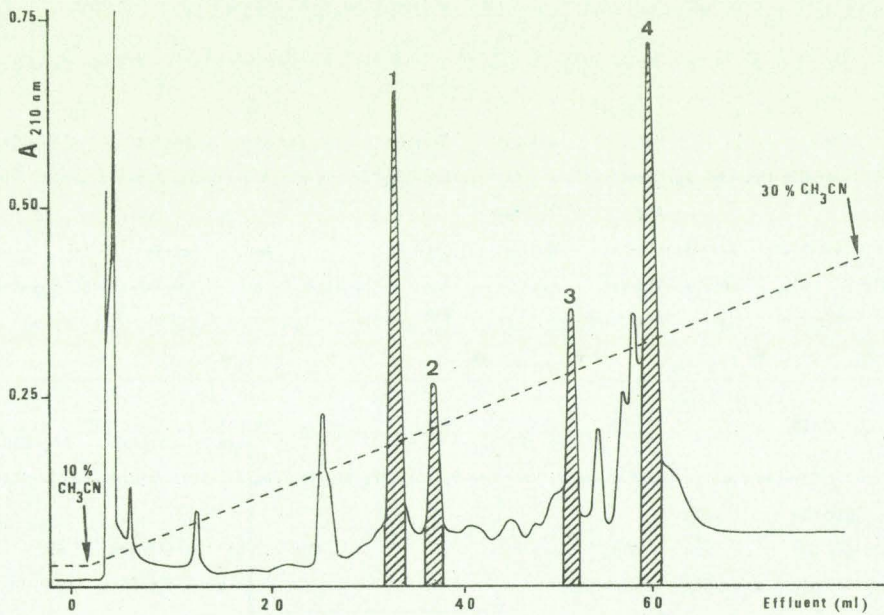


Fig. 3. Elution pattern of peptides generated by cleavage at aspartyl residues with 0.25 M acetic acid and fractionated by reverse-phase HPLC. The hydrolysate (70 nmol) was loaded on a column of C₁₈ microBondapak (0.39 × 30 cm) equilibrated as in Fig. 2. The column was eluted with a linear gradient (---) from 10% to 30% acetonitrile in 0.05% trifluoroacetic acid. Fractions of 1.2 ml were collected at a flow rate of 2 ml/min. The elution of peptides was monitored at 210 nm. The fractions were numbered by order of elution from 1 to 4

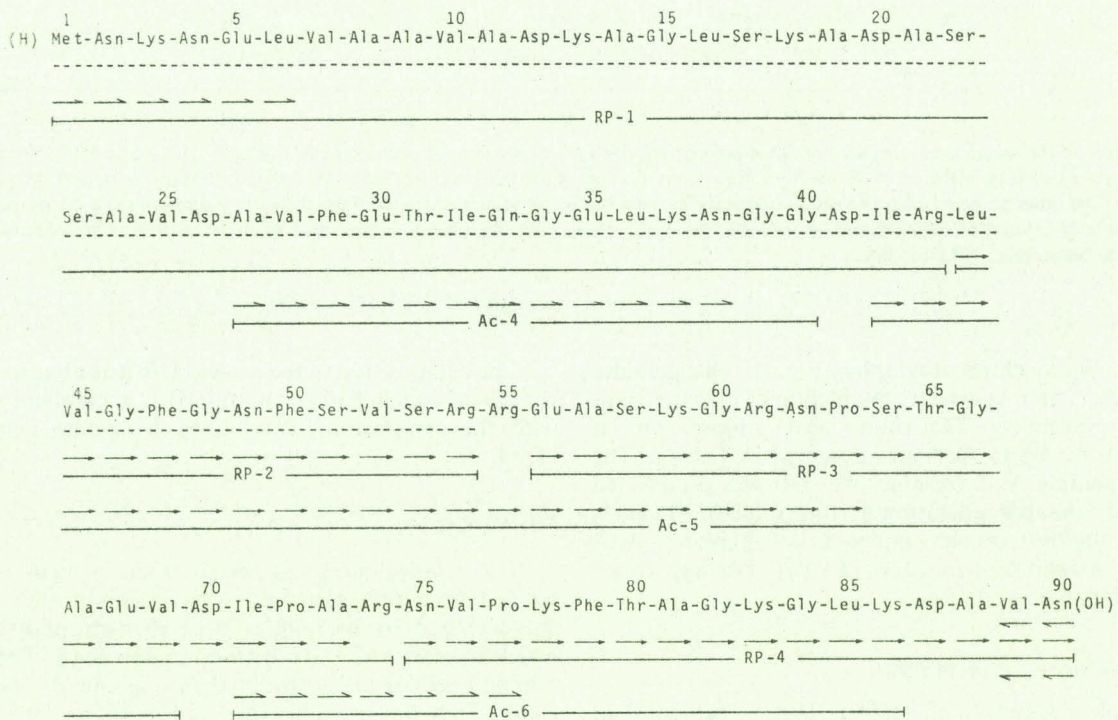


Fig. 4. Complete amino acid sequence of protein HRm. Methods used for determination of the sequence are indicated as following: ---, automated Edman degradation of the protein HRm; →, automated Edman degradation of peptides; →, manual Edman degradation of peptides; ←, carboxypeptidase hydrolysis

The protease from submaxillary gland of mouse failed to cleave the susceptible bond Arg⁶¹-Asn⁶². This bond was also resistant to trypsin and a tryptic peptide which covers the sequence 60–74 was therefore obtained in the tryptic hydrolysate of protein HRm (data not shown). The resistance of this bond to proteolytic cleavage for 4 h cannot be related to the nature of the residues adjacent to this bond (glycine and proline at positions 60 and 63 respectively) since in the

identical sequence Gly⁶⁰-Arg-Asn-Pro⁶³ of proteins HU-1 and HU-2 from *Escherichia coli*, the arginyl bond was cleaved by trypsin [9].

Cleavage at Aspartyl Residues

The peptides derived from hydrolysis of protein HRm with 0.25 M acetic acid were fractionated by reverse-phase

for DNA: during the isolation process, the proteins HRm, HU and HTa, of which the net charge is +2, +4 and +13 respectively, are eluted from a DNA-cellulose column with an NaCl concentration of approximately 0.12 M, 0.35 M and 0.75 M respectively [1, 22, 27].

Like the other bacteria *Bacillus subtilis* and *Pseudomonas aeruginosa*, *Rhizobium meliloti* contains a single HU-type protein whereas two polypeptide chains, HU-1 and HU-2 which differ from each other by 28 residues, are present in *Escherichia coli* [9, 11, 12]. Two HU-like proteins have also been observed in *Thermoplasma acidophilum* [20] and in the cyanobacteria *Synechocystis* [21], but the differences between the two polypeptide chains within each species are very few. Thus, in *T. acidophilum* the second form of protein HTa differs only from the first one by an additional residue at the amino terminus; in the protein HS from *Synechocystis*, of which the complete primary structure remains to be established, only two microheterogeneities have been observed in the amino-terminal sequence (27 identified residues). By comparison the corresponding sequences of proteins HU-1 and HU-2 differ by 12 residues.

The comparison of the sequence of protein HRm with that of other HU-type proteins (Fig. 5) shows that substitutions occur more frequently in the amino-terminal third of the molecule. On the other hand, two sequences, indicated by the boxes, appear almost invariable and may be important in the function of the protein. They contain residues which give to these sequences a high potential for β -turn conformation: glycine residues at positions 46, 48, 60 and proline residue at position 63. Indeed a β -turn conformation is predicted for these regions when a statistical predictive method [28] is applied to proteins HU-1 and HU-2 from *E. coli* [29]. In addition to the sequences mentioned above, five residues indicated by stars in Fig. 5 are strictly conserved in all sequences of HU-type proteins determined up to now. We must also emphasize that changes observed at positions 10, 17, 25, 29, 32, 42, 71 and 85 are uniquely conservative substitutions and that at 25 positions indicated by + in Fig. 5, only one change can be observed when comparing all the known sequences.

The degree of sequence homology between the HU-type proteins and protein HU-1 taken as reference was calculated as 87%, 52% and 26% for HPa, HRm and HTa respectively. This is in good agreement with the respective phylogenetic relationships of these bacteria evidenced by Fox et al. [30]. However these sequence data are not sufficient to make a valuable statement on the structural stability of the HU-type protein and need to be completed by structural study of HU-type proteins isolated from other bacteria chosen among the major lines of the prokaryotic descent described by Fox et al. [30].

The authors are grateful to Professor J. Guillaume (Université de Lille I) who provided the cultures of *Rhizobium meliloti*. They acknowledge with pleasure the skilful technical assistance of M. J. Dupire, T. Ernout and A. Héméz. They thank also Dr G. Briand and O. Hannedouche from the Service commun de Séquence Unité 16/124 Institut National de la Santé et de la Recherche Médicale (Lille) for the sequencer runs. This work was supported by grant from the Institut National de la Santé et de la

Recherche Médicale (CRL 80.20.35) and grants from the Centre National de la Recherche Scientifique (Action Thématique Programmée Chromatine 4203 and L.A. 268).

REFERENCES

1. Rouvière-Yaniv, J. & Gros, F. (1975) *Proc. Natl Acad. Sci. USA*, **72**, 3428–3432.
2. Varshavsky, A. J., Nedospasov, S. A., Bakayev, V. V., Bakayeva, T. G. & Georgiev, G. P. (1977) *Nucleic Acids Res.* **4**, 2725–2745.
3. Lathe, R., Buc, H., Lecocq, J. P. & Bautz, E. K. F. (1980) *Proc. Natl Acad. Sci. USA*, **77**, 3548–3552.
4. Hübscher, U., Lutz, H. & Kornberg, A. (1980) *Proc. Natl Acad. Sci. USA*, **77**, 5097–5101.
5. Pettijohn, D. E. (1976) *CRC Crit. Rev. Biochem.* **4**, 175–202.
6. Griffith, J. D. (1976) *Proc. Natl Acad. Sci. USA*, **73**, 563–567.
7. Laine, B., Sautière, P., Biserte, G., Cohen-Solal, M., Gros, F. & Rouvière-Yaniv, J. (1978) *FEBS Lett.* **89**, 116–120.
8. Mende, L., Timm, B. & Subramanian, A. R. (1978) *FEBS Lett.* **96**, 395–398.
9. Laine, B., Kmičik, D., Sautière, P., Biserte, G. & Cohen-Solal, M. (1980) *Eur. J. Biochem.* **103**, 447–461.
10. Rouvière-Yaniv, J. (1978) *Cold Spring Harbor Symp. Quant. Biol.* **42**, 439–447.
11. Hawkins, A. R. & Wootton, J. C. (1981) *FEBS Lett.* **130**, 275–278.
12. Nakayama, T. (1980) *Biochem. Biophys. Res. Commun.* **97**, 318–324.
13. Imber, R., Bächinger, H. & Bickle, T. A. (1982) *Eur. J. Biochem.* **122**, 627–632.
14. Haselkorn, R. & Rouvière-Yaniv, J. (1976) *Proc. Natl Acad. Sci. USA*, **73**, 1917–1920.
15. Searcy, D. G. (1975) *Biochim. Biophys. Acta*, **395**, 535–547.
16. Rouvière-Yaniv, J., Gros, F., Haselkorn, R. & Reiss, C. (1977) in *The Organisation and Expression of the Eukaryotic Genome* (Bradbury, E. M. & Javaherian, K., eds) pp. 211–231, Academic Press, New York.
17. Stein, D. B. & Searcy, D. G. (1978) *Science (Wash. DC)* **202**, 219–221.
18. Rouvière-Yaniv, J., Yaniv, M. & Germond, J. E. (1979) *Cell*, **17**, 265–274.
19. Searcy, D. G. & Stein, D. B. (1980) *Biochim. Biophys. Acta*, **609**, 180–195.
20. De Lange, R. J., Williams, L. C. & Searcy, D. G. (1981) *J. Biol. Chem.* **256**, 905–911.
21. Aitken, A. & Rouvière-Yaniv, J. (1979) *Biochem. Biophys. Res. Commun.* **91**, 461–467.
22. Laine, B., Bélaïche, D., Sautière, P. & Biserte, G. (1982) *Biochem. Biophys. Res. Commun.* **106**, 101–107.
23. Beringer, J. E. (1973) Ph.D. Thesis, University of East Anglia, UK.
24. Edman, P. & Henschen, A. (1975) in *Protein Sequence Determination* (Needleman, S. B., ed.) 2nd edn, pp. 232–279, Springer-Verlag, Berlin, Heidelberg, New York.
25. Hermann, J., Titani, K., Ericsson, L. H., Wade, R. D., Neurath, H. & Walsh, K. A. (1978) *Biochemistry*, **17**, 5672–5679.
26. Pisano, J. J. & Bronzert, T. J. (1969) *J. Biol. Chem.* **244**, 5597–5607.
27. De Lange, R. J., Green, G. R. & Searcy, D. G. (1981) *J. Biol. Chem.* **256**, 900–904.
28. Fasman, G. D., Chou, P. Y. & Adler, A. J. (1976) *Biophys. J.* **16**, 1201–1238.
29. Laine, B. (1980) Ph.D. Thesis, Université des Sciences et Techniques de Lille, France.
30. Fox, G. E., Stackebrandt, E., Hespell, R. B., Gibson, J., Maniloff, J., Dyer, T. A., Wolfe, R. S., Balch, W. E., Tanner, R. S., Magrum, L. J., Zablen, L. B., Blakemore, R., Gupta, R., Bonen, L., Lewis, B. J., Stahl, D. A., Luehrsen, K. R., Chen, K. N. & Woese, C. R. (1980) *Science (Wash. DC)* **209**, 457–463.

B. Laine, D. Bélaïche, and P. Sautière, Unité de Recherches Ultrastructurales et Biochimiques sur les Cellules Normales et Cancéreuses (Unité 124 de l'Institut National de la Santé et de la Recherche Médicale), Institut de Recherches sur le Cancer, Cité Hospitalière, 2 Place de Verdun, F-59020 Lille-Cedex, Nord, France

H. Khanaka, Laboratoire de Microbiologie, Université des Sciences et Techniques de Lille, F-59655 Villeneuve-d'Ascq, Nord, France

CHARACTERIZATION AND STRUCTURAL STUDY OF THE DNA-BINDING PROTEIN HRm FROM
Rhizobium meliloti

Bernard LAINE, Denise BELAICHE, Pierre SAUTIERE and Gérard BISERTE

Unité 124 INSERM, 2, Place de Verdun, BP 311, 59020 Lille Cédex, France

Received March 18, 1982

SUMMARY.- The protein HRm, a DNA-binding HU-like protein isolated from *Rhizobium meliloti*, is a single polypeptide chain (M_r 9 300) of 90 residues. Protein HRm exhibits lower affinity for double stranded DNA than *E. coli* protein HU. The sequence of the first 54 amino acid residues was established by automated Edman degradation of the protein. Many substitutions were observed by comparison with amino acid sequences of HU-like proteins from other bacteria. However a sequence of 7 residues is almost invariable and may be important in the function of the protein.

INTRODUCTION

In Prokaryotes the chromosomal DNA is packaged in a compact structure referred to as nucleoid (1) and, in electron microscopy, appears to be condensed in a beaded structure (2). The *E. coli* DNA-binding protein HU which has been shown to be associated with the nucleoid prepared at low ionic strength (3), forms *in vitro* with λ DNA a stable complex (4). Moreover protein HU can form nucleosome-like structures with circular double-stranded DNA (5). Therefore protein HU seems to be involved in the condensation of the chromosomal DNA. *E. coli* protein HU is a heterocomplex constituted of two polypeptide chains termed HU-1 and HU-2, of 90 residues each (6-8). Closely related proteins have been isolated from other bacteria (3, 9-12). Up to now, only the proteins HU-1 and HU-2 from *E. coli* (7, 8) and the protein HTa from *T. acidophilum* (13) have been completely sequenced. The partial sequences of the HU-like proteins from *P. aeruginosa* (12) and from the Cyanobacterium *Synechocystis* (14) have also been reported.

In this paper, we present the characterization and the structural study of protein HRm, a DNA-binding HU-like protein isolated from *Rhizobium meliloti*.

Abbreviations.- DFP, diisopropylfluorophosphate ; DMAA, dimethylallylamine ; EDTA, ethylene diamine tetraacetic acid ; HPLC, high pressure liquid chromatography ; M_r , relative molecular mass ; PTH, phenylthiohydantoin ; SDS, sodium dodecyl sulfate.

0006-291X/82/090101-07\$01.00/0

Copyright © 1982 by Academic Press, Inc.

All rights of reproduction in any form reserved.

MATERIALS AND METHODS

Rhizobium meliloti, strain 2011 Str 3 (streptomycin resistant mutant) was provided by J. Denarie (INRA, Versailles, France) and was grown as described in (15).

The protein was prepared by a procedure adapted from (16) and used for the preparation of *E. coli* protein HU (3). Bacteria (100 g) were suspended at 4°C in 400 ml of 20 mM Tris-HCl (pH 8.0), 1 mM Na₂ EDTA, 1.7 M NaCl, 0.1 mM DFP. Cells were disrupted in a French press with a pressure of 17 000 p.s.i. The high salt concentration is necessary to dissociate the nucleoprotein complex. The crude extract was centrifuged at 8 000 g for 1 h to remove the cell debris. An aqueous solution of polyethylene glycol 6 000 (30 % w:v) was slowly added to the supernatant to give a final polyethylene glycol concentration of 10 % (w:v). After gentle stirring for 30 min the suspension was centrifuged at 8 000 g for 1 h to remove the precipitated nucleic acids and the supernatant was dialysed three times against 5 liters of 0.05 M NaCl in 20 mM Tris-HCl (pH 8.0), 1 mM Na₂ EDTA, 1 mM 2-mercaptoethanol. The dialysed material, cleared by centrifugation at 8 000 g for 1 h, was made 10 % glycerol (v:v) and applied to a double-stranded DNA-cellulose column (10 x 2.5 cm) at a flow rate of 40 ml/h. The DNA-cellulose column was prepared as in (17) with chicken erythrocyte DNA. The column was washed with 0.05 M NaCl in 20 mM Tris-HCl (pH 8.0), 1 mM Na₂ EDTA, 10 % (v/v) glycerol. The column was then eluted with increasing concentrations of NaCl in the Tris-EDTA-glycerol buffer: 0.1 M NaCl, then a linear gradient from 0.1 to 0.175 M NaCl, and finally 0.4 and 2.0 M NaCl.

The protein HRm, eluted from the DNA-cellulose column with the linear gradient of NaCl, was purified by gel filtration chromatography on a Sephadex G-100 superfine column (90 x 2.5 cm) eluted with 0.01 M HCl saturated with chloroform.

Analytical electrophoresis was performed on polyacrylamide slab gel (120 x 80 x 1 mm) (18) in three different conditions: A, in SDS using the Laemmli buffer system (19); B, in 0.9 M acetic acid, 2.5 M urea (20) using a 17 % acrylamide concentration; C, in 0.9 M acetic acid, 6 M urea, and 0.38 % Triton X-100 using a 17 % acrylamide concentration (21).

Amino acid analyses were performed as described in (8).

Automated Edman degradation of the protein HRm was carried out in a Beckman 890 C Sequencer using a DMAA program (10 2974) in the presence of polybrene (22). PTH-amino acids were identified by HPLC on a column of C₁₈ micro Bondapak (Waters Associates) as described in (23). Alternatively the PTH-Phe and PTH-Ile as well as PTH-Met and PTH-Val which are not separated by the previous system were identified by gas-chromatography (24) on a Hewlett-Packard 5830 A unit. The PTH-Arg was identified by HPLC on a column of micro Bondapak Phenyl (Waters Associates) using the buffer system described in (23).

RESULTS AND DISCUSSION

The elution pattern of *R. meliloti* proteins fractionated on a DNA-cellulose column is presented in fig. 1-A. Fraction I corresponds to a large amount of material which has no affinity for the DNA and therefore is not retained by the column. As assessed by polyacrylamide gel electrophoresis (Fig. 1-B) the protein HRm is mainly eluted in the fraction III by a molarity in NaCl comprised between 0.10 and 0.15 M. The protein HRm exhibits a lower affinity for the DNA than *E. coli* protein HU which is eluted from the same column at about 0.35 M NaCl. *R. meliloti* protein HRm and *B. subtilis* protein HBs (11) exhibit a similar affinity for the DNA whereas protein HTa from *T. acidophilum* tight-

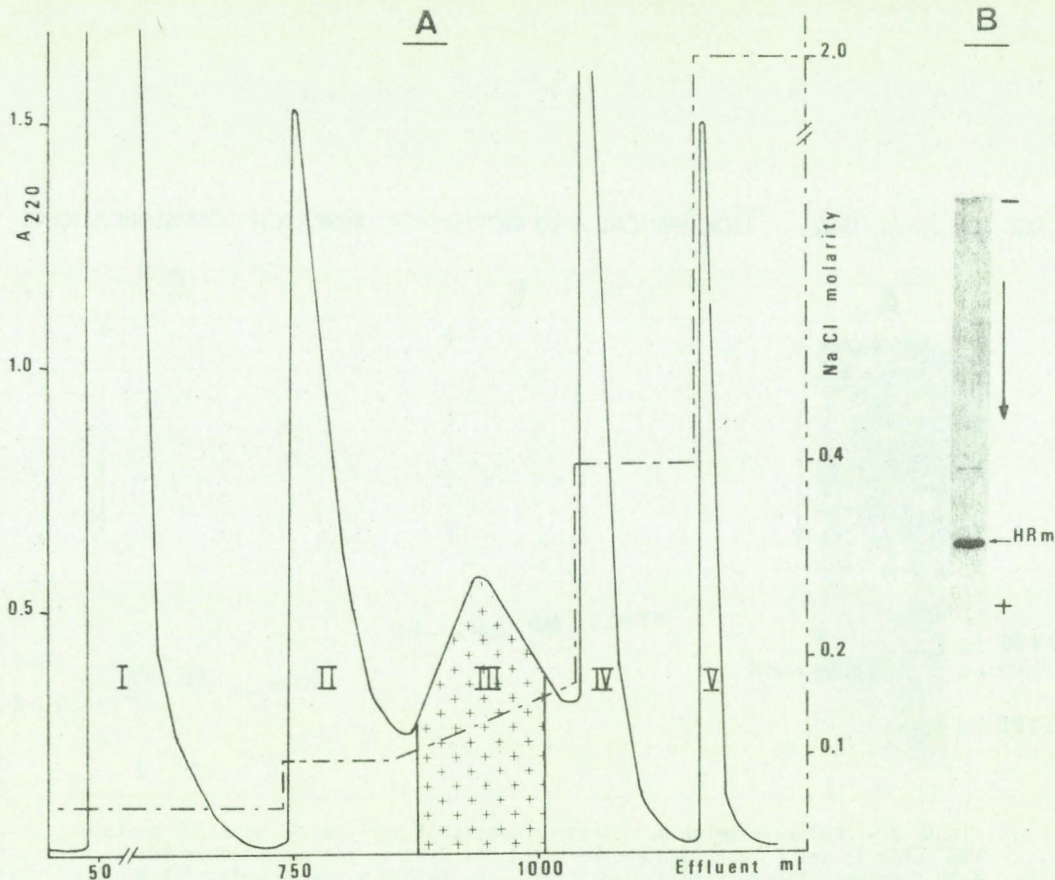


FIGURE 1.- A : Elution pattern of *R. meliloti* proteins fractionated on a DNA-cellulose column. An extract of *R. meliloti* bacteria prepared as described under "Materials and Methods" was applied to the column (10 x 2.5 cm) equilibrated with 0.05 M NaCl in 20 mM Tris-HCl (pH 8.0), 1 mM Na₂ EDTA, glycerol 10 % (v:v). The column was eluted at 40 ml/h with a gradient of NaCl in the same buffer as indicated by (—). The linear gradient of NaCl was from 0.1 to 0.175 M (2 x 100 ml). The effluent was monitored continuously at 220 nm (—) and fractions of 5 ml were collected. B : SDS gel electrophoresis (19) of fraction III. The experimental conditions were as indicated in Fig. 2-A.

ly binds to the DNA and is eluted from the DNA-cellulose column at about 0.75 M NaCl (25).

R. meliloti protein HRm has a M_r of 10 000 as estimated by SDS gel electrophoresis (Fig. 2-A). On polyacrylamide gel electrophoresis in 0.9 M acetic acid, 2.5 M urea, the protein HRm exhibits a cathodic mobility inferior to that of protein HU (Fig. 2-B) suggesting that protein HRm is less basic than protein HU. Moreover in the presence of Triton X-100 (Fig. 2-C) the protein HRm shows a single band whereas the protein HU from *E. coli* migrates as two bands corresponding to proteins HU-1 and HU-2.

The amino acid composition of the protein HRm from *R. meliloti* is presented in Table 1. On the basis of four residues of phenylalanine present in the

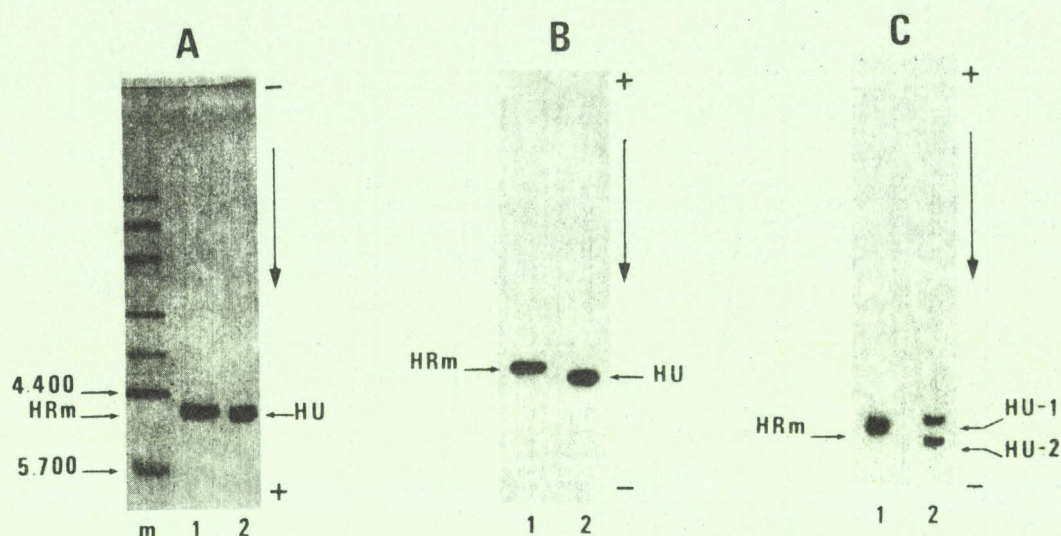


FIGURE 2.- Polyacrylamide gel electrophoresis of purified *R. meliloti* protein HRm (slot 1) and *E. coli* protein HU (slot 2). 2-A, in the presence of SDS on a 5-30 % polyacrylamide gradient gel (19). 2-B, in 0.9 M acetic acid, 2.5 M urea (20) on a gel containing 17 % acrylamide. 2-C, in 0.9 M acetic acid, 6 M urea and 0.38 % Triton X-100 (21) on a gel containing 17 % acrylamide. Electrophoresis were performed on a slab gel (18) at room temperature. In the case of Fig. 2-A samples (7 μ g) were treated with 5 % SDS, 2 % 2-mercaptoethanol for 2 min at 100°C and run at 40 mA for 150 min. Electrode buffer (Tris-glycine, pH 8.3) and gel buffer (Tris-HCl, pH 8.9) were made 0.1 % in SDS. m, molecular mass markers : Phosphorylase b (94,000), Bovine serum albumin (67,000), Ovalbumin (43,000), Carbonic anhydrase (30,000), Soybean trypsin inhibitor (20,100), α -Lactalbumin (14,400) and Insulin (5,700). In the case of Fig. 2-B and 2-C, samples (3 μ g) were dissolved in 0.01 M HCl, 8 M urea, 0.5 M 2-mercaptoethanol and were run at 22 mA for 3 h. The gels were stained with 0.5 % Coomassie blue R-250 in acetic acid/ethanol/water (1:2:7, v/v/v) for 15 h in the case of Fig. 2-A and for 3 h in the case of Fig. 2-B and 2-C. The gels were destained by diffusion in acetic acid/ethanol/water (7:20:73, v/v/v).

protein, the total number of amino acid residues is 90 which corresponds to a calculated M_r of 9 311. This value is in good agreement with that estimated by SDS gel electrophoresis. The protein is characterized by a high content of hydrophobic residues which represent 42 % of the total amino acids and is devoid of cysteine, tyrosine, histidine and tryptophan. Similar features were observed for proteins HU-1 and HU-2 from *E. coli* but protein HU-2 contains one histidine residue.

Automated Edman degradation of the *R. meliloti* protein HRm was performed in two different runs, with samples of 260 and 420 nM, respectively. The repetitive yield calculated from the alanine residues at positions 8 and 24 was 94 % in both experiments. Positive identifications were achieved through step 54. The sequence of the first 54 residues of the protein HRm thus determined by automated Edman degradation was confirmed by structural studies of the peptides isolated from a tryptic hydrolysate of the protein (Fig. 3). The identification of only one PTH amino acid at each step of the Edman degradation indi-

HRm	1	5	10	15	20	25	
	Met-Asn-Lys-Asn-Glu	Leu-Val-Ala-Ala-Val-Ala-Asp-Lys	Ala-Gly-Leu-Ser-Lys-Ala-Asp	Ala	Ser-Ser-Ala-Val-Asp-Ala		
HU-1	Met-Asn-Lys-Ser-Gln	Leu-Ile-Asp-Lys-Ile-Ala-Ala-Gly	Ala-Asp-Ile-Ser-Lys-Ala-Ala	Ala	Gly-Arg-Ala-Leu-Asp-Ala		
HU-2	Met-Asn-Lys-Thr-Gln	Leu-Ile-Asp-Val-Ile-Ala-Glu-Lys	Ala-Glu-Leu-Ser-Lys-Thr-Gln	Ala	Lys-Ala-Ala-Leu-Glu-Ser		
HPa	Met-Asn-Lys-Ser-Gln	Leu-Ile-Asp-Ala-Ile-Ala-Ala-Ser	Ala	? - ? - ? - Lys-Ala-Val	Ala	Gly-Lys-Ala-Leu-Asp-Ala	
HTa	Met-Val-Gly-Ile-Ser-Glu	Leu-Ser-Lys-Glu-Val-Ala-Lys-Lys	Ala-Asn-Thr-Thr-Gln-Lys-Val	Ala	Arg-Thr-Val-Ile-Lys-Ser		
HS	Met-Asn-Lys-Gly-Glu	Leu-Val-Asp-Ala-Val-Met-Ala-Lys	Ala	Asp-Thr-Val-Thr-Lys-Lys-Gln	Ala	Asp-Ala-Phe-Ile-Leu-Ala	
HRm	30	35	40	45	50	54	
	Val-Phe-Glu-Thr-Ile-Gln-Gly-Glu-Leu-Lys-Asn	Gly-Gly-Asp-Ile-Arg	Leu-Val	Gly-Phe-Gly	Asn-Phe	Ser-Val-Ser-Arg	
HU-1	Ile-Ile-Ala-Ser-Val-Thr-Glu-Ser-Leu-Lys-Glu	Gly-Asp-Asp-Val-Ala	Leu-Val	Gly-Phe-Gly	Thr-Phe	Ala-Val-Lys-Glu	
HU-2	Thr-Leu-Ala-Ala-Ile-Thr-Glu-Ser-Leu-Lys-Glu	Gly-Asp-Ala-Val-Gln	Leu-Val	Gly-Phe-Gly	Thr-Phe	Lys-Val-Asn-His	
HPa	Val-Ile-Glu-Ser-Val-Thr-Gly-Ala-Leu-Lys-Ala	Gly-Asp- ? - ? - ? - ?	Val	Gly-Phe-Gly	Thr-Phe	Ala-Val-Lys-Glu	
HTa	Phe-Leu-Asp-Glu-Ile-Val-Ser-Glu-Ala-Asn-Gly	Gly-Gln-Lys-Ile-Asn	Leu-Ala	Gly-Phe-Gly	Ile-Phe	Glu-Arg-Arg-Thr	

FIGURE 4.- Comparison of the amino-terminal amino acid sequences of DNA-binding HU-like proteins from different bacteria: HRm from *R. meliloti*, HU-1 and HU-2 from *E. coli* (8), HPa from *P. aeruginosa* (12), HTa from *T. acidophilum* (13), HS from the Cyanobacterium *Synechocystis* (14). Boxes indicate highly conserved residues.(?), not identified amino acid residues in *P. aeruginosa*.

