

50376
1990
43

69179

50376
1990
43

N° d'Ordre: 527

UNIVERSITE DES SCIENCES ET TECHNIQUES
DE LILLE FLANDRES-ARTOIS

THESE

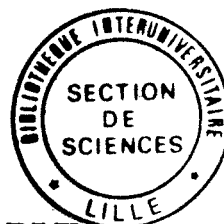
Présentée pour l'obtention du grade de

**DOCTEUR DE L'UNIVERSITE
DES SCIENCES ET TECHNIQUES DE LILLE**

OPTION BIOCHIMIE

PAR

Marie-Christine CHARTIER-HARLIN



**CONTRIBUTION A L'ETUDE DE LA TRANSITION
PROTEINE INTERMEDIAIRE ---> PROTAMINE
AU COURS DE LA SPERMIOGENESE DE LA SEICHE
*SEPIA OFFICINALIS.***

SOUTENUE LE 15 MAI 1990, DEVANT LA COMMISSION D'EXAMEN:

PRESIDENT:	Professeur	J. MONTREUIL
RAPPORTEURS:	Professeur	A. VERBERT
	Docteur	J. HARDY
EXAMINATEURS:	Docteur	J.C. D'HALLUIN
	Docteur	P. SAUTIERE

*A mon mari,
A mes parents,
A toute ma famille,
A mes amis,*

En témoignage de mon affection.

Ce travail est dédié à la mémoire du Professeur Gérard BISERTE

Que Messieurs les Professeurs Jean MONTREUIL, André VERBERT,
et Messieurs les Docteurs John HARDY, Jean-Claude D'HALLUIN
et Pierre SAUTIERE

soient assurés de ma profonde reconnaissance pour avoir
accepté de juger les travaux présentés dans cette thèse.

J'adresse un grand MERCI à toutes les personnes de l'URA 409 CNRS et de l'U 233 INSERM qui ont participé à l'élaboration de cette thèse, autant pour leur soutien scientifique que pour leur amitié.

Ce travail a été réalisé à l'Unité de Recherche Associée 409 du CNRS "Structures et Interactions des Protéines Associées à l'ADN des Eucaryotes et des Procaryotes" (Directeur : Pierre SAUTIERE) dans le groupe d'étude sur la spermatogenèse de la seiche, au sein de l'Institut de Recherches sur le Cancer de Lille (Directeur : Professeur Jacques VANLERENBERGHE).

La partie "Biologie Moléculaire" de ce travail a bénéficié du soutien logistique de l'Unité 233 INSERM (Directeur : Docteur Jean-Claude D'HALLUIN).

Les travaux de spectrométrie de masse ont été effectués par le Docteur Alain VAN DORSSELAER au Centre de Neurochimie, à Strasbourg.

Ce travail a bénéficié du soutien financier du Centre National de la Recherche Scientifique, de l'Université de Lille II et de la Fondation pour la Recherche Médicale.

En outre, ce travail n'aurait pu être mené à bien sans l'attribution d'une allocation de recherche par le Conseil d'Administration de l'Institut de Recherches sur le Cancer.

PUBLICATION

Cuttlefish spermatid specific protein variants T1 and T2 are putative precursors of protamine variants Sp1 and Sp2.

WOUTERS-TYROU, D., CHARTIER-HARLIN, M.C., MARTIN-PONTHIEU, A. and SAUTIERE, P.

J.Biol.Chem. (manuscrit en préparation).

COMMUNICATIONS PAR AFFICHES

Occurrence of a double transition in nuclear basic proteins during cuttlefish spermiogenesis.

I - Structural variants of the spermatid specific protein (Protein T) and of the protamine (Protein Sp).

D.WOUTERS-TYROU, A.MARTIN-PONTHIEU, M.C.CHARTIER-HARLIN and P.SAUTIERE

II - Expression of cytoplasmic mRNA from cuttlefish testis.
M.C.CHARTIER-HARLIN, A.MARTIN-PONTHIEU, D.WOUTERS-TYROU and P.SAUTIERE

First International Symposium on the cuttlefish *Sepia* - University of Caen - FRANCE, June 1-3, 1989.

ABRÉVIATIONS

ADN	:	acide désoxyribonucléique
ARN	:	acide ribonucléique
CLHP	:	chromatographie liquide haute pression
Da	:	dalton
ddNTP	:	didésoxyribonucléotide triphosphate
dNTP	:	désoxyribonucléotide triphosphate
DTT	:	dithiothréitol
EDTA	:	acide éthylènediaminetétraacétique
PEG	:	polyéthylèneglycol
PMSF	:	fluorure de phénylméthylsulfonyle
PTH	:	phénylthiohydantoïne
S	:	coefficient de Svedberg
SDS	:	dodécylsulfate de sodium
TFA	:	acide trifluoroacétique
TRIS	:	Tris (hydroxyméthyl)-aminométhane

TABLE DES MATIÈRES

	Pages
INTRODUCTION	1
<u>CHAPITRE I</u>	
GÉNÉRALITÉS	5
I - 1 LA SPERMATOGENESE	6
I - 2 LES TRANSITIONS PROTEIQUES DE LA SPERMIOGENESE	8
I - 2-1 Absence de transition	9
I - 2-2 Transition unique histone → protamine	10
I - 2-2-1 Les poissons téléostéens	10
I - 2-2-2 Les annélides et les mollusques bivalves	14
I - 2-3 Transition double histones → protéines intermédiaires → protamines	15
I - 2-3-1 Le modèle Roussette	15
I - 2-3-2 Le modèle Rat	17
I - 2-3-3 Le modèle Homme	21
I - 3 STRUCTURE ET EXPRESSION DES GENES DE PROTAMINES	25
I - 3-1 Organisation des gènes de protamines	25
I - 3-1-1 Les introns dans les gènes de protamines	26
I - 3-1-2 Les régions en amont et en aval des parties codantes	29
I - 3-1-3 Conformation de l'ADN des gènes de protamines	33
I - 3-2 Mécanismes de régulation transcriptionnelle et traductionnelle	33
I - 3-2-1 Mise en évidence de mécanismes de régulation	33
I - 3-2-2 Mécanismes potentiels de régulation	34
I - 3-3 Nature des codons utilisés	37
I - 3-4 Mécanismes de régulation post-traductionnelle des protamines	38

I - 4	LES PROTEINES BASIQUES ASSOCIEES A L'ADN AU COURS DE LA SPERMIOGENESE CHEZ LA SEICHE	40
I - 4-1	Intérêt de la seiche, <i>Sepia officinalis</i>	40
I - 4-2	Changements protéiques au cours de la spermiogenèse de la seiche	41
I - 4-2-1	La protéine T	43
I - 4-2-2	La protéine Sp	43

CHAPITRE II

	RÉSULTATS ET DISCUSSION	45
II - 1	DETERMINATION DE LA STRUCTURE PRIMAIRE DE LA PROTEINE T	46
II - 1-1	Préparation et purification de la protéine T	46
II - 1-1-1	Préparation	46
II - 1-1-2	Purification	46
II - 1-2	Composition en acides aminés de la protéine T	48
II - 1-3	Etudes structurales de la protéine T	52
II - 1-3-1	Dégradation automatique d'Edman de la protéine T	52
II - 1-3-2	Stratégie de coupure de la protéine T	54
II - 1-3-3	Hydrolyses enzymatiques de la protéine T	54
II - 1-4	Etudes structurales des variants T1 et T2	59
II - 1-4-1	Séparation des variants de la protéine T	59
II - 1-4-2	Dégradation automatique des variants séparés	59
II - 1-4-3	Stratégie de coupure des variants	62
II - 1-4-4	Hydrolyses enzymatiques des variants T1 et T2	63
II - 1-4-5	Séquences totales des protéines T1 et T2	71

II - 1-5	Prédictions de structure secondaire de la protéine T	79
II - 1-6	Comparaison des séquences des protéines T et Sp	83
II - 1-6-1	Comparaison des cartes peptidiques	83
II - 1-6-2	Comparaison des séquences	85
II - 2	APPROCHE GENETIQUE DE LA TRANSITION PROTEINE T → PROTAMINE Sp	88
II - 2-1	Extraction et purification des ARN messagers	89
II - 2-2	Traduction <i>in vitro</i> des ARN cytoplasmiques	89
II - 2-3	Enrichissement des ARNm de la protéine T	92
II - 2-4	Banques d'ADNc	97
II - 2-4-1	Synthèse de l'ADNc	97
II - 2-4-2	Clonage	100
II - 2-4-3	Criblage des banques d'ADNc	101
II - 2-4-4	Sous clonage de l'insert 5/14	109
II - 2-4-5	Séquençage de l'insert 5/14	109
II - 2-4-6	Etude du clone 5/20 correspondant à la protéine T2	112
II - 2-5	Discussion	116
	CONCLUSION ET PERSPECTIVES	118
	APPENDICE TECHNIQUE	122
I -	RECOLTE DU MATERIEL BIOLOGIQUE	123
II -	ETUDE BIOCHIMIQUE DE LA PROTEINE T SPECIFIQUE DES SPERMATIDES	123
II - 1	Isolement et purification des protéines basiques	123
II - 1-1	Isolement des protéines basiques	123
II - 1-2	Purification de la protéine T	125

II - 2	Analyses électrophorétiques des protéines	126
II - 2-1	Electrophorèse sur gel de polyacrylamide en milieu acide-urée	126
II - 2-2	Electrophorèse sur gel de polyacrylamide en présence de SDS	127
II - 2-3	Coloration et décoloration des gels	127
II - 3	Coupures enzymatiques de la protéine	128
II - 3-1	Coupure par l'endoprotéase Lys-C	128
II - 3-2	Coupure par la protéase d' <i>Astacus fluviatilis</i>	128
II - 3-3	Coupure par la chymotrypsine	128
II - 3-4	Coupure par l'élastase	129
II - 3-5	Traitement par la phosphatase alcaline	129
II - 4	Analyse des acides aminés et séquence protéique	130
III -	ETUDE DE LA PROTEINE T ET DE LA PROTAMINE DE SEICHE PAR LA BIOLOGIE MOLECULAIRE	131
III - 1	Etude des ARN messagers du testicule de seiche	131
III - 1-1	Extraction des ARN messagers cytoplasmiques	131
III - 1-2	Traduction <i>in vitro</i> des ARN messagers	133
III - 1-3	Enrichissement d'une classe d'ARN messagers sur gradient de saccharose	135
III - 2	Souches bactériennes et vecteurs de clonage	135
III - 3	Création de banques d'ADN complémentaire en plasmide et en phage	136
III - 3-1	Synthèse d'ADNc	137
III - 3-2	Analyse électrophorétique des ADNc	137
III - 3-3	Clonage de l'ADNc en phage λ gt 10 et en plasmide pUC 13	138

III - 4	Criblage de la banque d'ADNc plasmidique	139
III - 4-1	Méthode de préparation d'ADN plasmidique	139
III - 4-2	Sélection des clones positifs par hybridation et traduction d'ARNm	141
III - 5	Transfert d'ARNm et hybridation	144
III - 5-1	Marquage d'ADN avec une amorce aléatoire	144
III - 5-2	Transfert et hybridation	145
III - 6	Criblage de la banque de phages	146
III - 6-1	Méthodes de préparation de phages recombinants λ gt 10	146
III - 6-2	Hybridation des réplicats de boîtes de phages sur filtre de nitrocellulose	147
III - 6-3	Détermination rapide de la taille des inserts	149
III - 7	Sous clonage en plasmide pUC 13 et en phage M13	150
III - 7-1	Sous clonage de phage λ gt 10 en plasmide pUC 13	150
III - 7-2	Sous clonage de pUC 13 en phages M13mp18 et M13mp19	152
III - 8	Séquençage de l'ADN en double et simple brin	153
III - 8-1	Principe du séquençage d'ADN par la méthode de SANGER et COULSON (1975)	153
III - 8-2	Séquençage à partir de matrice double brin avec le vecteur pUC 13	154
III - 8-3	Séquençage à partir de matrice simple brin avec les vecteurs M13mp18 et M13mp19	156
	BIBLIOGRAPHIE	158

I N T R O D U C T I O N

Chez les eucaryotes, l'ADN est engagé dans une structure complexe appelée chromatine. La régulation de l'activité des gènes et le maintien de cette structure sont assurés par les protéines chromosomales, histones et non-histones.

Les histones forment un groupe homogène de cinq protéines basiques (H1, H2A, H2B, H3 et H4) qui ont peu varié au cours de l'évolution et dont la masse moléculaire est comprise entre 11.000 et 22.000 daltons. Elles sont caractérisées par leur richesse en acides aminés basiques qui représentent environ 25 à 30 % des résidus totaux de la molécule. Elles sont pauvres en acides aminés aromatiques ou soufrés et ne contiennent pas de tryptophane. Elles ont comme caractéristique structurale, une région hydrophobe, qui est le siège des interactions histone-histone alors que les régions amino- et carboxy-terminales basiques peuvent interagir avec l'ADN. Les histones jouent un rôle prépondérant dans l'organisation de la chromatine et en particulier dans l'organisation du nucléosome, unité structurale de base de la chromatine. Cette structure organisée autour de l'octamère d'histones (H2A, H2B, H3, H4)₂ (KORNBERG, 1974) est à l'origine de l'aspect en collier de perles de la chromatine lorsqu'elle est observée en microscopie électronique (OLINS et OLINS, 1974 ; OUDET *et al.*, 1974).

Les protéines non-histones forment une famille très hétérogène et sont présentes en quantité plus ou moins importante en fonction de l'espèce considérée. Elles se répartissent en deux groupes :

- Les protéines LMG (Low Mobility Group), insolubles dans l'acide trichloracétique à 2 %, de faible migration électrophorétique et de masse moléculaire comprise entre 30.000 et

150.000 daltons. Ce groupe renferme un grand nombre de protéines de structure, enzymes (polymérasés, kinases, phosphatases, méthylases...) ou protéines de régulation.

- Les protéines HMG (High Mobility Group), solubles dans l'acide trichloracétique à 2 %; elles ont une mobilité électrophorétique élevée et une masse moléculaire inférieure à 30.000 daltons. Ce sont essentiellement des protéines de structure.

Lors de la spermatogénèse, on observe dans la plupart des espèces animales, un remodelage de la forme des cellules et de leur noyau ainsi que de profonds changements dans l'organisation de la chromatine. Cette dernière passe d'un état diffus à un état fortement condensé. Cette condensation de la chromatine est liée à des changements importants dans la nature des protéines nucléaires basiques.

Au cours de la spermiogénèse, les histones de type somatique peuvent persister ou être remplacées partiellement ou totalement par des histones spécifiques du testicule ou du sperme. Dans la plupart des cas, les histones sont remplacées par des protéines plus basiques appelées protamines soit directement, soit après mise en place transitoire de protéines dites intermédiaires. Ces transitions seront illustrées par différents modèles. Nous accorderons une importance toute particulière aux modèles animaux dont la spermiogénèse présente une double transition - la Roussette, le Rat et l'Homme.