A comparison of the sequence of protein HRm with that of other HU-like DNA binding proteins shows many substitutions (Fig. 4) (8, 12-14). However 9 residues indicated by boxes appear highly conserved during the evolution. Particularly the sequence between residues 44 and 50 is almost invariable and may be important in the function of the protein. The two phenylalanine residues at positions 47 and 50 can be made more accessible for interaction since the adjacent glycine residues give to this sequence a high potential for β -turn conformation. Indeed a β -turn conformation is predicted for this region when a statistical predictive method (26) is applied to proteins HU-1 and HU-2 from *E. coli* (27).

ACKNOWLEDGEMENTS.- We thank Professor J. Guillaume for valuable discussion and for providing the cultures of *R. meliloti*. The skilful technical assistance of M.J. Dupire and A. Hémez is gratefully acknowledged. This work was supported by grant n° 80.20.35 of the Institut National de la Santé et de la Recherche Médicale and grant ATP Chromatin 4206 of the Centre National de la Recherche Scientifique.

REFERENCES

1. Pettijohn, D.E. (1976) CRC Critical Reviews in Biochemistry 4, 175-202.
2. Griffith, J.D. (1976) Proc. Natl. Acad. Sci. USA 73, 563-567.
3. Rouvière-Yaniv, J. (1978) Cold Spring Harbor Symp. Quant. Biol. 42, 439-447.
4. Rouvière-Yaniv, J. and Gros, F. (1975) Proc. Natl. Acad. Sci. USA 72, 3428-3432.
5. Rouvière-Yaniv, J., Yaniv, M. and Germond, J.E. (1979) Cell 17, 265-274.

6. Laine, B., Sautière, P., Biserte, G., Cohen-Solal, M., Gros, F. and Rouvière-Yaniv, J. (1978) FEBS-Lett. 89, 116-120.
7. Mende, L., Timm, B. and Subramanian, A.R. (1978) FEBS-Lett. 96, 395-398.
8. Laine, B., Kniécik, D., Sautière, P., Biserte, G. and Cohen-Solal, M. (1980) Eur. J. Biochem. 103, 447-461.
9. Haselkorn, R. and Rouvière-Yaniv, J. (1976) Proc. Natl. Acad. Sci. USA 73, 1917-1920.
10. Searcy, D.G. (1975) Biochim. Biophys. Acta 395, 535-547.
11. Nakayama, T. (1980) Biochem. Biophys. Res. Commun. 97, 318-324.
12. Hawkins, A.R. and Wootton, J.C. (1981) FEBS-Lett. 130, 275-278.
13. De Lange, R.J., Williams, L.C. and Searcy, D.G. (1981) J. Biol. Chem. 256, 905-911.
14. Aitken, A. and Rouvière-Yaniv, J. (1979) Biochem. Biophys. Res. Commun. 91, 461-467.
15. Beringer, J.E. (1973) Genetic studies with *Rhizobium leguminosarum*. Ph.D. Thesis, University of East Anglia, United Kingdom.
16. Humphries, P., McConnell, D.J. and Gordon, R.L. (1973) Biochem. J. 133, 201-203.
17. Litman, R.M. (1968) J. Biol. Chem. 243, 6222-6233.
18. Kerckaert, J.P. (1978) Anal. Biochem. 84, 354-360.
19. Laemmli, U.K. (1970) Nature 227, 680-685.
20. Panyim, S. and Chalkley, R. (1969) Arch. Biochem. Biophys. 130, 337-346.
21. Hardinson, R. and Chalkley, R. (1978) Methods in Cell Biol. 17, 235-251.
22. Hunkapiller, M.W. and Hood, L.E. (1978) Biochemistry 17, 2124-2133.
23. Hermann, J., Titani, K., Ericsson, L.H., Wade, R.D., Neurath, H. and Walsh, K.A. (1978) Biochemistry 17, 5672-5679.
24. Pisano, J.J. and Bronzert, T.J. (1969) J. Biol. Chem. 244, 5597-5607.
25. De Lange, R.J., Green, G.R. and Searcy, D.G. (1981) J. Biol. Chem. 256, 900-904.
26. Fasman, G.D., Chou, P.Y. and Adler, A.J. (1976) Biophys. J. 16, 1201-1238.
27. Laine, B. (1980) Isolement et études structurales de deux protéines se liant au DNA, les protéines HU-1 et HU-2 d'*Escherichia coli*. Ph.D. Thesis, Université des Sciences et Techniques de Lille, France.

IV - 1 DETERMINATION DE LA STRUCTURE PRIMAIRE DE LA PROTEINE

HRm de *Rhizobium meliloti*

La protéine HRm isolée de *Rhizobium meliloti* est une protéine de type HU qui se lie au DNA bactérien et intervient probablement dans sa condensation. La protéine HU isolée de *E. coli* stabilise le DNA contre la dénaturation thermique et stimule sa transcription par la RNA-polymérase *in vitro*. Cette protéine, dont la séquence est nettement distincte de celles des histones, est cependant capable d'organiser le DNA de SV 40 et de former des sous-unités structurales en apparence semblables aux nucléosomes rencontrés chez les Eucaryotes.

La protéine HRm est constituée d'une seule chaîne polypeptidique de 90 résidus d'acides aminés. Cette protéine est caractérisée par un taux élevé d'acides aminés hydrophobes ; elle est dépourvue de cystéine, d'histidine, de tyrosine et de tryptophane.

La stratégie utilisée pour déterminer la structure primaire de la protéine HRm est présentée dans la figure 29. Deux échantillons de protéine HRm, de 260 et 420 nanomoles respectivement, ont été soumis à la dégradation d'EDMAN automatisée (publication 1) (40). Le premier échantillon a été séquencé avec un programme DMBA en présence de polybrène, le second avec un programme DMAA en présence de polybrène. Les rendements obtenus lors de la première détermination, sont présentés dans le tableau VI.

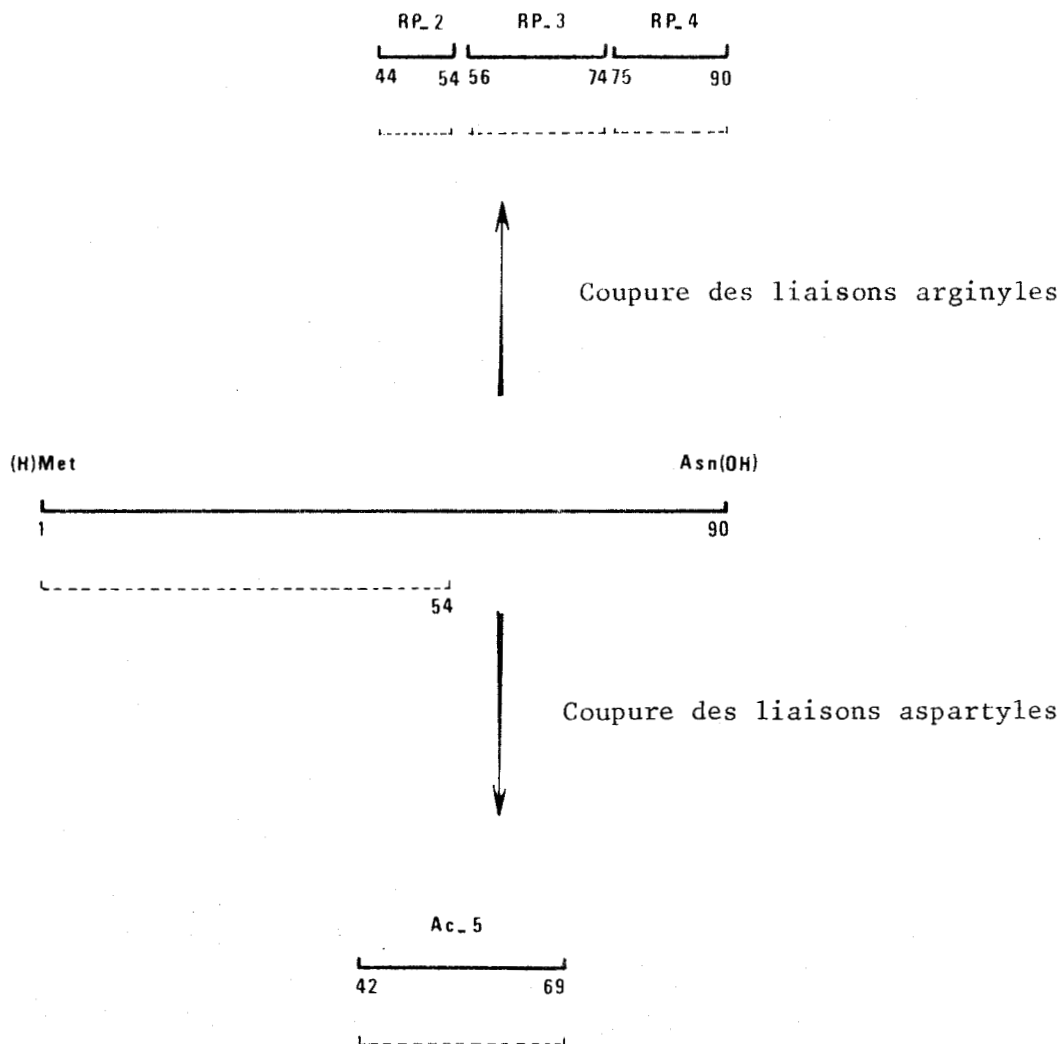


FIGURE 29 STRATEGIE UTILISEE POUR LA DETERMINATION DE LA SEQUENCE EN ACIDES AMINES DE LA PROTEINE HRm.

..... dégradation d'EDMAN automatisée de la protéine et des fragments obtenus par coupure des liaisons arginyles de la protéine avec la protéase Arg-C de glande sous-maxillaire de Souris (fragments RP) ou par coupure des liaisons aspartyles de la protéine avec l'acide acétique dilué (fragment Ac).

TABLEAU VI

RENDEMENTS OBTENUS LORS DE LA DEGRADATION D'EDMAN AUTOMATISEE,
AVEC UN PROGRAMME DMBA, DE LA PROTEINE HRm (260 NANOMOLES)

N° Résidu	Résidu	Rendements en nanomoles		
		R-1(1)	R(2)	R+1(3)
1	Met		n.d(4)	n.d
2	Asn		92,8	31,4
3	Lys		69,3	22,5
4	Asn	31,4	90,3	55,4
5	Glu	1,3	87	17,8
6	Leu		115,6	23,7
7	Val		n.d	n.d
8	Ala	10,6	119,1	144,8
9	Ala	119,1	144,8	35,4
10	Val	n.d	n.d	n.d
11	Ala	35,4	108,3	38
12	Asp		59,4	40,9
13	Lys	1,4	43,7	23,1
14	Ala	15	82,6	35,8
15	Gly	9	82,7	29,8
16	Leu		47,5	18,2
17	Ser		17,7	4,3
18	Lys	1,8	29,7	20,5
19	Ala	8,8	56,6	33,4
20	Asp	1,9	38,1	20,3
21	Ala	33,4	46,7	37,3
22	Ser		9	11,3
23	Ser	9	11,3	6
24	Ala	20,9	33,1	30
25	Val	n.d	n.d	n.d

(2) R = rendement en nanomoles de la PTH-AA obtenue au cours du cycle de dégradation.

(1) R-1 = rendement de cette PTH-AA au cycle précédent.

(3) R+1 = rendement de cette PTH-AA au cycle suivant.

(4) n.d = identifiée mais non dosée.



On peut observer qu'à partir du douzième cycle, le rendement a chuté de moitié et qu'au cycle suivant "l'overlap" d'acide aspartique est très important (68 p.100). Cette dernière observation n'est pas isolée car les résidus d'acide aspartique créent assez fréquemment des "overlaps" élevés dans les cycles suivants. On peut également remarquer que les rendements obtenus pour les résidus de sérine sont faibles mais il faut se souvenir que la PTH-sérine n'est pas stable (figure 13, p.63).

Cependant, malgré la baisse de rendement au douzième résidu, nous avons pu établir la séquence N-terminale de la protéine HRm jusqu'au résidu 25. Au delà du résidu 25 l'identification de la PTH-amino acide libérée à chaque cycle est devenue impossible.

Les rendements obtenus lors de la seconde détermination, effectuée avec un programme DMAA, sont présentés dans le tableau VII.

Bien que le DMBA soit un tampon de couplage recommandé pour les protéines et les peptides hydrophobes (30), nous lui avons préféré au cours de ce second passage le DMAA qui est éliminé par simple évaporation sous vide ce qui a pour effet de limiter les lavages par le benzène et donc de diminuer les pertes de film protéique. De plus, le tampon DMBA salit très fortement toute l'enceinte de la coupelle alors que ce n'est pas le cas du tampon DMAA.

TABLEAU VII

DEGRADATION D'EDMAN AUTOMATISEE AVEC UN PROGRAMME DMAA DE LA PROTEINE HRm (420 NANOMOLES)

N° Résidu	Résidu	R-1 (1)	Rendements en nanomoles R(2)	R+1 (3)
1	Met		n.d(4)	
2	Asn		254,5	28,8
3	Lys		264,5	28,8
4	Asn	28,8	250,7	31,3
5	Glu		195,3	30,3
6	Leu		251	19,2
7	Val		n.d	
8	Ala	15,3	243,7	268,4
9	Ala	243,7	268,4	50,6
10	Val		n.d	n.d
11	Ala	50,6	210,3	47,8
12	Asp		126,9	44,3
13	Lys		114,1	42,3
14	Ala	24,6	174,4	54,2
15	Gly	12,6	143,3	42,6
16	Leu		122,6	26,2
17	Ser		27,1	1,9
18	Lys		85,8	34,9
19	Ala	22,2	127,4	52,3
20	Asp		83,9	49,2
21	Ala	52,3	118,3	56
22	Ser		22,1	26,4
23	Ser	22,1	26,4	9,6
24	Ala	32	89,7	52,9
25	Val		n.d	n.d
26	Asp		46,6	33,4
27	Ala	29,2	74,3	47,8
28	Val		n.d	n.d
29	Phe		49,4	69
30	Glu		20,1	16,2
31	Thr		12	7,1
32	Ile	40	45,4	27,8
33	Gln	12,4	15	23,7
34	Gly	15,5	37,3	30,1
35	Glu		15,2	
36	Leu		18,4	21,7
37	Lys	18,4	29,4	17,7
38	Asn		27,7	16,8
39	Gly	19,9	30,8	38,1
40	Gly	30,8	38,1	28,7
41	Asp	7,4	18,7	13,9
42	Ile		19,4	22,7
43	Arg	25,3	47,5	40,8
44	Leu		7,8	10,6
45	Val		n.d	
46	Gly	15	27,9	21
47	Phe	15,1	16,7	38,4
48	Gly	21	27	24,3
49	Asn	12,8	20,3	19
50	Phe	32	40,3	15,7
51	Ser		0,7	
52	Val		n.d	n.d
53	Ser		0,8	0,5
54	Arg		n.d	

Voir la légende du tableau VI page 115.



Au cours de ce second passage au séquenceur, nous avons pu identifier les 49 premiers résidus de la protéine HRm. On peut noter comme au premier passage, que le rendement est fortement abaissé à partir du résidu d'acide aspartique en position 12 ; au 13ème cycle l'"overlap" d'acide aspartique est encore important (34 p.100), mais il a été diminué de moitié par rapport à la première détermination.

Pour les deux différents passages de la protéine HRm au séquenceur, nous avons calculé le rendement répétitif qui est plus correct que le rendement individuel de chaque PTH-AA. Ce rendement est calculé à partir d'acides aminés comme l'alanine, la glycine, la leucine et la lysine. Il est donné par la formule suivante :

$$Y_R = \left(\frac{B}{A} \right)^{\frac{1}{n_B - n_A}}$$

dans laquelle Y_R représente le rendement répétitif, B le rendement individuel de l'acide aminé X à l'étape B , A le rendement individuel de ce même acide aminé à l'étape A (B étant supérieur à A), n_B et n_A la position de l'acide aminé X dans la chaîne polypeptidique ($n_B > n_A$).

Plus le rendement répétitif est élevé, plus grand est le nombre de résidus identifiés dans la séquence. Ainsi, le rendement répétitif calculé à partir des résidus d'alanine 8 et 24 est de 92,3 p.100 pour le premier passage de la protéine HRm

au séquenceur, et de 94,5 p.100 pour le second. Le choix de la PTH-alanine pour le calcul de ces rendements, nous a été dicté par le fait que c'est la PTH-AA qui a l'un des meilleurs coefficients de réponse en HPLC, et qui se détruit le moins facilement. L'amélioration du rendement répétitif doit être un souci constant pour qui séquence une protéine car dans cet exemple précis, il a été possible avec un rendement répétitif de 94,5 p.100, d'identifier sans ambiguïté les 49 premiers résidus de la protéine HRm alors qu'avec un rendement répétitif de 92,3 p.100 seuls les 25 premiers résidus purent être identifiés. Au second passage, des indications très intéressantes ont même été obtenues jusqu'au résidu 54, indications qui ont été confirmées grâce à un recouvrement de la partie N-terminale de la protéine par des peptides obtenus par hydrolyse trypsique. Cette extension de séquence jusqu'au résidu 54 nous a également permis de choisir la stratégie pour déterminer le reste de la structure primaire de la protéine. La détermination de la séquence des 54 premiers résidus de la protéine HRm nous a fourni de précieuses indications qui nous ont guidée dans le choix des méthodes de coupure de la protéine pour l'obtention des fragments couvrant la région C-terminale. Sur les 5 résidus d'arginine présents dans la protéine HRm, 2 avaient été identifiés aux positions 43 et 54. L'hydrolyse de la protéine par la protéase Arg-C de glande sous-maxillaire de Souris nous a permis d'obtenir 3 fragments intéressants, désignés par les lettres RP- et numérotés suivant leur position dans la séquence : RP-2 (44-54), RP-3 (56-74) et RP-4 (75-90).

Par ailleurs le fragment Ac-5 (42-69) obtenu par coupure de la protéine au niveau des résidus d'acide aspartique avec l'acide acétique 0.25N à 100°C pendant 15 heures nous a permis les recouvrements nécessaires pour aligner les peptides RP-2 et RP-3.

Le peptide RP-2 (résidus 44 à 54) (tableau 3, publication 2) a été séquencé entièrement avec un programme quadrol 0.1M en présence de polybrène. Son rendement répétitif, calculé à partir des résidus de glycine en position 3 et 5 est de 91 p.100.

Le peptide RP-3 (résidus 56 à 74) (tableau 4, publication 2), a été séquencé avec un programme DMAA en présence de polybrène. Le rendement répétitif calculé à partir des résidus d'alanine 57 et 67 est de 83,2 p.100. A partir du 5ème cycle, le rendement a fortement diminué et l'"overlap" de glycine au 6ème cycle est très important (84 p.100). Ceci est peut être dû au fait que le résidu de glycine est suivi d'un résidu d'arginine qui est un acide aminé à chaîne latérale encombrante. Le rendement de la dégradation a ensuite diminué progressivement ceci étant dû à un entraînement progressif, par les solvants de lavage, du peptide résiduel dont la taille diminue à chaque cycle de dégradation. Néanmoins, 18 des 19 résidus du peptide ont été identifiés.

En ce qui concerne le peptide RP-4 (résidus 75 à 90), il a été soumis à la dégradation automatique d'EDMAN avec un programme DMAA en présence de polybrène. Les résultats obtenus

sur ce peptide sont présentés dans le tableau 5 de la publication 2. Son rendement répétitif calculé à partir des résidus d'alanine en position 81 et 88 est de 85,5 p.100. Un double clivage, par l'acide heptafluorobutyrique, a été effectué après 3 cycles de dégradation pour cliver complètement le résidu de proline attendu en position 3 de la séquence du peptide RP-4, par analogie avec les structures primaires connues d'autres protéines de type HU.

Parmi les peptides obtenus par coupure de la protéine à l'acide acétique seul le peptide Ac-5 (42-69) a été séquencé par la méthode de dégradation automatique d'EDMAN. Ce peptide a été séquencé en programme DMAA en présence de polybrène (tableau 6, publication 2). Le rendement répétitif calculé à partir des résidus de glycine en position 46 et 60 est de 78 p.100. Les 22 premiers résidus du peptide qui en compte 28 ont pu être ainsi identifiés.

Nous voyons donc que grâce aux résultats obtenus avec la dégradation d'EDMAN automatisée à partir de la protéine native et de quatre de ses fragments, nous avons pu déterminer la séquence complète de la protéine HRm de *Rhizobium meliloti* (figure 4, publication 2) (41).

Les travaux concernant la détermination de la structure primaire d'histones H2A d'Invertébrés marins ont fait l'objet des publications suivantes :

Publication 3

The amino-acid sequence of histone H2A from cuttlefish *Sepia officinalis*.

WOUTERS-TYROU D., MARTIN-PONTHIEU A., BRIAND G., SAUTIERE P. and BISERTE G. (1982)

Eur.J.Biochem., 124, 489-498.

Publication 4

Primary structure of histone H2A from gonads of the starfish *Asterias rubens*.

MARTINAGE A., BELAÏCHE D., DUPRESSOIR T. and SAUTIERE P. (1983)

Eur.J.Biochem., 130, 465-472.

Publication 5

Primary structure of histone H2A from nucleated erythrocyte of the marine worm *Sipunculus nudus*.

Presence of two forms of H2A in the sipunculid chromatin.

KMIECIK D., COUPPEZ M., BELAÏCHE D. and SAUTIERE P. (1983)

Eur.J.Biochem., 135, 113-121.

The Amino-Acid Sequence of Histone H2A from Cuttlefish *Sepia officinalis*

Danièle WOUTERS-TYROU, Annie MARTIN-PONTHIEU, Gilbert BRIAND, Pierre SAUTIERE, and Gérard BISERTE

Unité 124 de l'Institut National de la Santé et de la Recherche Médicale, Lille

(Received December 3, 1981 / February 3, 1982)

The amino acid sequence of cuttlefish testis histone H2A (124 residues) was established from structural data obtained by automated sequencing of large peptides generated by the cleavage of the protein with V8 staphylococcal protease or by limited chymotryptic hydrolysis.

Compared to the calf thymus homologous histone, cuttlefish H2A shows 14 substitutions (most of them conservative) and 5 deletions.

Extensive evolutionary changes were mainly observed in the basic amino-terminal and carboxy-terminal regions of the molecule, which are the primary DNA-binding sites. Few punctual changes are observed in the central region (residues 18–118), which interacts strongly with histone H2B to form the dimer H2A-H2B.

In the core particle histones H2A and H2B appear more variable throughout evolution than histones H3 and H4 (for recent reviews, see [1,2]). Structural studies of histones H2B from different species widely separated on the evolutionary scale [3–11] have shown that extensive structural changes (point mutations, deletions and insertions) occur mainly in the amino-terminal part of the molecule. On the other hand, the carboxy two-thirds of the protein have been highly conserved during evolution.

Similarly the structural variations of histone H2A have been studied in our laboratory for many years, in a number of species taken among vertebrates: calf [12,13], rat [14], chicken [15]; and marine invertebrates: sea-urchin [16], starfish, cuttlefish and sipunculus. These studies show that most of the changes occur in the highly basic amino-terminal and carboxy-terminal parts of the protein: the central region of the molecule (residues 18–118), which is involved in strong protein-protein interactions between histones H2A and H2B, is highly conserved from species to species.

Cuttlefish and sipunculus are located on the branch of an evolutionary tree which led to insects whereas echinoderms are located on the branch which led to vertebrates [17].

Cuttlefish (*Sepia officinalis*) is a cephalopod which exhibits an intermediary type of evolution. It is, in some ways, closer to vertebrates than to other marine invertebrates, such as echinoderms: the testis is well differentiated and the organisation of the eye, for example, is highly elaborated.

Here we present the amino acid sequence of histone H2A from cuttlefish testis. The alignment of the 124 residues of cuttlefish histone H2A was established from the structural data obtained essentially from large peptides generated by hydrolysis of the protein with V8 staphylococcal protease and chymotrypsin.

Abbreviations. TosLysCH₂Cl, *N*-tosyllysyl chloromethane; PhMeSO₂F, phenylmethylsulfonyl fluoride; *i*Pr₂P-F, diisopropylfluorophosphate; dansyl, 5-dimethylaminonaphthalene-1-sulphonyl.

Enzymes (IUB Recommendations 1978). *Staphylococcus aureus* protease (EC 3.4.21.19); α -chymotrypsin (EC 3.4.21.1); clostripain (EC 3.4.22.8); thermolysin (EC 3.4.24.4); aminopeptidase (microsomal) (EC 3.4.11.2); carboxypeptidase A (EC 3.4.17.1); carboxypeptidase B (EC 3.4.17.2).

MATERIALS AND METHODS

Materials

Mature cuttlefish testes were frozen in solid CO₂ after excision and stored at –20°C until use.

Staphylococcus aureus V8 protease was from Miles laboratories. Clostripain was purchased from Precibio. Thermolysin and TosLysCH₂Cl-chymotrypsin were from Merck. Carboxypeptidase A (treated with PhMeSO₂F) was obtained from Worthington. Carboxypeptidase B was from Boehringer and aminopeptidase M (microsomal) from Rohm and Haas.

Dimethylallylamine, propane-1-ol and dithioerythritol were sequanal grade from Pierce. Benzene and chloro-1-butane for sequential analysis were purchased from S.D.S. (Peypin, France). Phenylisothiocyanate, heptane and heptafluorobutyric acid were sequanal grade from Beckman. Polybrene was purchased from Aldrich and the dipeptide Gly-Gly from Sigma. All other reagents were of the highest purity available.

Isolation of Histone H2A

Chromatin was isolated from testes as described by Goodwin et al. [18], in presence of 0.1 mM *i*Pr₂P-F, and the crude fraction F2a2 was extracted from chromatin either by method 2 of Johns [19] or by the mixture ethanol/guanidine chloride [20]. The extracts were further fractionated by acetone precipitation. Histone H2A was isolated from fraction F2a2 by gel filtration chromatography on a Bio-Gel P10 column (150 × 5 cm) equilibrated and eluted with 0.01 M HCl, saturated with chloroform.

The purity of the protein was assessed by polyacrylamide gel electrophoresis in 0.9 M acetic acid containing 6.25 M urea [21], using a 17% acrylamide concentration.

Amino acid analyses of histone samples were performed as described previously [22].

Structural Studies

Identification of the Blocking Group at the Amino-Terminal End. Hydrazinolysis of the amino-terminal tripeptide followed

by dansylation at low pH was used to identify the blocked amino terminus as described in [23]. The presence of dansyl-acetylhydrazide was established by thin-layer chromatography on micropolyamide sheets [24].

Carboxy-Terminal Analysis. Histone H2A (2 mg, 140 nM) was dissolved in 1.2 ml 0.2 M *N*-methylmorpholine acetate, pH 8.5, and digested at 37°C with carboxypeptidase B (43 µg). At 15 min and 120 min, aliquots were taken off, acidified with glacial acetic acid and freeze-dried. At 120 min, carboxypeptidase A was added (50 µg). At 135, 150, 180 and 240 min, aliquots were taken off and treated as above. The released amino acids were analysed on the amino acid analyser.

Enzymatic Hydrolyses. Histone H2A (100 mg, 7 µmol) was digested with V8 *S. aureus* protease in 0.05 M ammonium acetate (20 ml), at pH 4.0, and 37°C for 4.30 h or 24 h, using an enzyme-to-substrate ratio of 1:50 (w/w). The digestion was stopped by adding 1 ml glacial acetic acid. The NH₂-terminal and COOH-terminal fragments obtained were further hydrolysed by clostripain and thermolysin respectively. Hydrolysis by clostripain was performed in 0.1 M potassium phosphate, pH 7.6, containing 194 mM dithiothreitol, at 25°C, for 4 h, using an enzyme-to-substrate ratio of 1:150 (w/w). Hydrolysis was stopped by lowering the pH to 3.5 with formic acid. Hydrolysis by thermolysin was carried out in 0.1 M *N*-methylmorpholine acetate, pH 8.0, at 40°C for 30 min, using an enzyme-to-substrate ratio of 1:100 (w/w). The chymotryptic digestion of histone H2A was performed in 0.5 M NaCl, 0.1 M ammonium bicarbonate, pH 8.3, at room temperature, for 1 h, using an enzyme-to-substrate ratio of 1:1000 (w/w). The enzymatic digests were freeze-dried and submitted to fractionation.

Fractionation of Peptides. The peptides obtained by hydrolysis of histone H2A with V8 staphylococcal protease were separated by chromatography on a Biogel P10 column (150 × 5 cm) equilibrated and eluted with 0.01 M HCl, saturated with chloroform. The peptides obtained by hydrolysis of the NH₂-terminal and COOH-terminal fragments of histone H2A with clostripain or thermolysin, were separated by ion-exchange chromatography on a Chromobeads P column (Technicon Corporation) as described in [16]. The chymotryptic hydrolysate was fractionated on a Sephadex G-100 column (180 × 5 cm) equilibrated and eluted with 0.01 M HCl, saturated with chloroform.

Automated Sequence Analysis. Automated Edman degradation of large peptides was performed in a Beckman 890 C Sequencer using a dimethylbenzylamine program (082773) [25] or a dimethylallylamine program (102974) in the presence of polybrene [26, 27].

Identification of Phenylthiohydantoin Derivatives of Amino Acids. The phenylthiohydantoin derivatives of amino acids were identified by high-pressure liquid chromatography on a column of C₁₈ Microbondapak (Waters Associates), using a Hewlett Packard liquid chromatograph 1084 B [25]. The phenylthiohydantoin derivatives of methionine, valine, phenylalanine and isoleucine were separated by high-pressure liquid chromatography on a column of cyanopropylsilane, according to a method derived from [28] or identified by gas chromatography as described in [29]. Small peptides were sequenced by manual Edman degradation [30], with identification of phenylthiohydantoin derivatives by high-pressure liquid chromatography or dansylation of the residual peptide [31]. Amide groups were determined as described in [16].

Nomenclature of Peptides

Peptides obtained by cleavage of the protein at glutamyl residues with the V8 staphylococcal protease are designated SP-. Chymotryptic, thermolysin and clostripain peptides were designated respectively by Ch-, Th- or Cl-.

Peptides were numbered according to their position in the sequence of protein.

RESULTS

Cuttlefish crude fraction F2a2 contains essentially histone H2A, slightly contaminated by histones H3 and H4, as shown on polyacrylamide gel electrophoresis (Fig. 1B, slot 3). When submitted to chromatography on a Biogel P10 column, F2a2 is resolved into three peaks (Fig. 1A). The peak 1 corresponds to histone H2A, free of contamination (Fig. 1B, slot 4). Contrary to calf histones H2A and H2B, cuttlefish histones H2A and H2B are separated from each other of polyacrylamide gel electrophoresis in 6.25 M urea, 0.9 M acetic acid [21] (Fig. 1B, slots 1 and 2).

As shown in Table 1, cuttlefish H2A contains less threonine and histidine and more serine than calf H2A [13]. Similar characteristics are also exhibited by sea-urchin histone H2A [16]. On the other hand, the presence of a methionine residue and the lack of a leucine residue suggest a substitution Leu → Met as was observed for vertebrate histone variant H2A2 [32, 33]. Cuttlefish histone H2A lacks a free amino-terminal group, as established by dansylation [31]. Similar observation has been made for all histones H2A isolated from animal species [1, 2]. The blocking group was identified as acetyl after hydrazinolysis followed by dansylation of the amino terminal tripeptide *N*-acetyl-Ser-Gly-Arg generated by hydrolysis of the fragment SP-1 with clostripain.

The carboxy-terminal sequence of cuttlefish histone H2A was determined as (Ala. Lys)-Ala-Lys(OH) after digestion of the protein with carboxypeptidases B and A.

With the technology of automated sequencing, the strategy used for the elucidation of the sequence of cuttlefish histone H2A was intended to produce a minimum of large adjacent or overlapping fragments by highly specific or limited enzymatic cleavages, in opposition to the previous strategy adopted in our laboratory for sequencing histones H2A from different sources [13–16]. Thus, cuttlefish histone H2A was first cleaved at glutamyl residues by the V8 staphylococcal protease, as shown in Fig. 2. Five peptides were obtained, which were separated on a Bio-Gel P10 column (Fig. 3). Their amino acid composition is presented in Table 1.

The peptide SP-1 contains most of the glycine and serine residues present in the H2A molecule. These characteristics allowed us to assign unequivocally the peptide SP-1 to the amino-terminal position in the sequence of the protein. Surprisingly the aminoterminal end of this peptide was found partially deacetylated after a 4-h hydrolysis of the protein with the V8 staphylococcal protease. Similarly the amino-terminal end of histone H4, which is blocked by acetylation, was found partially deacetylated by the V8 staphylococcal protease [34].

From these observations it was concluded that a deacetylase activity, strictly limited to the α-acetylamino group, was present in the V8 staphylococcal protease. It was therefore attractive for structural purposes to hydrolyse the cuttlefish histone H2A with the protease over a period of 24 h in order to obtain the highest rate of deacetylation of the amino-terminal fragment SP-1. This fragment was successfully

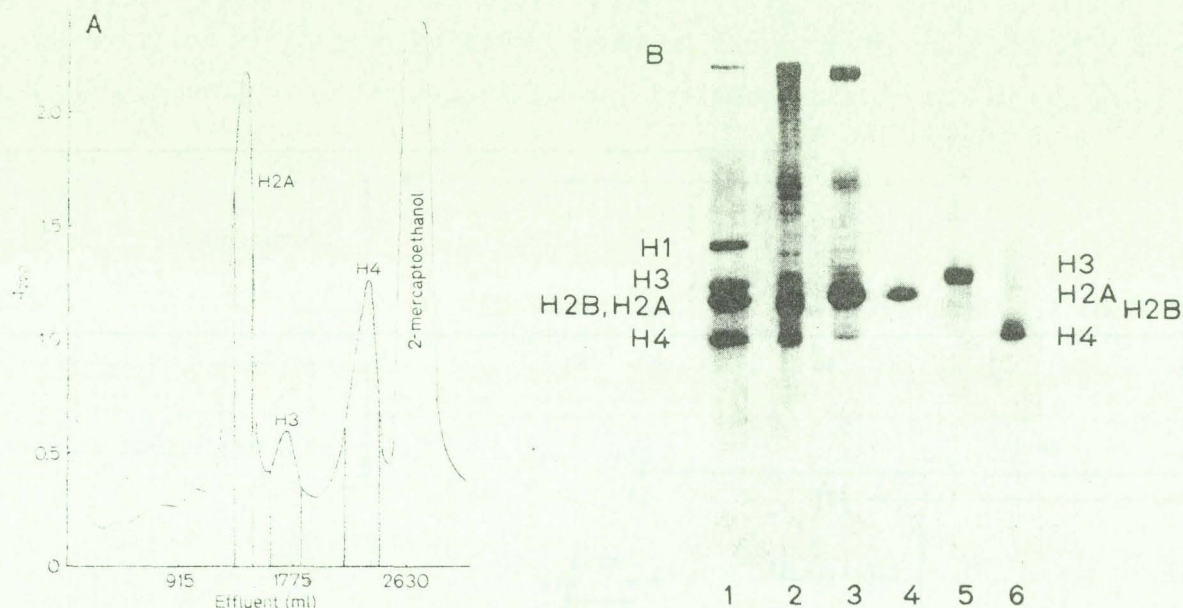


Fig. 4. Biogel P10 chromatography of crude fraction F2a2 from cuttlefish testis chromatin. (B) Electrophoretic analysis of fractions obtained after Biogel P10 chromatography of crude fraction F2a2 from cuttlefish testis chromatin. (A) 150 mg material were dissolved in 6 ml 6 M guanidine chloride, 0.1 M Tris, pH 8.0, 5% mercaptoethanol and left under nitrogen at room temperature for 24 h before chromatography. The solution was then centrifuged and applied to a Biogel P10 column (150 × 5 cm) equilibrated and eluted with 0.01 M HCl, saturated with chloroform. The flow rate was 60 ml/h and 17-ml fractions were collected. The absorbance at 220 nm was measured. (B) Electrophoresis was performed on polyacrylamide slab gel (15 × 10 cm) at 22 mA for 3 h, at pH 3.2, in 6.25 M urea [21]. Migration was from top to bottom. Slabs were stained for 60 min with 0.5% Coomassie blue R in acetic acid ethanol/water (1:2:7, by vol.) and destained by diffusion in the mixture acetic acid/ethanol/water (7:20:73, by vol.). Slot 1, whole histone from calf thymus; slot 2, whole histone from cuttlefish testis; slot 3, crude histone fraction F2a2; slots 4–6, peaks 1–3 respectively from Fig. 1A.

Table 1. Amino acid composition of cuttlefish histone H2A and of peptides used for the sequence determination (expressed as number of residues per mole of protein or peptide)

Values in parentheses are residues per molecule derived from the final sequence. Values for threonine and serine, in the protein, are zero-time extrapolations

Amino acids	Histone H2A from cuttle- fish testis	SP-1 (residues 1–55)	SP-2 (residues 56–63)	SP-3 (residues 64–124)	SP-3a (residues 64–91)	SP-3b (residues 92–124)	Ch-1 (residues 39–124)	Cl-12 (residues 35–41)
Aspartic acid	8.9 (8)	1.4 (1)		7.2 (7)	5.1 (5)	2.0 (2)	7.0 (7)	0.8 (1)
Threonine	3.5 (3)	0.9 (1)		2.0 (2)		1.8 (2)	1.8 (2)	
Serine	6.8 (6)	3.1 (4)		2.2 (2)	1.0 (1)	1.1 (1)	2.1 (2)	
Glutamic acid	11.2 (11)	2.9 (3)	2.1 (2)	6.5 (6)	3.0 (3)	3.1 (3)	9.5 (10)	0.9 (1)
Proline	5.4 (5)	2.0 (2)		3.0 (3)	1.2 (1)	1.7 (2)	3.9 (4)	
Glycine	14.2 (14)	8.7 (10)		4.4 (4)	1.2 (1)	3.0 (3)	6.4 (6)	1.0 (1)
Alanine	15.8 (16)	5.2 (6)	1.9 (2)	7.9 (8)	3.6 (4)	4.0 (4)	14.6 (15)	1.0 (1)
Valine	9.0 (9)	4.2 (5)	1.0 (1)	3.1 (3)		3.3 (3)	6.5 (7)	
Methionine	1.0 (1)	0.7 (1)					0.7 (1)	
Isoleucine	6.0 (6)	0.9 (1)		4.0 (5)	2.2 (3)	2.0 (2)	3.9 (5)*	
Leucine	14.9 (15)	4.0 (4)	1.9 (2)	9.4 (9)	3.5 (3)	5.7 (6)	12.2 (12)	
Tyrosine	3.2 (3)	1.8 (2)	1.2 (1)				1.8 (2)	0.7 (1)
Phenylalanine	1.2 (1)	1.0 (1)					0.0	
Histidine	2.1 (2)	0.8 (1)		1.0 (1)	1.0 (1)		1.0 (1)	
Lysine	13.6 (13)	7.0 (6)		8.5 (7)	2.2 (2)	4.8 (5)	7.0 (7)	1.1 (1)
Arginine	10.6 (11)	6.5 (7)		4.0 (4)	4.0 (4)		5.2 (5)	1.3 (1)
Number of residues	124	55	8	61	28	33	86	7

* This value is low due to incomplete cleavage of an Ile-Ile bond.

submitted to automated Edman degradation over 37 steps (Table 2, Fig. 5).

Furthermore, the sequence of the fragment SP-1 was confirmed by the structural data obtained from peptides generated by hydrolysis of SP-1 with clostripain (Table 3). In that

experiment the clostripain, of which the specificity should have been limited to arginyl bonds, was found to cleave the arginyl and lysyl bonds as well.

The peptides Cl-10 and Cl-13, corresponding to the sequences 29–31 and 42–55 respectively, were lost during

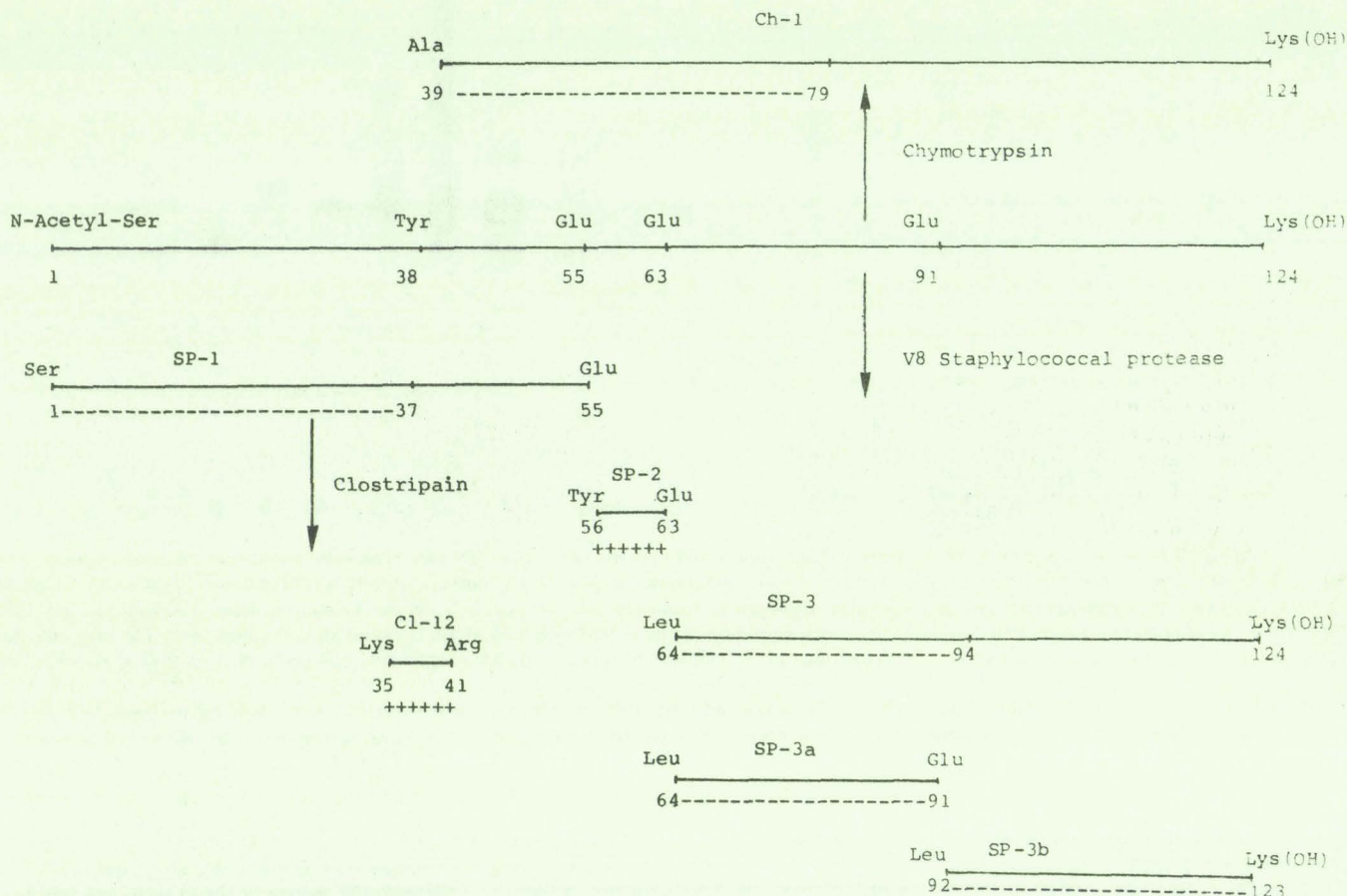


Fig. 2. Strategy used for the determination of the amino acid sequence of the cuttlefish histone H2A. SP, V8 staphylococcal protease peptides; Ch, chymotryptic peptide; Cl, clostripain peptide. (---) Automated Edman degradation; (++) manual Edman degradation

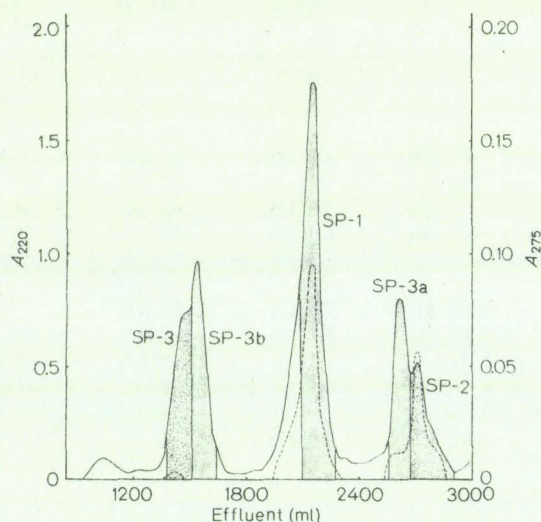


Fig. 3. Elution pattern of peptides generated by digestion of cuttlefish histone H2A with *Staphylococcus aureus* V8 protease (SP-peptides) and fractionated on Bio-Gel P10. The column (150 × 5 cm) was equilibrated and eluted with 0.01 M HCl, saturated with chloroform. Fractions of 12 ml were collected at a flow rate of 72 ml/h. Absorbance was measured at 220 nm (—) and 275 nm (---)

the fractionation of the clostripain hydrolysate by ion-exchange chromatography on Chromobeads P.

The peptide SP-2 (Fig. 5) containing one residue of tyrosine probably originates from the central hydrophobic region

of the histone molecule, by comparison with the sequence of calf histone H2A [13].

The peptide SP-3 (61 residues) corresponds to the carboxy-terminal half of cuttlefish histone H2A. It was found in the 4-h V8 staphylococcal protease hydrolysate of the protein, together with peptides SP-3a (28 residues) and SP-3b (33 residues), generated by a partial cleavage of the glutamyl bond at position 91 (Fig. 2 and 3, Table 1). The carboxy-terminal position of peptide SP-3b in the sequence of the cuttlefish histone H2A was ascertained by digestion of the peptide with carboxypeptidases B and A; 2 mol lysine and 2 mol alanine mol peptide were released. The partial sequence deduced from these results was as follows: (Lys, Ala)-Ala-Lys(OH); it corresponds to the carboxy-terminal sequence of the protein.

The peptides SP-3, SP-3a and SP-3b were submitted to automated Edman degradation (Tables 4–6 and Fig. 5). In addition, the peptide SP-3 was hydrolysed with thermolysin. The thermolysin peptides were fractionated on a Chromobeads P column. Seven main peptides, accounting for 35 of the 61 residues present in the peptide SP-3, were obtained, together with eight derivative peptides produced by partial cleavage of susceptible bonds (Table 7 and Fig. 5). These data allowed us to align the 61 C-terminal residues of protein (residues 64–124).

Chymotryptic Peptide Ch-1

The only region of the protein which remained unknown was the C-terminal part (residues 42–56) of the N-terminal

Table 2. Automated Edman degradation of peptide SP-1

Double identification of phenylthiohydantoin of amino acids was performed by high-pressure liquid chromatography on columns of C₁₈ Microbondapak and of cyanopropylsilane [25, 28]

Cycle number	Residue number	Residue	Yield nmol
1	1	Ser	48
2	2	Gly	175
3	3	Arg	—
4	4	Gly	171
5	5	Lys	97
6	6	Gly	97
7	7	Gly	105
8	8	Lys	80
9	9	Val	125
10	10	Lys	77
11	11	Gly	81
12	12	Lys	69
13	13	Ser	11
14	14	Lys	54
15	15	Thr	15
16	16	Arg	86
17	17	Ser	9
18	18	Ser	17
19	19	Arg	84
20	20	Ala	37
21	21	Gly	38
22	22	Leu	—
23	23	Gln	44
24	24	Phe	23
25	25	Pro	19
26	26	Val	78
27	27	Gly	32
28	28	Arg	38
29	29	Ile	7
30	30	His	8
31	31	Arg	37
32	32	Leu	16
33	33	Leu	26
34	34	Arg	31
35	35	Lys	13
36	36	Gly	21
37	37	Asn	7

Table 4. Automated Edman degradation of peptide SP-3

TLC, thin-layer chromatography [35, 36]; HPLC, see legend to Table 2

Cycle number	Residue number	Residue	Yield nmol	Identification
1	64	Leu	445	TLC, HPLC
2	65	Ala	385	TLC, HPLC
3	66	Gly	345	TLC, HPLC
4	67	Asn	254	TLC, HPLC
5	68	Ala	295	TLC, HPLC
6	69	Ala	251	TLC, HPLC
7	70	Arg	—	TLC, HPLC
8	71	Asp	270	TLC, HPLC
9	72	Asn	204	TLC, HPLC
10	73	Lys	160	TLC, HPLC
11	74	Lys	172	TLC, HPLC
12	75	Ser	22	TLC, HPLC
13	76	Arg	—	TLC, HPLC
14	77	Ile	143	TLC, HPLC
15	78	Ile	150	TLC, HPLC
16	79	Pro	55	TLC, HPLC
17	80	Arg	—	TLC, HPLC
18	81	His	—	TLC, HPLC
19	82	Leu	31	TLC, HPLC
20	83	Gln	45	TLC, HPLC
21	84	Leu	45	TLC, HPLC
22	85	Ala	24	TLC, HPLC
23	86	Ile	11	TLC, HPLC
24	87	Arg	—	TLC, HPLC
25	88	Asn	23	TLC, HPLC
26	89	Asp	16	TLC, HPLC
27	90	Glu	16	TLC, HPLC
28	91	Glu	25	TLC, HPLC
29	92	Leu	10	TLC, HPLC
30	93	Asn	—	TLC, HPLC
31	94	Lys	5	TLC, HPLC
32	95	Leu	12	TLC, HPLC
33	96	Leu	15	TLC, HPLC

Table 3. Amino acid composition of peptides obtained by hydrolysis of peptide SP-1 with clostripain

Amino acids	Cl-1	Cl-1a	Cl-2	Cl-5	Cl-3	Cl-4	Cl-6	Cl-7	Cl-8	Cl-9	Cl-11	Cl-12
Aspartic acid												0.8
Threonine								1.0				
Serine	0.9	0.9					1.1		1.9			
Glutamic acid										1.1		0.9
Proline										1.1		
Glycine	1.1	1.1	1.1		2.1					2.1		1.0
Alanine										1.0		1.0
Valine						1.0				0.9		
Methionine												
Isoleucine												
Leucine										0.9	1.8	
Tyrosine												0.7
Phenylalanine										0.9		
Lysine			1.0		1.0	1.0	1.0					1.1
Histidine												
Arginine	0.9	1.1						1.1	1.1	0.9	1.0	1.3
Total	3	3	2		3	2	2	2	3	9	3	6
N-Terminal	blocked	Ser	Gly		Gly	Val	Ser	The	Ser	Ala	Leu	Lys

Table 5. Automated Edman degradation of peptide SP-3a
GC, identification by gas chromatography [29]; HPLC, see legend to Table 2

Cycle number	Residue number	Residue	Yield nmol	Identification
1	64	Leu	546	HPLC
2	65	Ala	598	HPLC
3	66	Gly	550	HPLC
4	67	Asn	543	HPLC
5	68	Ala	573	HPLC
6	69	Ala	554	HPLC
7	70	Arg	362	HPLC
8	71	Asp	442	HPLC
9	72	Asn	504	HPLC
10	73	Lys	510	HPLC
11	74	Lys	433	HPLC
12	75	Ser	165	HPLC
13	76	Arg	280	HPLC
14	77	Ile	286	HPLC, GC
15	78	Ile	303	HPLC, GC
16	79	Pro	424	HPLC
17	80	Arg	145	HPLC
18	81	His	100	HPLC
19	82	Leu	210	HPLC
20	83	Gln	188	HPLC
21	84	Leu	191	HPLC
22	85	Ala	133	HPLC
23	86	Ile	110	HPLC, GC
24	87	Arg	129	HPLC
25	88	Asn	69	HPLC
26	89	Asp	53	HPLC

peptide SP-1. In addition, we were lacking overlapping sequences to align the peptides SP-1, SP-2 and SP-3 (see Fig. 2).

Since we had observed previously that chymotrypsin failed to cleave the susceptible bonds in the central part (residues 40–85) of the calf H2A molecule [13], the cuttlefish H2A was submitted to a limited hydrolysis with chymotrypsin in order to generate a fragment overlapping the unknown sequence of the protein. Hydrolysis was therefore performed at high ionic strength (0.5 M), which allows the protein to have a maximum of secondary and tertiary conformation (globular structure).

The elution diagram of the limited chymotryptic hydrolysate fractionated on Sephadex G-100 is presented in Fig. 4. Fraction 3 corresponds to a large peptide. Ch-1, which appears slightly contaminated by a larger fragment on polyacrylamide gel electrophoresis (gel not shown).

The characteristics of the amino acid composition of the peptide Ch-1 (Table 1) clearly indicate that this peptide (containing 86 residues) corresponds approximately to the C-terminal two-thirds of the protein. 41 steps of automated Edman degradation were performed on the peptide Ch-1. The obtention of two phenylthiohydantoin of amino acids at each step, in the approximate ratio of 3:1, revealed that the peptide Ch-1 was contaminated by a minor peptide Ch-1a obtained by the cleavage of the protein at histidine 30. Since the amino-terminal sequence of peptide Ch-1a was known from peptide SP-1, the sequence of peptide Ch-1 could be easily deduced (Table 8).

From these results, the peptide Ch-1 was shown to be generated by the chymotryptic cleavage of the tyrosyl bond at position 38 in the cuttlefish histone H2A (Fig. 5). Thus

Table 6. Automated Edman degradation of peptide SP-3b
AA, identification after back-hydrolysis of amino acid phenylthiohydantoin; HPLC, GC, see legends to Tables 2 and 5

Amino acids	Residue number	Residue	Yield nmol	Identification
1	93	Leu	317	HPLC
2	94	Asn	460	HPLC
3	95	Lys	374	HPLC
4	96	Leu	264	HPLC
5	97	Leu	235	HPLC
6	98	Ser	37	HPLC
7	99	Gly	264	HPLC
8	100	Val	235	HPLC, GC
9	101	Thr	69	HPLC
10	102	Ile	87	HPLC, GC
11	103	Ala	172	HPLC
12	104	Gln	102	HPLC
13	105	Gly	141	HPLC
14	106	Gly	151	HPLC
15	107	Val	182	HPLC, GC
16	108	Leu	89	HPLC
17	109	Pro	99	HPLC
18	110	Asn	45	HPLC
19	111	Ile	27	HPLC, GC
20	112	Gln	66	HPLC
21	113	Ala	78	HPLC
22	114	Val	84	HPLC, GC
23	115	Leu	36	HPLC, GC
24	116	Leu	43	HPLC
25	117	Pro	23	HPLC
26	118	Lys	26	HPLC
27	119	Lys	28	HPLC
28	120	Thr	4	HPLC, AA
29	121	Gln	—	HPLC, AA
30	122	Lys	15	HPLC
31	123	Ala	12	HPLC

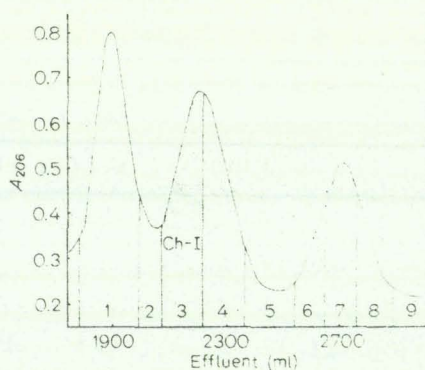


Fig. 4. Elution pattern of peptides generated by a limited chymotryptic hydrolysis of cuttlefish histone H2A. The peptides are fractionated on a column of Sephadex G-100 (180 × 5 cm) equilibrated and eluted with 0.01 M HCl, saturated with chloroform. Fractions of 16 ml were collected at a flow rate of 48 ml/h

the structural study of the peptide Ch-1 gave us the necessary informations to establish the complete amino-acid sequence of cuttlefish histone H2A (Fig. 5).

DISCUSSION

The amino acid sequence of cuttlefish histone H2A (124 residues) shows, by comparison with that of calf thymus

Table 7. Amino acid composition of thermolysin peptides obtained from peptide SP-3

Amino acids	Th-1	Th-2	Th-3	Th-3a	Th-3b	Th-3c	Th-4	Th-4a	Th-4b	Th-4c	Th-5	Th-6	Th-6a	Th-7	Th-7a
Aspartic acid			2.9	2.7	1.0	1.2									
Threonine							0.9	0.8		1.1		1.0	1.0		
Serine							1.0	0.8	1.0						
Glutamic acid	1.0		2.0	1.9							1.0	1.0	1.0		
Proline												1.2	1.2		
Glycine							1.2	1.1	1.0						
Alanine		1.1									1.0		1.0	1.9	1.0
Valine							1.0	0.8		1.0		0.6			
Methionine															
Isoleucine			0.9	0.8							0.9				
Leucine	1.1	1.0	2.0	1.2	0.8	1.7	1.6	1.0	1.0			1.7	1.7		
Tyrosine															
Phenylalanine															
Lysine			1.1	1.0	0.9	1.0						3.1	3.1	1.0	1.2
Histidine															
Arginine			1.0	1.2											
Total	2	2	10	9	3	4	6	5	3	2	3	9	9	3	2
N-Terminal	Leu	Leu	Ile	Ile	Leu	Leu	Leu	Leu	Leu	Val	Ile	Val	Leu	Ala	Ala
Amino-peptidase	Gln														

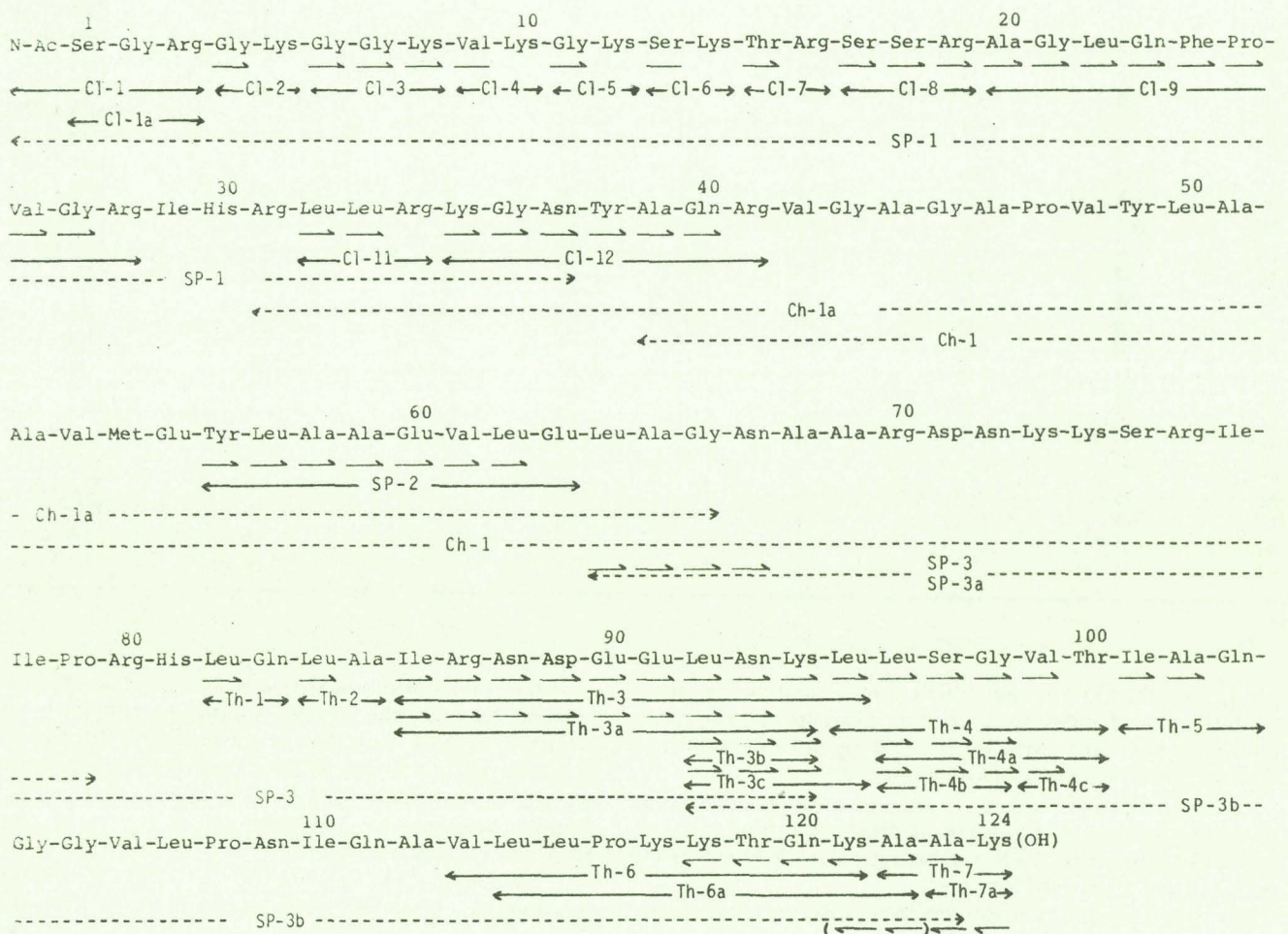


Fig. 5. Alignment of peptides and amino acid sequence of cuttlefish histone H2A. Methods used for determination of the sequence of peptides are indicated as follows: →, manual Edman degradation; —, dansylation; ←, carboxypeptidase hydrolysis; ---, automated Edman degradation

Table 8. Automated Edman degradation of peptide Ch-1
See legend to Table 2

Cycle number	Major peptide Ch-1			Minor peptide Ch-1a			Identification
	residue number	residue	yield	residue number	residue	yield	
			nmol			nmol	
1	39	Ala	144	31	Arg	51	HPLC
2	40	Gln	164	32	Leu	56	HPLC
3	41	Arg	151	33	Leu	64	HPLC
4	42	Val	119	34	Arg	65	HPLC
5	43	Gly	111	35	Lys	47	HPLC
6	44	Ala	99	36	Gly	59	HPLC
7	45	Gly	115	37	Asn	30	HPLC
8	46	Ala	114	38	Tyr	41	HPLC
9	47	Pro	48	39	Ala	39	HPLC
10	48	Val	119	40	Gln	47	HPLC
11	49	Tyr	53	41	Arg	52	HPLC
12	50	Leu	55	42	Val	52	HPLC
13	51	Ala	106	43	Gly	62	HPLC
14	52	Ala	151	44	Ala	—	HPLC
15	53	Val	110	45	Gly	45	HPLC
16	54	Met	95	46	Ala	41	HPLC
17	55	Glu	41	47	Pro	25	HPLC
18	56	Tyr	37	48			HPLC
19	57	Leu	57	49	Tyr	32	HPLC
20	58	Ala	66	50	Leu	30	HPLC
21	59	Ala	132	51	Ala	—	HPLC
22	60	Glu	48	52	Ala	88	HPLC
23	61	Val	115	53	Val	—	HPLC
24	62	Leu	51	54	Met	37	HPLC
25	63	Glu	24	55	Glu	—	HPLC
26	64	Leu	58	56	Tyr	24	HPLC
27	65	Ala	47	57	Leu	22	HPLC
28	66	Gly	27	58	Ala	20	HPLC
29	67	Asn	14	59	Ala	22	HPLC
30	68	Ala	53	60	Glu	—	HPLC
31	69	Ala	64	61	Val	29	HPLC
32	70	Arg	26	62	Leu	24	HPLC
33	71	Asp	24	63	Glu	20	HPLC
34	72	Asn	29	64	Leu	25	HPLC
35	73	Lys	20	65	Ala	16	HPLC
36	74	Lys	31	66	Gly	17	HPLC
37	75	Ser	5	67			HPLC
38	76	Arg	28	68	Ala	19	HPLC
39	77	Ile	9	69	Ala	22	HPLC
40	78	Ile	20	70	Arg	15	HPLC
41	79	Pro	5	71	Asp	9	HPLC

H2A, 14 substitutions and 5 deletions (Table 9). Extensive evolutionary changes were mainly observed in the basic amino-terminal and carboxy-terminal regions of the protein which are the primary DNA-binding sites.