Ces modèles nous serviront de points de comparaison avec la Seiche dont la spermiogénèse présente également une double transition :

1ère transition

2ème transition

histones → protéine intermédiaire (protéine T) → protamine (protéine Sp)

L'étude de ces transitions a fait l'objet de nos recherches, menées à la fois par une approche biochimique impliquant la caractérisation physico-chimique et la détermination de la structure primaire de la protéine T et une approche génétique impliquant l'isolement des gènes codant pour les protéines T et Sp.

CHAPITRE I

G É N É R A L I T É S

I - 1 LA SPERMATOGENÈSE

La spermatogenèse est un processus complexe de différenciation cellulaire qui, partant d'une cellule somatique diploïde, aboutit à la formation d'une cellule haploïde ayant une morphologie caractéristique et une fonction hautement spécialisée : la fécondation de l'ovule.

La description anatomique précise des étapes successives de la spermatogenèse n'a été réalisée que ces dernières années par des études en microscopie électronique.

Chez l'homme, la formation des spermatozoïdes ou spermatogenèse s'effectue dans les testicules de manière permanente à partir de la puberté. Avant la puberté, chaque testicule contient de fins pelotons de longs tubes séminifères. Ces tubes sont limités par une membrane externe qui renferme de grosses cellules diploïdes appelées gonocytes primordiaux. Un des signes cellulaires observable en microscopie électronique, du déclenchement de la puberté est un changement de morphologie du gonocyte et de la forme de son noyau. A ce stade le gonocyte devient une spermatogonie souche. Les tubes séminifères sont tapissés de spermatogonies qui vont entrer successivement en spermatogenèse. Cette dernière se déroule de façon centripète, jusqu'à la lumière du tube séminifère où sont libérés les spermatozoïdes. Ce processus s'étale sur quelques semaines.

Les différentes étapes de la spermatogenèse sont présentées dans la figure 1. Par mitose, la spermatogonie souche donne des spermatocytes primaires. Lorsque ces derniers at-

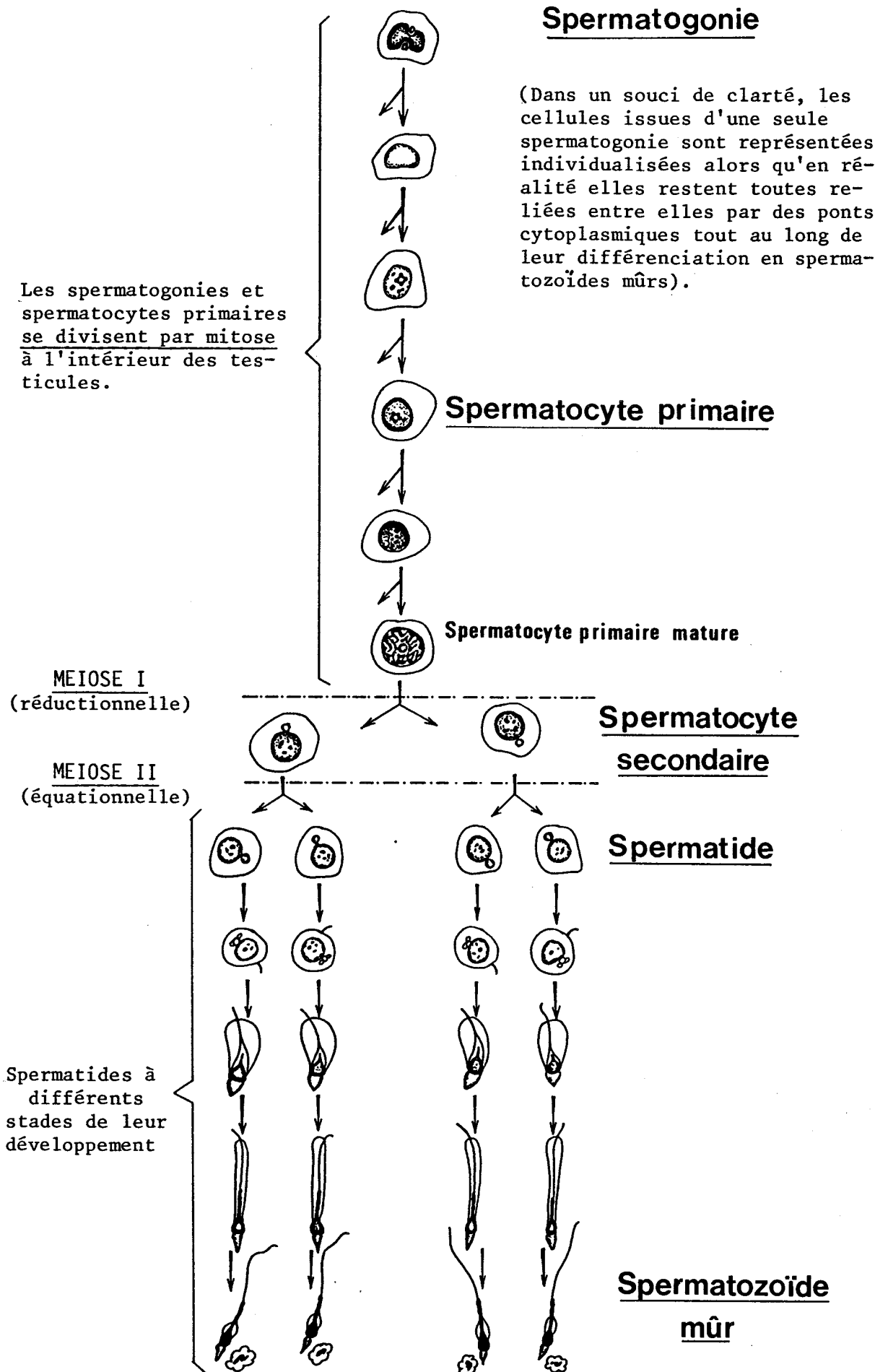


FIGURE 1 - Les différentes étapes de la spermatogenèse humaine.

teignent leur taille maximale, ils subissent de manière synchrone une méiose qui va permettre le passage de la diploïdie à l'haploïdie et qui aboutit à la formation de deux spermatoocytes secondaires. Le spermatoocyte secondaire se divise rapidement pour donner naissance à deux spermatides qui ne se divisent plus. Ces spermatides vont subir une série de transformations pour aboutir aux spermatozoïdes. Cette phase finale de la spermatogenèse, qui dure environ trois semaines chez l'homme, constitue la spermiogenèse.

La spermiogenèse est la phase de la spermatogenèse qui a retenu notre attention, et qui est à l'origine des travaux présentés dans cette thèse. En effet, c'est au cours de la spermiogenèse que l'on assiste à un remaniement profond de l'organisation de la chromatine qui passe d'un état diffus dans les spermatides rondes à un état fortement condensé dans les spermatozoïdes mûrs. Ce remaniement est lié à des changements ou transitions dans la nature des protéines associées à l'ADN. En particulier, les histones de type somatique, ou spécifiques du testicule, sont directement remplacées par des protéines très basiques appelées protamines, ou remplacées de façon transitoire par des protéines spécifiques des spermatides qui sont ensuite elles-mêmes remplacées par des protamines.

I - 2 LES TRANSITIONS PROTÉIQUES DE LA SPERMIOGENÈSE

La nature et le nombre de protéines basiques associées à la chromatine des spermatozoïdes varient largement parmi les différents organismes (CHEVAILLIER, 1983 ; SUBIRANA, 1983 ; POCCIA, 1986). Dès 1969, BLOCH en a proposé une répartition en quatre classes majeures :

- les protamines riches en arginine (protamines des poissons téléostéens)
- les protamines riches en arginine et cystéine (protamines des mammifères)
- les protamines riches en arginine et lysine (protamines des mollusques bivalves)
- les histones spécifiques ou de type somatique

Actuellement, avec la découverte chez certaines espèces animales d'une étape intermédiaire dans le remplacement des histones par les protamines, trois situations peuvent être observées dans la spermiogenèse animale : (I) absence de transition, (II) transition unique, (III) double transition.

I - 2-1 ABSENCE DE TRANSITION

Chez certains animaux appartenant à des ordres différents, la condensation de l'ADN au cours de la spermatogenèse s'effectue en ne faisant intervenir comme protéines basiques, que des histones. Celles-ci peuvent être de type somatique ou spécifiques du sperme (où des changements apparaissent alors au niveau des familles d'histones H1 et H2B).

En effet, les histones du sperme sont parfois de type somatique : c'est le cas du poisson rouge *Carassius auratus* (MUNOZ-GUERRA *et al.*, 1982). Dans d'autres espèces comme la grenouille léopard *Rana pipiens* ainsi que la carpe *Ctenopharyngodon idella* (SUBIRANA, 1983), certaines sous-fractions d'histone H1 spécifiques du sperme ont néanmoins une composition identique à des sous-fractions d'histone H1 d'origine somatique.

Par contre, dans le sperme des échinodermes tels que l'oursin et l'étoile de mer, les histones H1 et H2B sont plus riches en arginine que leurs homologues d'origine somatique, alors que les histones H2A, H3 et H4 sont de type somatique (VON HOLT *et al.*, 1979).

Chez ces différents animaux, la condensation de l'ADN ne fait intervenir que des modifications au sein d'une même classe de protéines, les histones. Il ne s'agit donc pas d'une véritable transition protéique faisant intervenir d'autres catégories de protéines, comme cela se passe dans la plupart des espèces.

I - 2-2 TRANSITION UNIQUE HISTONES → PROTAMINES

I - 2-2-1 Les poissons téléostéens

Chez les poissons téléostéens, les histones sont remplacées au cours de la spermiogénèse par des protamines.

Les protamines sont de petite taille (environ 30 résidus) et riches en arginine (60 à 70 %) (ANDO *et al.*, 1973). La nature des acides aminés constitutifs est peu diversifiée et l'arginine est le seul acide aminé basique de ces protéines à l'exception de celles de l'esturgeon et du brochet où les rapports lysine/arginine sont respectivement de 0.42 et 0.05 (YULIKOVA *et al.*, 1979 ; SPECKERT *et al.*, 1983).

Le nom de chacune de ces protamines rappelle souvent l'espèce ou le genre du poisson dont elle a été isolée : salmine du saumon, iridine de la truite...

Les séquences en acides aminés de ces protamines présentent toutes la même sous-unité structurale fondamentale constituée par un pentapeptide de structure X-(Arg)₄ où X représente Ala, Gly, Ile, Pro, Ser, Thr ou Val. Cet enchaînement est généralement rencontré trois à quatre fois dans la séquence de ces protamines (Figure 2). Cette observation a amené BLACK et DIXON (1967) à proposer un schéma dans lequel la clupéine Z du hareng dériverait d'un pentapeptide ancestral Ala-Arg₄ par le jeu de duplications et d'insertions de gènes.

Parmi les poissons téléostéens, c'est chez la truite arc-en-ciel, *Salmo gairdnerii*, que la spermatogenèse a été étudiée de manière approfondie. Dans la chromatine du testicule mature, les protéines basiques sont essentiellement représentées par un groupe de six protamines très riches en arginine (70 %) qui présentent une parenté structurale très étroite (Mc KAY *et al.*, 1986) (Figure 3). On trouve également en faible quantité des histones et des protéines HMG.

Ce mécanisme de transition protéique directe, histones → protamines est pratiquement total et fait appel à des modifications post-traductionnelles de ces protéines. En effet, les histones sont synthétisées et phosphorylées dans les cellules diploïdes et ne le sont pratiquement plus dans les spermatides. LOUIE et DIXON (1972a, 1972b) suggèrent que la phosphorylation des histones serait requise pour une liaison correcte des histones à l'ADN. Au stade spermatide, les histones sont graduellement éliminées de la chromatine et remplacées par les protamines qui sont nouvellement synthétisées et fortement phosphorylées.

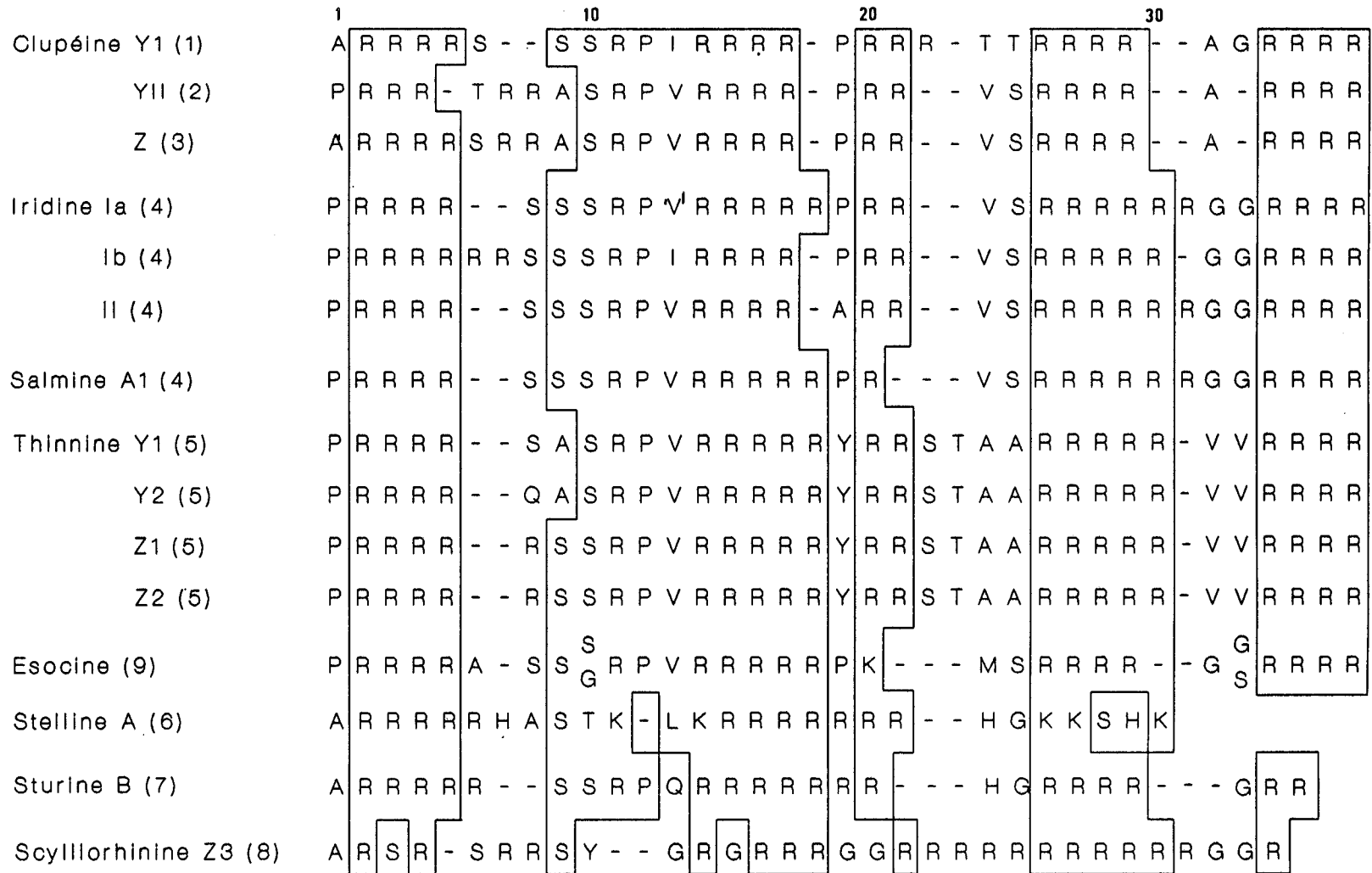


FIGURE 2: Comparaison des séquences en acides aminés des protamines de poissons. Les zones d'identité structurale sont encadrées selon CHEVAILLIER (1983) (1) ANDO et SUZUKI (1967), (2) ANDO et SUZUKI (1966), (3) ANDO *et al.* (1962), ANDO et WATANABE (1969), (5) BRETZEL (1972, 1973), (6) YULIKOVA *et al.* (1976), (8) SAUTIERE *et al.* (1981), (9) SPECKERT *et al.* (1983)



FIGURE 3 - Séquences en acides aminés des six protamines de truite (Mc KAY *et al.*, 1986) Les régions encadrées correspondent aux séquences communes aux six protamines.