All substitutions, except 4, are conservative. A residue of valine replaces a residue of alanine at position 10. This is the first time that such a substitution is observed in the NH₂-terminal region of histones.

The three other non-conservative substitutions involve the replacement of charged residues by neutral residues and are located at positions 41 (Glu → Gln), 99 (Lys → Gly) and 121 (Glu → Gln). Similarly, in sea-urchin H2A, the polar residues at positions 41, 99 and 121 were replaced by neutral residues [16]. Only the modification at position 99 (Lys → Gly) involves the change of two bases in the nucleotidic triplet of the gene.

Four substitutions and one deletion occur in the NH₂-terminal region of the protein (residues 1–18) but the overall electric charge of the region is not modified. The deletion of a residue of glutamine at position 6 has already been observed in sea-urchin H2A [16,37] and results in a total identity between the NH₂-terminal sequences of histones H2A and H4 (residues 1–8).

Few punctual changes are observed in the central region (residues 18–118). The mutation which occurs at position 30 (Val → Ile) has already been observed in wheat histone variants H2A1, H2A2 and H2A3 [38]. A residue of methionine was found at position 55, instead of a leucine residue, whereas in vertebrates histone variant H2A2 [32,33], or in histone H2A from sea-urchin embryo [39], a change Leu → Met occurred at position 51. At position 59, the charge Thr → Ala introduces a second sequence Leu-Leu-Ala-Ala.

Table 9. Comparison of the sequences of cuttlefish testis and calf thymus histones H2A. Deletions are indicated by —. Numbering in this table and in the text refers to sequence positions in calf thymus histone H2A

Species	Residues at position								
	6	10	11	12	14	30	41	55	59
Calf thymus [13]	Gln-Gly-Gly-Lys-Ala-Arg-Ala-Lys-Ala-Lys-					-Val-	-Glu-	-Leu-	-Thr-
Cuttlefish testis		Gly-Gly-Lys-Val-Lys-Gly-Lys-Ser-Lys-				-Ile-	-Gln-	-Met-	-Ala
Calf thymus [13]	62	76		98	99	121		129	
	-Ile-	-Thr-		-Gly-Lys-		-Glu-Ser-His-His-Lys-Ala-Lys-Gly-Lys(OH)			
Cuttlefish testis	-Val-	-Ser-		-Ser-Gly-		-Gln-	-----Lys-Ala-	-----Ala-Lys(OH)	

which is then found at positions 49–52 and 56–59 in the sequence of the cuttlefish histone H2A. At position 76 the replacement of a threonine residue by a serine residue introduces a potential site of phosphorylation in the basic cluster Lys-Lys-Ser-Arg (residues 74–77). In this sequence the serine residue is placed in such a way that it could be phosphorylated by a cyclic-AMP-dependent protein kinase [40, 41]. We must emphasize that two sites of phosphorylation have been identified in the highly basic amino-terminal sequence of calf histone H2A, at serine residues 1 and 19 [42].

Two substitutions and four deletions occur in the COOH-terminal region of the molecule (residues 118–124). They do not, however, modify the overall electrical charge of this region. The deletion of the two residues of histidine at positions 123 and 124 has also been found in histone H2A from different marine invertebrates: sea-urchins [16, 37], starfish, sipunculids, crustaceans (P. Sautière, unpublished results).

From our results and previous work [12–16, 37, 38, 43], we can conclude that the structural variability of histone H2A is limited to the NH₂-terminal (residues 1–18) and the COOH-terminal (residues 118–124) regions of the molecule. However, the variability of these highly basic regions does not appear to alter their biological function that are electrostatic interactions with the phosphate groups of DNA and even with non-histone phosphoproteins.

Moreover, the sites of the post-transcriptional modifications which can affect the histone H2A and modulate the histone-DNA interactions are conserved: acetylations on lysine-5 [44], phosphorylation on serine-1 and serine-19 [42] and ubiquitin binding on lysine-119 [45, 46]. On the other hand, in spite of the complete identity of the amino-terminal sequences of cuttlefish histones H2A and H4, the lysine residues at position 5 and 12, which are acetylated in histone H4 to an extent of 25.6% and 35.6%, respectively [34], were not found acetylated in histone H2A.

The few punctual changes observed in the central region (residues 18–118) are generally conservative and should not affect the conformation of the protein. As predicted from Chou and Fasman calculations [47], this region is highly structured (α helix and β sheets conformations) and interacts strongly with histone H2B to form the dimer H2A-H2B.

In all H2A studied at present time the sequence 21–29 is highly conserved and was shown to interact with the sequence 35–43 of histone H2B [48]. In fact from circular dichroism and NMR spectroscopy studies, the region 18–31 of histone H2A appears to be essential for the complex formation (M. Couppez and P. Sautière, unpublished results). The compact globular structure acquired by the histone H2A, when it is placed in a high-ionic-strength medium (e.g. 0.5 M), makes inaccessible to trypsin or chymotrypsin most of the

cleavage sites in the sequence 18–118. However, in cuttlefish H2A the bonds His³⁰-Arg³¹ and Tyr³⁸-Ala³⁹ were partially hydrolyzed by chymotrypsin. Indeed, these two bonds are located next to a β turn [47] and, therefore, are exposed to the enzyme. On the other hand the tyrosyl bonds at positions 49 and 56 are buried in the tertiary structure of the molecule and protected from chymotryptic hydrolysis as observed previously in calf histone H2A [13].

The authors are grateful to the *Institut Scientifique et Technique des Pêches Maritimes* and to Dr Alain Richard (*Institut de Biologie Maritime et Régionale*, 62930 Wimereux) who provided the cuttlefish. They acknowledge with pleasure the skillful technical assistance of Mrs D. Bélaïche, M. J. Dupire, T. Ernout and A. Hemez. This work was supported by grant 77.5.048.2 from the *Institut National de la Santé et de la Recherche Médicale* and by grant *ATP Chromatine* 2875 from the *Centre National de la Recherche Scientifique*.

REFERENCES

- Isenberg, I. (1979) *Annu. Rev. Biochem.* **48**, 159–191.
- von Holt, C., Strickland, W. N., Brandt, W. F. & Strickland, M. S. (1979) *FEBS Lett.* **100**, 201–218.
- Iwai, K., Hayashi, H. & Ishikawa, K. (1972) *J. Biochem. (Tokyo)* **72**, 357–367.
- Kootstra, A. & Bailey, G. S. (1976) *FEBS Lett.* **68**, 76–78.
- Strickland, M., Strickland, W. N., Brandt, W. F. & von Holt, C. (1977) *Eur. J. Biochem.* **77**, 263–275.
- Strickland, W. N., Strickland, M., Brandt, W. F. & von Holt, C. (1977) *Eur. J. Biochem.* **77**, 277–286.
- Strickland, M., Strickland, W. N., Brandt, W. F., Von Holt, C., Wittman-Liebold, B. & Lehman, A. (1978) *Eur. J. Biochem.* **89**, 443–452.
- Kootstra, A. & Bailey, G. S. (1978) *Biochemistry*, **17**, 2504–2510.
- Martínage, A., Mangeat, P., Sautière, P., Marchis-Mouren, G. & Biserte, G. (1979) *Biochimie (Paris)* **61**, 61–69.
- Van Helden, P., Strickland, W. N., Brandt, W. F. & von Holt, C. (1979) *Eur. J. Biochem.* **93**, 71–78.
- Elgin, S. C. R., Schilling, J. & Hood, L. E. (1979) *Biochemistry*, **18**, 5679–5685.
- Sautière, P., Tyrou, D., Laine, B., Mizon, J., Lambelin-Breynaert, M. D., Ruffin, P. & Biserte, G. (1972) *C.R. Hebd. Seances Acad. Sci.* **274**, 1422–1425.
- Sautière, P., Tyrou, D., Laine, B., Mizon, J., Ruffin, P. & Biserte, G. (1974) *Eur. J. Biochem.* **41**, 563–576.
- Laine, B., Sautière, P. & Biserte, G. (1976) *Biochemistry*, **15**, 1640–1645.
- Laine, B., Kmiecik, D., Sautière, P. & Biserte, G. (1978) *Biochimie (Paris)* **60**, 147–150.
- Wouters, D., Sautière, P. & Biserte, G. (1978) *Eur. J. Biochem.* **90**, 231–239.
- Dayhoff, M. O. (1972) *Atlas of Protein Sequence and Structure*, p. 15. National Biomedical Research Foundation, Silver Springs, MD.

18. Goodwin, G. H., Nicolas, R. H. & Johns, E. W. (1975) *Biochim. Biophys. Acta*, **405**, 280–291.
19. Johns, E. W. (1964) *Biochem. J.* **92**, 55–59.
20. Johns, E. W. (1967) *Biochem. J.* **105**, 611–614.
21. Panyim, S. & Chalkley, R. (1969) *Arch. Biochem. Biophys.* **130**, 337–346.
22. Martinage, A., Sautière, P., Kerckaert, J. P. & Biserte, G. (1976) *Biochim. Biophys. Acta*, **420**, 37–41.
23. Schmer, G. & Kriel, G. (1969) *Anal. Biochem.* **29**, 186–192.
24. Doolittle, R. F. (1977) in *Advanced Methods in Protein Sequence Determination* (Needleman, S. B., ed.) pp. 38–54, Springer, Berlin, Heidelberg, New York.
25. Hermann, J., Titani, K., Ericsson, L. H., Wade, R. D., Neurath, H. & Walsh, K. A. (1978) *Biochemistry*, **17**, 5672–5679.
26. Tarr, G. E., Beecher, J. F., Bell, M. & McKeaw, D. (1978) *Anal. Biochem.* **84**, 622–627.
27. Klapper, D. G., Wilde, C. E. & Capra, J. D. (1978) *Anal. Biochem.* **85**, 126–131.
28. Johnson, N. D., Hunkapiller, M. W. & Hood, L. E. (1979) *Anal. Biochem.* **100**, 335–338.
29. Pisano, J. J. & Bronzert, T. J. (1969) *J. Biol. Chem.* **244**, 5597–5607.
30. Edman, P. & Henschen, A. (1975) in *Protein Sequence Determination* (Needleman, S. B., ed.) 2nd edn, pp. 232–279, Springer, Berlin, Heidelberg, New York.
31. Hartley, B. S. (1970) *Biochem. J.* **119**, 805–822.
32. Blankstein, L. A., Stollar, B. D., Franklin, S. G., Zweidler, A. & Levy, S. B. (1977) *Biochemistry*, **16**, 4557–4562.
33. West, M. H. P. & Bonner, W. M. (1980) *Biochemistry*, **19**, 3238–3245.
34. Wouters-Tyrou, D., Martin-Ponthieu, A., Sautière, P. & Biserte, G. (1981) *FEBS Lett.* **128**, 195–200.
35. Cohen-Solal, M. & Bernard, J. L. (1973) *J. Chromatogr.* **80**, 140–143.
36. Roseau, G. & Pantel, P. (1969) *J. Chromatogr.* **44**, 392–395.
37. Strickland, W. N., Strickland, M. S., De Groot, P. C. & von Holt, C. (1980) *Eur. J. Biochem.* **109**, 151–158.
38. Rodrigues, J. de A., Brandt, W. F. & von Holt, C. (1979) *Biochim. Biophys. Acta*, **578**, 196–206.
39. Schaffner, W., Kunz, G., Daetwyler, H., Felford, J., Smith, H. O. & Birnstiel, M. L. (1978) *Cell*, **14**, 655–671.
40. 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.
41. MacLeod, A. R., Wong, N. C. W. & Dixon, G. H. (1977) *Eur. J. Biochem.* **78**, 281–291.
42. Martinage, A., Mangeat, P., Laine, B., Couppez, M., Sautière, P., Marchis-Mouren, G. & Biserte, G. (1980) *FEBS Lett.* **118**, 323–329.
43. Bailey, G. S. & Dixon, G. H. (1973) *J. Biol. Chem.* **248**, 5463–5472.
44. Candido, E. P. M. & Dixon, G. H. (1972) *J. Biol. Chem.* **247**, 3868–3873.
45. Goldknopf, I. L. & Busch, H. (1977) *Proc. Natl. Acad. Sci. U.S.A.* **74**, 864–868.
46. Hunt, L. T. & Dayhoff, M. O. (1977) *Biochem. Biophys. Res. Commun.* **74**, 650–655.
47. Chou, P. Y. & Fasman, G. D. (1974) *Biochemistry*, **13**, 211–221.
48. Delange, R. J., Williams, L. C. & Martinson, H. G. (1979) *Biochemistry*, **18**, 1942–1946.

D. Wouters-Tyrou, A. Martin-Ponthieu, G. Briand, P. Sautière, and G. Biserte, Unité de Recherches Ultrastructurales et Biochimiques sur les Cellules Normales et Cancéreuses (Unité 124 de l'Institut National de la Santé et de la Recherche Médicale), Institut de Recherches sur le Cancer, Cité Hospitalière, 2 Place de Verdun, F-59020 Lille-Cedex, Nord, France

Primary Structure of Histone H2A from Gonads of the Starfish *Asterias rubens*

Ariette MARTINAGE, Denise BELAICHE, Thierry DUPRESSOIR, and Pierre SAUTIERE

Unité 124 de l'Institut National de la Santé et de la Recherche Médicale, Lille

(Received August 25, November 2, 1982) — EJB 5936

The complete amino acid sequence (124 residues) of histone H2A from gonads of the starfish *Asterias rubens* has been established from automated sequence analyses of large fragments obtained by staphylococcal protease digestion of histone H2A and by limited hydrolysis of H2A-H2B complex with mouse submaxillary gland protease and from structural studies of peptides generated by enzymatic hydrolyses of these fragments or of the protein.

By comparison with calf homologous histone, the starfish histone H2A shows 5 deletions and 12 substitutions. Half of the substitutions are non-conservative.

Microheterogeneities were found at positions 18, 40 and 50 and result in the existence of at least two variants of starfish gonad histone H2A.

Structural studies of histones H2A from different species widely separated on the evolutionary scale [1-6] have shown that extensive structural modifications (point mutations, deletions and insertions) mainly occur in the highly basic N-terminal and C-terminal parts of the protein. The central region of the molecule (residues 18-118), which strongly interacts with the region 32-125 of histone H2B to form the dimer H2A-H2B in the nucleosome, appears generally conserved during evolution.

The higher degree of variability of histone H2B in Echinidae [7], one of the three families of the echinoderms has led us to investigate the primary structure of both histones H2A and H2B in Asterinidae.

In this paper we report the complete amino acid sequence of histone H2A from gonads of the starfish *Asterias rubens*.

MATERIALS AND METHODS

Materials and Animals

Marine starfish were obtained from the Institut de Biologie Maritime de Wimereux (France). After excision, gonads were frozen in solid CO₂ and kept at -20°C until use.

Staphylococcus aureus V8 protease was from Miles laboratories. Trypsin (treated with TosPheCH₂Cl), carboxypeptidases A and B (treated with PhMeSO₂F) and pepsin were obtained from Worthington. Thermolysin was from Merck. Arg-C endoproteinase from the submaxillary glands of mice was purchased from Boehringer.

Dimethylallylamine, propane-1-ol and dithioerythritol were sequential grade from Pierce. Benzene and chloro-1-butane for sequential analysis were purchased from SDS (Peypin, France). Phenylisothiocyanate, heptane and hepta-

Abbreviations TosPheCH₂Cl, N-tosylphenylalanyl chloromethane; PhMeSO₂F, phenylmethylsulfonyl fluoride.

Enzymes (IUB Recommendations 1978). *Staphylococcus aureus* protease (EC 3.4.21.19); trypsin (EC 3.4.21.4); pepsin (EC 3.4.23.1); thermolysin (EC 3.4.24.4); endoproteinase Arg-C (EC 3.4.-.-); carboxypeptidase A (EC 3.4.17.1); carboxypeptidase B (EC 3.4.17.2).

fluorobutyric acid were sequential grade from Beckman. Polybrene was purchased from Aldrich and the dipeptide Gly-Gly from Sigma. All other reagents were of the highest purity available.

Isolation of Histone H2A and of Histone Complex H2A-H2B

Chromatin was isolated from gonads as described in [8]. The crude fraction F2a2 was obtained from chromatin according to [9]. The crude fraction F2b was extracted from chromatin by method 2 of Johns [10]. Histone H2A was isolated from fraction F2a2 by gel filtration chromatography on Bio-Gel P10 (150 × 5 cm) equilibrated and eluted with 0.01 M HCl saturated with chloroform.

The dimer H2A-H2B was obtained from fraction F2b by ion-exchange chromatography on Bio-Rex 70 (100 × 2.5 cm) [11]. The column was equilibrated with 0.1 M sodium phosphate buffer (pH 6.8) containing 8% guanidine chloride and eluted with a linear gradient of the latter from 8% to 14% in the same buffer.

Analytical electrophoresis of histone H2A and of complex H2A-H2B in polyacrylamide slab gels was carried out at pH 3.2 in 6.25 M urea according to Panyim and Chalkley [12] using a 17% acrylamide concentration.

Amino acid analyses of histone samples were performed on a Beckman 119 CL amino acid analyser after hydrolysis in 6 M HCl (1 ml/mg protein) in a vacuum at 110°C for 24 h and 72 h with one drop of 1% phenol to avoid excessive degradation of tyrosine.

Structural Studies

Identification of the Blocking Group at the Amino-terminal End. This identification was made by high-resolution nuclear magnetic resonance spectroscopy of the amino-terminal tryptic peptide (residues 1-3) of histone H2A with a Bruker 270-MHz Fourier transform spectrometer. Spectra were obtained in ²H₂O at pH ≈ 2 and pH ≈ 11 in 5-mm tubes.

Carboxy-terminal Analysis. Histone H2A (1.4 mg, 100 nmol) was dissolved in 0.5 ml 0.2 M N-methylmorpho-

line acetate pH 8.0 and digested at 40°C with carboxypeptidase B (40 µg). At 30 min an aliquot was taken off, acidified with glacial acetic acid and freeze-dried. Carboxypeptidase A (40 µg) was then added. Aliquots were taken off at 45 min, 60 min and 90 min and treated as above. The released amino acids were analysed on the amino acid analyser. Digestion of peptides with carboxypeptidase A and/or B was performed in the same conditions.

Enzymatic Hydrolyses. Histone H2A (70 mg, 5 µmol) was hydrolysed with V8 staphylococcal protease in 0.05 M ammonium acetate (14 ml) at pH 4.0 and 37°C for 15 h, using an enzyme-to-substrate ratio of 1:50 (w/w). The N-terminal fragment (residues 1–55) so obtained was further hydrolysed by thermolysin in 0.1 M *N*-methylmorpholine acetate at pH 8.0 and 40°C for 2 h using an enzyme-to-substrate ratio of 1:60 (w/w). Hydrolysis of the protein (40 mg ≈ 3 µmol) with TosPheCH₂Cl-trypsin was carried out in 0.1 M *N*-methylmorpholine acetate at pH 8.0 and 37°C for 2 h using an enzyme-to-substrate ratio of 1:50 (w/w). Histone H2A (60 mg ≈ 4 µmol) was hydrolysed with pepsin in 5% formic acid at 37°C for 1 h using an enzyme-to-substrate ratio of 1:50 (w/w). The complex H2A-H2B was submitted to a limited hydrolysis with Arg-C endoproteinase in 0.1 M ammonium bicarbonate pH 8.3 containing 0.5 M NaCl at room temperature for 5 h using an enzyme-to-substrate ratio of 1:5000 (w/w). The enzymatic hydrolysates were freeze-dried and submitted to fractionation.

Fractionation of Peptides. The peptides obtained by hydrolysis of histone H2A with V8 staphylococcal protease and by hydrolysis of the complex H2A-H2B with Arg-C endoproteinase were separated by chromatography on a Bio-Gel P10 column (150 × 5 cm) equilibrated and eluted with 0.01 M HCl saturated with chloroform. When necessary, peptides were further purified by chromatography on a Sephadex

G-50F column (200 × 2.5 cm) equilibrated and eluted with 0.01 M HCl saturated with chloroform. Thermolysin, trypsin and pepsin hydrolysates were fractionated by ion-exchange chromatography on a Chromobeads P column (Technicon Corporation) with pyridine formate and pyridine acetate buffers [6]. The purity of the peptides was checked by polyacrylamide slab gel electrophoresis at pH 2.7 in 2.5 M urea [13] using a 17% acrylamide concentration, or by paper chromatography and electrophoresis [14].

Sequence of Peptides. Small peptides were sequenced by manual Edman degradation [15] and large peptides by automated Edman degradation using a Beckman 890 C sequencer with a dimethylallylamine program (102974) in the presence of polybrene [16,17]. Identification and quantification of phenylthiohydantoin derivatives of amino acids were performed by high-pressure liquid chromatography on a column of C18 µBondapak (Waters Associates) using a Hewlett-Packard liquid chromatograph 1084B [18]. Phenylthiohydantoin derivatives of Met, Val, Phe and Ile were identified on a Hewlett-Packard 5830A gas chromatograph [19].

Nomenclature of Peptides. Peptides generated by cleavage of the protein at glutamyl residues with the V8 staphylococcal protease were designated SP. Tryptic, thermolysin and pepsin peptides were designated T, Th and P respectively. Peptides obtained by limited hydrolysis of the complex H2A-H2B with the Arg-C endoproteinase were designated RP.

All peptides but pepsin peptides were numbered according to their position in the amino acid sequence of the protein.

RESULTS

The crude fraction F2a2 isolated from starfish gonad chromatin according to Johns [9] contains essentially histone

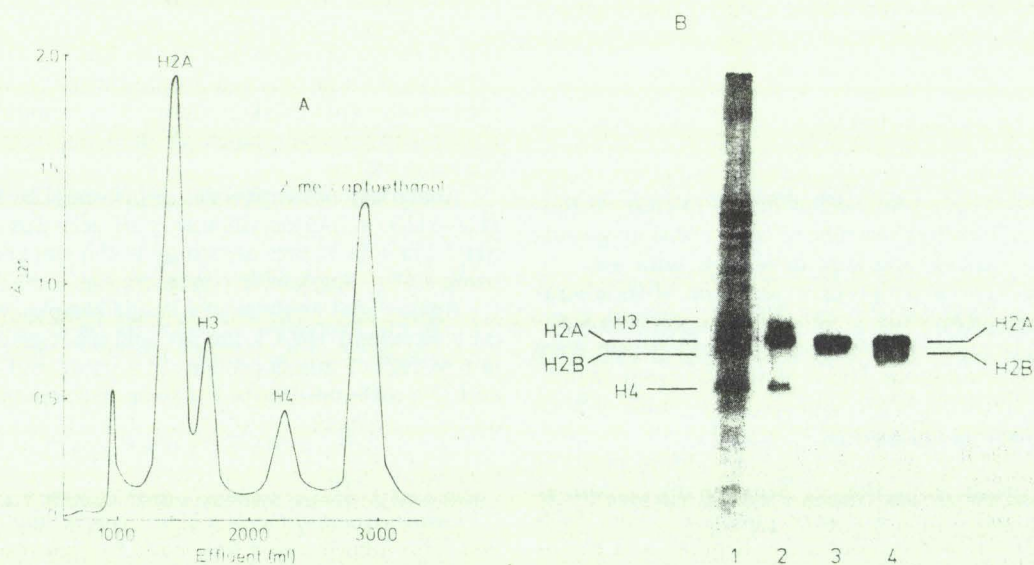


Fig. 1. (A) Biogel P10 chromatography of crude fraction F2a2 from starfish gonad chromatin. (B) Electrophoretic analysis of fractions obtained after chromatography on Biogel P10 and on Biorex 70 of crude fractions F2a2 and F2b from starfish gonad chromatin. (A) 150 mg of crude fraction F2a2 was dissolved in 6 ml 6 M guanidine chloride, 0.1 M Tris, pH 8.0, 5% mercaptoethanol and left under nitrogen at room temperature for 24 h before chromatography. The solution was then centrifuged and applied to a Biogel P10 column (150 × 5 cm) equilibrated and eluted with 0.01 M HCl saturated with chloroform. The flow rate was 60 ml/h and 17-ml fractions were collected. (B) Electrophoresis was performed on polyacrylamide slab gels (15 × 10 cm) at 22 mA for 3 h at pH 3.2 in 6.25 M urea [12]. Migration was from top to bottom. Slabs were stained for 60 min with 0.5% Coomassie blue R in acetic acid/ethanol/water (1:2:7, by vol.) and destained by diffusion in the mixture acetic acid/ethanol/water (7:20:73, by vol.). Lane 1: crude histone from starfish gonad; slot 2: crude histone fraction F2a2; slot 3: H2A from Fig. 1A; slot 4: dimer H2A-H2B obtained by ion-exchange chromatography on Biorex 70.

Table 1. Amino acid composition of starfish histone H2A and of peptides used for the sequence determination

Results are expressed as the number of residues/mol protein or peptide. Values in parentheses are residues/mol derived from the final sequence. Values for aspartic acid, glutamic acid and serine, in the protein, are zero-time extrapolations

Amino acid	Histone H2A		SP-1	SP-2	SP-3	SP-3b	RP-1
	calf [2]	starfish	(residues 1-55)	(residues 56-63)	(residues 64-124)	(residues 92-124)	(residues 35-124)
Aspartic acid	8	9.0 (9-10) ^a	1.2 (1-2) ^a		7.7 (8)	2.3 (2)	9.0 (9)
Threonine	5	2.9 (3)			2.8 (3)	1.9 (2)	3.0 (3)
Serine	4	4.2 (4-5) ^a	3.0 (3-4) ^a		1.1 (1)	1.1 (1)	1.8 (1)
Glutamic acid	12	10.4 (10)	2.5 (3)	2.1 (2)	5.8 (5)	2.8 (3)	9.3 (9)
Proline	5	4.8 (5)	2.2 (2)		3.1 (3)	2.1 (2)	3.5 (4)
Glycine	14	12.3 (13)	8.7 (9)		4.3 (4)	3.2 (3)	7.3 (7)
Alanine	16	18.8 (20-21) ^a	9.2 (9-10) ^a	2.0 (2)	8.8 (9)	5.0 (5)	16.7 (16)
Valine	8	7.6 (8)	5.0 (5)		3.0 (3)	3.0 (3)	5.9 (6)
Methionine		0.8 (1)	1.0 (1)				0.5 (1)
Isoleucine	6	4.6 (5)		0.8 (1)	3.9 (4)	2.0 (2)	4.7 (5)
Leucine	16	13.1 (13-14) ^a	3.1 (2-3) ^a	1.9 (2)	8.3 (9)	6.0 (6)	11.9 (12)
Proline	3	2.7 (3)	1.6 (2)	0.9 (1)			2.7 (3)
Phenylalanine	1	2.0 (2)	2.2 (2)				
Histidine	4	2.1 (2)	1.0 (1)		1.0 (1)		1.1 (1)
Lysine	14	11.5 (12)	4.9 (5)		6.8 (7)	5.0 (5)	8.5 (8)
Arginine	12	11.2 (12)	7.5 (8)		3.8 (4)		5.3 (5)
Number of residues	129	124	55	8	61	34	90

^a The values for aspartic acid, serine, alanine and leucine take into account the presence of microheterogeneities at positions 18 and 50 in the sequence (Fig. 7)



Fig. 2. 270-MHz NMR spectra of the N-terminal tryptic peptide Ac-Ser-Gly-Arg (residues 1-3) of starfish histone H2A in ²H₂O. (A) At pH 2.0; (B) at pH 11.

H2A slightly contaminated by histones H3 and H4 and non-histone proteins, as shown by polyacrylamide gel electrophoresis (Fig. 1B, slot 2).

Histone H2A was obtained free of contamination (Fig. 1B, slot 3) by chromatography of F2a2 on a Bio-Gel P10 column (Fig. 1A).

The dimer H2A-H2B was eluted at 11% guanidium chloride concentration upon ion-exchange chromatography of crude fraction F2b [11] on Bio-Rex 70 (Fig. 1B, slot 4). This complex was used to obtain by limited proteolytic cleavage

large fragments of histones H2A and H2B without prior isolation of individual histones. Contrary to calf homologous histones, starfish histones H2A and H2B were well separated from each other by polyacrylamide gel electrophoresis in 6.25 M urea at pH 3.2 [20] (Fig. 1B, slots 1 and 4).

Table 1 shows the amino acid composition of starfish histone H2A. By comparison with calf homologous histone, the starfish H2A is characterized by a higher content of alanine and phenylalanine, a lower content of threonine, glutamic acid, isoleucine, leucine and lysine, the presence of a methio-

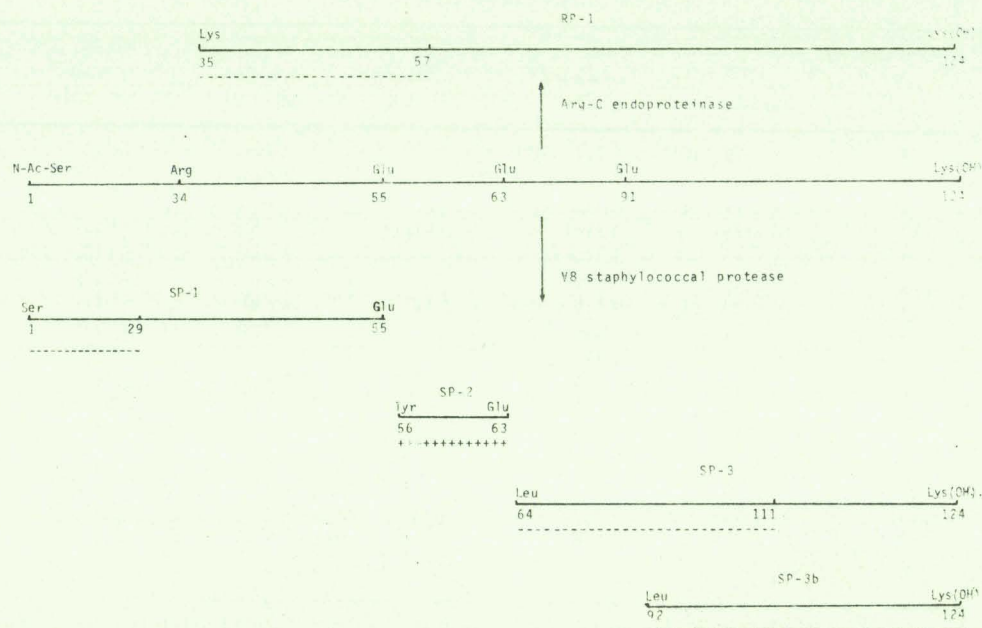


Fig. 3. Strategy used for the determination of the amino acid sequence of the starfish histone H2A. SP, V8 staphylococcal protease peptide; RP, submaxillary gland of mouse protease (Arg-C endoprotease) peptides. (---) Automated Edman degradation; (+++) manual Edman degradation

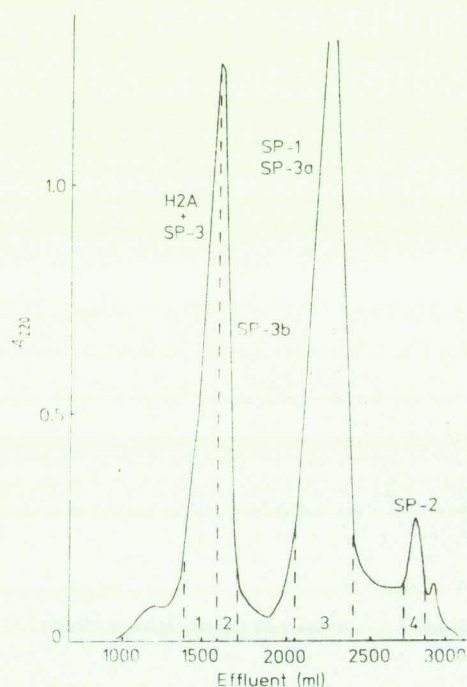


Fig. 4. Elution pattern of staphylococcal protease peptides fractionated on Bio-Gel P10. The column (150 × 5 cm) was equilibrated and eluted with 0.01 M HCl saturated with chloroform. Fractions of 12 ml were collected at a flow rate of 72 ml/h

nine residue and the lack of two histidine residues. Starfish histone H2A is devoid of a free amino group as established by dansylation. The blocking group was identified as acetyl by nuclear magnetic resonance spectroscopy of the amino-terminal tripeptide Ac-Ser-Gly-Arg generated by tryptic hydrolysis of the protein (Fig. 2). Digestion of starfish histone H2A with carboxypeptidases B and A led to establish the carboxy-terminal sequence as Ala-Lys(OH).

Table 2. Amino acid composition of thermolysin peptides obtained from peptide SP-1 and used to elucidate its sequence

Amino acids	Th-5 18-21	Th-8 29-32	Th-9 33-38	Th-10 39-41	Th-11 42-49	Th-12 50-52
Aspartic acid			0.9 (1)			
Threonine						
Serine						
Glutamic acid				1.2 (1)		
Proline					1.9 (1)	
Glycine	1.1 (1)		0.9 (1)		2.0 (2)	
Alanine	2.0 (2)			1.1 (1)	2.0 (2)	2.0 (2)
Valine		0.8 (1)			1.7 (2)	
Methionine						
Isoleucine						
Leucine			0.8 (1)			1.0 (1)
Tyrosine			0.8 (1)		0.4 (1)	
Phenylalanine		1.1 (1)				
Histidine		0.9 (1)				
Lysine			1.5 (1)			
Arginine	1.2 (1)	1.1 (1)	1.1 (1)	1.0 (1)		

With the method of automated Edman degradation, the strategy chosen for the sequence determination of starfish histone H2A was first intended to generate a minimum of large adjacent or overlapping fragments by highly specific or limited enzymatic cleavages. In the second place, hydrolysis of the starfish histone H2A with trypsin or pepsin yielded small peptides, which were essentially used to confirm the structural data provided by the automated sequencing of the large fragments.

Thus, the cleavage of starfish histone H2A at glutamyl residue by the V8 staphylococcal protease yielded five peptides (Fig. 3), which were obtained in pure form through chromatographic separations on Bio-Gel P10 (Fig. 4) and Sephadex G-50. Their amino acid compositions are given in Table 1. Peptides SP-1, SP-2 and SP-3 together account for the total

Table 3. Amino acid composition of tryptic peptides obtained from starfish histone H2A and used to elucidate its amino-terminal sequence (residues 1–41)

Amino acids	T-1 1–3	T-2 4–5	T-3 6–8	T-4 9–10	T-5, T-6 11–12 13–14	T-7 15–16	T-8a 17–19	T-8b 17–19	T-9 20–28	T-10 29–31	T-11 32–34	T-12 35–41
Aspartic acid												1.0 (1)
Threonine												
Serine	0.9 (1)					0.8 (1)	1.6 (2)	0.7 (1)				
Glutamic acid									1.1 (1)			1.0 (1)
Proline									1.0 (1)			
Glycine	1.1 (1)	1.1 (1)	2.1 (2)						2.1 (2)			1.0 (1)
Alanine				1.0 (1)	1.0 (1)			1.0 (1)	1.0 (1)			1.1 (1)
Valine									1.1 (1)	0.9 (1)		
Methionine												
Isoleucine												
Leucine									1.0 (1)		1.1 (1)	
Tyrosine												0.8 (1)
Phenylalanine									1.0 (1)		0.7 (1)	
Histidine										0.8 (1)		
Lysine		1.0 (1)	1.0 (1)		1.0 (1)							1.1 (1)
Arginine	1.0 (1)			1.0 (1)		1.0 (1)	1.0 (1)	1.0 (1)	1.0 (1)	1.0 (1)	1.0 (1)	1.1 (1)

Table 4. Amino acid composition of pepsin peptides obtained from starfish histone H2A and used to elucidate its sequence

Amino acids	P-1a 39–50	P-1b 39–50	P-2 53–55	P-3 60–64	P-4 85–92	P-5 108–112	P-6 113–124
Aspartic acid	0.8 (1)				2.1 (2)	1.0 (1)	
Threonine							1.0 (1)
Serine							
Glutamic acid	1.4 (1)	1.0 (1)	1.0 (1)	2.0 (2)	2.2 (2)	1.0 (1)	
Proline	1.0 (1)	0.9 (1)				1.0 (1)	0.9 (1)
Glycine	1.9 (2)	1.9 (2)					
Alanine	3.0 (3)	2.9 (3)			1.0 (1)	1.0 (1)	3.1 (3)
Valine	1.7 (2)	1.8 (2)	0.9 (1)				0.9 (1)
Methionine			0.5 (1)				
Isoleucine				1.0 (1)	0.9 (1)	1.0 (1)	
Leucine		1.1 (1)		1.8 (2)	1.1 (1)		2.1 (2)
Tyrosine	0.8 (1)	0.8 (1)					
Phenylalanine							
Histidine							
Lysine							4.2 (4)
Arginine	1.3 (1)	1.0 (1)			1.0 (1)		

number of residues present in the protein. The partial cleavage of the Glu⁹¹-Leu⁹² bond in the C-terminal fragment SP-3 yielded peptides SP-3a (residues 64–91) and SP-3b (residues 92–124). Since the peptide SP-1 contained most of the serine, glycine and arginine residues present in the H2A molecule, it clearly originated from the amino-terminal region of the protein (residues 1–55). The *N*-acetylserine residue present at its amino-terminal end was found partially deacetylated. Indeed, the deacetylation of the α -acetylamino group, by a deacetylase activity present in the V8 staphylococcal protease, has already been observed in V8 staphylococcal protease hydrolysates of histones H4 and H2A [6,21]. The peptide SP-1 was, therefore, submitted successfully to automated Edman degradation over 29 steps (Fig. 7). The complete sequence of the peptide SP-1 was achieved through structural studies of thermolysin peptides generated from SP-1 (Table 2) and of peptides obtained by hydrolysis of the protein with trypsin or pepsin (Tables 3 and 4, Fig. 7).

The tryptic peptide T-1 lacked a free α -amino group. Its sequence was determined as described previously [2]. The

presence of an acetyl group at the amino-terminal end of the peptide T-1 was established by NMR spectroscopy in ²H₂O at pH 2.0 and pH \approx 11 (Fig. 2).

Among the other tryptic peptides, two originate from the same part of the molecule and result from a microheterogeneity at position 18, where a serine and an alanine residue are found, generating the tryptic peptides Ser-Ser¹⁸-Arg and Ser-Ala¹⁸-Arg. The presence of an alanine residue at position 18 was confirmed by the thermolysin peptide Th-5 Ala¹⁸-Arg-Ala-Gly²¹ (Table 2 and Fig. 7). Two of the peptides obtained by pepsin digestion of the protein, P-1a and P-1b (Fig. 5), have a similar amino acid composition except for aspartic acid and leucine (Table 4). Both peptides were obtained in equal amounts and, as shown by automated Edman degradation (Fig. 7) differ of each other at positions 40 and 50:

P-1a: Ala-Glu⁴⁰-Arg-Val-Gly-Ala-Gly-Ala-Pro-Val-Tyr-Asn⁵⁰

P-1b: Ala-Gln⁴⁰-Arg-Val-Gly-Ala-Gly-Ala-Pro-Val-Tyr-Leu⁵⁰

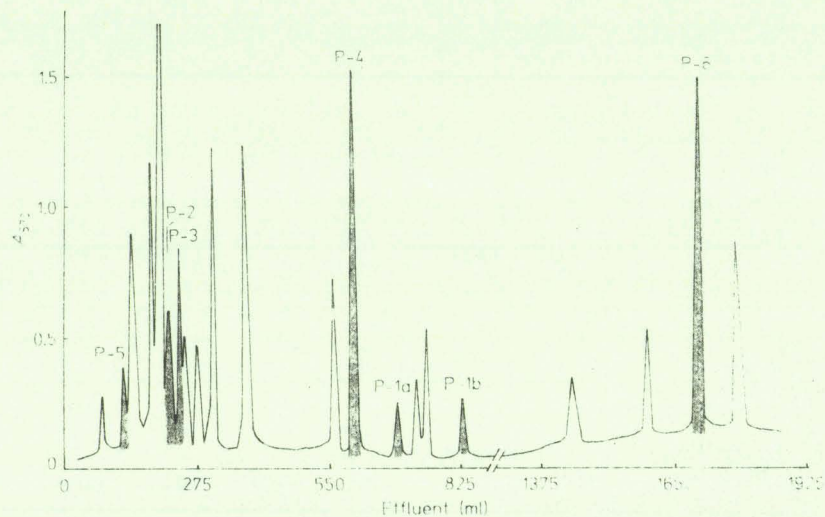


Fig. 5. Elution pattern of pepsin peptides from starfish histone H2A fractionated on Chromobeads P. The column (70 × 0.635 cm) was eluted with pyridine formate and pyridine acetate buffers used in a gradient of ionic strength and pH. Fractions of 5.5 ml were collected at a flow rate of 33 ml/h. Elution of the peptides was monitored automatically by ninhydrin reaction. Peptides used for the determination of the histone primary structure were numbered according to their position in the amino acid sequence of the protein.

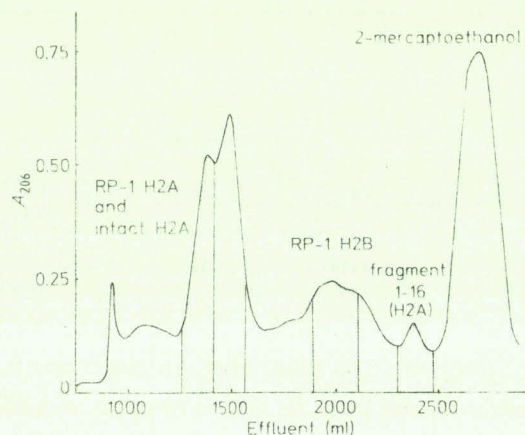


Fig. 6. Elution pattern of Arg-C endoproteinase hydrolysate of the complex H2A-H2B. The peptides were fractionated on a column of Bio-Gel P10 (150 × 5 cm) equilibrated and eluted with 0.01 M HCl, saturated with chloroform. Fractions of 17 ml were collected at a flow rate of 70 ml/h.

Among the other pepsin peptides, the peptide P-2 is of interest since it ascertains the presence of a methionine residue at position 54 in the sequence of the protein.

The sequence of the central peptide SP-2 (residues 55–63) was established by manual Edman degradation (Fig. 7).

The sequence of the carboxy-terminal half of the starfish histone H2A was entirely elucidated from the data provided by the automated Edman degradation of peptides SP-3 and SP-3b (Table 1 and Fig. 7).

Arg-C Endoproteinase Digestion of the Dimer H2A-H2B

The limited hydrolysis of the dimer H2A-H2B by the submaxillary gland of mouse protease, highly specific for arginyl bonds, was used first with the aim of producing a large fragment of histone H2A, overlapping the staphylococcal protease fragments SP-1, SP-2 and SP-3. Indeed, the rationale of this method was to generate simultaneously large fragments of histones H2A and H2B for structural studies and to define the interacting sequences within both proteins. When

the histone dimer H2A-H2B is placed in a buffered solution at slightly alkaline pH (pH 8.3) and high ionic strength (0.5 M) the strong interactions between the hydrophobic regions of the proteins delimit a protected domain in both sequences inaccessible to proteolytic cleavage by trypsin or related enzymes.

The fractionation of the limited enzymatic digest on Bio-Gel P10 allowed us to isolate a large fragment of histone H2A (RP-1 H2A, residues 35–124) together with a large peptide of histone H2B (RP-1 H2B, residues 24–125) (Fig. 6).

Purification of RP-1 H2A was achieved by gel filtration chromatography on Sephadex G-100 SF. Its amino acid composition (Table 1) is consistent with that of a fragment of 90 residues corresponding to the C-terminal three-quarters of the protein. 23 steps of automated Edman degradation were performed and gave us the necessary overlaps to determine the complete amino acid sequence of starfish histone H2A (Fig. 7).

DISCUSSION

The amino acid sequence of starfish gonad histone H2A (124 residues) presents, by comparison with that of calf homologous histone, twelve substitutions and five deletions (Table 5). Five deletions and four substitutions occur in the N-terminal and C-terminal basic regions; two substitutions and one deletion are observed in the sequence 1–19; two substitutions and four deletions are observed in the sequence 118–124. The deletion of the glutamine residue at position 6 leads to a total identity between the N-terminal sequences of histones H2A and H4. The substitution at position 19 (Ser → Ala) results in the loss of a phosphorylation site [22]. This phosphorylation site is also lost in the variants of *Triticum* H2A [7], where serine is replaced by leucine or valine. Loss of a phosphorylation site involves the intervention of other process of regulation of the DNA-histone interactions in the nucleosome. The acetylation might very well be this other process since, in addition to the lysine residue at position 5, the lysine residue at position 8 can be acetylated.

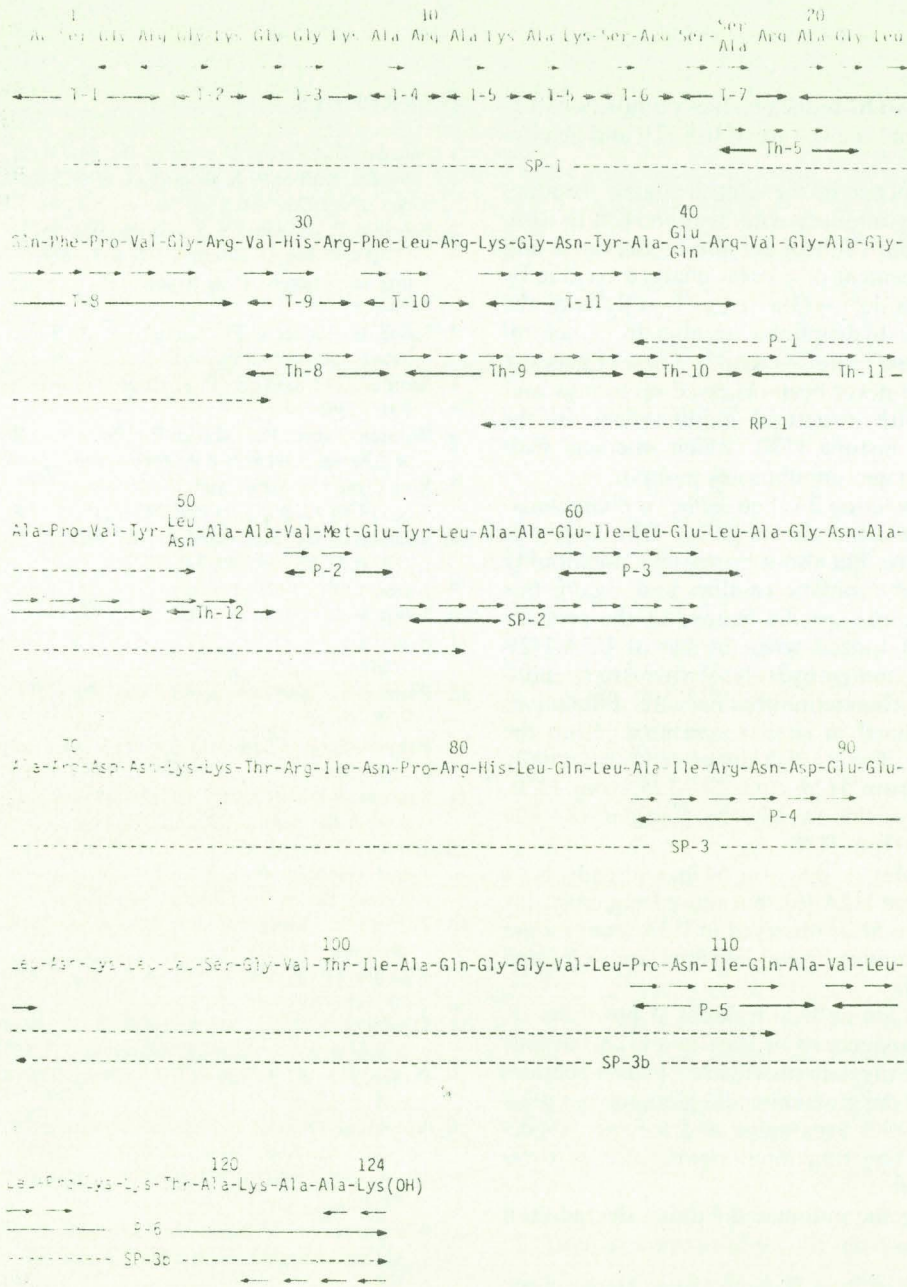


Fig. 7. Alignment of peptides and amino acid sequence of starfish histone H2A. Methods used for determination of the sequence of peptides are indicated as follows: — manual Edman degradation; ← carboxypeptidase hydrolysis; --- automated Edman degradation.

Table 7. Comparison of the sequences of histone H2A from two echinoderms (starfish and sea-urchin) and from calf

Species	Residues at position						
	6	11	16	20	33	41	51
Calf thymus [2]	Gln-Gly-Gly-Lys-Ala-Arg		Thr-Arg-Ser-Ser-Arg		Leu	Glu	Leu
Starfish gonads			Ser-	Ser-Ala	Phe	Glu/Gln	Leu/Asn
Sea-urchin gonads [5]		-Ala- Gly-Lys	Ser-		Phe	Asn	Leu
Calf thymus	88 Leu	89 Thr	90 Ile	98 Gly-Lys Arg	121 Glu-Ser	122 His-His-Lys-Ala-Lys-Gly-Lys(OH)	129
Starfish gonads		Met	Ala	Asn Ser-Gly	Ala	Lys-Ala	Ala-Lys(OH) 124
Sea-urchin gonads			Ala	Ile Gly	Gly-Ser	Lys-Ser	Ser-Lys(OH) 125

Deletions are indicated by —. Numbering in this table and in the text refers to sequence positions in calf thymus histone H2A.

The deletion of the two histidine residues (at positions 123 and 124 in the calf histone) is a feature of fish [23] and invertebrate H2A [5,6,24].

Eight substitutions occur in the central region (residues 20–117), which strongly interacts with histone H2B to form the dimer H2A-H2B. Four of them are non-conservative and involve either the replacement of a polar charged residue by a neutral polar residue (Glu⁴¹→Gln⁴⁰; Lys⁹⁹→Gly⁹⁸) or the replacement of a bulky hydrophobic residue by a neutral polar residue (Leu⁵¹→Asn⁵⁰; Ile⁷⁹→Asn⁷⁸). These changes at positions 51 and 79 have never been observed up to now and could be associated with structural modifications in the hydrophobic region of histone H2B, which interacts with histone H2A (A. Martinage, unpublished results).

The substitution at position 33 (Leu→Phe) is characteristic of histone H2A from echinoderms [5,24]. Obviously this substitution is conservative but also it brings up a potentiality of interaction with other aromatic residues and, again, this could be related with some specific feature of the primary structure of histone H2B. Indeed, when the starfish H2A-H2B dimer is submitted to a limited hydrolysis with Arg-C endoproteinase, the histone fragments obtained are different in length from those obtained in similar conditions from the calf H2A-H2B complex. The starfish dimer yields essentially the peptides 35–124 from H2A and 24–125 from H2B, whereas the calf dimer yields mainly the peptides 18–129 from H2A and 32–125 from H2B.

The change Leu→Met at position 54 has already been found in cuttlefish histone H2A [6]; but more frequently this change occurs at position 51 as observed in H2A from mouse ascite tumor cells [25], mouse Friend leukemia cells [26] and sea-urchin embryos [27].

The presence of two amino acid residues at positions 18, 40 and 50 reflects the existence of at least two H2A variants in starfish gonads. From the data provided by pepsin peptides P-1a and P-1b (Table 4) the glutamate and glutamine at position 40 are associated with asparagine and leucine, respectively, at position 50. The functional significance of these variants remains unclear.

The data concerning the automated Edman degradation of peptides are available from the authors upon request.

The authors are grateful to Dr A. Richard (Station Marine, 62930 Wimereux, France) who provided the starfish gonads, and to Dr C. Crane-Robinson who performed the NMR spectrum of the amino terminal tripeptide. They acknowledge with pleasure the skilful technical assistance of M. J. Dupire, T. Ernout and A. Hémez. They thank also Dr G. Briand and O. Hannedouche from the *Service commun de Séquence U. 16/U. 124 Institut National de la Santé et de la Recherche Médicale* (Lille) for the sequencer runs. This work was supported by grant 77.5.048.2 from the *Institut National de la Santé et de la Recherche Médicale* and grants from the *Centre National de la Recherche Scientifique (ATP Chromatine 4203 and L.A. 268)*.

A. Martinage, D. Bélaiche, T. Dupressoir, and P. Sautière,
Groupe de Recherches Ultrastructurales et Biochimiques sur les Cellules Normales et Cancéreuses
(Unité 124 de l'Institut National de la Santé et de la Recherche Médicale), Institut de Recherches sur le Cancer, Cité Hospitalière,
2 Place de Verdun, F-59020 Lille-Cedex, Nord, France

REFERENCES

- Sautière, P., Tyrrou, D., Laine, B., Mizon, J., Lambelin-Breynaert, M. D., Ruffin, P. & Biserte, G. (1972) *C R Hebd Seances Acad Sci*, **274**, 1422–1425.
- Sautière, P., Tyrrou, D., Laine, B., Mizon, J., Ruffin, P. & Biserte, G. (1974) *Eur. J. Biochem.* **41**, 563–576.
- Laine, B., Sautière, P. & Biserte, G. (1976) *Biochemistry*, **15**, 1640–1645.
- Laine, B., Kmiećik, D., Sautière, P. & Biserte, G. (1978) *Biochimie (Paris)* **60**, 147–150.
- Wouters, D., Sautière, P. & Biserte, G. (1978) *Eur. J. Biochem.* **90**, 231–239.
- Wouters-Tyrrou, D., Martin-Ponthieu, A., Briand, G., Sautière, P. & Biserte, G. (1982) *Eur. J. Biochem.* **124**, 489–498.
- Von Holt, C., Strickland, W. N., Brandt, W. F. & Strickland, M. S. (1979) *FEBS Lett.* **100**, 201–218.
- Strickland, M., Strickland, W. N., Brandt, W. F. & Von Holt, C. (1974) *FEBS Lett.* **40**, 346–348.
- Johns, E. W. (1967) *Biochem. J.* **105**, 611–614.
- Johns, E. W. (1964) *Biochem. J.* **92**, 55–59.
- Skandrani, F., Mizon, J., Sautière, P. & Biserte, G. (1972) *Biochimie (Paris)* **54**, 1267–1272.
- Panyim, S., Blek, D. & Chalkley, R. (1971) *J. Biol. Chem.* **246**, 4206–4215.
- Panyim, S. & Chalkley, R. (1969) *Arch. Biochem. Biophys.* **130**, 337–346.
- Sautière, P., Moschetto, Y., Dautrevaux, M. & Biserte, G. (1970) *Eur. J. Biochem.* **12**, 222–226.
- Edman, P. & Henschen, A. (1975) in *Protein Sequencing, Determination* (Needleman, S. B., ed.) 2nd edn, pp. 252–279. Springer-Verlag, Berlin, Heidelberg, New York.
- Tarr, G. E., Beecher, J. F., Bell, M. & McKeay, D. (1975) *Anal. Biochem.* **84**, 622–627.
- Kiapper, D. G., Wilde, C. E. & Capra, J. D. (1975) *Anal. Biochem.* **85**, 126–131.
- Hermann, J., Titani, K., Ericsson, L. H., Wade, R. D., Neurath, H. & Walsh, K. A. (1978) *Biochemistry*, **17**, 5672–5679.
- Pisano, J. J. & Bronzert, T. J. (1969) *J. Biol. Chem.* **244**, 5597–5607.
- Vanhoutte-Durand, G., Mizon, J., Sautière, P. & Biserte, G. (1977) *Comp. Biochem. Physiol.* **57B**, 121–126.
- Wouters, D., Martin-Ponthieu, A., Sautière, P. & Biserte, G. (1981) *FEBS Lett.* **128**, 195–200.
- Martinage, A., Mangeat, P., Laine, B., Couppez, M., Sautière, P., Marchis-Mouren, G. & Biserte, G. (1980) *FEBS Lett.* **118**, 323–329.
- Bailey, G. S. & Dixon, G. H. (1973) *J. Biol. Chem.* **248**, 5463–5472.
- Strickland, W. N., Strickland, M. S., De Groot, P. C. & Von Holt, C. (1980) *Eur. J. Biochem.* **109**, 151–158.
- Franklin, S. G. & Zweidler, A. (1977) *Nature (Lond.)* **266**, 273–275.
- Blankstein, L. A., Stollar, B. D., Franklin, S. G., Zweidler, A. & Levy, S. B. (1977) *Biochemistry*, **16**, 4557–4562.
- Schaffner, W., Kunz, G., Daetwyler, H., Telford, J., Smith, H. O. & Birnstiel, M. L. (1978) *Cell*, **14**, 655–671.

Primary structure of histone H2A from nucleated erythrocyte of the marine worm *Sipunculus nudus*

Presence of two forms of H2A in the sipunculid chromatin

Daniel KMIECIK, Maurice COUPPEZ, Denise BELAICHE, and Pierre SAUTIERE

Unité 124 de l'Institut National de la Santé et de la Recherche Médicale, Lille

(Received March 17/June 7, 1983) – EJB 83 0263

The complete amino acid sequence (123 residues) of histone H2A from erythrocytes of the marine worm *Sipunculus nudus*, has been established from data provided by automated sequence analysis of large fragments generated by V8 staphylococcal protease digestion of histone H2A and by limited hydrolysis of the protein with α -chymotrypsin and from structural studies of tryptic peptides of the protein.

By comparison with calf homologous histone, the sipunculid histone H2A shows 6 deletions and 13 substitutions. Six of the substitutions are non-conservative.

Most of the evolutionary changes are mainly observed in the basic amino-terminal and carboxy-terminal regions of the molecule, which are the primary DNA-binding sites. Few conservative point changes are observed in the central region (residues 18–118) which interacts strongly with histone H2B to form the dimer H2A-H2B. 60% of the H2A molecules were found phosphorylated on the amino-terminal residue, *N*-acetyl-serine. The high content of phosphorylated histone H2A in the sipunculid erythrocyte chromatin could probably be related to smaller repeat length (177 ± 5 base pairs) of nucleosomal DNA and to nuclear inactivation and chromatin condensation.

Sipunculid sea worms form an isolated ancestral phylum (Fig. 1) whose origin dates back to the Cambrian period. The sipunculid erythrocytes are nucleated and contain hemerythrin, a respiratory pigment which differs from that of vertebrates.

The first investigations on nuclear basic proteins of *Sipunculus nudus* erythrocyte chromatin were intended to look for the presence of an erythrocyte-specific histone homologous to the avian erythrocyte histone H5 and for evolutionary variations in histones H2A and H2B [1, 2]. Since sipunculid H2A and H2B were found to be markedly different from the homologous histones in vertebrates [3–5], it was therefore of interest to investigate the primary structure of both histones.

According to Mazen et al. [2], the differences could probably be associated with a smaller repeat length of the nucleosomal DNA (177 ± 5 base pairs). Moreover the sequence determination of sipunculid H2A takes place in a study of the structural variations of histone H2A undertaken in our laboratory in organisms widely separated on the evolutionary scale [3–9]. In this paper we present the

Abbreviations. TosLysCH₂Cl, *N*-tosyllysyl chloromethyl ketone; TosPheCH₂Cl, *N*-tosylphenylalanyl chloromethyl ketone; PhMeSO₂F, phenylmethylsulfonyl fluoride; *i*Pr₂PF, diisopropylfluorophosphate; dansyl-, 5-dimethylaminonaphthalene-1-sulfonyl.

Enzymes. *Staphylococcus aureus* protease (EC 3.4.21.19); α -chymotrypsin (EC 3.4.21.1); trypsin (EC 3.4.21.4); carboxypeptidase A (EC 3.4.17.1); carboxypeptidase B (EC 3.4.17.2).

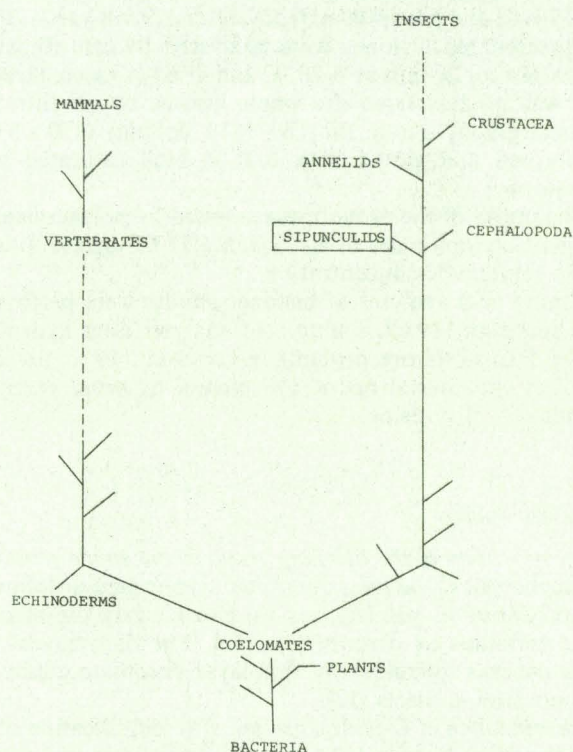


Fig. 1. Phylogenetic position of the sipunculids

complete amino acid sequence of sipunculid erythrocyte H2A established from structural data obtained essentially from large peptides generated by hydrolysis of the protein with V8 staphylococcal protease and chymotrypsin.

MATERIALS AND METHODS

Materials

Staphylococcal V8 protease was from Miles laboratories. Trypsin (treated with TosPheCH₂Cl) and carboxypeptidase A (treated with PhMeSO₂F) were obtained from Worthington, α -chymotrypsin (treated with TosLysCH₂Cl) was from Merck. Carboxypeptidase B (treated with iPr₂PF) was from Boehringer.

Dimethylallylamine, propane-1-ol and dithioerythritol were sequanal grade from Pierce. Benzene and chloro-1-butane for sequential analysis were purchased from SDS (Peypin, France). Phenylisothiocyanate, heptane and heptafluorobutyric acid were sequanal grade from Beckman. Polybrene was purchased from Aldrich and the dipeptide Gly-Gly from Sigma. All other reagents were of the highest purity available.

Isolation of histone H2A

Worms were dug out of the sludge of river mouths on the Brittany coast. As soon as the worms were collected, nucleated erythrocytes were separated from the coelomic liquid by centrifugation. Nuclei were prepared as described by Mazen et al. [1] in the presence of 0.5 mM PhMeSO₂F.

Histones were extracted from the nuclear pellet with 0.25 M HCl and precipitated with acetone (9 vol.) at -20 °C. The precipitated histones were recovered by centrifugation at 4000 \times g for 20 min at -20 °C and dried *in vacuo*. Histone H2A was isolated from the whole histone by gel filtration chromatography on a Bio-Gel P-10 column (150 \times 5 cm) equilibrated and eluted with 0.01 M HCl saturated with chloroform.