L'acétylation sur les groupements ϵ -NH₂ des résidus de lysine de la région amino-terminale des histones pourrait intervenir dans le mécanisme de substitution (CANDIDO et DIXON, 1972 ; LOUIE et DIXON, 1972 c).

I - 2-2-2 Les annélides et les mollusques bivalves

Chez les annélides et les mollusques bivalves, la transition histone \rightarrow protamine n'est que partielle. En effet, on observe encore à des degrés divers le maintien d'histones dans le spermatozoïde mature. En exemple, nous pouvons citer :

- *Platynereis dumerilii* (annélides) où l'on trouve :
 - . des core histones de type somatique,
 - . quatre histones H1 spécifiques dont deux sont caractérisées par leur faible poids moléculaire (\sim 13.000 Da),
 - . deux protamines de poids moléculaire d'environ 7.000 Da, contenant 50% d'arginine et représentant seulement 15 % des protéines basiques totales de la chromatine (SELLOS et KMIECIK, 1985).
- Le cas de la moule *Mytilus edulis* (mollusque bivalve) où le spermatozoïde contient une faible quantité d'histones de type somatique (environ 15 %) et 72 % d'une protamine de 90 résidus contenant 30 % d'arginine et 24,3 % de lysine (AUSIO et SUBIRANA, 1982).

Dans la chromatine du sperme de ces espèces, la persistance des histones s'accompagne du maintien de la structure nucléosomale de la chromatine, dans un rapport qui est fonction de la quantité d'histones présentes.

I - 2-3 TRANSITION DOUBLE HISTONES → PROTEINES INTERMEDIAIRES
→ PROTAMINES

Ce schéma se retrouve chez la roussette et les mammifères.

I - 2-3-1 Le modèle Roussette

La roussette présente la particularité d'avoir une différenciation dorso-ventrale des cellules germinales au cours de la spermatogenèse qui s'accompagne d'une zonation du testicule en trois parties A, B, C, correspondant chacune à une population de cellules parvenues au même stade de différenciation.

- La zone A renferme des spermatogonies et des spermatoocytes. L'organisation de la chromatine est assurée uniquement par des histones.
- La zone B renferme des spermatides rondes et en début d'élongation. Les protéines basiques nucléaires sont surtout des histones, mais deux protéines intermédiaires apparaissent dans les spermatides en voie d'élongation.
- La zone C contient des spermatides en élongation ou allongées ainsi que des spermatozoïdes immatures. La chromatine est fortement condensée et les protéines basiques associées à l'ADN sont les protéines S1 et S2, mais surtout les protamines Z1, Z2, Z3 et S4 (Figure 4) (GUSSE et CHEVAILLIER, 1981 ; CHAUVIERE *et al.*, 1983).

Les structures primaires de ces différentes protéines ont été déterminées (SAUTIERE *et al.*, 1981, 1984 ; MARTINAGE *et al.*, 1985 ; CHEVAILLIER *et al.*, 1987 ; CHAUVIERE *et al.*, 1987, 1989).




	SPERMATIDES (Zone B du testicule)			SPERMATOZOÏDES (Zone C du testicule et épiddyme)	
	Stade 1	Stade 2	Stade 3	Stade 4	Stade 5
Organisation des noyaux	Contenu hétérogène nucléole, particules nucléaires	Contenu homogène	Début de l'élongation	Milieu de l'élongation	Maximum de l'élongation Noyaux biréfringents
Chromatine	Fractions condensées et dispersées	Dispersée	Orientation des fibres le long du grand axe du noyau	Agrégation des fibres de chromatine en un réseau pseudolamellaire	Homogène - fortement condensée - torsadée
Structures répétitives	192 pb	"Aspect "collier de perles"		165 pb	123 pb (?)
Accessibilité DNase	-----				
Histones					
Protéines intermédiaires					
Protamines					
				Réduites, phosphorylées (SH) (Solubles ATCA)	- Déphosphorylées - Oxydées (S-S) (Insolubles ATCA sauf Z3)

FIGURE 4 - SPERMIOGENESE DE LA ROUSSETTE d'après M.GUSSE et Ph.CHEVAILLIER (1981)

Les protamines de roussette Z1, Z2 et S4 sont caractérisées par des séquences répétitives (ARKRPKCKK)₂ et peuvent être rapprochées des protamines de mammifères, selon la classification de BLOCH (1969), puisqu'elles contiennent de la cystéine. La protamine Z3 est proche des protamines de poissons téléostéens (cf figure 2). En effet, elle est caractérisée par un taux élevé d'arginine et un petit nombre d'acides aminés constitutifs où les acides aminés dicarboxyliques, hydrophobes et aromatiques sont très peu représentés.

Les protéines intermédiaires S1 et S2 diffèrent à la fois des histones et des protamines. Elles présentent une forte homologie structurale (environ 70 %), un centre acide et une distribution asymétrique des résidus basiques et hydrophobes (Figure 5).

Ce mécanisme de transition double faisant intervenir des protéines intermédiaires se retrouve également chez les mammifères.

I - 2-3-2 Le modèle Rat

Dans les spermatides rondes, premier stade de la spermiogénèse, les histones de type somatique ou spécifiques du testicule (BUCCI *et al.* 1982 ; RAO *et al.* ,1983) et les protéines de type HMG (BUCCI *et al.* , 1984) représentent la totalité de l'équipement en protéines basiques associées à l'ADN (Figure 6). La chromatine a une apparence granulo-filamenteuse qui correspond à la structure nucléosomale classique (GRIMES *et al.*, 1977).

Protéines intermédiaires

S1 T-K-S-R-V-R-N-R-R-S-R-P-R-R-R-Y-G-R-R-M-R-R-T-R---C-R-R-K-G-
 + + * + + + * + + + * + * + + + * + + +
 S2 V-K-S-R-Y-H-Q-R-Q-Y-R-A-R-K-R-Y-A-R---A-R-R-T-K-K-P-K-R-R-P-

 S1 R-R-I-S-R-R-P-R-R-T-T-Y-R-R-R-V-R-K-I-V-H-L-K-R-R-S-R-P-R-D-
 + * + + + * + * * + * + * #
 S2 K-P-P-R-K-L-R-Y-A-P-S---K-K-Q-P---K-I-M-K-L-K-----L---D-

 S1 --E-I-D-N---L-K-V-K-N-N-R-R-L-N-E-S-L-K-Q-H-R-L-P-M-R-V-P-V
 # * # * + * + + * # * + + * * + *
 S2 N-E-V-D-N-T-L-K-A-K-N---K-S-L-N-E-A-L-K-N---R-L-S-L-R-K-H-V

FIGURE 5 - Séquences en acides aminés des protéines intermédiaires S1 et S2 de Roussette (CHAUVIERE *et al.*, 1987, 1989). Les séquences ont été alignées de façon à avoir un maximum d'homologie. Les acides aminés acides, basiques et hydrophobes communs aux deux protéines ont été symbolisés par #, + et * de façon à faire ressortir leur distribution dans la chaîne polypeptidique.

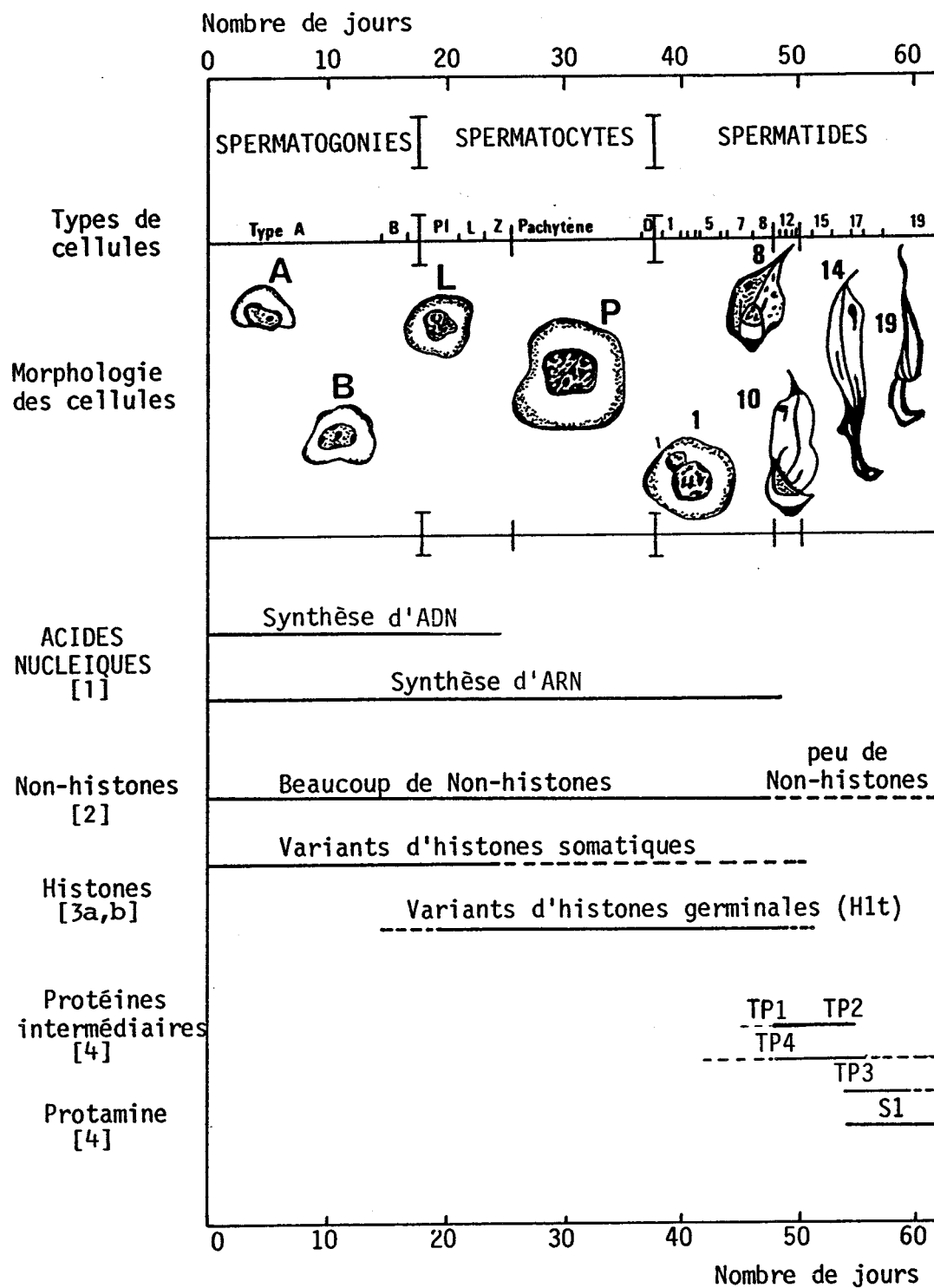


FIGURE 6 - SCHEMA DE LA SPERMATOGENESE CHEZ LE RAT

[1] MEISTRICH *et al.*, 1981 ; [2] PLATZ *et al.*, 1975 ; [3a] BUCCI *et al.*, 1982 ; [3b] RAO *et al.*, 1983 ; [4] GRIMES *et al.*, 1977.

(Pl, pré-leptotène ; L, leptotène ; Z, zygotène ; P, pachytène ; D, diplotène sont des stades successifs typiques de la prophase de la méiose).

Ensuite, en l'espace de deux jours, le noyau se condense, s'allonge et devient résistant à la sonication. A partir de ce moment, il n'y a plus d'activité transcriptionnelle décelable. Cela implique que toutes les protéines qui sont synthétisées après ce stade proviennent d'ARN messagers qui ont été stockés sous forme de ribonucléoprotéine messagère (RNPM) (MALI *et al.*, 1988).

Les histones et les protéines HMG disparaissent et sont remplacées par deux protéines TP1 et TP2 (GRIMES *et al.*, 1977). Ces deux protéines resteront les protéines basiques majeures jusqu'à l'apparition de la protamine S1. A cet instant, les protéines TP1 et TP2 disparaissent du noyau et on observe la présence éphémère de deux protéines basiques nucléaires mineures TP3 et TP4.

L'apparition de la protamine S1 coïncide avec une nouvelle condensation de la chromatine qui devient alors résistante aux nucléases.

Les protéines TP1 et TP2 ne sont ni des produits de dégradation des histones, ni des précurseurs de la protamine S1 comme le montre la structure primaire des protéines TP1 (KISTLER *et al.*, 1975) et TP2 (LUERSEN *et al.*, 1989). Ces deux protéines pourraient être impliquées dans la condensation de la chromatine, le modelage de la tête du spermatozoïde et l'arrêt de la transcription. SINGH et RAO (1987) ont montré que la protéine TP1 déstabilise le nucléosome *in vitro*. Cette protéine serait donc capable de modifier la structure du nucléosome de façon à faciliter le remplacement des histones.