The purity of the protein was assessed by polyacrylamide slab gel electrophoresis at pH 3.2 in 6.25 M urea [10] using a 17% acrylamide concentration.

Amino acid analyses of histone samples were performed on a Beckman 119 CL amino acid analyser after hydrolysis in 6 M HCl (1 ml/mg protein) *in vacuo* at 110 °C for 24 h and 72 h with one drop of 1% phenol to avoid excessive degradation of tyrosine.

Structural studies

Identification of the blocking group at the amino terminus. Hydrazinolysis of the amino-terminal tryptic peptide followed by dansylation at pH 3.0, was used to identify the blocked amino terminus as described in [11]. The dansylhydrazide derivative was identified by thin-layer chromatography on micropolyamide sheets [12].

Identification of O-phosphoserine. For identification of O-phosphoserine, histone H2A and the amino-terminal tryptic peptides were hydrolysed for 2 h at 110 °C under vacuum with 6 M HCl [13]. The hydrolysates were submitted to electrophoresis at 2500 V for 2 h on Whatman 3MM paper at pH 1.9 (acetic acid/formic acid/water, 35:10:400, by

vol.). O-Phosphoserine and O-phosphothreonine (Sigma) were used as markers.

Carboxy-terminal analysis. Histone H2A (0.84 mg, 60 nmol) was dissolved in 0.15 ml of 0.2 M N-methylmorpholine acetate pH 8.5 and digested at 37 °C with carboxypeptidase B (0.2 μ g). At 30 min, an aliquot (0.05 ml) was taken off, acidified with glacial acetic acid and freeze-dried. Carboxypeptidase A (1.7 μ g) was then added. At 60 min and 120 min following addition of carboxypeptidase A, aliquots (0.05 ml) were taken off and treated as above. The released amino acids were analysed on the amino acid analyser.

Enzymatic hydrolyses. Histone H2A (70 mg, 5 μ mol) was hydrolysed with V8 *Staphylococcus aureus* protease in 0.05 M ammonium acetate (14 ml), at pH 4.0, and 37 °C for 20 h using an enzyme/substrate ratio of 1/50 (w/w). The hydrolysate was then freeze-dried. The limited chymotryptic hydrolysis of histone H2A (20 mg, 1.5 μ mol) was performed in 0.5 M NaCl, 0.1 M ammonium bicarbonate pH 8.3, at 19 °C for 150 min using an enzyme/substrate ratio of 1/1000 (w/w). After lowering the pH at 3.0, the hydrolysate was freeze-dried. The tryptic digestion of histone H2A (40 mg, \approx 3 μ mol) was carried out in 0.1 M ammonium bicarbonate pH 8.0 at 37 °C for 4 h using an enzyme/substrate ratio of 1/50 (w/w). The digestion was stopped by lowering the pH at 3.5 with formic acid. The insoluble material formed upon acidification of the tryptic hydrolysate and corresponding to large hydrophobic peptides was collected by centrifugation. The pellet was washed twice in 10% formic acid, dissolved in 0.1 M NH₄OH and freeze-dried. The soluble fraction of the tryptic digest consisting mainly of small hydrophilic peptides was freeze-dried.

Fractionation of peptides. The peptides obtained by hydrolysis of histone H2A with V8 staphylococcal protease were separated by chromatography on a Bio-Gel P 10 column (150 \times 5 cm) equilibrated and eluted with 0.01 M HCl saturated with chloroform. The chymotryptic hydrolysate was fractionated on a Sephadex G-100 column (200 \times 2.5 cm) equilibrated and eluted with 0.01 M HCl, saturated with chloroform. The large hydrophobic tryptic peptides were separated by reverse-phase liquid chromatography on a column of C₁₈ μ Bondapak (Waters Associates) using a Waters Associates ALC/GPC 244 apparatus equipped with a U6K sample injector and a Cecil 212 variable-wavelength ultraviolet monitor. The column was equilibrated in 0.05% trifluoroacetic acid containing 10% acetonitrile (v/v). The peptides were eluted with a linear gradient of 10–50% of acetonitrile in 0.05% trifluoroacetic acid. The small hydrophilic tryptic peptides were separated by ion-exchange chromatography on a Chromobeads P column (Technicon Corporation) as described in [6].

Sequence analysis. Automated Edman degradation of large peptides was carried out in a Beckman 890 C sequencer using a dimethylbenzylamine program (082773) [14] or a dimethylallylamine program (102974) in the presence of polybrene [15, 16]. Small peptides were sequenced by manual Edman degradation [17]. The phenylthiohydantoin derivatives of amino acids were identified by high-pressure liquid chromatography as described in [7].

Nomenclature of peptides. Peptides obtained by cleavage of the protein at glutamyl residues with the V8 staphylococcal protease were designated SP-. Chymotryptic and tryptic peptides were designated by Ch- and T- respectively. Peptides were numbered according to their position in the sequence of the protein.

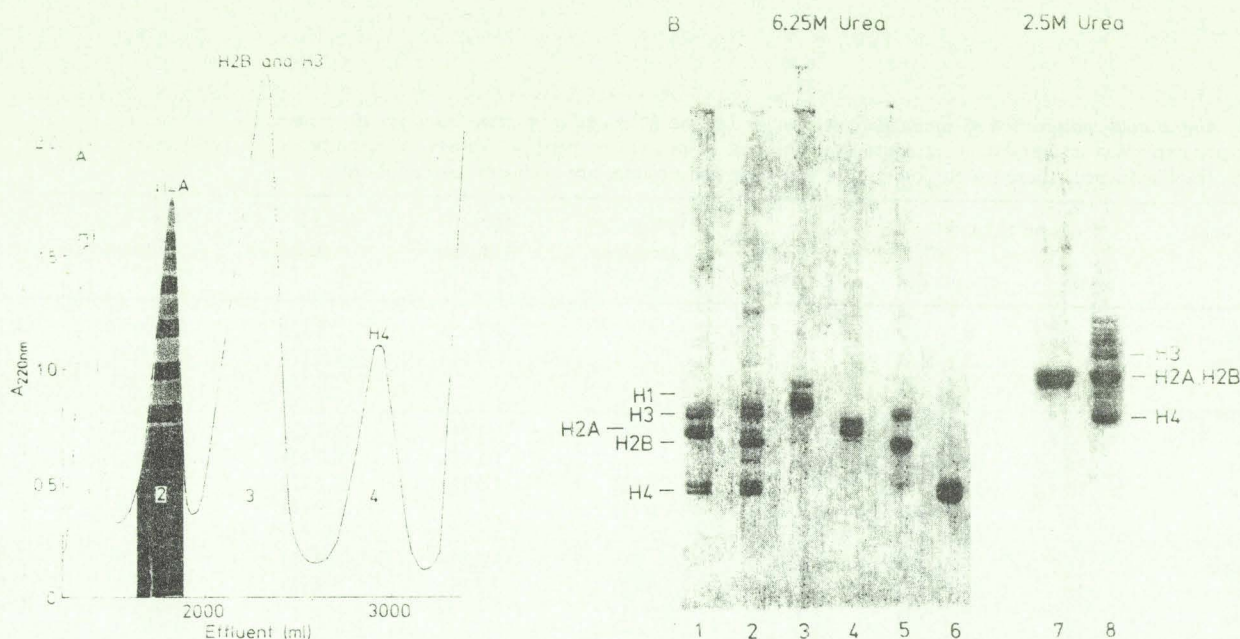


Fig. 2. (A) Bio-Gel P10 chromatography and (B) electrophoretic analysis of whole histone from sipunculid erythrocyte. (A) 100 mg of material were dissolved in 6 ml 6 M guanidine chloride, 0.1 M Tris, pH 8.0, 5% 2-mercaptoethanol and left under nitrogen at room temperature for 24 h before chromatography. The solution was then centrifuged and applied to a Bio-Gel P10 column (150 × 5 cm) equilibrated and eluted with 0.01 M HCl saturated with chloroform. The flow rate was 60 ml/h and 18 ml — fractions were collected. (B) Fractions obtained in A were subjected to electrophoresis on polyacrylamide slab gels (15 × 10 cm) at 22 mA for 3 h at pH 3.2 in 6.25 M urea (lanes 1–6) and at pH 2.7 in 2.5 M urea (lanes 7, 8) [10]. Migration was top to bottom. Slabs were stained for 60 min with 0.5% Coomassie blue R in acetic acid/ethanol/water (1:2:7, by vol.) and destained by diffusion in the mixture acetic acid/ethanol/water (7:20:73, by vol.). Lane 1, whole histone from calf thymus; lanes 2 and 8, whole histone from sipunculid erythrocyte; lane 3, peak 1 from A; lanes 4 and 7, peak 2 from A; lane 5, peak 3 from A; lane 6, peak 4 from A.

RESULTS

Bio-Gel P10 chromatography of whole histone from *Sipunculus* erythrocyte yielded four fractions (Fig. 2A). Histone H2A was obtained free of contamination in fraction 2 (Fig. 2B, slot 4). Histones H1 and H4 were also obtained directly in pure form from fractions 1 and 4 respectively (Fig. 2B, slots 3 and 6). Fraction 3 is a mixture of histones H2B and H3 (Fig. 2B, slot 5). On polyacrylamide gel electrophoresis at pH 3.2 in 6.25 M urea, histone H2A appears as two bands stained with approximately the same intensity. At this concentration of urea, the electrophoretic migration of H2A is intermediate between that of H3 and H2B. In 2.5 M urea at pH 2.7, H2A comigrates with H2B (Fig. 2B, slots 7 and 8). Of the two bands of *Sipunculus* H2A which are observed in gel electrophoresis at acidic pH (Fig. 2B, slots 4 and 7), the slow-moving one corresponds to mono-phosphorylated H2A, whereas the fast-moving band represents the non-phosphorylated H2A. The phosphate group was shown to be located on the serine residue at the amino terminus of the protein, following the isolation of the amino-terminal tryptic peptide T-1a, AcSer(P)-Gly-Arg.

Table 1 shows the amino acid composition of sipunculid histone H2A. By comparison with calf homologous histone, the sipunculid histone H2A is characterized by a higher content of serine and isoleucine, a lower content of threonine, glutamic acid, alanine, valine, leucine, lysine and arginine, the presence of one residue of methionine and the lack of two residues of histidine.

Sipunculid histone H2A has a blocked amino group as established as acetyl by hydrazinolysis followed by dansylation of the tripeptide AcSer-Gly-Arg T-1b generated by tryptic

hydrolysis of the protein. The carboxy-terminal sequence of histone H2A was established as -Lys-Ser-Lys after digestion with carboxypeptidases B and A.

With the method of automated sequencing, the strategy used for the determination of the primary structure of sipunculid histone H2A was firstly intended to produce a minimum of large adjacent or overlapping fragments by highly specific or limited enzymatic cleavages (Fig. 3). Secondly, tryptic hydrolysis of the protein generated small peptides (see Table 2) which were used to confirm the structural data provided by the automated sequencing of the large fragments.

Thus, the cleavage of sipunculid histone H2A at glutamyl residues yielded four fragments which were separated on Bio-Gel P10 (Fig. 4). Their amino acid compositions are presented in Table 1. Fragments SP-1, SP-2, SP-3 and SP-4 together accounted for the total number of residues present in the protein. Since the peptide SP-1 contains most of the serine, glycine and arginine residues found in histone H2A, it obviously originates from the amino-terminal region of the protein (residues 1–55). As observed previously in V8 staphylococcal protease hydrolysates of histones H4 and H2A, the *N*-acetylserine residue present at the amino terminus of the fragment SP-1 was found to be partially deacetylated. The fragment SP-1 was therefore successfully submitted to automated Edman degradation over 33 steps (Table 3, Fig. 8). The sequence of the fragment SP-2 (residues 56–63) was established by manual Edman degradation (Fig. 8). The peptides SP-3 (residues 64–91) and SP-4 (residues 92–123) were completely sequenced by automated Edman degradation (Tables 4 and 5, Fig. 8). The carboxy-terminal position of SP-4 was ascertained by digestion of the peptide with

Table 1. Amino acid composition of sipunculid erythrocyte histone H2A and of peptides used for the sequence determination. Results are expressed as number of residues per molecule of protein or peptide. Values in parentheses are residues per molecule derived from the final sequence; these for threonine and serine, in the protein, are zero-time extrapolations

Amino acid	Histone H2A		SP-1 (residues 1-55)	SP-2 (residues 56-63)	SP-3 (residues 64-91)	SP-4 (residues 92-123)	Ch-1 (residues 31-123)	Ch-2 (residues 85-123)
	calf	sipunculid						
Aspartic acid	8	8.2 (8)	1.2 (1)		4.7 (5)	2.1 (2)	7.9 (8)	4.2 (4)
Threonine	5	3.5 (3)			0.9 (1)	1.9 (2)	2.9 (3)	2.0 (2)
Serine	4	7.4 (7)	5.0 (5)			2.0 (2)	2.7 (2)	2.0 (2)
Glutamic acid	12	11.5 (11)	2.6 (3)	2.2 (2)	3.1 (3)	3.3 (3)	10.2 (10)	5.9 (5)
Proline	5	5.2 (5)	2.2 (2)		1.1 (1)	2.7 (2)	3.8 (4)	2.0 (2)
Glycine	14	13.7 (14)	10.3 (10)		1.5 (1)	3.0 (3)	7.6 (7)	3.4 (3)
Alanine	16	14.7 (15)	6.6 (7)	1.9 (2)	4.0 (4)	2.2 (2)	13.2 (13)	3.4 (3)
Valine	8	7.1 (7)	2.3 (3)	1.0 (1)		2.9 (3)	5.4 (6)	3.0 (3)
Methionine	—	0.9 (1)	0.2 (1)				0.6 (1)	
Isoleucine	6	6.6 (7)	1.8 (2)		3.2 (3)	1.9 (2)	5.0 ^a (6)	3.0 (3)
Leucine	16	14.5 (15)	4.2 (4)	2.0 (2)	3.2 (3)	5.9 (6)	13.7 (14)	7.0 (6)
Tyrosine	3	2.5 (3)	1.8 (2)	0.9 (1)			2.9 (3)	
Phenylalanine	1	1.2 (1)	1.1 (1)					
Histidine	4	2.1 (2)	1.0 (1)		0.9 (1)		1.1 (1)	
Lysine	14	13.0 (13)	6.6 (6)		2.3 (2)	4.9 (5)	8.8 (8)	5.7 (5)
Arginine	12	10.7 (11)	7.3 (7)		4.0 (4)		7.0 (7)	1.2 (1)
Number of residue	129	123	55	8	28	32	93	39

^a This value is low due to incomplete cleavage of an Ile-Ile bond.

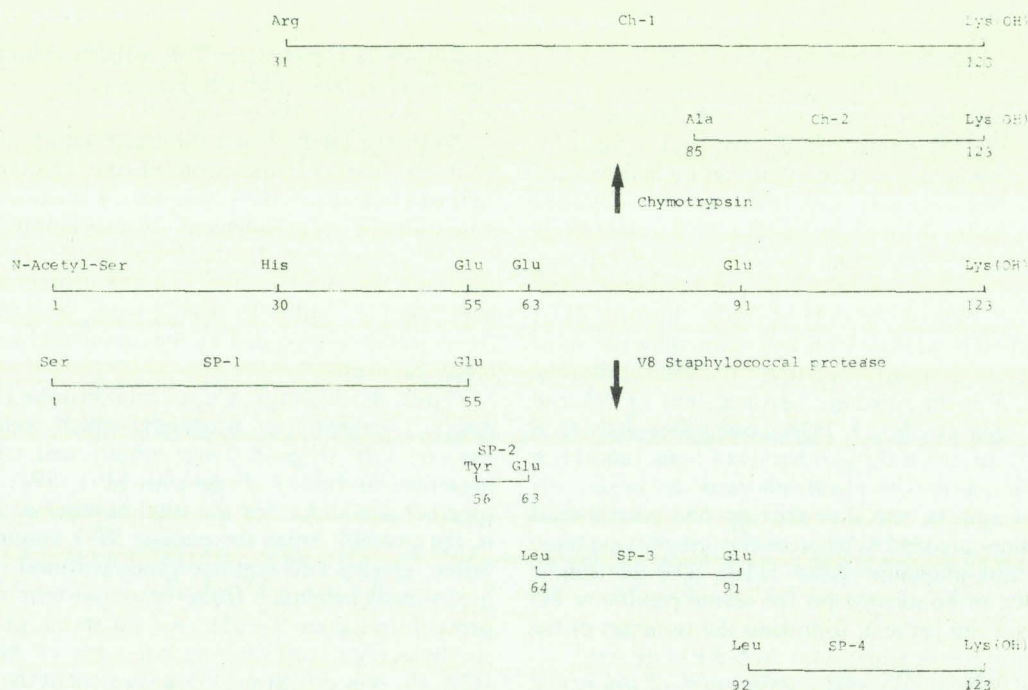


Fig. 3. Strategy used for the determination of the amino acid sequence of the sipunculid erythrocyte histone H2A

carboxypeptidases B and A. Its carboxy-terminal sequence was established as -Lys-Ser-Lys; it corresponds to the carboxy-terminal sequence of the protein.

The alignment of fragments SP-1, SP-2 and SP-3 was obtained from a structural study of the large chymotryptic peptides generated by limited hydrolysis of the sipunculid histone H2A with chymotrypsin. Hydrolysis was carried out at high ionic strength (0.5 M NaCl) which allows the protein

to have a maximum of secondary and tertiary conformations (globular structure), thus making most of the potential cleavage sites located in the organized domain of the histone H2A inaccessible to the enzyme. The fractionation of the limited chymotryptic hydrolysate on Sephadex G-100 yielded two large peptides Ch-1 and Ch-2 which were eluted in peaks 2 and 3 respectively (Fig. 5). Peptide Ch-1 although slightly contaminated by a smaller peptide, was used without

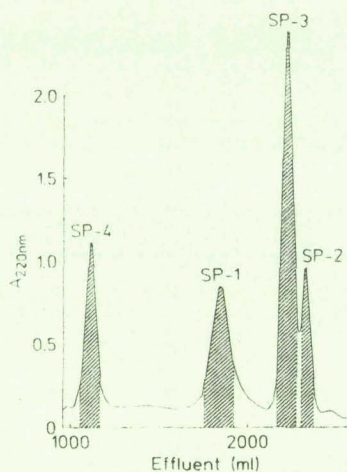


Fig. 4. Elution of pattern of V8 staphylococcal protease fragments of sipunculid histone H2A fractionated on Bio-Gel P10. The column (150 × 5 cm) was equilibrated and eluted with 0.01 M HCl saturated with chloroform. Fractions of 18 ml were collected at a flow rate of 72 ml/h

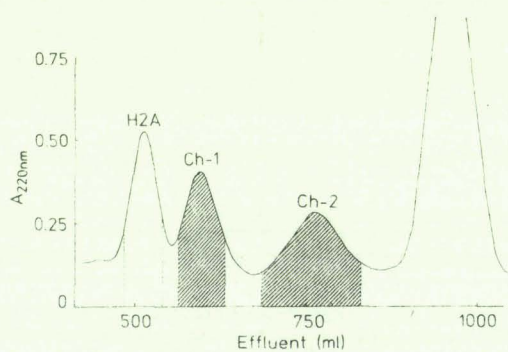


Fig. 5. Elution pattern of peptides generated by limited chymotryptic hydrolysis of sipunculid histone H2A. The peptides were fractionated on a column of Sephadex G-100 (200 × 2.5 cm) equilibrated and eluted with 0.01 M HCl saturated with chloroform. Fractions of 4.7 ml were collected at a flow rate of 14.2 ml/h

further purification. Peptide Ch-2 was purified by chromatography on a carboxymethyl-cellulose column as described in [18]. Their amino acid compositions are given in Table 1. Peptide Ch-1 (93 residues) corresponds to the C-terminal two-thirds of the H2A molecule (residues 31–123). 32 steps of automated Edman degradation were performed on peptide Ch-1 (Table 6). The identification of two phenylthiohydantoin derivatives of amino acids at each step of the degradation revealed that peptide Ch-1 was contaminated by a peptide (peptide Ch-1a, residues 34–123) obtained from the partial cleavage of the protein at leucine 33 (Fig. 8). Peptide Ch-2 (39 residues) clearly originates from the carboxy-terminal part of histone H2A as shown by its high content in aspartic acid and glutamic acid and its low content in arginine. It is generated by partial cleavage of the protein at leucine 84. Its amino-terminal sequence established by ten steps of automated Edman degradation (Table 7), provided the data necessary to align fragments SP-3 and SP-4 (Fig. 8).

All the data obtained from automated Edman degradation of staphylococcal protease fragments and of chymotryptic fragments were confirmed by structural studies of tryptic

Table 5. Automated Edman Degradation of peptide SP-3

Cycle number	Residue number	Residue	Yield (nmol)	Identification
1	92	Leu	827	HPLC
2	93	Asn	631	HPLC
3	94	Lys	941	HPLC
4	95	Leu	903	HPLC
5	96	Leu	76	HPLC
6	97	Ser	195	HPLC
7	98	Val	-	HPLC
8	99	Val	493	HPLC, GC
9	100	Thr	113	HPLC
10	101	Ile	435	HPLC
11	102	Ala	996	HPLC
12	103	Gln	555	HPLC
13	104	Gly	921	HPLC
14	105	Gly	714	HPLC
15	106	Val	36	HPLC
16	107	Leu	137	HPLC
17	108	Pro	97	HPLC
18	109	Asn	75	HPLC
19	110	Ile	81	HPLC
20	111	Gln	385	HPLC
21	112	Val	127	HPLC
22	113	Val	263	HPLC
23	114	Leu	159	HPLC
24	115	Leu	218	HPLC
25	116	Pro	80	HPLC
26	117	Lys	96	HPLC
27	118	Lys	111	HPLC
28	119	Thr	43	HPLC
29	120	Gln	124	HPLC
30	121	Lys	66	HPLC
31	122	Ser	13	HPLC
32	123	Lys	43	HPLC

See legend of Table 1.

Table 6. Automated Edman Degradation of peptide Ch-1

Cycle number	Major peptide Ch-1			Minor peptide Ch-1a			Identification
	Residue number	Residue	Yield (nmol)	Residue number	Residue	Yield (nmol)	
1	31	Arg	296	34	Arg	-	HPLC
2	32	Leu	40	35	Lys	30	HPLC
3	33	Leu	39	36	Gly	33	HPLC
4	34	Arg	150	37	Asn	21	HPLC
5	35	Lys	46	38	Tyr	41	HPLC
6	36	Gly	46	39	Ala	49	HPLC
7	37	Asn	37	40	Glu	21	HPLC
8	38	Tyr	48	41	Arg	83	HPLC
9	39	Ala	53	42	Ile	-	HPLC, GC
10	40	Glu	-	43	Gly	48	HPLC
11	41	Arg	120	44	Ala	33	HPLC
12	42	Ile	32	45	Gly	28	HPLC, GC
13	43	Gly	48	46	Ala	48	HPLC
14	44	Ala	-	47	Pro	-	HPLC
15	45	Gly	41	48	Val	-	HPLC, GC
16	46	Ala	44	49	Tyr	24	HPLC
17	47	Pro	31	50	Leu	22	HPLC, GC
18	48	Val	-	51	Ala	27	HPLC, GC
19	49	Tyr	52	52	Ala	43	HPLC, GC
20	50	Leu	49	53	Val	-	HPLC, GC
21	51	Ala	56	54	-	-	HPLC, GC
22	52	Ala	41	55	Glu	-	HPLC, GC
23	53	Val	-	56	Tyr	13	HPLC, GC
24	54	-	-	57	Leu	37	HPLC, GC
25	55	Glu	13	58	Ala	-	HPLC
26	56	Tyr	22	59	Ala	13	HPLC
27	57	Leu	29	60	Glu	-	HPLC, GC
28	58	Ala	25	61	-	-	HPLC
29	59	Ala	27	62	Leu	-	HPLC
30	60	Glu	14	63	Glu	-	HPLC, GC
31	61	Val	-	64	Leu	-	HPLC, GC
32	62	Leu	26	65	Ala	16	HPLC

See legend of Table 3.

Table 2. Amino acid composition of peptide 6b

Peptide	Residue	Yield (nmol)	Identification
1	Ala	17	HPLC, GC
	Ile	-	HPLC, GC
2	Asn	13	HPLC
	Asp	8	HPLC
3	Asp	-	HPLC
	Glu	-	HPLC
4	Glu	-	HPLC
	Leu	-	HPLC
5	Asn	-	HPLC
	Lys	-	HPLC

See legend of Table 2.

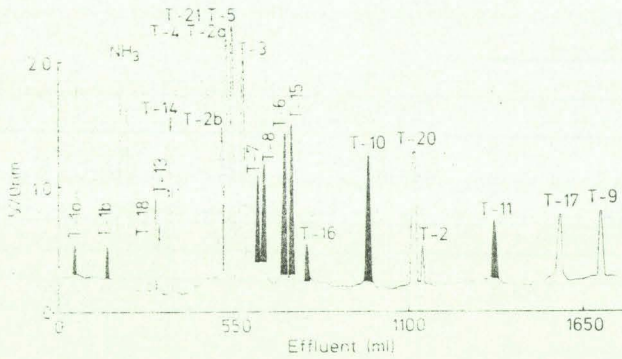


Fig. 6. Elution pattern of tryptic peptides from sipunculid histone H2A. The peptides were fractionated on a column of Chromobeads P (60 x 0.35 cm) with pyridine/formate and pyridine/acetate buffers [6]. Fractions of 5.5 ml were collected at 33 ml/h. Peptides are numbered according to their position in the sequence of the protein. Hatched peaks corresponds to arginine-containing peptides

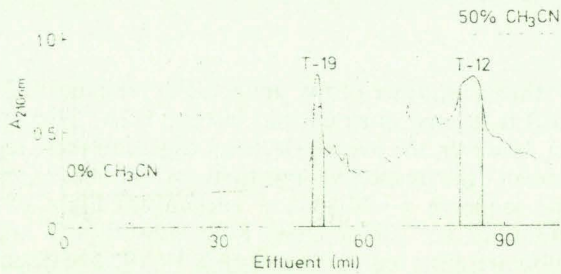


Fig. 7. Elution pattern of large hydrophobic tryptic peptides of histone H2A fractionated by reverse-phase high-pressure liquid chromatography. The insoluble material formed upon acidification of the tryptic hydrolysate (1 mg) was loaded on a column of C₁₈ μ Bondapak (0.39 x 30 cm) equilibrated with 0.05% trifluoroacetic acid. The column was eluted with a linear gradient (---) from 0–50% acetonitrile in 0.05% trifluoroacetic acid. Fractions of 1.2 ml were collected at a flow rate of 2 ml/min. The elution of peptides was monitored at 210 nm. The peptides were numbered according to their position in the sequence of the molecule.

peptides generated from the native protein. The elution diagram of the small tryptic peptides separated on Chromobeads P is shown in Fig. 6. The separation of the large hydrophobic tryptic peptides by reverse-phase liquid chromatography is presented in Fig. 7. The amino acid composition of tryptic peptides is given in Table 2. Two peptides indicated in Fig. 6 by T-1a and T-1b were found to

Table 8. Comparison of the sequences of histones H2A from marine invertebrates and from calf. Deletions are indicated by ---. Numbering in this table and in the text refers to sequence positions in histone H2A

Calf thymus [3]	6	Gln-Gly-Gly-Lys-Ala-Arg-Ala-Lys-Ala-Lys-Ala-Lys-Thr-Arg-Ser-Ser-Arg	20	Val	30	Val	40	Glu	41	Glu	43	Val	44	Leu	51	Leu	55	Leu	59	Thr	62	Ile	66	Thr	76	Ile	79	Ile	98	Gly-Arg	99	Lys	121	Glu-Ser-His-His-Lys-Ala-Lys-Gly-Lys	129
<i>Sipunculus</i> erythrocyte	6	Gly-Gly-Lys-Ala-Lys-Gly-Lys-Ser-Lys-Ser-Arg-Ser-Ser-Arg	15	Gly	19	Ile	29	Ile	32	Leu	42	Ile	44	Leu	50	Met	54	Met	58	Ala	61	Val	66	Thr	75	Ile	78	Ile	97	Ser-Gly	120	Gln	123	Lys-Ser-Lys	
Cuttlefish testis [7]	6	Gly-Gly-Lys-Val-Lys-Gly-Lys-Ser-Lys-Thr-Arg-Ser-Ser-Arg	15	Gly	19	Ile	29	Ile	32	Leu	42	Val	44	Leu	50	Met	54	Met	58	Ala	61	Val	66	Thr	75	Ile	78	Ile	97	Ser-Gly	120	Gln	123	Lys-Ser-Lys	
Starfish gonads [9]	6	Gly-Gly-Lys-Ala-Arg-Ala-Lys-Ser-Arg-Ser-Arg-Ser-Arg	15	Gly	19	Ile	29	Ile	32	Leu	42	Val	44	Leu	50	Met	54	Met	58	Ala	61	Val	66	Thr	75	Ile	78	Ile	97	Ser-Gly	120	Gln	123	Lys-Ser-Lys	
Sea urchin gonads [6]	6	Gly-Ala-Lys-Gly-Lys-Ala-Lys-Ser-Arg-Ser-Ser-Arg	15	Gly	19	Ile	29	Ile	32	Leu	42	Val	44	Leu	50	Met	54	Met	58	Ala	61	Val	66	Thr	75	Ile	78	Ile	97	Ser-Gly	120	Gln	123	Lys-Ser-Lys	

through a single step of fractionation of whole histone by chromatography on Bio-Gel P10. On polyacrylamide gel electrophoresis at acidic pH sipunculid H2A appears as two bands stained with approximately the same intensity. Phosphorylation of the amino-terminal *N*-acetyls erine residue of histone H2A has been observed *in vivo*, in cultured cells [22] and regenerating rat liver [23]. *In vitro*, serine-1 and serine-19 were found phosphorylated at the same extent in calf and rat chloroleukemia H2A [24]. Generally it has been suggested that H2A phosphorylation might be correlated with heterochromatin condensation [25, 26]. The phosphorylation of sipunculid erythrocyte H2A could probably be related to nuclear inactivation and chromatin condensation following the observation that H2A is the only histone in avian erythrocyte which continues to be phosphorylated during the late stages of red cell maturation [27]. However, to our knowledge, the physiological phosphorylation of histone H2A to such an extent (60%) has never been observed previously. On the other hand this high rate of phosphorylation of sipunculid H2A, together with variations in the primary sequence of H2A and H2B could be related to a smaller repeat length (177 ± 5 base pairs) of nucleosomal DNA in the chromatin of *Sipunculus nudus* as suggested previously by Mazen et al. [2].

The authors are indebted to D. Sellos and Dr Y. Le Gal (Laboratoire de Biologie Marine, Collège de France, Concarneau, France) for help in collecting sipunculid worms. They acknowledge with pleasure the skilful technical assistance of M. J. Dupire, T. Ernout and A. Hémez. They thank also G. Briand and O. Hannedouche from the *Service Commun de Séquence Unité 16/ Unité 124 INSERM* (Lille) for the sequencer runs. The authors are grateful to M. Blanquet for editorial assistance. This work was supported by grant 77.5.048.2 from the *Institut National de la Santé et de la Recherche Médicale* (INSERM) and grants from the *Centre National de la Recherche Scientifique* (ATP Chromatine 4203 and L.A. 268).

REFERENCES

- Mazen, A. & Champagne, M. (1976) *Exp. Cell. Res.* **103**, 119–125.
- Mazen, A., Champagne, M., Wilhelm, M. & Wilhelm, F. X. (1978) *Exp. Cell. Res.* **117**, 431–438.
- Sautière, P., Tyrou, D., Laine, B., Mizon, J., Ruffin, P. & Biserte, G. (1974) *Eur. J. Biochem.* **41**, 563–576.
- Laine, B., Sautière, P. & Biserte, G. (1976) *Biochemistry*, **15**, 1640–1645.
- Laine, B., Kmićek, D., Sautière, P. & Biserte, G. (1978) *Biochimie (Paris)* **60**, 147–150.
- Wouters, D., Sautière, P. & Biserte, G. (1973) *Eur. J. Biochem.* **90**, 231–239.
- Wouters-Tyrou, D., Martin-Ponthieu, A., Briand, G., Sautière, P. & Biserte, G. (1982) *Eur. J. Biochem.* **124**, 489–498.
- Chauvière, M., Sautière, P., Couppez, M. & Chevallier, P. (1982) *Comp. Biochem. Physiol.* **72 B**, 393–399.
- Martinage, A., Bélaïche, D., Dupressoir, T. & Sautière, P. (1983) *Eur. J. Biochem.* **130**, 465–472.
- Panyim, S. & Chalkley, R. (1969) *Arch Biochem. Biophys.* **130**, 337–346.
- Schmer, G. & Kriel, G. (1969) *Anal. Biochem.* **29**, 186–192.
- Doolittle, R. F. (1977) in *Advanced Methods in Protein Sequence Determination* (Needleman, S. B., ed.), pp. 38–54, Springer-Verlag, Berlin, Heidelberg, New York.
- Cohen, P., Rylatt, D. B. & Nimmo, G. A. (1977) *FEBS Lett.* **76**, 182–186.
- Hermann, J., Titani, K., Ericsson, L. H., Wade, R. D., Neurath, H. & Walsh, K. A. (1978) *Biochemistry*, **17**, 5672–5679.
- Tarr, G. E., Beecher, J. F., Bell, M. & McKean, D. (1978) *Anal. Biochem.* **84**, 622–627.
- Klapper, D. G., Wilde, C. E. & Capra, J. D. (1978) *Anal. Biochem.* **85**, 126–131.
- Edman, P. & Henschen, A. (1975) in *Protein Sequence Determination* (Needleman, S. B., ed.) 2nd ed, pp. 232–279, Springer-Verlag, Berlin, Heidelberg, New York.
- Briand, G., Kmićek, D., Sautière, P., Wouters, D., Borie-Loy, O., Biserte, G., Mazen, A. & Champagne, M. (1980) *FEBS Lett.* **112**, 147–151.
- Bailey, G. S. & Dixon, G. H. (1973) *J. Biol. Chem.* **248**, 5463–5472.
- Bradbury, E. M., Cary, P. D., Crane-Robinson, C., Rattle, H. W. E., Boublik, M. & Sautière, P. (1975) *Biochemistry*, **14**, 1876–1885.
- Rodrigues, J. de A., Brandt, W. F. & Von Holt, C. (1979) *Biochim. Biophys. Acta*, **578**, 196–206.
- Sung, M. T. & Dixon, G. H. (1970) *Proc. Natl Acad. Sci. USA*, **67**, 1616–1623.
- Sung, M. T., Dixon, G. H. & Smithies, O. (1971) *J. Biol. Chem.* **246**, 1358–1364.
- Martinage, A., Mangeat, P., Laine, B., Couppez, M., Sautière, P., Marchis-Mouren, G. & Biserte, G. (1980) *FEBS Lett.* **118**, 323–329.
- Gurley, L. R. & Walters, R. A. (1973) *Biochem. Biophys. Res. Commun.* **55**, 697–703.
- Halleck, M. S. & Gurley, L. R. (1980) *Exp. Cell Res.* **125**, 377–388.
- Ruiz-Carillo, A., Wangh, L. J. & Allfrey, V. G. (1976) *Arch. Biochem. Biophys.* **174**, 273–290.

D. Kmićek, M. Couppez, D. Bélaïche, and P. Sautière,

Unité de Recherche Ultrastructurales et Biochimiques sur les Cellules Normales et Cancéreuses

(Unité 124 de l'Institut National de la Santé et de la Recherche Médicale),

Institut de Recherches sur le Cancer de Lille, Cite Hospitalière, 2 Place Verdun, F-59020 Lille-Cedex, Nord, France

IV - 2 DETERMINATION DE LA STRUCTURE PRIMAIRE D'HISTONES H2A
ISOLEES D'INVERTEBRES MARINS

Dans notre laboratoire, la détermination de la séquence en acides aminés des histones H2A isolées de différents organismes animaux s'appuyait au début des années 70, sur l'étude des peptides tryptiques et chymotryptiques obtenus à partir de la protéine native ainsi que sur l'étude des peptides obtenus par hydrolyse de la protéine native avec la thermolysine. A cette époque la protéase staphylococcique (S.aureus souche V8) était à peine découverte (16).

La commercialisation de cette enzyme vers 1978 et l'acquisition d'un séquenceur par notre laboratoire nous ont amenés à choisir pour l'étude structurale d'histone H2A une nouvelle stratégie, fondée sur l'obtention de gros fragments peptidiques que l'on pouvait ensuite soumettre à la dégradation d'EDMAN automatisée.

Ces gros fragments sont essentiellement produits par hydrolyse, avec la protéase staphylococcique, des liaisons glutamyles situées aux positions 55, 63 et 91 de la séquence de l'histone H2A. Cette stratégie a été appliquée à la détermination de la structure primaire des histones H2A de Seiche, d'Etoile de mer et de Siponcle.

Les commentaires qui vont suivre, porteront sur l'étude structurale de l'histone H2A de testicule de Seiche, étude à laquelle nous avons apporté notre assistance technique.

La stratégie utilisée pour élucider la séquence de cette protéine est schématisée dans la figure 30.

L'étude structurale des fragments peptidiques obtenus par hydrolyse de la protéine avec la protéase staphylococcique (fragments SP-1 (1-55), SP-2 (56-63), SP-3 (64-124), SP-3a (64-91) et SP-3b (92-124) a été complétée par celle du fragment Ch-1 (39-124) obtenu par hydrolyse chymotrypsique limitée de l'histone H2A.

Tous les gros fragments peptidiques ont été partiellement ou complètement séquencés par la dégradation automatisée d'EDMAN. La déacétylation partielle de l'extrémité N-terminale de l'histone H2A produite par une activité déacylase présente dans différents lots de protéase staphylococcique, nous a permis d'aligner les 37 premiers résidus du peptide SP-1 par dégradation automatisée.

Sans cette déacétylation il eût fallu recourir à la fragmentation du peptide SP-1 (1-55) par des hydrolyses enzymatiques appropriées (avec la trypsine et la thermolysine) génératrices de petits peptides.

Les résultats obtenus pour la dégradation automatisée du peptide SP-1, avec un programme DMAA en présence de polybrène, sont présentés dans le tableau VIII. Le rendement répétitif de ce peptide calculé à partir des résidus de lysine 5 et 14 est de 93,8 p.100. Au cours de l'identification des PTH-AA de ce peptide, nous avons remarqué qu'il y avait eu une coupure secondaire mineure entre la lysine 12 et la sérine 13. Nous avons

FIGURE 30 STRATEGIE UTILISEE POUR LA DETERMINATION DE LA
STRUCTURE PRIMAIRE DE L'HISTONE H2A DE TESTICULE
DE SEICHE *Sepia officinalis*.

----- dégradation d'EDMAN automatisée des fragments obtenus par coupure des liaisons glutamyles de la protéine avec la protéase staphylococcique (*S.aureus* souche V8) (fragments SP), ou par coupure avec la chymotrypsine (fragment Ch) ou la clostripaïne (fragment Cl).

+***** dégradation manuelle d'EDMAN des peptides obtenus par coupure avec la protéase staphylococcique V8, ou la clostripaïne.

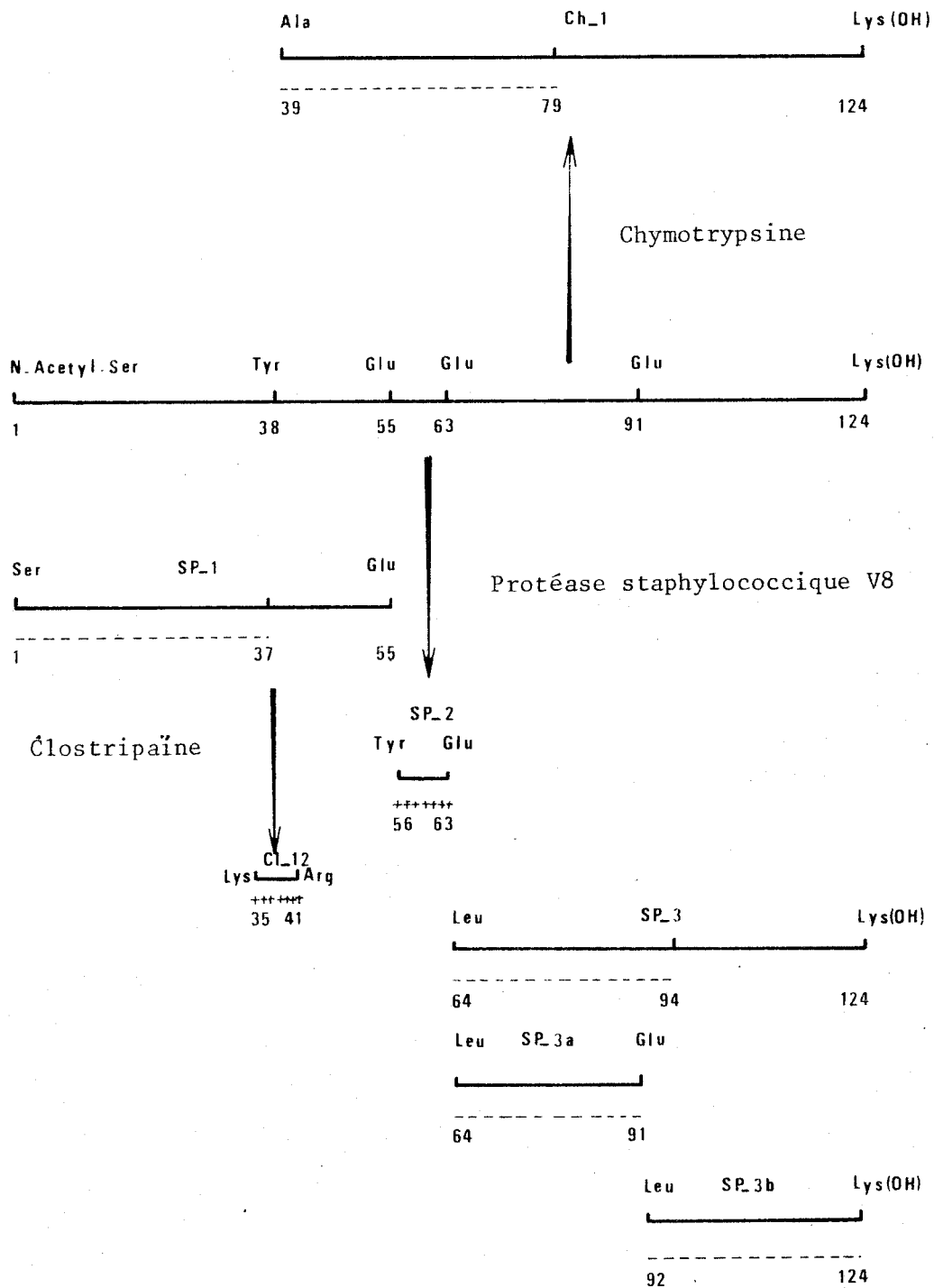


TABLEAU VIII

DEGRADATION D'EDMAN AUTOMATISEE DU PEPTIDE SP-1 (350 NANOMOLES)
AVEC UN PROGRAMME DMAA

N° Résidu	Résidu	Rendements en nanomoles		
		R-1 (1)	R(2)	R+1 (3)
1	Ser		48,4	7,3
2	Gly	59,9	175,3	36,1
3	Arg	42,9	n.d(4)	54,8
4	Gly	36,1	170,9	46
5	Lys	21,9	96,7	29,4
6	Gly	46	97,1	105,1
7	Gly	97,1	105,1	28,8
8	Lys	17,3	79,6	20,8
9	Val	n.d	n.d	n.d
10	Lys	20,8	76,6	22,5
11	Gly	25,2	81	23,5
12	Lys	22,2	69,5	15,5
13	Ser	3,3	11,5	
14	Lys	15,5	54,3	18,1
15	Thr		15,2	
16	Arg	34,2	86,4	62,7
17	Ser	1,4	9,3	8,3
18	Ser	9,3	8,3	3,9
19	Arg	45,1	83,6	52,5
20	Ala	9,1	37	8,2
21	Gly	19,1	37,6	21
22	Leu	13,4	n.d	12,2
23	Gln	30,8	47,6	34,2
24	Phe	7,4	23,4	4,1
25	Pro	1,5	19	7,6
26	Val	n.d	n.d	n.d
27	Gly	14,6	31,8	20,6
28	Arg	25,4	37,9	39,2
29	Ile		7,4	2,9
30	His	1,9	8,1	6,9
31	Arg	27,5	37,3	31,3
32	Leu	2,4	15,9	25,9
33	Leu	15,9	25,9	11,2
34	Arg	18,4	31,1	23,4
35	Lys	3,9	12,6	9,3
36	Gly	10,9	21	17,6
37	Asn	1,6	7,4	

- Voir la légende du tableau VI page 115.



donc séquencé les deux peptides en même temps ce qui n'a pas facilité les opérations d'identification.

Par ailleurs, la présence d'un résidu d'histidine au 30ème cycle induit dès l'étape de couplage avec le phényliso-thiocyanate en milieu diméthylallylamine, un clivage précoce de la chaîne peptidique. Le mécanisme de cette réaction catalysée par le noyau imidazole de l'histidine est présenté dans la figure 31. Il s'ensuit qu'au 30ème cycle on a non seulement un couplage avec le résidu d'histidine, mais aussi un couplage partiel avec le résidu suivant. Ce phénomène se traduit par un "overlap" important qui enlève rapidement toute possibilité d'identification des PTH-AA libérées.

Le peptide SP-3 (64-124) a été séquencé avec la diméthylbenzylamine comme tampon de couplage. Le rendement répétitif de ce peptide, calculé à partir des résidus d'alanine en position 2 et 22 est de 87,1 p.100. La séquence des 33 premiers résidus a pu être établie.

Le peptide SP-3a (64-91) a été soumis à la dégradation automatique avec un programme DMAA en présence de polybrène. Les résultats sont présentés dans le tableau IX. Le rendement répétitif calculé à partir des résidus d'alanine 2 et 20 est de 92,8 p.100. Le rendement de la dégradation est assez stable jusqu'au 16ème cycle puis chute assez fort au 17ème cycle, et nous voyons que malgré le double clivage effectué pour le résidu de proline en position 16, nous retrouvons encore 25 p.100 de ce résidu au 17ème cycle. Ce phénomène est sans doute lié à

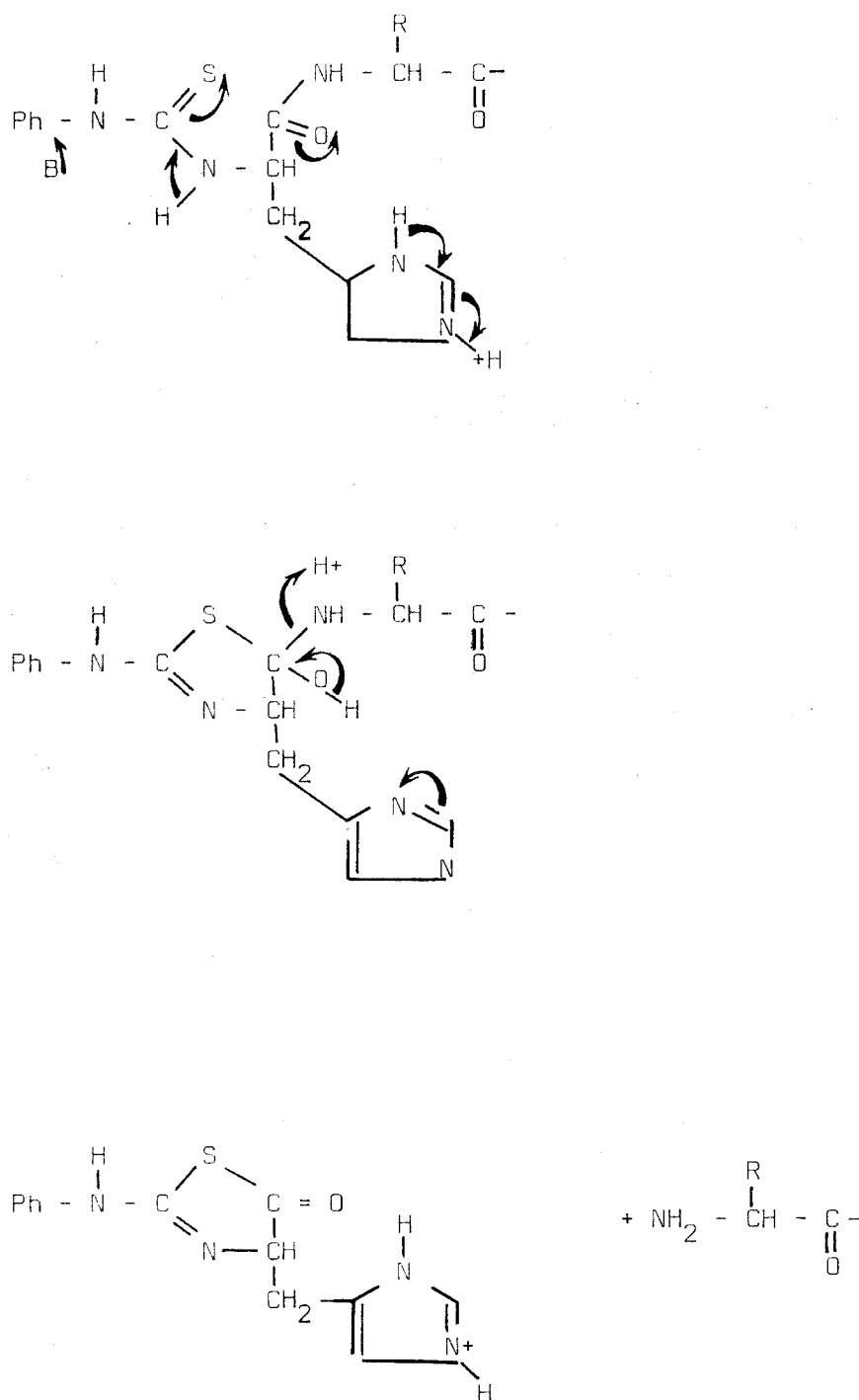


FIGURE 31 - MECANISME D'ACTION DE LA REACTION DE CLIVAGE DE
LA PHENYLTHIOCARBAMYL-HISTIDINE



TABLEAU IX

DEGRADATION D'EDMAN AUTOMATISEE DU PEPTIDE SP-3a (560 NANOMOLES)
AVEC UN PROGRAMME DMAA

N° Résidu	Résidu	Rendements en nanomoles		
		R-1(1)	R(2)	R+1(3)
64	Leu		545,7	36,7
65	Ala	34,6	597,8	47,9
66	Gly	17,4	550	41,3
67	Asn		542,9	61,2
68	Ala	17,1	573,1	554,3
69	Ala	573,1	554,3	41
70	Arg	24,3	361,7	175,4
71	Asp		441,9	169,9
72	Asn		473,6	163,5
73	Lys		509,6	443
74	Lys	509,6	443	73,3
75	Ser		165	
76	Arg		280,3	145,6
77	Ile		286,4	302,7
78	Ile	286,4	302,7	62,8
79	Pro		424,3	107,4
80	Arg		145,5	95
81	His		n.d(4)	n.d
82	Leu		209,6	72,4
83	Gln		188,5	82,2
84	Leu	72,4	190,6	62,5
85	Ala		132,9	59,5
86	Ile		109,9	38,4
87	Arg	28,4	128,7	75,9
88	Asn		68,9	
89	Asp	14,2	52,6	

- Voir la légende du tableau VI page 115.



la présence d'un résidu d'arginine adjacent au résidu de proline et qui provoque un encombrement stérique important. Néanmoins, malgré cette baisse de rendement au 17^{ème} cycle nous avons pu déterminer la séquence complète de ce peptide (28 résidus).

Le peptide SP-3b (92-124) a été séquencé avec un programme DMAA en présence de polybrène. Les résultats de la dégradation sont présentés dans le tableau X . Le rendement répétitif de ce peptide calculé à partir des résidus de leucine 92 et 115 est de 92 p.100. Un double clivage a été effectué pour les résidus de proline en positions 108 et 116. Après la proline 108, l'"overlap" de proline au cycle suivant est de 24 p.100. Après la proline 116, l'"overlap" de proline au cycle suivant s'élève à 43 p.100 malgré le double clivage. L'importance de cet "overlap" est vraisemblablement liée à la présence de 2 résidus de lysine, placés immédiatement après le résidu de proline 116, dont l'encombrement stérique important gêne le clivage de la proline .

L'alignement des peptides SP- et les compléments d'information nécessaires à l'élucidation de la séquence complète de l'histone H2A de Seiche ont été obtenus par l'étude structurale du peptide chymotrypsique Ch-1.

Ce peptide a été produit par hydrolyse limitée de l'histone H2A avec la chymotrypsine. L'histone H2A placée en milieu de force ionique élevée (0.5M) acquiert un maximum de structure

TABLEAU X

DEGRADATION D'EDMAN AUTOMATISEE DU PEPTIDE SP-3b (450 NANOMOLES)
AVEC UN PROGRAMME DMAA

N° Résidu	Résidu	Rendements en nanomoles		
		R-1(1)	R(2)	R+1(3)
92	Leu		316,8	122
93	Asn		460,4	54,3
94	Lys	24,7	373,6	45
95	Leu	63	264	235,4
96	Leu	264	235,4	11,3
97	Ser		37,5	
98	Gly	25,5	264,1	107,8
99	Val		n.d(4)	n.d
100	Thr		69,1	8,6
101	Ile		87	10,6
102	Ala	10,8	172,3	37,7
103	Gln		102,5	29,2
104	Gly	5,2	140,8	150,8
105	Gly	140,8	150,8	49
106	Val		n.d	n.d
107	Leu		88,9	19,4
108	Pro		99,1	23,7
109	Asn		45,3	23
110	Ile		27,1	22,7
111	Gln		89,1	20,5
112	Ala		78	21,8
113	Val		n.d	n.d
114	Leu		36,5	43,4
115	Leu	36,5	43,4	13,5
116	Pro		23,2	10
117	Lys		26,2	27,9
118	Lys	26,2	27,9	17,3
119	Thr		3,7	
120	Gln		n.d	n.d
121	Lys	11,8	15	10,9
122	Ala	4,1	12,1	14,1
123	Ala	12,1	14,1	10,2

- Voir la légende du tableau VI page 115.



globulaire. Dans ces conditions, la plupart des sites potentiels de clivage deviennent inaccessibles à l'enzyme, et la coupure majeure se situe au niveau du résidu de tyrosine 38. Une coupure mineure intervient au niveau du résidu d'histidine 30. Les peptides Ch-1 (39-124) et Ch-1a (31-124) n'ont pu être séparés et ont été soumis ensemble à la dégradation d'EDMAN automatisée avec un programme DMBA en présence de polybrène (tableau XI) . Le rendement répétitif du peptide Ch-1 calculé à partir des résidus d'alanine 39 et 69 est de 97,3 p.100. La quantification en chromatographie liquide haute performance des PTH-AA libérées après chaque cycle de dégradation a permis d'établir un rapport de 3:1 pour les deux peptides dans le mélange. Le peptide Ch-1 est nettement le composant majeur du mélange.

Ainsi, avec les données structurales acquises sur les fragments SP- et la comparaison avec la séquence de l'histone H2A de Veau, nous avons pu déterminer sans ambiguïté la séquence N-terminale (41 résidus) du peptide Ch-1.

TABLEAU XI

DEGRADATION D'EDMAN AUTOMATISEE DU PEPTIDE Ch-1 (195 NANOMOLES)
 AVEC UN PROGRAMME DMBA

N° Résidu	Résidu	Rendements en nanomoles		
		R-1(1)	R(2)	R+1(3)
39	Ala		144,3	11,6
40	Gln		163,6	
41	Arg		151	164
42	Val		n.d(4)	n.d
43	Gly		110,8	59
44	Ala		99,5	31,6
45	Gly	59	114,7	25,3
46	Ala	31,6	113,8	77,1
47	Pro		47,7	28,8
48	Val		n.d	n.d
49	Tyr		50,3	34,7
50	Leu		54,9	38,7
51	Ala	16,5	106,1	151,2
52	Ala	106,1	151,2	73,8
53	Val		n.d	n.d
54	Met		n.d	n.d
55	Glu		41	21,9
56	Tyr	14,3	37,5	62,8
57	Leu	15,1	56,8	60,1
58	Ala	40,8	66,2	132,1
59	Ala	66,2	132,1	88,5
60	Glu		24	22,8
61	Val		n.d	n.d
62	Leu		51,3	39,5
63	Glu		24	29,9
64	Leu	39,5	57,9	52,1
65	Ala	33,2	47,4	60,1
66	Gly		26,7	24
67	Asn		13,6	16,5
68	Ala	52,7	53,2	64,3
69	Ala	53,2	64,3	54,5
70	Arg		25,8	
71	Asp		24,3	
72	Asn		29	17,2
73	Lys		20,1	31,3
74	Lys	20,1	31,3	29,7
75	Ser		5,1	2,7
76	Arg	16,6	27,8	27,6
77	Ile		8,7	20,4
78	Ile	8,7	20,4	11,2
79	Pro		5,4	

- Voir la légende du tableau VI page 115.



IV - 3 DETERMINATION DE LA STRUCTURE PRIMAIRE DE PROTAMINES
DE MAMMIFERES ET DE POISSONS

Ce travail s'inscrit dans une thématique pluridisciplinaire* qui a pour objet l'étude de l'organisation de la chromatine du spermatozoïde mûr ou en cours de différenciation dans quelques espèces animales.

IV - 3-1 Protamines de Mammifères

Les protamines de mammifères (environ 50 résidus) ont une taille moléculaire supérieure à celle des protamines de poissons (environ 30 résidus) et renferment des acides aminés aromatiques et de la cystéine alors que les protamines de poissons en sont généralement dépourvues.

IV - 3-1-1 Structure primaire de la protamine de sperme
de Bélier (*Ovis aries*)

La protamine de Bélier est un polypeptide riche en arginine constitué de 50 résidus d'acides aminés (tableau XII).

La structure primaire de la protamine de Bélier a été complètement déterminée par dégradation d'EDMAN automatisée

* Cette thématique regroupe les laboratoires suivants :
Ph.Chevallier, UER de Biologie Cellulaire, Créteil ;
Y.Le Gal et D.Sellos, Laboratoire de Biologie Marine,
Concarneau ; M.Loïr, INRA, Rennes.

TABLEAU XII

COMPOSITION EN ACIDES AMINES DE LA PROTAMINE DE BELIER

	mol./100 mol.	Nombre de résidus par molécule de protamine
S-carboxyméthylcystéine	15,5	7,7 (7)
Thréonine (a)	5,4	2,7 (3)
Sérine (a)	4,0	2,0 (2)
Acide glutamique	2,0	1,0 (1)
Glycine	2,4	1,2 (1)
Alanine	1,6	0,8 (1)
Valine (b)	5,3	2,6 (3)
Leucine	2,0	1,0 (1)
Tyrosine	3,5	1,7 (2)
Phénylalanine	1,8	0,9 (1)
Histidine	1,9	0,9 (1)
Arginine	54,8	27,4 (27)
Total		(50)

Les quantités de chaque acide aminé ont été calculées à partir de deux analyses après hydrolyse de 24 et 72 heures. Le nombre de résidus a d'abord été calculé sur la base d'un résidu de leucine.

(a) = Valeurs obtenues par extrapolation linéaire au temps zéro de l'hydrolyse.

(b) = Valeur obtenue de l'hydrolysats de 72 heures.

- Les chiffres entre parenthèses correspondent au nombre de résidus déterminés par la séquence.



avec un programme DMAA en présence de polybrène (tableau XIII) Deux passages au séquenceur ont été effectués : le premier avec la protamine, réduite et alkylée par l'iodoacétamide (tableau XIII), le second après réduction et blocage des groupes thiols par la 4-vinyl-pyridine. Les résidus de cystéine ont donc été identifiés sous forme de PTH-S-carboxamidométhyl-cystéine au cours du premier passage et sous forme de PTH-S-éthyl-pyridine cystéine au cours du second passage. La séquence C-terminale de la protamine a été également établie par une étude cinétique de l'hydrolyse de la protamine par les carboxypeptidases A et B, ainsi que par l'étude structurale du peptide chymotrypsique C-terminal (résidus 43-50) de la protamine.

Pour éliminer au maximum les "overlaps" créés par les séquences riches en résidus d'arginine rencontrées habituellement dans les protamines, un programme comportant un double clivage après chaque cycle a été utilisé du 15ème au 28ème cycle inclus.

La structure primaire de la protamine de Bélier présentée dans la figure 32, est caractérisée par l'accumulation de résidus d'arginine dans la partie centrale de la molécule, où on observe trois séquences polyarginine (séquences 15-21, 23-28 et 31-36). Les séquences 13-20 et 21-28 sont identiques : la présence de séquences répétitives est fréquemment observée dans les protamines.

La comparaison de nos résultats avec ceux de COELINGH et coll. (15) confirme l'existence de variations entre les protamines de deux espèces (le Bélier et le Taureau) appartenant à

TABLEAU XIII

DEGRADATION D'EDMAN AUTOMATISEE DE LA PROTAMINE DE BELIER (257 NANOMOLES)

AVEC UN PROGRAMME DMAA

N° Résidu	Résidu	Rendements en nanomoles		
		R-1(1)	R(2)	R+1(3)
1	Ala		98,7	5,8
2	Arg	17,8	102,2	33,3
3	Tyr		117,1	19,9
4	Arg	33,3	106,8	36,9
5	Cys		n.d(4)	n.d
6	Cys	n.d	n.d	n.d
7	Leu		70,1	11,5
8	Thr		21,1	
9	His		n.d	n.d
10	Ser		29,3	10,6
11	Arg	24,7	60,4	39,6
12	Ser	10,6	16	
13	Arg	39,6	127,5	81,6
14	Cys		n.d	n.d
15	Arg	81,6	117,8	115,2
16	Arg	117,8	115,2	120,1
17	Arg	115,2	120,1	104,6
18	Arg	120,1	104,6	116,5
19	Arg	104,6	116,5	110,9
20	Arg	116,5	110,9	113,7
21	Arg	110,9	113,7	70,5
22	Cys		n.d	n.d
23	Arg	70,5	86,3	79,7
24	Arg	86,3	79,7	89,6
25	Arg	79,7	89,6	85,7
26	Arg	89,6	85,7	77,8
27	Arg	85,7	77,8	75
28	Arg	77,8	75	50,6
29	Phe		20,4	8
30	Gly	1,3	21,9	10
31	Arg	40,6	52,8	51,2
32	Arg	52,8	51,2	55
33	Arg	51,2	55	56,5
34	Arg	55	56,5	51,6
35	Arg	56,5	51,6	52,8
36	Arg	51,6	52,8	45
37	Val		n.d	n.d
38	Cys		n.d	n.d
39	Cys	n.d	n.d	n.d
40	Arg	34,9	43,4	48
41	Arg	43,4	48	43,3
42	Tyr	1,2	15,9	10,1
43	Thr	1,2	4,4	3,2
44	Val		n.d	n.d
45	Val	n.d	n.d	n.d
46	Arg	n.d	n.d	31,8
47	Cys		n.d	n.d
48	Thr		5,3	2,5
49	Arg	29,5	32,2	30,7
50	Gln		2,6	

- Voir la légende du tableau VI page 115.