Des travaux récents, menés à partir d'ARNm ont montré que le spermatozoïde du rat contenait une seconde protamine (BOWER *et al.*, 1987). Ces auteurs ont appelé les deux protamines P1 et P2, et ont signalé qu'elles étaient retrouvées chez la souris

et chez le hamster. La protamine P2 de rat n'est présente qu'en faible quantité dans le spermatozoïde (BOWER *et al.*, 1987). Ceci explique qu'elle n'ait pas été signalée au cours des études antérieures. Elle présente une étroite parenté structurale (90,7 %) avec la protamine mP2 de la souris (TANHAUSER et HECHT, 1989). Chez cette dernière, la protamine mP2 serait synthétisée sous forme d'un précurseur (p-mP2) (Figure 7 page 23), contrairement à la protamine mP1 (YELICK *et al.*, 1987).

L'existence de précurseurs potentiels n'a pas été décrite chez tous les Mammifères, mais semble exister chez l'Homme.

I - 2-3-3 Le modèle Homme

Généralement, la chromatine du spermatozoïde des mammifères ne contient que des protamines à l'exclusion de toute autre protéine basique (COELINGH *et al.*, 1972 ; KISTLER *et al.*, 1976 ; CALVIN, 1976 ; TOBITA *et al.*, 1983 ; SAUTIERE *et al.* 1984 ; MAZRIMAS *et al.*, 1986 ; LEE *et al.*, 1986, 1987a).

Le spermatozoïde humain est un cas particulier puisque sa chromatine renferme à la fois des protamines, qui sont les constituants majeurs (70 %), des protéines intermédiaires (10 à 20 %) et des histones de type somatique ou spécifiques du testicule (10 à 20 %).

Les protamines, au nombre de quatre, ont été appelées HP1, HP2, HP3 et HP4. Elles sont riches en arginine (~ 50 %) et en cystéine (~ 10 %), caractéristiques communes à toutes les protamines de mammifères, et se répartissent en deux classes suivant leur taux de tyrosine ou d'histidine (GUSSE *et al.*, 1986).

La protamine HP1 (50 résidus), riche en tyrosine (8%) montre une étroite parenté structurale avec la protamine de type P1, qui représente l'unique composant basique ou le composant basique majeur du spermatozoïde de mammifères (Mc KAY *et al.*, 1985 ; AMMER *et al.*, 1986 ; LEE *et al.*, 1987b). Les protamines HP2 (57 résidus) et HP3 (54 résidus) qui appartiennent à la famille des protamines riches en histidine, présentent de très fortes homologies structurales et ne diffèrent l'une de l'autre que par la présence d'un tripeptide supplémentaire à l'extrémité N-terminale de la protamine HP2. La comparaison de leur structure primaire avec celle des autres protamines de mammifères montre l'existence d'homologies de séquence avec la protamine mP2 de souris (Figure 7) mais peu d'homologie avec les protamines de type P1.

Les protéines intermédiaires HPI1, HPS1 et HPS2 possèdent une région C-terminale commune de 54 résidus, identique à la séquence de la protamine HP3, avec une extension N-terminale de 47 résidus pour HPI1, 15 pour HPS1 et 12 pour HPS2. Par sa relative richesse en acide glutamique et par l'absence de cystéine, la structure primaire de l'extension N-terminale de chacune de ces protéines contraste avec celle de la région basique C-terminale commune aux trois protéines.

Ces protéines intermédiaires pourraient être considérées comme les précurseurs des protamines de type P2 (SAUTIERE *et al.*, 1988) (Figure 8). Ce phénomène est similaire à celui qui a été observé chez la souris (Chapitre I - 2-3-2) (YELICK *et al.*, 1987 ; DOMENJOURD *et al.*, 1988). Cependant le mécanisme exact de cette transition reste encore à élucider.

Nous avons vu dans les parties 1 et 2 de ce chapitre, la grande diversité des protéines nucléaires basiques présentes dans le spermatozoïde des espèces animales, ainsi que les dif-

```

HPI1   V-R-Y-R-V-R-S-L-S-E-R-S-H-E-V-Y-R-Q-Q-L-H-G-Q-E---Q-G-H-H-G-Q-E-
      : : : : : : : : : : : : : : : : : : : : : : : : : :
pmP2   V-R-Y-R-M-R-S-P-S-E-G-P-H-Q-G-P-G-Q---D-H-E-R-E-E-Q-G-----Q---
      : : : : : : : : : : : : : : : : : : : : : : : : : :
p2 Rat V-R-Y-R-M-R-S-P-S-E-G-Q-H-Q-G-P-G-Q---D-H-E-R-E-E-Q-G-----Q---

HPI1   E-Q-G-L-S-P-E-H-V-E-V-Y-E-R-T-H---G-Q-S-H-Y-R-R-R-H-C-S-R-R-R-L-
      : : : : : : : : : : : : : : : : : : : : : : : : : :
pmP2   G-Q-G-L-S-P-E-R-V-E-D-Y-G-R-T-H-R-G-H-H-H-H-R-H-R-R-C-S-R-Q-R-L-
      : : : : : : : : : : : : : : : : : : : : : : : : : :
p2 Rat G-Q-E-L-S-P-E-R-V-E-D-Y-G-R-T-E-R-G-H-H-H-H-R-H-R-R-C-----Q-R-L-

HPI1   H-R-I-H-R-R-Q-H-R-S-C-R-R-R-K-R-R-S-C-R-H-R-R-R-H-R-R-G-C-R-----
      : : : : : : : : : : : : : : : : : : : : : : : : : :
pmP2   H-R-I-H-Q-R---R-R-S-C-R-R-R-R-R-H-S-C-R-H-R-R-R-H-R-R-G-C-R-R-S-
      : : : : : : : : : : : : : : : : : : : : : : : : : :
p2 Rat H-R-I-H-Q-R---R-R-S-C-R-R-R-R-R-H-S-C-R-H-R-R-R-H-R-R-G-C-R-R-S-

HPI1   -----T-R-K-R-T-C-R-R-H
      : : : : :
pmP2   R-R-R-R-R-C-R-C-R-Q-C-R-R-H-H-H
      : : : : : : : : : :
p2 Rat R-R-R-R-S-C-R-C-R-Q-C-R-W-H-Y-Y

```

FIGURE 7: Analogies structurales entre les précurseurs potentiels de protamines de type P2. Les protéines présentées sont: la protéine HPI1 humaine (DOMENJOURD *et al.*, 1988; MARTINAGE *et al.*, 1990) la protéine pmP2 de souris (YELICK *et al.*, 1987) et la protamine P2 de rat (TANHAUSER *et al.*, 1989). Les séquences ont été alignées de façon à présenter un maximum d'analogies (symbolisées par :) après traitement informatique par le programme PCBOSS.

► indique le début de la séquence des protamines P2 de la souris et de l'homme.

HPI1 1 V-R-Y-R-V-R-S-L-S-E-R-S-H-E-V-Y-R-Q-Q-L-H-G-Q-E-Q-G-H-H-G-Q-E-E-Q -G-L-S-P-E-H-V-E-V-Y-E-R-T-H-G-Q-S-H-Y-R.

HPS1 1 E-G-L-S-P-E-H-V-E-V-Y-E-R-T-H-G-Q-S-H-Y-R

HPS2 1 S-P-E-H-V-E-V-Y-E-R-T-H-G-Q-S-H-Y-R

HP2 1 R-T-H-G-Q-S-H-Y-R

HP3 1 G-Q-S-H-Y-R

HP3 10 G-Q-S-H-Y-R-R-R-H-C-S-R-R-R-L-H-R-I-H-R-R-Q-H-R-S-C-R-R-R-K-R-R-S -C-R-H-R-R-R-H-R-R-G-C-R-T-R-K-R-T-C-R-R

FIGURE 8: Extension amino-terminale des protéines Intermédiaires humaines par rapport aux protéines HP2 et HP3. Seule la séquence de la protéine la plus courte (protéine HP3) est donnée dans son intégralité, cette séquence est commune à toutes les autres protéines.

férentes transitions qui peuvent intervenir dans le remplacement des histones par les protamines au cours de la spermiogenèse. L'étude de ces mécanismes de transition a aussi été abordé par des méthodes complémentaires de l'analyse des protéines, c'est-à-dire l'étude de la structure et de l'expression des gènes de protamines.

I - 3 STRUCTURE ET EXPRESSION DES GÈNES DE PROTAMINES

Les premiers travaux effectués dans ce domaine ont porté sur l'étude des ARN messagers des protamines chez la truite. Six protamines ont été mises en évidence par le groupe de DIXON au Canada (LOUIE et DIXON, 1972 ; DAVIES *et al.*, 1976 ; GEDAMU et DIXON, 1976, 1979 ; GEDAMU *et al.*, 1977, 1981 ; IATROU *et al.*, 1979).

L'obtention d'ADNc a permis de déterminer un grand nombre de structures primaires de protamines et protéines intermédiaires, tant chez les mammifères que chez les oiseaux (KLEENE et FLYNN, 1987 ; KLEENE *et al.*, 1988 ; HEIDARAN et KISTLER, 1987a ; LEE *et al.*, 1986 ; LEE *et al.*, 1987b ; OLIVA et DIXON, 1989). Elle a aussi permis de définir les différents types d'organisation des gènes de protamine selon les espèces étudiées.

I - 3-1 ORGANISATION DES GENES DE PROTAMINES

Contrairement à ceux des histones, les gènes de protamines n'apparaissent pas groupés sur un chromosome mais disséminés (AIKEN *et al.*, 1983). Le nombre de copies par génome haploïde diffère selon l'animal étudié, une chez la souris et

le taureau (JOHNSON *et al.*, 1988 ; KRAWETZ *et al.*, 1987), deux chez le coq (OLIVA et DIXON, 1989) et de l'ordre de 20 chez la truite (DIXON *et al.*, 1986).

I - 3-1-1 Les introns dans les gènes de protamines

Les gènes de protamines de la truite ne possèdent pas d'intron (AIKEN *et al.*, 1983). Ils partagent cette particularité avec les gènes des histones (SCHAFFNER *et al.*, 1978) et de certaines protéines non basiques telles que l'interféron humain (NAGATA *et al.*, 1980).

Cependant chez les mammifères, les gènes de protamines renferment un intron unique (KRAWETZ *et al.*, 1988 ; JOHNSON *et al.*, 1988) dont la taille est relativement constante. Ainsi les gènes de protamines mp1 et mp2 de souris ont respectivement un intron de 94 et de 105 nucléotides (Figure 9). Chez le taureau, le gène de la protamine P1 possède un intron de 101 nucléotides. Dans tous les cas décrits, les introns tendent à séparer le domaine C-terminal très variable des protamines des mammifères du domaine N-terminal très conservé (KRAWETZ *et al.*, 1988).

Les gènes de protamines des oiseaux ne possèdent pas d'intron, ce qui tend à les rapprocher des poissons, cependant les séquences N-terminales des protamines des oiseaux sont proches de celles des mammifères (Figure 10).

Les connaissances sur les protamines montrent clairement qu'il y a eu divergence de leurs gènes au cours de l'évolution. Cependant ces connaissances sont encore trop fragmentaires pour définir précisément la phylogénie de ces gènes.

I - 3-1-2 Les régions en amont et en aval des parties codantes

Les connaissances de l'organisation générale des gènes de protamines en amont et en aval des séquences codantes proviennent des travaux effectués sur la souris (JOHNSON *et al.*, 1989), le taureau (KRAWETZ *et al.*, 1987), le coq (OLIVA et DIXON, 1989) et la truite (DIXON *et al.*, 1986). Ces études réalisées à partir de banques génomiques permettent de localiser les éléments de séquences nucléotidiques pouvant jouer un rôle dans l'expression des parties codantes des gènes. Un exemple caractéristique d'organisation de ces séquences est présenté dans la figure 11.

Les caractéristiques générales des gènes de protamines sont les suivantes :

Dans la région 5' non codante, on remarque la séquence consensus d'initiation de la traduction des cellules eucaryotes (A C C A T G G) déterminée par KOZAK (1986). On trouve également une séquence T A T A A qui aide l'initiation de la transcription par l'ARN polymérase II.

D'autres éléments de régulation sont parfois présents : une séquence C A A T existe dans le gène mP2 de souris, mais pas dans le gène mP1. Cette séquence existe dans le gène de protamines P1 de taureau, mais pas dans les multiples gènes de protamines de truite.

Lorsque cette séquence est présente on trouve parfois immédiatement en amont de celle-ci une séquence G C qui, selon DYNAN et TJIAN (1985) pourrait jouer un rôle d'activateur ("enhancer") par fixation du facteur de transcription Sp-1.

L'analyse des gènes de protamines de souris a permis de montrer qu'il existe des séquences répétées (Figure 11) et des séquences qui se trouvent dans d'autres gènes exprimés dans le testicule (Figure 12).

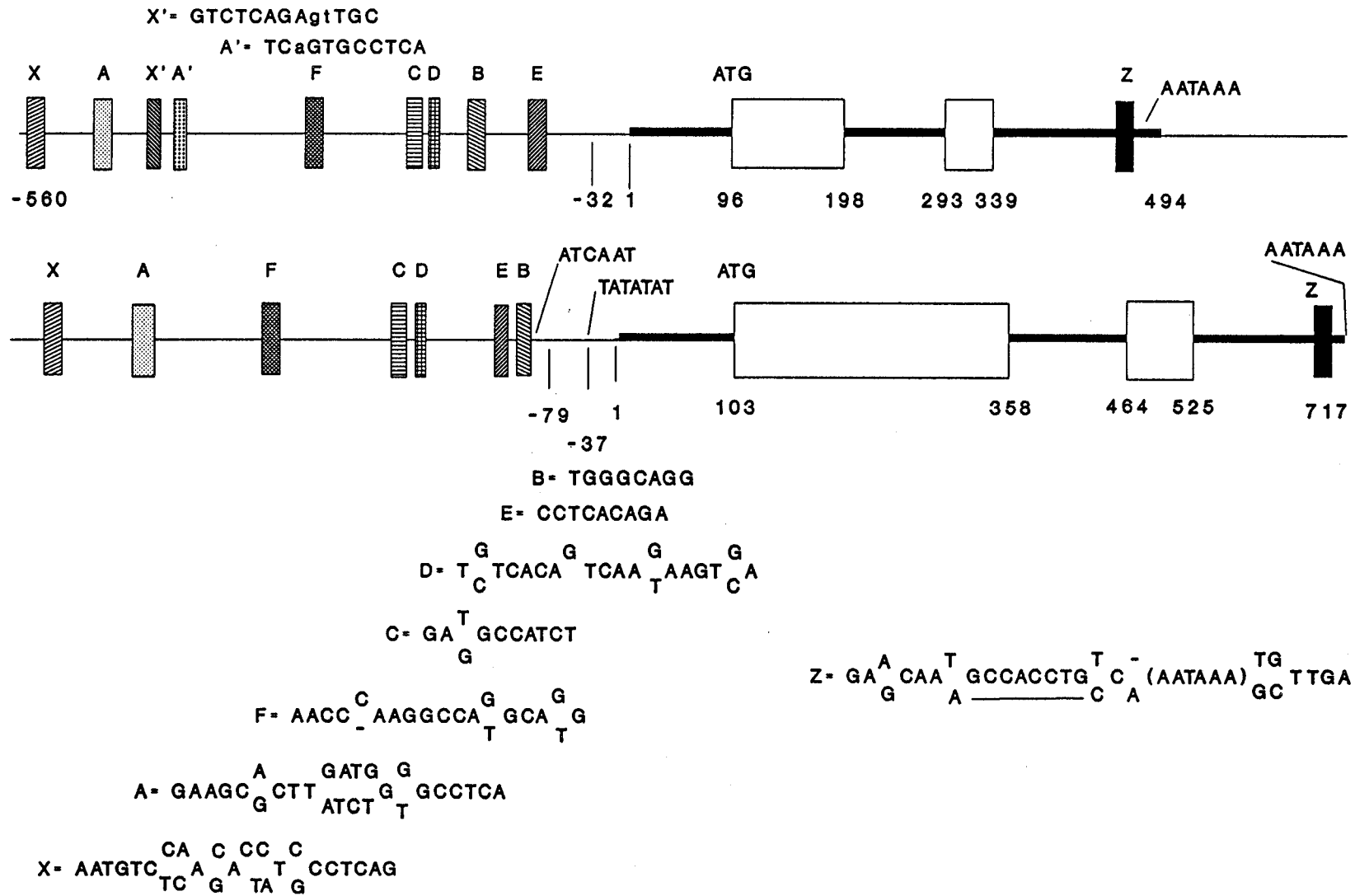


FIGURE 11: Séquences nucléotidiques conservées dans les gènes de protamine de souris mP1 et mP2



Séquence codante pour la protéine



Séquence transcrite en ARNm

Selon JOHNSON *et al.* 1988

Séquence	Position dans le gène de souris		séquence retrouvée dans le gène de		
	mP1	mP2			
AGCACACAT	-338	-276	l'histone H1t de rat		
GGTGGGCAG	-143	+20	l'histone H1t de rat		
CCTGGAG	-10	-253	la protéine PGK2 humaine		
AGCC ^A _C CA	-510	-311	-18	+121	la protéine PKG2 humaine (retrouvée 2 fois)
GCCTCAAT	-475	+66	la protéine PGK2 humaine		
G ^A _T CCCTCTG	-434	-602	la protéine PGK2 humaine		
CCCTC ^A _T GC	-369	-600	la protéine PGK2 humaine		
T _G GGGTGGG	-455	-140	-379	-72	la protéine PKG2 humaine (retrouvée 2 fois)
AAAG ^A _G GGACCA	-74	-243	la protamine de truite		
GA ^T _G GCCATCT	-199	-200	la protamine de truite		

FIGURE 12: Séquences nucléotidiques communes aux les gènes de protamines de souris mP1 et mP2 et d'autres gènes spécifiques du testicule d'après JOHNSON *et al.* (1988).

Des sites de fixation pour les facteurs de transcription tels que Sp-1, Ap1 (DYNAN et TJIAN, 1983 ; MITCHELL et TJIAN, 1989) et CREB (FINK *et al.*, 1988) ont été décrits dans la région 5' non codante de protamine de coq.

Une séquence potentielle de régulation spécifique des protamines de type P1 est suggérée par KRAWETZ et DIXON (1988). Cette séquence est retrouvée en 5' des gènes de protamines de souris, de taureau, de truite et de coq (Figure 13).

Souris	T T C T T G A C T G T G
Taureau	T T A A T T A A T G T G
Truite	C T A T T T A A A A T G
Coq	C C C T T T A A T - T G
Consensus	Y t M t T t A a t R T G

FIGURE 13 - Séquences potentielles de régulation des gènes spécifiques des protamines de type P1.

Les minuscules représentent les bases communes à 3 des 4 espèces étudiées (Y = pyrimidine C ou T ; R = purine A ou G; M = A ou C).

Dans la région 3' non traduite, on retrouve le signal de polyadénylation (A A T A A A ou A A A T A A A) et, quelques nucléotides en amont de ce signal, une séquence G C C A C C T G qui est présente dans les gènes de mammifères, chez la souris (LEE *et al.*, 1987a), le taureau (KRAWETZ *et al.*, 1987), l'homme (LEE *et al.*, 1987b) et le rat (TANHAUSER et HECHT, 1989). Cette séquence est aussi retrouvée dans une protéine de sécrétion androgénodépendante de l'épidydime de rat (JOHNSON *et al.*, 1988) et est répétée quatre fois dans un "activateur" des chaînes lourdes d'immunoglobine. Selon cet auteur, elle pourrait avoir un rôle d'"activateur" transcriptionnel des gènes de protamines ou intervenir dans la réduction de taille de la queue poly-A des ARNm ou au niveau de la régulation traductionnelle de ces ARNm.

I - 3-1-3 Conformation de l'ADN des gènes de protamines

Les quelques travaux portant sur l'organisation des gènes de protamines mentionnent un fort potentiel de forme Z dans les gènes de protamines.

Chez la truite, une conformation Z a été caractérisée *in vitro* dans la région 3' non codante où se situe une séquence répétée de bases puriques et pyrimidiques alternées. Cette forme Z est aussi retrouvée dans la région 3' des gènes de protamines de souris avant le signal de polyadénylation à la fois pour les protamines mp1 et mp2. Chez le coq, il y a accumulation de régions en forme Z potentielle en 3' et aussi entre la séquence TATA et le CRE (OLIVA et DIXON, 1989) en 5'. Ces auteurs suggèrent qu'il pourrait y avoir une corrélation entre les régions en forme Z et la régulation de la transcription.

I - 3-2 MECANISMES DE REGULATION TRANSCRIPTIONNELLE ET TRADUCTIONNELLE

I - 3-2-1 Mise en évidence de mécanismes de régulation

L'expression des gènes de protamines ne prend place que dans les cellules de la lignée germinale. Ces gènes constituent donc un excellent modèle pour l'étude des mécanismes de l'expression des gènes conduisant à une différenciation terminale. Les gènes de protamines sont en effet soumis à différents types de régulations transcriptionnelles et traductionnelles.

Chez la truite, il y a arrêt de l'expression des gènes d'histones et transcription des gènes des protamines au cours de la méiose (spermatocytes I et II). Les ARNm de protamines sont ensuite stockés dans le cytoplasme sous forme de particules inactives : les ribonucléoprotéines (RNPM) (SINCLAIR et DIXON, 1982). L'activation de ces RNPM n'a lieu que trois à quatre semaines plus tard au cours du stade spermatide. Il y a donc un contrôle transcriptionnel au stade spermatocyte et un contrôle traductionnel au stade spermatide. Il est possible qu'un tel contrôle traductionnel élimine la nécessité d'une étape de régulation dans la synthèse des ARNm (AIKEN *et al.*, 1983).

Chez le rat, on trouve des protéines intermédiaires dont les ARNm apparaissent au stade spermatide tardive, mais ne sont traduits que trois à quatre jours plus tard (HEIDARAN et KISTLER, 1987b). Cette étape de régulation traductionnelle se produit donc à un stade ultérieur par rapport à ce qui se passe chez la truite.

Chez le rat, le hamster et la souris, il y a deux variants de protamines P1 et P2 qui sont transcrits et traduits. Cependant, la quantité de variants synthétisés reflète des niveaux différentiels d'ARNm. De telles différences pourraient être dues à une transcription différentielle des gènes de protamines ou encore à des différences de stabilité des ARNm selon les espèces (HECHT, 1987).

I - 3-2-2 Mécanismes potentiels de régulation

Les résultats décrits précédemment indiquent bien la présence de phénomènes de régulation à la fois au niveau transcriptionnel et traductionnel. Cependant leurs mécanismes sont encore mal connus :

* En ce qui concerne les mécanismes de transcription des gènes de protamines, l'étude des séquences signalées dans le paragraphe I - 3-1-2 devrait apporter des éléments de réponse quant à la régulation de la transcription.

* En ce qui concerne les mécanismes de régulation traductionnelle, il est vraisemblable que ceux-ci soient liés aux caractéristiques structurales des ARNm.

HECHT (1987) indique que les ARNm des protamines mP1 et mP2 de souris sont synthétisés et stockés sous forme de ribonucléoprotéines où la taille des ARNm varie de 580 à 830 nucléotides. Immédiatement avant ou pendant la traduction, l'ARNm est partiellement déadénylé. Cette réduction de la queue poly-A de 160 à 30 nucléotides environ est commune à toutes les espèces de protamines. Cela n'est cependant pas un phénomène spécifique de la spermiogenèse et n'explique pas le mécanisme de régulation traductionnelle.

Les ARNm de protamines se distinguent aussi par un taux élevé de structures secondaires qui doivent jouer un rôle dans la structure de la ribonucléoprotéine messagère et par conséquent dans le stockage des ARNm. L'existence de ces structures secondaires a été mise en évidence :

- par des études de dénaturation thermique. En effet une partie de cet ARN ne se dissocie qu'à haute température (76°C) (DAVIES *et al.*, 1976).
- par la présence de régions résistantes à la coupure par la ribonucléase T1 (DAVIES et DIXON, 1979).
- par des expériences d'extension d'amorce avec la reverse transcriptase

Ces données sont en accord avec les résultats obtenus par les méthodes statistiques de calcul de structures secondaires qui indiquent des appariements importants de bases (Figure 14).

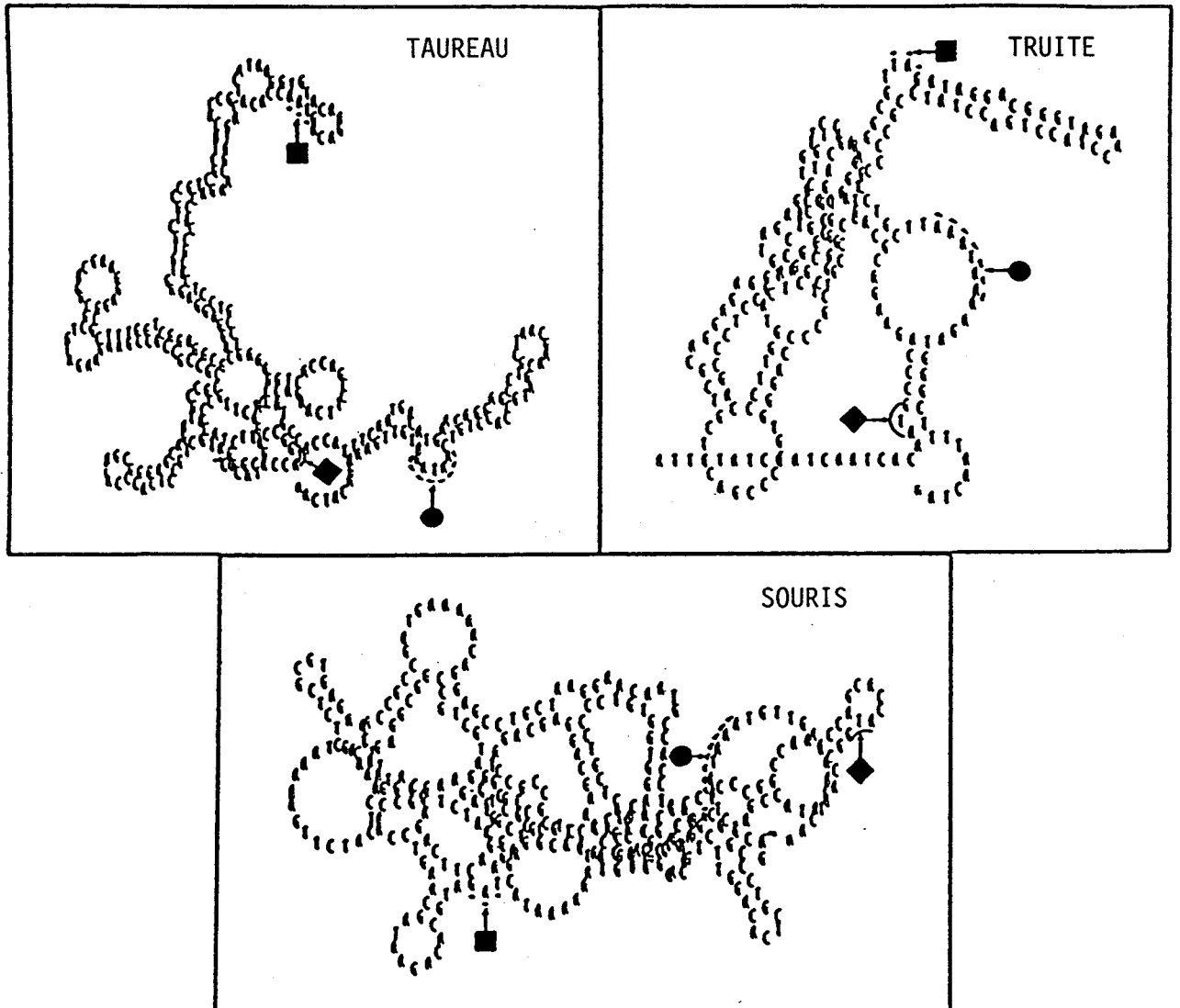


FIGURE 14 - Représentation des structures secondaires prédites des ARNm de protamines de différentes espèces selon KRAWETZ *et al.* (1987). Les codons d'initiation, de terminaison et les signaux de polyadénylation sont indiqués par : ◆, ■ et ● respectivement.

De plus, l'organisation des ARNm de protamines présente une dissymétrie des parties non traduites 5' et 3'.

- En effet, les parties 5' non traduites sont relativement courtes et de taille inégale : 14 à 19 nucléotides chez la truite (GREGORY *et al.*, 1982 ; STATES *et al.*, 1982), 94 pour le taureau, 91 pour le rat et environ 113 pour la souris (KLEENE *et al.*, 1985).
- Par contre les segments non codants en 3' sont plus longs, mais similaires en longueur (environ 150 nucléotides) dans les différentes espèces. KRAWETZ et ses collaborateurs (1987) suggèrent qu'il pourrait y avoir un contrôle traductionnel dans cette région 3' des ARNm. En effet si l'on regarde la position qu'occupe cette région dans les prédictions de structures secondaires (Figure 14), on s'aperçoit que les codons de terminaison et les signaux de polyadénylation sont relativement exposés alors que les codons initiateurs de traduction sont partiellement appariés (2 bases sur 3) et la troisième initie une boucle.

Il pourrait donc y avoir passage d'une forme d'ARNm "non traductible" avec des structures secondaires importantes, à une forme d'ARNm "traductible" après intervention d'un ou plusieurs facteurs trans-ou cis-activateurs qui interféreraient sur ces structures secondaires.