(H) Ala-Arg-Tyr-Arg-Cys-Cys-Leu-Thr-His-Ser-Arg-Ser-Arg-Cys-Arg-Arg-Arg-Arg-

Arg-Arg-Arg-Cys-Arg-Arg-Arg-Arg-Arg-Phe-Gly-Arg-Arg-Arg-Arg-Arg-Arg-

Val-Cys-Cys-Arg-Arg-Tyr-Thr-Val-Val-Arg-Cys-Thr-Arg-Gln (OH)

FIGURE 32 SEQUENCE EN ACIDES AMINES DE LA PROTAMINE DE SPERME DE BELIER



une même famille zoologique (Bovidés) (figure 33). Ces variations ne concernent pas toutefois la région centrale (résidus 13-36) extrêmement riche en arginine, mais seulement les régions N- et C-terminales moins basiques. Chez les Mammifères la région C-terminale est plus variable que la région N-terminale. Entre les protamines de Bélier et de Taureau, on observe un changement non conservatif en position 11 (Arginine_{Bélier} — Glycine_{Taureau}), une délétion de 3 résidus (Cys₃₉ - Arg - Arg₄₁ -) dans la protamine de Taureau, un changement conservatif en position 45 (Valine_{Bélier} — Isoleucine_{Taureau}).

IV - 3-1-2 Séquence N-terminale de la protamine de sperme d'Etalon (*Equus equus*)

La protamine d'Etalon est un polypeptide riche en arginine constitué de 50 résidus d'acides aminés (tableau XIV). Elle se différencie de la protamine de Bélier par l'absence de glycine, de phénylalanine et d'histidine.

Un premier essai de détermination de séquence a été réalisé avec le séquenceur équipé du piège à froid Beckman, des nouvelles vannes de délivrance et de la conversion automatique.

312 nanomoles de protamine S-carboxyméthylée ont été soumises à la dégradation d'EDMAN automatisée avec un programme quadrol 0.33M en présence de polybrène. Les 30 premiers résidus de la protéine ont été déterminés. Les résultats obtenus lors de ce premier passage au séquenceur sont présentés dans le tableau XV.

TABLEAU XIV

COMPOSITION EN ACIDES AMINES DE LA PROTAMINE D'ETALON

	mol./100 mol.	Nombre de résidus par molécule de protamine
S-carboxyméthylcystéine	15,8	7,7(8)
Thréonine (a)	2,1	1,0(1)
Sérine (a)	8,0	3,9(4)
Acide glutamique	6,2	3,0(3)
Alanine	1,4	0,7(1)
Valine (b)	6,0	2,9(3)
Leucine	2,1	1,0(1)
Tyrosine	3,9	1,9(2)
Arginine	54,5	26,7(27)
Total		(50)

Les quantités de chaque acide aminé ont été calculées à partir de deux analyses effectuées après hydrolyse de 24 et 72 heures. Le nombre de résidus a d'abord été calculé sur la base d'un résidu de leucine.

(a) = Valeurs obtenues par extrapolation linéaire au temps zéro de l'hydrolyse.

(b) = Valeur obtenue de l'hydrolysate de 72 heures.

- Les chiffres entre parenthèses représentent l'entier le plus proche.



TABLEAU XV

DEGRADATION D'EDMAN AUTOMATISEE DE LA PROTAMINE D'ETALON (312 NANOMOLES)AVEC UN PROGRAMME QUADROL

N° Résidu	Résidu	Rendements en nanomoles		
		R-1(1)	R(2)	R+1(3)
1	Ala		122,6	17,1
2	Arg	10,5	137,8	38,1
3	Tyr		87	17,6
4	Arg	38,1	125,6	36,5
5	Cys		n.d(4)	n.d
6	Cys	n.d	n.d	n.d
7	Arg	18,3	86,9	28,5
8	Ser		n.d	n.d
9	Gln Glu		40 35,8 4,2	14,4
10	Ser	n.d	n.d	n.d
11	Gln Glu	14,4	31,9 28,5 3,4	13,9
12	Ser	n.d	n.d	n.d
13	Arg		39	19,2
14	Cys		n.d	n.d
15	Arg	19,2	20,6	20,7
16	Arg	20,6	20,7	21
17	Arg	20,7	21	19,3
18	Arg	21	19,3	18
19	Arg	19,3	18	15,6
20	Arg	18	15,6	14,1
21	Arg	15,6	14,1	8,4
22	Cys		n.d	n.d
23	Arg	8,4	9,5	8,8
24	Arg	9,5	8,8	10
25	Arg	8,8	10	8,6
26	Arg	10	8,6	7,7
27	Arg	8,6	7,7	7,6
28	Arg	7,7	7,6	7,8
29	Ser		n.d	n.d
30	Val		1,3	0,2



On y observe que l'accumulation de résidus d'arginine, en positions 15 à 21 puis 23 à 28, provoque des "overlaps" importants. Ceux-ci créent un "bruit de fond" élevé, et de ce fait nous n'avons pu déterminer que les 30 premiers résidus de la protamine. On peut également remarquer qu'en positions 9 et 11, le taux de PTH-glutamine est de 90 p.100. Le faible taux de désamidation (10 p.100 environ) est dû à l'utilisation de la conversion automatique. En conversion manuelle, le taux de désamidation atteint fréquemment 40 p.100. Toutefois, la conversion automatique ne nous a donné jusqu'à présent de bons résultats qu'avec un programme quadrol, et il paraît évident que le programme quadrol n'est pas le mieux adapté à la dégradation d'EDMAN d'une protamine, quand on compare les résultats de la dégradation automatisée de la protamine d'Etalon avec ceux de la dégradation d'EDMAN automatisée de la protamine de Bélier réalisée avec un programme DMAA et la conversion manuelle.

La séquence N-terminale (30 résidus) de la protamine d'Etalon est présentée dans la figure 34. Cette séquence montre une homologie très grande avec les séquences de protamines de Bélier et de Taureau, en particulier dans les zones 1-6 et 10-28 (figure 33).

(H) Ala-Arg-Tyr-Arg-Cys-Cys-Arg-Ser-Gln-Ser-
Gln-Ser-Arg-Cys-Arg-Arg-Arg-Arg-Arg-Arg-
Arg-Cys-Arg-Arg-Arg-Arg-Arg-Arg-Ser-Val-

FIGURE 34 SEQUENCE N-TERMINALE DE LA PROTAMINE
MAJEURE DU SPERME D'ETALON.



IV - 3-2 Protamines de Poissons. Structure primaire de la
scylliorhinine 4 du sperme de Roussette (*Scylliorhinus caniculus*)

Le testicule de la Roussette a une organisation régulière où l'on distingue trois domaines à l'intérieur desquels les cellules sont pratiquement au même stade de différenciation. Au cours de la spermiogénèse les histones sont remplacées par des protéines de caractère plus basique, appelées protéines intermédiaires S1 et S2 qui sont à leur tour remplacées dans les spermatozoïdes mûrs par quatre protamines Z1, Z2, Z3 et S4 désignées sous le nom de scylliorhines.

La scylliorhinine 4 est un polypeptide de 32 résidus d'acides aminés (tableau XVI) dont la composition présente des caractéristiques inhabituelles par rapport aux autres protamines de poissons connues à ce jour. La scylliorhinine 4 contient en effet deux fois plus de lysine que d'arginine. Elle est totalement dépourvue d'acides aminés dicarboxyliques, d'acides aminés hydroxylés et d'acides aminés aromatiques. Elle renferme de la cystéine alors que les protamines de poissons en sont généralement dépourvues.

La structure primaire de la scylliorhinine 4 a été complètement déterminée par dégradation d'EDMAN automatisée avec un programme DMAA en présence de polybrène (tableau XVII). Le rendement répétitif calculé à partir des résidus d'alanine en

TABLEAU XVI

COMPOSITION EN ACIDES AMINES DE LA SCYLLIORHININE 4 DE ROUSSETTE

	mol./100 mol.	Nombre de résidus par molécule
S-carboxyméthylcystéine	15,5	5,3 (4)
Proline	6,6	2,2 (2)
Glycine	2,9	1,0 (1)
Alanine	10,2	3,5 (3)
Valine	3,0	1,0 (1)
Lysine	40,7	13,7 (14)
Arginine	21,1	7,0 (7)
Total		32

Les quantités de chaque acide aminé ont été calculées à partir d'au moins cinq analyses effectuées sur des hydrolysats de 24 heures. Le nombre de résidus a d'abord été calculé sur la base d'un résidu de glycine.

Les chiffres entre parenthèses correspondent au nombre de résidus déterminés par la séquence.



TABLEAU XVII

DEGRADATION D'EDMAN AUTOMATISEE DE LA PROTAMINE S4 (SCYLLIORHININE 4)DE SPERME DE ROUSSETTE *Scylliorhinus caniculus* (450 NANOMOLES) AVECUN PROGRAMME DMAA

N° Résidu	Résidu	Rendements en nanomoles		
		R-1(1)	R(2)	R+1(3)
1	Gly		228,7	24,5
2	Cys		n.d	n.d
3	Lys	17	237,9	254,3
4	Lys	237,9	254,3	50,1
5	Arg	13,6	244,5	76,7
6	Lys	50	228,9	59
7	Ala		207,7	28
8	Arg	40,1	217,6	76,8
9	Lys	26,6	176,8	58,8
10	Arg	76,8	226,3	77,6
11	Pro		91,2	19,4
12	Lys	25	117,6	31,8
13	Cys		n.d	n.d
14	Lys	31,8	87,2	89
15	Lys	87,2	89	40,1
16	Ala		110,9	22,9
17	Arg	10,5	93,9	36,1
18	Lys	24,6	71,6	32,9
19	Arg	36,1	97,2	46,1
20	Pro		41,8	17,5
21	Lys	18,4	42,6	22,1
22	Cys		n.d	n.d
23	Lys	22,1	44,7	26,8
24	Arg	14,8	43,5	55,8
25	Arg	43,5	55,8	33,4
26	Lys	16,2	31,5	18
27	Val	14,2	43	26,4
28	Ala		35,9	14,4
29	Lys	13,3	24,1	25,6
30	Lys	24,1	25,6	27,2
31	Lys	25,6	27,2	16,1
32	Cys		n.d	

- Voir la légende du tableau VI page 115.



position 7 et 16 est de 94,1 p.100. Le rendement est stable jusqu'au 10ème cycle puis chute environ de moitié au cycle suivant. En dépit du triple clivage réalisé pour le résidu de proline en position 11, un "overlap" de 25 p.100 a été obtenu au cycle suivant, vraisemblablement à cause de la séquence Lys-Cys-Lys-Lys adjacente au résidu de proline. Pour la proline en position 20, malgré le double clivage que nous avons effectué, nous avons un "overlap" de 46 p.100 au cycle suivant mais, comme pour la proline précédente, la proximité de 3 résidus de lysine doit gêner fortement le clivage de cette proline qui est de plus située près de l'extrémité C-terminale de la protamine. Cependant, malgré la chute de rendement au 10ème cycle, les "overlaps" créés par les deux résidus de proline et les nombreux doublets ou triplets de lysine, la protamine S4 (scylliorhinine 4) (32 résidus) a pu être séquencée entièrement par dégradation d'EDMAN automatisée. La séquence C-terminale de la protamine a également été élucidée par l'étude structurale du peptide C-terminal (résidus 27-32) obtenu par hydrolyse de la protamine avec la thermolysine.

La séquence en acides aminés de la scylliorhinine 4 (figure 35), est caractérisée par une distribution régulière des résidus basiques le long de la chaîne peptidique par groupes de 3 ou 4 résidus séparés par un acide aminé hydrophobe (séquences 3-6, 8-10, 17-19, 23-26, 29-31). Elle est en outre caractérisée par la succession de deux séquences identiques de neuf résidus (séquences 6-14 et 15-23).

(H) Gly-Cys-Lys-Lys-Arg-Lys-Ala-Arg-Lys-Arg-
Pro-Lys-Cys-Lys-Lys-Ala-Arg-Lys-Arg-Pro-
Lys-Cys-Lys-Arg-Arg-Lys-Val-Ala-Lys-Lys-
Lys-Cys (OH)

FIGURE 35 SEQUENCE EN ACIDES AMINES DE LA
SCYLLIORHININE 4 DE ROUSSETTE

Bien que la scylliorhinine 4 présente de fortes homologues structurales avec la scylliorhinine 3 et la clupéine Z que l'on peut considérer comme la protamine de référence chez les poissons (figure 36), elle se démarque nettement des protamines de poissons connues à ce jour, par la présence de lysine et de cystéine, deux acides aminés dont les protamines de poissons sont habituellement dépourvues.

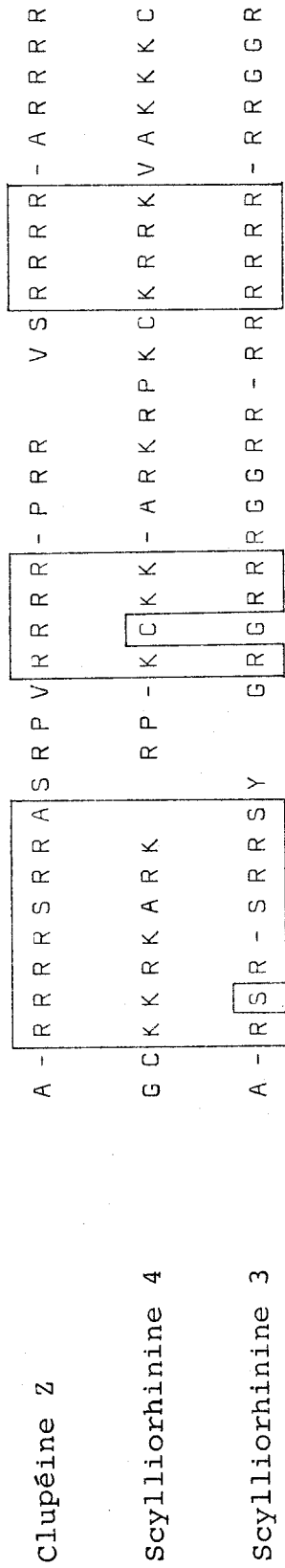


FIGURE 36 HOMOLOGIES DE SEQUENCE ENTRE LES PROTAMINES DE POISSONS

Les séquences ont été alignées pour un maximum d'homologie.
 Les cadres correspondent à une très forte homologie.



CONCLUSION

Dans ce mémoire, nous avons présenté la dégradation d'EDMAN automatisée, des protéines et des peptides, réalisée en phase liquide à l'aide du séquenceur Beckman 890 C.

Le grand pouvoir de résolution et le seuil élevé de la sensibilité de la chromatographie liquide haute performance utilisée pour identifier et doser les phénylthiohydantoïnes d'acides aminés nous ont amenée à apporter à l'appareil les modifications techniques suivantes : installation du piège à froid, changement des vannes de délivrance, réalisation et installation d'un système de conversion automatique.

Ces modifications concourent à améliorer le rendement de la dégradation et donc à accroître le nombre de résidus d'acides aminés identifiables dans la chaîne peptidique.

Elles ont en outre pour effet de :

- prolonger la durée de vie des pompes à vide.
- minimiser les quantités de solvants et de réactifs nécessaires à la dégradation et de ce fait d'avoir moins de perte de la protéine (ou du peptide) lors des étapes de lavage.
- limiter au maximum la destruction des anilino-thiazolinones d'acides aminés.
- améliorer la sensibilité de l'appareil.

Dans un avenir proche, nous nous proposons d'asservir les différentes fonctions du séquenceur à un microprocesseur pour remplacer le programme à bande perforée.

Il reste enfin à améliorer par distillation, ou par passage sur colonne absorbante, la pureté des réactifs et des solvants vendus très chers dans la qualité "Sequanal grade". Mais il faut savoir que ces opérations de purification sont très contraignantes à la fois sur le plan technique et sur le plan de la sécurité.

Cependant, l'avènement de la microméthode de séquence en phase "gazeuse" ouvre un champ d'investigations nouvelles dans des domaines tels que la biologie du développement et la neurobiologie où les quantités de protéine disponibles pour l'étude structurale sont généralement très faibles et incompatibles avec la méthode de séquence en phase liquide.

BIBLIOGRAPHIE

- | | Pages |
|--|--------|
| <p>1 - AFRICA, B. and CARPENTER, F.H. (1966)</p> <p style="padding-left: 2em;">Alkaline hydrolysis of phenylthiohydantoins of amino acids.</p> <p style="padding-left: 2em;"><u>Biochem. Biophys. Res. Commun.</u>, <u>24</u>, 113-119.</p> | 91 |
| <p>2 - BEGG, G.S. and MORGAN, F.J. (1976)</p> <p style="padding-left: 2em;">A non-volatile buffer with improved performance in automated protein sequencing.</p> <p style="padding-left: 2em;"><u>FEBS Letters</u> <u>66</u>, 243-245.</p> | 44 |
| <p>3 - BHOWN, A.S., CORNELIUS, T.W., MOLE, J.E., LYNN, J.D. and TIDWELL, W.A. and BENNETT, J.C. (1980)</p> <p style="padding-left: 2em;">A simple modification on the vacuum system of the Beckman automated sequencer to improve the efficiency of Edman degradation.</p> <p style="padding-left: 2em;"><u>Anal. Biochem.</u>, <u>102</u>, 35-38.</p> | 57 |
| <p>4 - BHOWN, A.S., MOLE, J.E. and BENNETT, J.C. (1981)</p> <p style="padding-left: 2em;">An improved procedure for high-sensitivity microsequencing : use of aminoethyl amino-propyl glass beads in the Beckman sequencer and ultrasphere ODS-column for PTH-amino acid identification.</p> <p style="padding-left: 2em;"><u>Anal. Biochem.</u>, <u>110</u>, 355-359.</p> | 71, 95 |
| <p>5 - BHOWN, A.S., BENNETT, J.C., MORGAN, P.H. and MOLE, J.E. (1981)</p> <p style="padding-left: 2em;">Use of fluorescamine as an effective blocking reagent to reduce the background in protein sequence analyses by the Beckman automated sequencer.</p> <p style="padding-left: 2em;"><u>Anal. Biochem.</u>, <u>112</u>, 158-162.</p> | 47 |

- 6 - BRANDT, W.F. and VON HOLT, C. (1972) 45
 The complete amino acid sequence of histone F3 from Chicken erythrocytes.
FEBS Letters 23, 357-363.
- 7 - BRANDT, W.F., EDMAN, P., HENSCHEN, A. and VON HOLT, C. (1976) 35
 Abnormal behaviour of proline in the isothiocyanate degradation.
Hoppe Seyler's Z.Phys.Chem., 357, 1505-1508.
- 8 - BRAUER, A.W., MARGOLIES, M.N. and HABER, E. (1975) 44
 The application of 0.1M quadrol to the microsequence of proteins and sequence of tryptic peptides.
Biochemistry 14, 3029-3035.
- 9 - BRAUNITZER, G. and SCHRANK, B. (1970) 45
 3-dimethylamino-1-propin als Puffer-substanz.
Hoppe Seyler's Z.Phys.Chem., 351, 417.
- 10 - BRAUNITZER, G., SCHRANK, B. and RUHFUS, A. (1970) 41, 42
 Versuche zum vollständigen and automatischen abbau von peptiden nach der quadrol methode.
Hoppe Seyler's Z.Phys.Chem., 351, 1589-1590.
- 11 - BRAUNITZER, G., SCHRANK, B., PETERSEN, S. and PETERSEN, U. (1973) 41, 42
 Über ϵ -markierung von peptiden. Automatische sequenzanalyse des insulins.
Hoppe Seyler's Z.Phys.Chem., 354, 1563-1566.
- 12 - CHANG, J.Y., BRAUER, D. and WITTMANN-LIEBOLD, B. (1978) 17
 Microsequence analysis of peptides and proteins using 4-NN-dimethyl-aminoazobenzene 4'-isothiocyanate/phenylisothiocyanate double coupling method.
FEBS Letters 93, 205-214.

- 13 - CHANG, J.Y., LEHMANN, A. and WITTMANN-LIEBOLD, B. 17
(1980)
Analysis of DABTH-AA by HPLC.
Anal.Biochem., 102, 380-383.
- 14 - COHEN-SOLAL, M. and BERNARD, J.L. (1973) 87
Miniature thin-layer chromatography of
phenylthiohydantoin amino-acids. Appli-
cation to automatic Edman degradation.
J.Chromtogr., 80, 140-143.
- 15 - COELINGH, J.P., MONFOORT, C.H., ROZIJN, T.H., 136
GEVERS LEUVEN, J.A., SCHIPHOF, R., STEYN-PARVE, E.,
BRAUNITZER, G., SCHRANK, B. and RUHFUS, A. (1972)
The complete amino acid sequence of the
basic nuclear protein of Bull spermatozoa.
Biochim.Biophys.Acta 285, 1-14.
- 16 - DRAPEAU, G.R., BOILY, Y. and HOUMARD, J. (1972) 123
Purification and properties of an extra-
cellular protease of *Staphylococcus aureus*.
J.Biol.Chem., 247, 6720-6726.
- 17 - DWULET, F.E. and GURD, F.R.N. (1976) 43
A comparison of sulfonated phenyl-isothio-
cyanates for reducing losses of lysine-
containing peptides during automated se-
quencing.
Anal.Biochem., 76, 530-538.
- 18 - EDMAN, P. (1949) 10
A method for the determination of the
amino acid sequences in peptides.
Arch.Biochem.Biophys., 22, 475
- 19 - EDMAN, P. (1950) 10
Method for determination of the amino
acid sequence in peptides.
Acta Chem.Scand., 4, 283

- 20 - FAIRWELL, T. and BREWER, H.B. (1979) 62
Automated Edman degradations : studies
with a large sequencer cup and high-speed
drive.
Anal.Biochem., 99, 242-248.
- 21 - FOHLMAN, J., RASK, L. and PETERSON, P.A. (1980) 95
High-pressure liquid chromatographic
identification of phenylthiohydantoin
derivatives of all twenty common amino
acids.
Anal.Biochem., 106, 22-26.
- 22 - FOSTER, J.A., BRUENGER, E., HU, C.C., ALBERTSON, K. 43
and FRANZBLAU, C. (1973)
A new technique for automated sequencing
of non-polar peptides.
Biochem.Biophys.Res.Comm., 53, 70-74.
- 23 - FRANK, G. and STRUBERT, W. (1973) 93
Identification of PTH (phenylthiohydan-
toin) amino acids by high-performance
liquid chromatography.
Chromatographia 6, 522-524.
- 24 - FRANK, G. and ZUBER, H. (1976) 39,40
The amino acid sequence of Monellin.
Hoppe Seyler's Z. Phys.Chem., 357, 585-592.
- 25 - FRANK, G. (1979) 47
A cheap and simple method to achieve an
maintain the necessary purity of reagents
and solvents for automated amino acid se-
quence determination with the sequanator.
Hoppe Seyler's Z.Phys.Chem., 360, 997-999.
- 26 - GARRICK, M.D. and SLOAN, R.L. (1977) 43
Preparative Biochem., 7, 111-128.
- 27 - GRAY, W.R. and HARTLEY, B.J. (1963) 14
A fluorescent end-group reagent for
proteins and peptides.
Biochem.J., 89, 379

- 28 - HENDERSON, L.E., COPERLAND, T.D. and OROSZLAN, S. 95
(1980)
Separation of amino acid phenylthiohydantoin-
ins by high-performance liquid chromatography on phenylalkyl support.
Anal. Biochem., 102, 1-7.
- 29 - HERMANN, J., TITANI, K., ERICSSON, L.H., WADE, R.D. 96
NEURATH, H. and WALSH, K.A. (1978)
Amino acid sequence of two cyanogen bromide
fragments of glycogen phosphorylase.
Biochemistry 17, 5672-5679.
- 30 - HERMODSON, M.A., ERICKSSON, L.H., TITANI, K., 44,116
NEURATH, H. and WALSH, K.A. (1972)
Application of sequanator analyses to the
study of proteins.
Biochemistry 11, 4493-4502.
- 31 - HEWICK, R.M., HUNKAPILLER, M.W., HOOD, L.E. and 49,83,84
DREYER, W.J. (1981)
A gas-liquid solid phase peptide and
protein sequanator.
J. Biol. Chem., 256, 7990-7997.
- 32 - HIRS, C.H.W., MOORE, S. and STEIN, W.H. (1960) 3
J. Biol. Chem., 235, 633-647.
- 33 - HUNKAPILLER, M.W. and HOOD, L.E. (1978) 40,72
Direct microsequence analysis of poly-
peptides using an improved sequanator, a
non protein carrier (polybrene), and high
pressure liquid chromatography.
Biochemistry 17, 2124-2133.
- 34 - HUNKAPILLER, M.W. and HOOD, L.E. (1983) 95
Analysis of PTH by ultrasensitive gra-
dient HPLC.
in Methods in Enzymology 91, 486-493.

- 35 - INGLIS, A.S., RUBIRA, M.R. and STRIKE, P.M. (1981) 62
The protein sequanator : a new extraction principle and a metal reaction cup.
J.of Biomedical and Biophys.Methods, 4, 279-285.
- 36 - JOHNSON, N.D., HUNKAPILLER, M.W. and HOOD, L.E. 95
(1979)
Analysis of phenylthiohydantoin amino acid by high-performance liquid chromatography on Dupont Zorbax cyanopropylsilane columns.
Anal.Biochem., 100, 335-338.
- 37 - KLAPPER, D.G., WILDE, C.E. and CAPRA, J.D. (1978) 40
Automated amino acid sequence of small peptides utilizing polybrene.
Anal.Biochem., 85, 126-131.
- 38 - KOPEYAN, C., MIRANDA, F. and ROCHAT, H. (1975) 39
Amino-acid sequence of toxin III of *Naja haje*.
Eur.J.Biochem., 58, 117-122.
- 39 - KULBE, K.D. (1974) 88
Micropolyamide thin-layer chromatography of phenylthiohydantoin amino acids (PTH) at subnanomolar level. A rapid microtechnique for simultaneous multisample after automated Edman degradation.
Anal.Biochem., 59, 564-573.
- 40 - LAINE, B., BELAÏCHE, D., SAUTIERE, P. and BISERTE, G. 113
(1982)
Characterization and structural study of the DNA-binding protein HRm from *Rhizobium meliloti*.
Biochem.Biophys.Res.Commun., 106, 101-107.
- 41 - LAINE, B., BELAÏCHE, D., KHANAKA, H. and SAUTIERE, P. 121
(1983)
Primary structure of the DNA-binding protein HRm from *Rhizobium meliloti*.
Eur.J.Biochem., 131, 325-331.

- 42 - LAURSEN, R.A. (1969) 49
 An isotope dilution procedure for detection and quantitative estimation of phenylthiohydantoins released in the Edman degradation.
Biochem. Biophys. Res. Commun., 37, 663-667.
- 43 - LAURSEN, R.A. (1971) 49
 Solid-phase Edman degradation : an automatic peptide sequencer.
Eur. J. Biochem., 20, 89-102.
- 44 - LIGHT, A. (1972) 9
 Leucine aminopeptidase in sequence determination of peptide.
 in Methods in Enzymology 25, 253-262.
- 45 - Mc DONALD, J.K., CALLAHAN, P.X. and ELLIS, S. (1972) 9
 Preparation and specificity of dipeptidyl aminopeptidase I.
 in Methods in Enzymology 25, 272-281.
- 46 - MOORE, S., COLE, R.D., GUNDLACH, H.G. and STEIN, W.H. 3
 (1960)
 in Proteins Proceedings of the 4th International Congress Biochemistry, Vienna 1958. Vol. 8 (H. Neurath and H. Tuppy eds), Pergamon, 1960, p.52.
- 47 - NIALL, H.D., KEUTMANN, H.T., COPP, D.H. and POTTS, J.T. (1969) 44
Proc. Nat. Acad. Sci. USA, 64, 771
- 48 - NIALL, H.D., JACOBS, J.W., VAN RIETSCHOTEN, J. and TREGGAR, G.W. (1974) 39
 Protected Edman degradation : a new approach to microsequence analysis of proteins.
FEBS Letters, 41, 62-64.

- 49 - OROSZLAN, S., COPELAND, T., SUMMERS, M. and SMYTHERS, G. (1975) 39
in Solid phase methods in protein sequence analysis (Laursen A.Ed.), 179. Pierce chemical Co Press ROCKFORD III.
- 50 - PISANO, J.J. and BRONZERT, T.J. (1969) 90
Analysis of amino acid phenylthiohydantoin by gas chromatography.
J.Biol.Chem., 244, 5597-5607.
- 51 - ROCHAT, H., BECHIS, G., KOPEYAN, C., GREGOIRE, J. and VAN RIETSCHOTEN, J. (1976) 39
Use of parvalbumin as a protecting protein in the sequanator ; an easy and efficient way for sequencing small amounts of peptides.
FEBS Letters 64, 404-408.
- 52 - ROSEAU, G. and PANTEL (1969) 87
Révélation colorée des spots de phénylthiohydantoïnes d'acides aminés.
J.Chromatogr., 44, 392-395.
- 53 - SANGER, F. (1949) 3
Biochem.J., 44, 126
- 54 - SHIVELY, J.E., HAWKE, D. and JONES, B.N. (1982) 74
Microsequence analysis of peptides and proteins.
III Artefacts and the effects of impurities on analysis.
Anal.Biochem., 120, 312-322.
- 55 - SILVER, J. and HOOD, L.E. (1974) 39,40
Automated microsequence analysis in the presence of a synthetic "carrier".
Anal.Biochem., 60, 285-292.

- 56 - SMITHIES, O., GIBSON, D., FANNING, E.M., GOODFIESH, R.M., GILMAN, J.G. and BALLANTYNE, D.L. (1971) 91
 Quantitative procedures for use with the Edman-Begg sequanator. Partial sequences of two unusual immunoglobulin light chains, Rzf and Sac.
Biochemistry 10, 4912-4921.
- 57 - TARR, G.E., BEECHER, J.F., BELL, M. and Mc KEAW, D. (1978) 40
 Polyquaternary amines prevent peptide.
Anal. Biochem., 84, 622-627.
- 58 - THOMSEN, J., KRISTIANSEN, K., BRUNFELDT, K. and SUNBY, F. (1972) 44
 The amino acid sequence of human glucagon.
FEBS Letters 21, 315-319.
- 59 - VAN ORDER, H.O. and CARPENTER, F.H. (1964) 91
 Hydrolysis of phenylthiohydantoins of amino acids.
Biochem. Biophys. Res. Commun., 14, 399-403.
- 60 - WATERFIELD, M.D. and BRIDGEN, J. (1975) 39
 in Instrumentation in amino acid sequence analysis. Perham R.H. editor, Acad. Press., London-USA.
- 61 - WILLIAMS, D.H., BRADLEY, C.V., SANTIKAR, N.S. and BOJESEN, G. (1982) 92
 Fast atom bombardment mass spectrometry. A new technique for the determination of molecular weights and amino acid sequences of peptides.
Biochem. J., 201, 105-117.
- 62 - WITTMANN-LIEBOLD, B. (1973) 56
 The miniature diaphragm valves are described by GRAFFUNDER et al in U.S. Patent N° 4,168,724 owned by Max-Plank-Gesellschaft, Federal Republic of Germany.
Hoppe Seyler's Z. Phys. Chem., 354, 1415

- 63 - WITTMANN-LIEBOLD, B., GRAFFUNDER, H. and KOHLS, H. (1976) 65, 66, 69
A device coupled to a modified sequanator for the automated conversion of anilino-thiazolinones into PTH amino acids.
Anal.Biochem., 75, 621-633.
- 64 - WITTMANN-LIEBOLD, B. (1980) 33
11 : Current advances in sequencing as applied to the structure determination of ribosomal proteins.
in Polypeptide hormones. ed.by BEERS R.F Jr and BASSETT E.G. Raven Press, New-York.
- 65 - WITTMANN-LIEBOLD, B. (1981) 36, 46, 57, 65
Microsequencing by manual and automated methods as applied to ribosomal proteins.
in Chemical synthesis and sequencing of peptides and proteins. ed.by LIU, SCHECHTER, HEINRIKSON, CONDLIFFE by Elsevier North-Holland.
- 66 - ZIMMERMAN, C.L., APELLA, E. and PISANO, J.J. (1976) 95
Advances in the analysis of amino acids phenylthiohydantoin by HPLC.
Anal.Biochem., 75, 77-85.