I - 3-3 NATURE DES CODONS UTILISES

D'une manière générale, il semble qu'il existe des codons préférentiels pour les protamines selon les espèces. Ainsi, chez la truite, la souris, la roussette et le coq (MORI *et al.*, 1985 ; GEDAMU *et al.*, 1981 ; KLEENE *et al.*, 1985 ; BERLOT-PICARD

et al., 1986 ; OLIVA et DIXON, 1989), sur les six possibilités de codons pour l'arginine (CGU, CGG, CGA, CGC, AGA et AGG) les codons les plus abondants sont CGC (88,6 % des codons arginine chez le coq) et AGG. Les quatre autres possibilités ont des fréquences différentes. Par exemple, le codon CGA n'est pas utilisé dans les gènes de protamines de coq et de roussette, mais est présent dans ceux de truite et de souris.

L'agencement des codons dans les multipléts d'arginine rencontrés dans les séquences de protamines, est différent selon l'espèce étudiée. Ainsi chez la roussette (protamine Z2), il y a une alternance des codons CGC et AGG ; chez le coq, seul CGC est présent tandis que dans les protamines des autres espèces tous les codons sont utilisés.

Pour les autres acides aminés, les codons ont des fréquences totalement différentes selon l'espèce (Figure 15).

I - 3-4 MECANISMES DE REGULATION POST-TRADUCTIONNELLE DES PROTAMINES

Dans la plupart des protamines, la séquence en acides aminés ne commence pas par une méthionine. Cependant, les études réalisées chez la truite par WIGLE et DIXON (1970) et GILMOUR et DIXON (1972) ont montré par des expériences *in vivo* et *in vitro* qu'il y a un résidu de méthionine incorporé à l'extrémité N-terminale des protamines nouvellement synthétisées, mais que ce résidu est éliminé lorsque la maturation de la spermatide est finie. On est donc en présence d'un mécanisme de régulation post-traductionnelle qui implique l'excision de la méthionine "initiatrice" de la traduction.

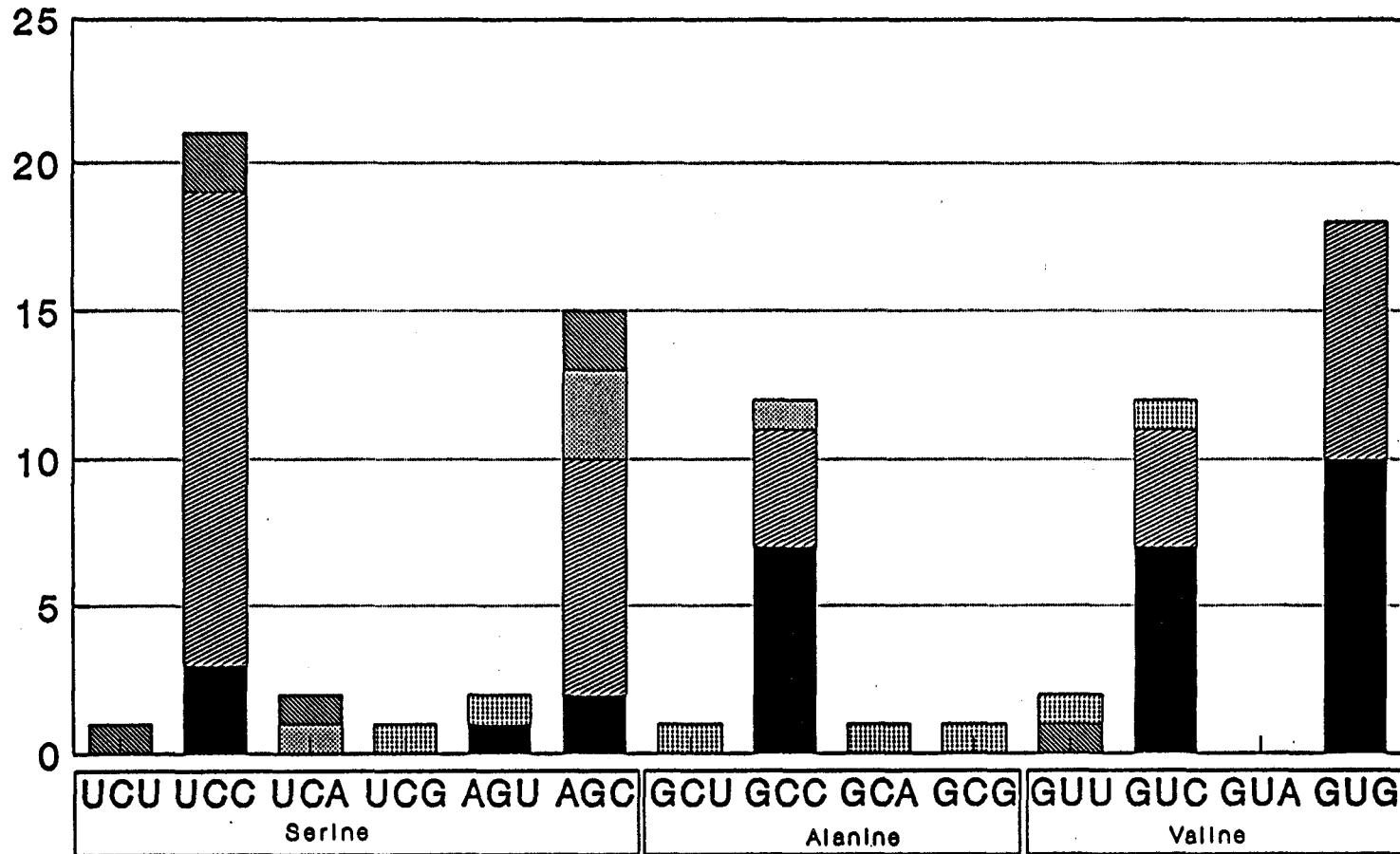


FIGURE 15: Fréquence de quelques codons utilisés dans les gènes de protamines de différentes espèces. ■ protamine de Truite selon MORI *et al.* (1985), ▨ protamine de truite selon GEDAMU *et al.* (1981), ▩ protamine mP1 de souris selon KLEENE *et al.* (1985), ▤ protamine mP2 de souris selon YELICK *et al.* (1987), ▧ protamine Z1 de roussette selon BERLOT-PICARD *et al.* (1986).

Récemment un ADNc codant pour la protéine de type P2 de la souris (mp2) a été identifié. Cette protéine serait synthétisée sous la forme d'une protéine précurseur plus importante qui serait ensuite maturée pour donner la protamine P2 dans le spermatozoïde (YELICK *et al.*, 1987).

De plus, lors de la réalisation de la carte génomique des protamines mp1 et du précurseur de mp2, JOHNSON et ses collaborateurs (1988) n'ont pas mentionné la présence d'un gène pour la protamine P2 "maturée".

Ces travaux vont dans le sens d'un mécanisme de régulation post-traductionnelle par coupure protéolytique d'un précurseur.

I - 4 LES PROTÉINES BASIQUES ASSOCIÉES À L'ADN AU COURS DE LA SPERMIOGENÈSE CHEZ LA SEICHE

I - 4-1 INTERET DE LA SEICHE, *SEPIA OFFICINALIS*

La seiche, *Sepia officinalis*, est un Invertébré marin qui appartient à la classe des Mollusques Céphalopodes. Cet animal présente l'avantage de posséder un testicule bien différencié qui renferme tous les types de cellules germinales - de la spermatogonie à la spermatide allongée.

Le choix de la seiche est aussi lié au fait qu'elle partage certains caractères évolués avec les Vertébrés supérieurs:

- Contrairement à la plupart des autres Invertébrés marins, elle a un mode de fécondation interne avec la présence d'un bras copulateur appelé hectocotyle.

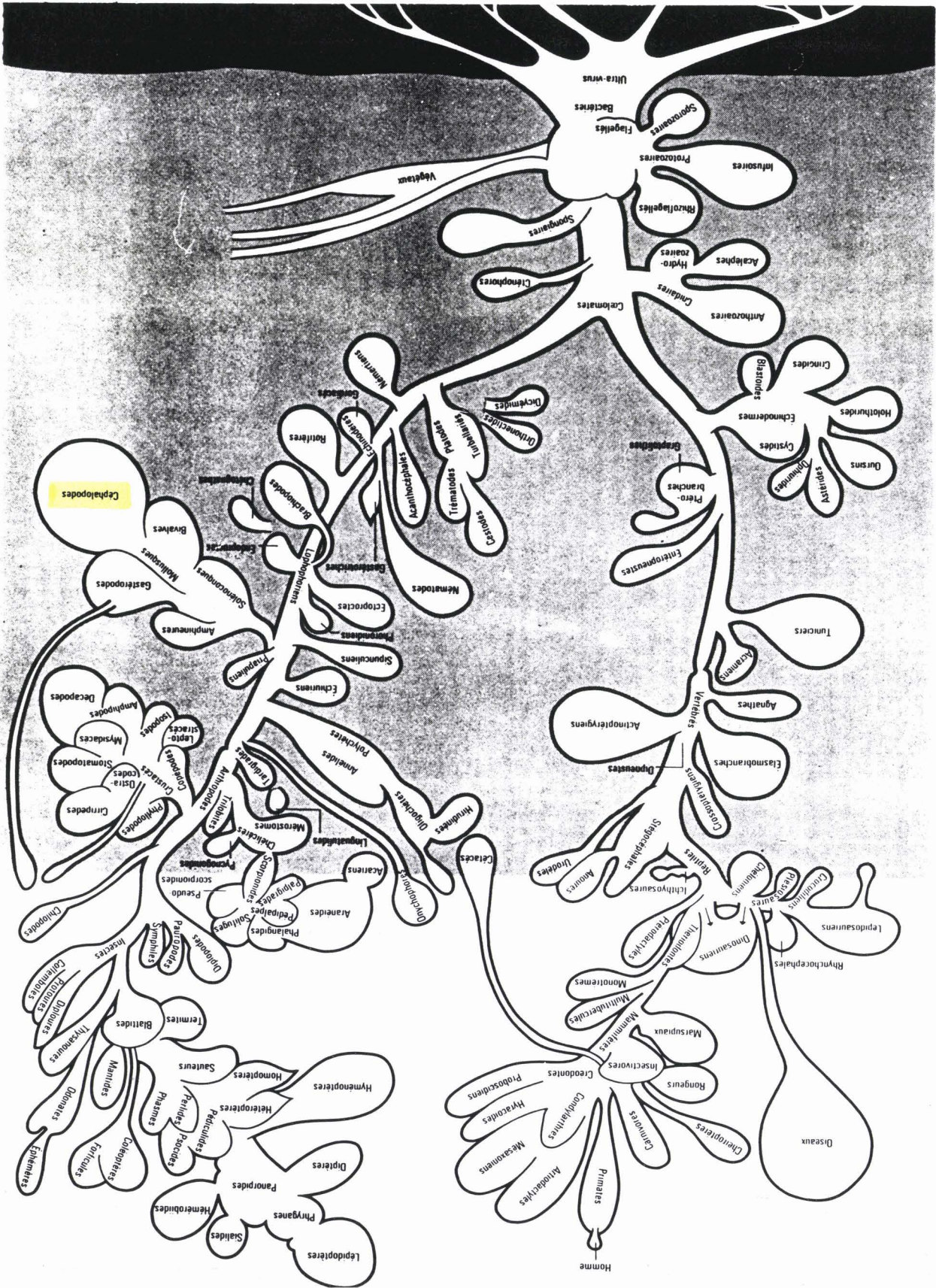
- De plus, la maturation génitale est sous contrôle hormonal et il y a un mécanisme de rétrocontrôle des gonades sur la glande neurosécrétrice.
- Par ailleurs, l'organisation de l'oeil est comparable à celle des mammifères, et l'ensemble des diverses fonctions de régulation assurées par le système "Cerveau-glande optique" est comparable à l'unité fonctionnelle de l'hypothalamus-hypophyse des Vertébrés (RICHARD, 1971).

Chez la seiche, la spermiogenèse est caractérisée par une double transition protéique nucléaire : histones → protéine intermédiaire → protamine. Cette double transition est également observée chez le calmar (SUBIRANA *et al.*, 1973 ; KHRAPUNOV *et al.*, 1988). Cette caractéristique s'ajoute à celles qui déjà rapprochent les Céphalopodes des Vertébrés et les placent au sommet de leur phylum dans l'évolution phylogénétique (Figure 16).

I - 4-2 CHANGEMENTS PROTEIQUES AU COURS DE LA SPERMIOGENESE DE LA SEICHE

A côté des histones de type somatique (WOUTERS-TYROU *et al.*, 1981, 1982 ; COUPPEZ *et al.*, 1987), deux protéines très basiques ont été isolées à partir de la chromatine du testicule (WOUTERS-TYROU *et al.*, 1988). La protéine T est une protéine intermédiaire qui apparaît dans les spermatides rondes et qui disparaît des spermatides allongées où elle est remplacée par une protamine, appelée protéine Sp (ROUSSEAU *et al.*, 1988). Dans les spermatides allongées, la protamine est phosphorylée à des degrés divers. La protéine Sp déphosphorylée est le composant protéique basique majeur associé à l'ADN dans le spermatozoïde mûr.

FIGURE 16 - Arbre phylogénétique du règne animal d'après CUENOT



I - 4-2-1 La protéine T

La protéine T est une protéine très riche en arginine (environ 60%). Elle est dépourvue de cystéine, d'isoleucine, de phénylalanine et de tryptophane (WOUTERS-TYROU *et al.*, 1988). Cette protéine existe sous différentes formes phosphorylées dans les spermatides rondes et en cours d'élongation.

I - 4-2-2 La protéine Sp

La protéine Sp a une composition en acides aminés très particulière : Ser₇, Pro₁, Tyr₅, Arg₄₃ ou 44. Elle doit sa particularité à son taux très élevé d'arginine (77 %) et au petit nombre d'acides aminés constitutifs. Ce dernier caractère joint à l'absence de cystéine est retrouvé dans les protamines de poissons téléostéens où le taux d'arginine est toutefois moins élevé (~ 60 %). Par ailleurs, la richesse en tyrosine de la protamine Sp et la distribution des résidus d'arginine et des acides aminés hydroxylés sont des caractéristiques structurales que l'on retrouve dans les protamines P1 de mammifères (MARTIN-PONTHIEU *et al.*, 1990).

De plus, sa composition en acides aminés s'intègre dans celle de la protéine T, ce qui suggère que cette dernière pourrait être un précurseur de la protamine (WOUTERS-TYROU *et al.*, 1988).

Afin d'infirmer ou de confirmer cette hypothèse, nous avons participé à la détermination de la structure primaire de la protéine T.

Cependant, il faut noter que, dans tous les cas où une hypothèse de précurseur protéique des protamines a été évoquée, le mécanisme de coupure protéique n'a jamais été élucidé et l'enzyme ou les enzymes impliquées n'ont jamais été caractérisées. Il nous a donc semblé indispensable d'entreprendre une étude au niveau génétique, car nous ne pouvions exclure la possibilité d'une régulation génique (gènes différents, expression en ARNm différente ou maturation différentielle par épissage de l'ARNm).

Cette dernière étude a nécessité l'introduction des techniques de biologie moléculaire au sein du laboratoire dont les activités étaient jusqu'à présent plus orientées sur l'étude des structures et des interactions des protéines de la chromatine avec l'ADN.

CHAPITRE II

RÉSULTATS ET DISCUSSION

II - 1 DÉTERMINATION DE LA STRUCTURE PRIMAIRE DE LA PROTÉINE T

II - 1-1 PRÉPARATION ET PURIFICATION DE LA PROTÉINE T

II - 1-1-1 Préparation

Les protéines basiques majeures sont extraites de la chromatine de testicule de seiche par l'HCl 0.4 N. Elles ont une migration électrophorétique très similaire à celle des histones de type somatique (Figure 17 - piste 3). Cet extrait contient une protéine appelée protéine T dont la migration électrophorétique est intermédiaire entre celle de l'histone H4 (M_r 11.282) (Figure 17 - piste 5) et celle de la protamine de bélier ($M_r \sim 6700$) (Figure 17 - piste 2).

Les protéines basiques résiduelles sont alors extraites par une solution de GuCl 5 M en présence de 2-mercaptoéthanol. Ce deuxième extrait (Figure 17 - piste 4) renferme des traces d'histones et une quantité appréciable de protéine T, mais son composant majeur est une protéine plus petite, plus basique dont la mobilité électrophorétique est similaire à celle de la protamine de bélier. Cette protéine est identique à la protamine isolée du sperme de la seiche (Figure 17 - piste 7) et a été appelée protéine Sp (WOUTERS-TYROU *et al.*, 1988).

II - 1-1-2 Purification

* A partir de l'extrait HCl 0.4 N

La protéine T est séparée des histones par fractionnement CLHP sur une colonne de C8 ultrapore (Beckman).

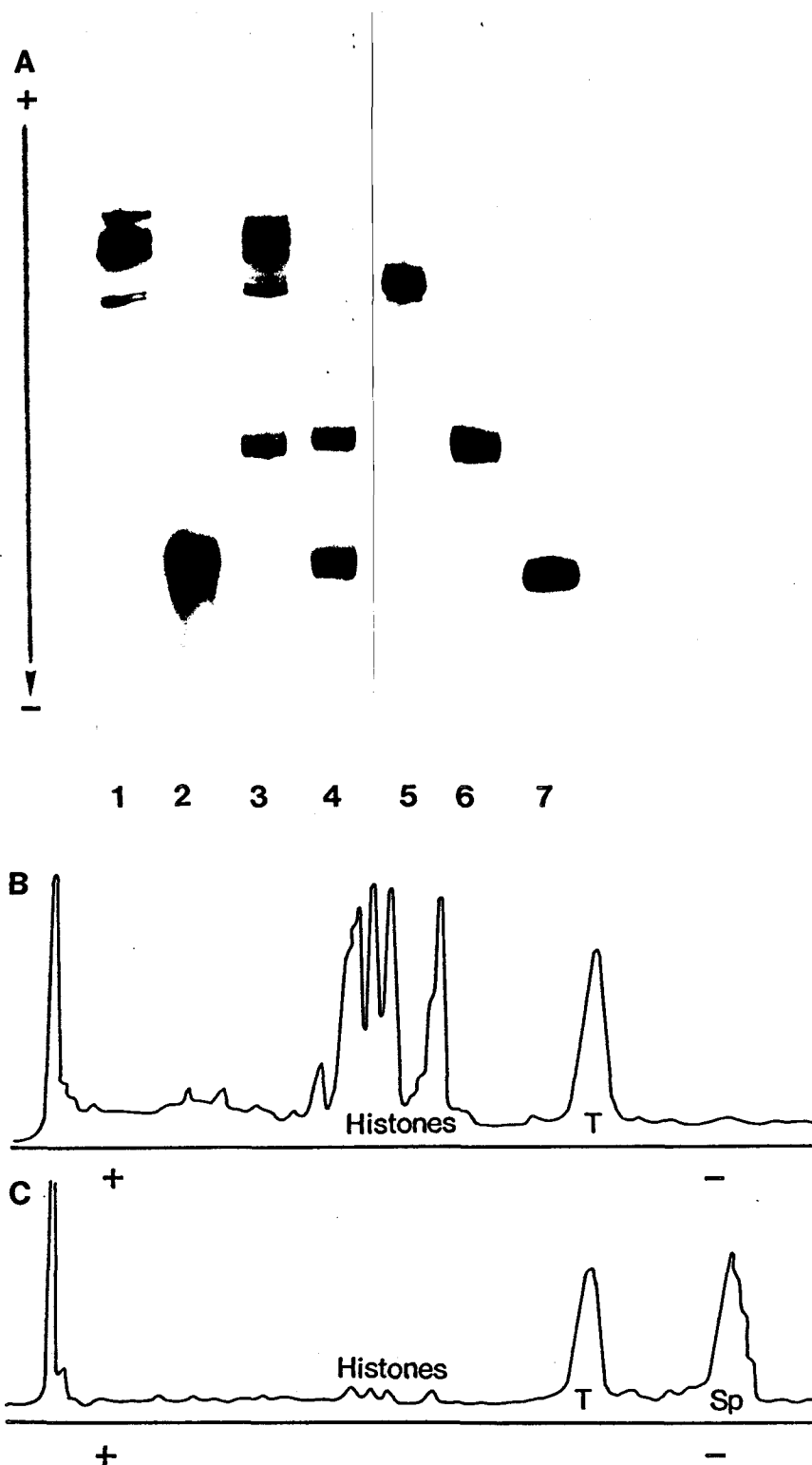


FIGURE 17: A: Analyse électrophorétique des protéines basiques nucléaires spécifiques de la spermiogenèse de la seiche *Sepia officinalis*. L'électrophorèse est réalisée sur un gel de polyacrylamide (17%) en milieu: acide acétique (0,9 M), urée (6,25 M). Les échantillons sont dissous dans: HCl 0,01 M/urée 8 M/2-mercaptoéthanol 0,5 M, 24 h avant dépôt sur le gel. Après 1,5 h de préélectrophorèse, la migration est réalisée sous un ampérage constant de 22 mA pendant 2,5 h. Le gel est ensuite coloré par la technique de STEPHANO *et al.* (1986). 1: témoin d'histones totales de veau, 2: témoin de protamine de bélier, 3: extrait HCl 0,4 N de la chromatine de testicule de seiche, 4: extrait GuCl 5 M de la chromatine de testicule de seiche, 5: histone H4 de seiche, 6: protéine T purifiée, 7: protéine Sp purifiée. B et C: enregistrement densitométrique (Spectrophotomètre Vernon PHI-6) des pistes 3 et 4.

Le diagramme d'élution est présenté dans la figure 18-A. Les différents composants de l'extrait sont répartis dans trois groupes de pics. La protéine T est éluee dans le premier groupe. L'hétérogénéité de la fraction est liée à l'existence de variants structuraux et de formes phosphorylées. Les différents variants de l'histone H1 sont élués dans le groupe intermédiaire. Le dernier groupe de pics correspond aux core-histones qui sont éluées dans l'ordre suivant : H2B, H2A, H4 et H3 (Figure 18-B).

* A partir de l'extrait GuCl 5 M

La protéine T est séparée de la protéine Sp par chromatographie en phase inverse sur une colonne C18 μ Bondapak. Trois fractions protéiques sont isolées (Figure 19) qui correspondent respectivement à la protéine Sp (pic 1), la protéine T (pic 2) et les histones résiduelles H3 et H4 (pic 3).

II - 1-2 COMPOSITION EN ACIDES AMINES DE LA PROTEINE T

La caractéristique majeure de la composition en acides aminés de la protéine T, présentée dans le Tableau I, est le taux très élevé d'arginine (environ 59 %). Les autres acides aminés les plus représentés sont la sérine (~ 10 %) et la tyrosine ($\sim 6,5$ %).

La protéine T occupe une position intermédiaire entre l'histone H4 et la protamine Sp de seiche, à la fois du point de vue du nombre d'acides aminés constitutifs (H4 : 16 > T : 13 > Sp : 4) et de la richesse en acides aminés basiques (H4 : 23,6 % > T : 63,7 % > Sp : 77,4 %). Par contre, la composition en acides aminés de la protéine T diffère de celle des protéines de type

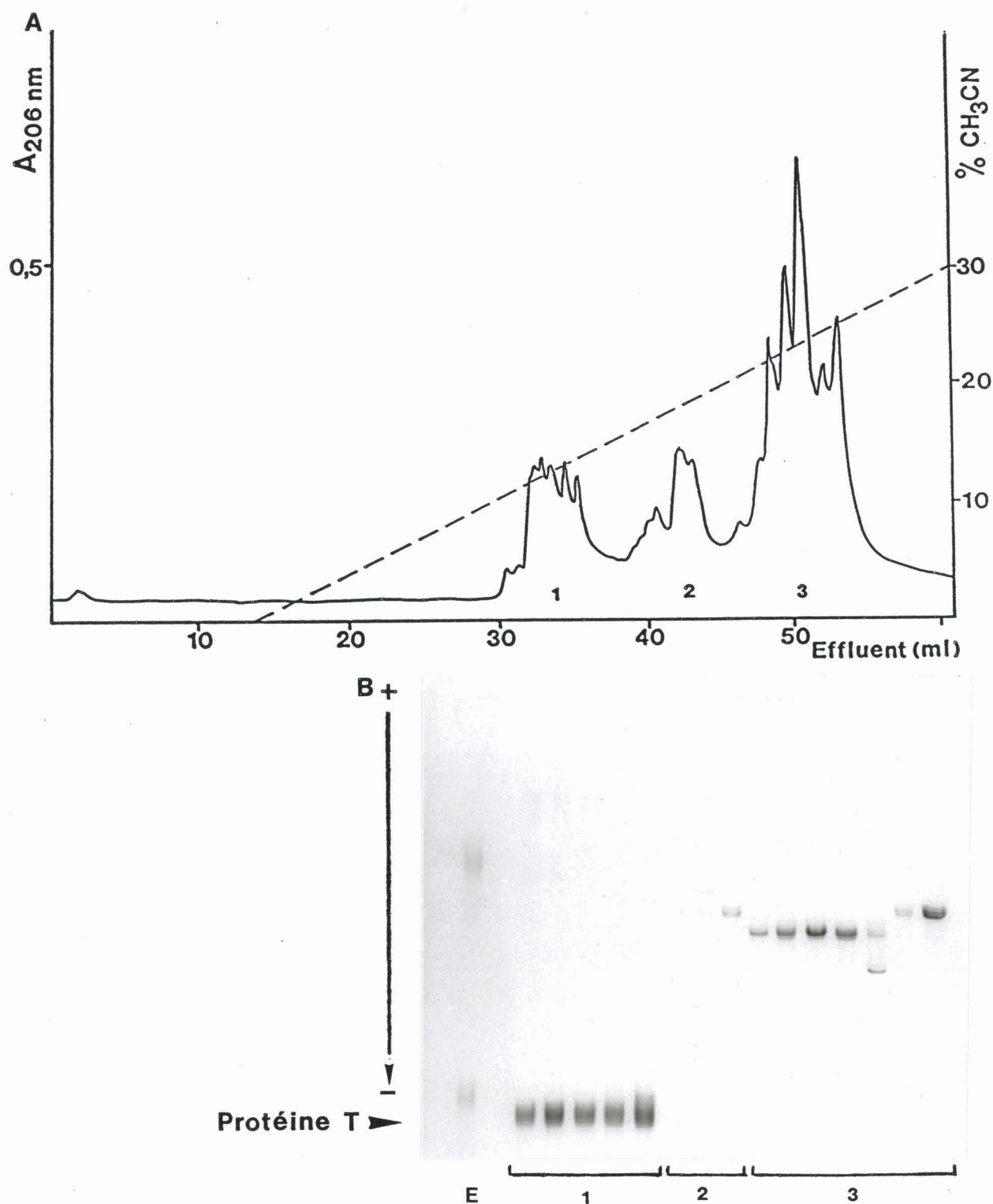


FIGURE 18: A: Purification de la protéine T à partir de l'extrait HCl 0,4 N de la chromatine de testicule de seiche. La séparation des fractions est réalisée sur colonne C_8 Ultrapore (4,6 x 250 mm) équilibrée et éluée par un gradient d'acétonitrile dans le TFA 0,1 % (---) à un débit de 1 ml/mn. L'éluion des pics est suivie à 206 nm (—). B: Trois fractions 1, 2 et 3, constituées de plusieurs sous-fractions, ont été récupérées et analysées sur gel de polyacrylamide dans les conditions décrites dans la légende de la figure 17 : piste E extrait HCl 0,4 N de la chromatine, les autres pistes portent le nom de la fraction dont elles proviennent.

TABLEAU I

Composition en acides aminés de la protéine T. Les compositions de la protamine Sp et de l'histone H4 de Seiche, des protéines TP1 et TP2 spécifiques des spermatozoïdes de Rat ainsi que des protéines intermédiaires S1 et S2 de Roussette sont données pour comparaison.

	S e i c h e			R a t			Roussette	
	Protéine T (1)	Protamine Sp (1)	Histone H4 (2)	TP1 (3)	TP2 (4)	S1 (5)	S2 (5)	
Asp	2.1 (1-2)		5.4 (5)	7.4 (4)	4.3 (4)	7.1 (7)	6.7 (7)	
Thr	0.6 (1)		6.3 (7)	3.7 (2)	6.6 (8)	4.3 (4)	2.1 (2)	
Ser	9.6 (8-9)	11.4 (6)	2.2 (2)	14.8 (8)	14.4 (21)	5.4 (5)	3.8 (4)	
Glu	2.6 (1)		6.4 (6)		4.8 (4)	3.1 (3)	5.0 (5)	
Pro	1.4 (1)	2.2 (1)	1.2 (1)	1.8 (1)	12.0 (14)	4.7 (5)	5.7 (6)	
Gly	3.1 (2)		16.5 (17)	11.1 (6)	5.8 (8)	2.2 (2)	1.0 (0)	
Ala	3.3 (2-3)		6.9 (7)	3.7 (2)	4.7 (5)		6.5 (6)	
Cys/2					4.4 (5)	1.2 (1)		
Val	1.7 (1-2)		9.0 (9)	1.8 (1)	2.8 (2)	4.7 (5)	3.2 (3)	
Met	1.0 (3)		0.7 (1)	1.8 (1)	3.1 (2)	1.9 (2)	0.8 (1)	
Ile			6.0 (6)		0.3 (-)	2.8 (3)	1.0 (1)	
Leu	4.3 (3-4)		8.1 (8)	5.5 (3)	2.3 (1)	4.9 (5)	7.4 (8)	
Tyr	6.5 (5)	9.0 (5)	3.8 (4)	3.7 (2)	1.8 (2)	2.9 (3)	3.5 (4)	
Phe			2.0 (2)		0.2 (-)			
His			1.9 (2)	5.5 (3)	8.3 (12)	2.9 (3)	1.8 (2)	
Lys	5.0 (4)		10.8 (11)	18.5 (10)	9.9 (9)	7.6 (8)	16.9 (18)	
Arg	58.7 (43-44)	77.4 (43-44)	12.8 (14)	20.4 (11)	14.2 (19)	31.2 (31)	12.8 (13)	

Les résultats sont exprimés en moles/100 moles. Les chiffres entre parenthèses correspondent aux résidus effectivement présents dans la séquence. (1) WOUTERS-TYROU *et al.* (1988), (2) Résultats non publiés WOUTERS-TYROU *et al.* (1975), (3) KISTLER *et al.* (1975), (4) GRIMES *et al.* (1977); LUERSEN *et al.* (1989), (5) CHAUVIERE *et al.* (1987, 1989).

intermédiaire que l'on peut trouver chez les mammifères ou la roussette par l'absence d'histidine et par un taux plus élevé d'arginine. Ce taux est très proche de celui trouvé dans les protamines des poissons téléostéens (Mc KAY *et al.*, 1986).

II - 1-3 ETUDES STRUCTURALES DE LA PROTEINE T

II - 1-3-1 Dégradation automatique d'Edman de la protéine T

La protéine T a été soumise à la dégradation automatique d'Edman. Les 50 cycles réalisés nous ont permis de déterminer sans ambiguïté les 39 premiers résidus d'acides aminés et d'obtenir des indications de séquence jusqu'au résidu 50 (Figure 20).

Ces premiers résultats démontrent une asymétrie de la protéine puisque les résidus hydrophobes sont accumulés dans la région N-terminale (résidus 1 à 19).

De plus l'obtention de deux PTH d'acides aminés en positions 7 (Ser ou Thr), 12 (Ala ou Val) et 16 (Glu ou Asp) montre la présence d'au moins deux variants structuraux de la protéine T.

Les résultats de la dégradation d'Edman suggèrent qu'il existe deux chaînes polypeptidiques dont l'une serait dépourvue de méthionine en position N-terminale puisqu'on observe un décalage d'un acide aminé à chaque cycle. Toutefois, à ce stade de la détermination de la séquence de la protéine T, il ne nous est pas possible d'attribuer la délétion du résidu de méthionine N-terminal à l'un des deux variants signalés précédemment.

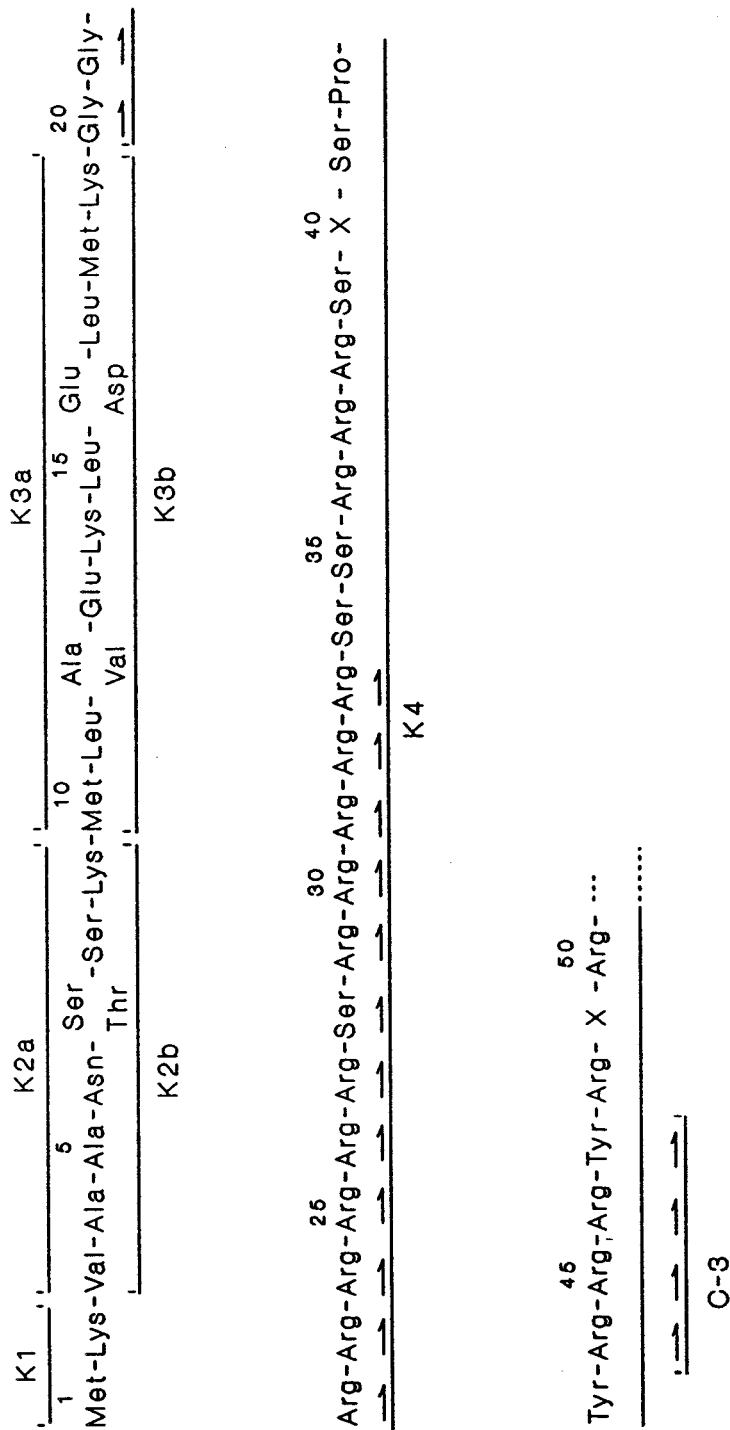


FIGURE 20: Séquence amino-terminale du mélange des variants de la protéine T déterminée par la dégradation automatique d'Edman (résidus 1 à 50). Les lettres K et C représentent respectivement les peptides obtenus par coupure avec l'endoprotéinase Lys-C et la chymotrypsine. Les acides aminés identifiés par la dégradation d'Edman dans les peptides sont représentés par: →. X indique un acide aminé non identifié.

II - 1-3-2 Stratégie de coupure de la protéine T

Les caractéristiques de la composition en acides aminés et de la séquence N-terminale de la protéine T nous ont conduit à adopter la stratégie suivante pour l'étude structurale de cette protéine :

- la présence de 4 résidus de lysine suggère l'utilisation de l'endoprotéase Lys-C. En effet, les 4 résidus de lysine ont été placés dans les 20 premiers résidus de la protéine, la coupure par l'endoprotéase Lys-C permet d'obtenir le fragment C-terminal (résidus 20 à la fin) à partir duquel pourra se poursuivre l'étude structurale de la protéine T.

- la présence de 5 résidus de tyrosine dans la composition en acides aminés nous a incité à hydrolyser la protéine par la chymotrypsine. Les résidus de tyrosine sont localisés dans la moitié carboxy-terminale de la protéine (La première tyrosine est en position 43) et les peptides chymotrypsiques devraient nous permettre d'accéder à la séquence C-terminale de la protéine.

II - 1-3-3 Hydrolyses enzymatiques de la protéine T

* *Endoprotéase Lys-C*

Le diagramme d'élution des peptides obtenus par action de l'endoprotéase Lys-C sur la protéine T est représenté dans la figure 21. Les compositions en acides aminés de chacune des fractions sont rassemblées dans le Tableau II.

Ces peptides couvrent toute la chaîne protéique. Ils ont permis de confirmer la séquence N-terminale établie par la dégradation automatique d'Edman de la protéine entière et de

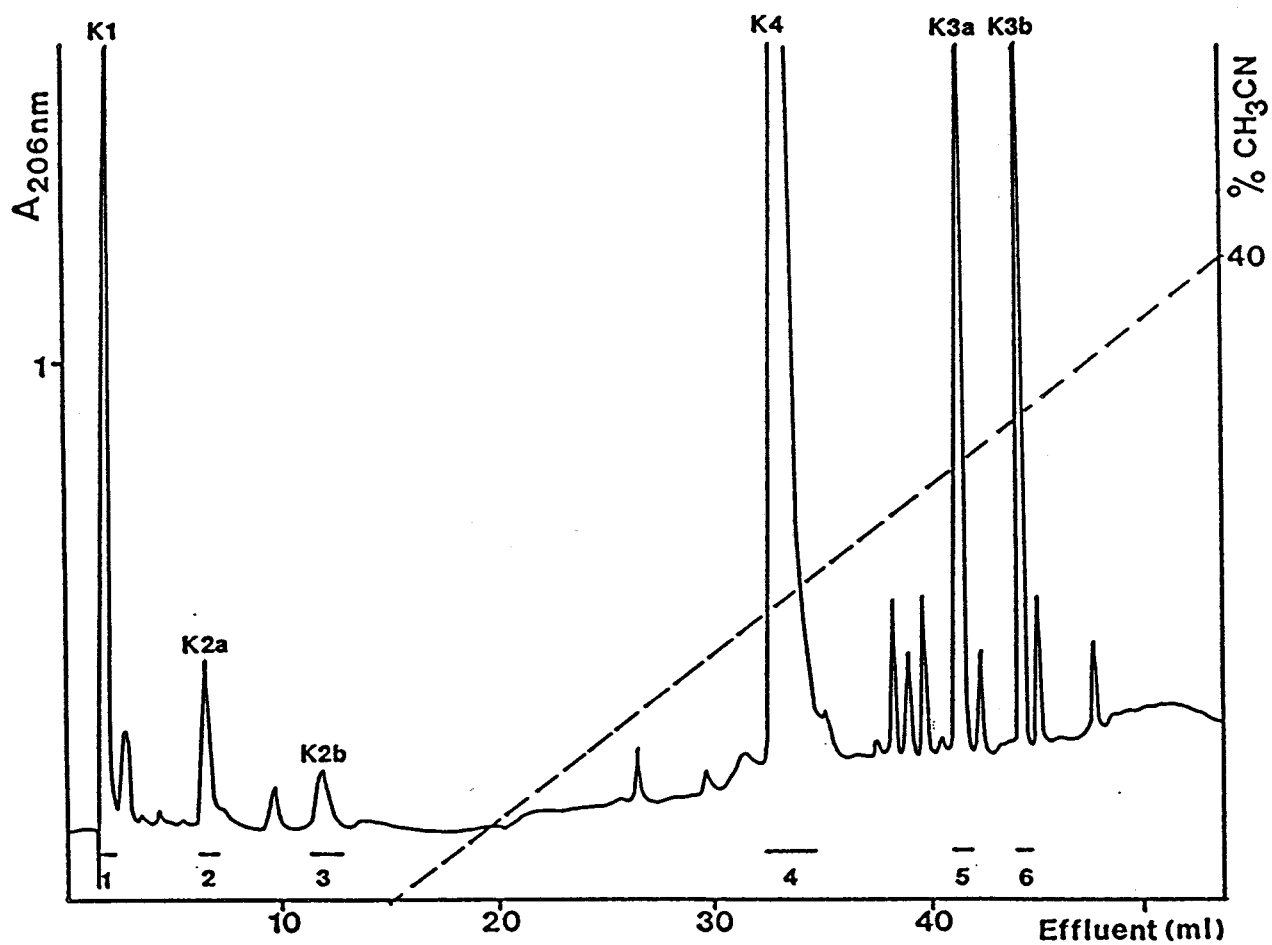


FIGURE 21: Séparation des peptides obtenus après action de l'endoprotéinase Lys-C sur la protéine T. La séparation est réalisée sur une colonne C₁₈ µBondapak dans les conditions décrites dans la légende de la figure 18. Les peptides contenus dans les différentes fractions sont symbolisés par la lettre K et numérotés en fonction de leur position dans la séquence de la protéine.

TABLEAU II

Composition en acides aminés des peptides obtenus après coupure de la protéine T par l'endoprotéase Lys-C.

	K1 1-2	K2a 3-9(T1)	K2b 3-9(T2)	K3a 10-19(T1)	K3b 10-19(T2)	K4 20 à la fin
Acide aspartique		1.0 (1)	1.0 (1)		1.2 (1)	
Thréonine			0.9 (1)			
Sérine		1.6 (2)	0.9 (1)			5.6 (7)
Acide glutamique				2.2 (2)	1.5 (1)	
Proline						1.0 (1)
Glycine						2.0 (2)
Alanine		1.8 (2)	1.8 (2)	1.0 (1)		
Valine		0.9 (1)	1.0 (1)		1.0 (1)	
Méthionine	0.9 (1)			1.4 (2)	0.7 (1)	
Leucine				2.7 (3)	2.9 (4)	
Tyrosine						3.8 (5)
Lysine	1.5 (1-2)	1.1 (1)	1.0 (1)	2.0 (2)	1.7 (2)	
Arginine						43.4 (43-44)
Total	2 ou 3	7	7	10	10	58 ou 59
N° des fractions	1	2	3	6	7	5

Les résultats sont exprimés en moles/100 moles. Aucune correction n'a été faite pour les pertes occasionnées lors de l'hydrolyse. Les chiffres entre parenthèses correspondent aux nombres de résidus déduits de la séquence. Les peptides contenus dans les différentes fractions sont symbolisés par la lettre K et numérotés en fonction de leur position dans la séquence de la protéine.

mettre en évidence la présence de variants structuraux. Deux types de peptides ont été obtenus qui appartiennent sans ambiguïté à deux chaînes différentes présentes dans le rapport 2 : 1. La chaîne majeure, appelée protéine T1, renferme un résidu de sérine en position 7, un résidu d'alanine en position 12 et un résidu d'acide glutamique en position 16, alors que dans le variant mineur, protéine T2, on trouve à ces mêmes positions des résidus de thréonine, de valine et d'acide aspartique respectivement (Figure 20).

La fraction K4 renferme les peptides C-terminaux de chacun des variants (Figure 20). Cette fraction a été soumise à la dégradation automatique d'Edman, mais seuls les 14 premiers résidus ont pu être identifiés.

* *Chymotrypsine*

Le diagramme d'élution de l'hydrolysate chymotrypsique de la protéine T est complexe (Figure 22). En effet, l'enzyme a coupé, non seulement au niveau des liaisons tyrosyles, mais encore au niveau des liaisons leucyles, méthionyles et asparaginyles.

A l'hétérogénéité liée à la large spécificité de l'enzyme, s'ajoute celle qui résulte de la présence de deux variants structuraux de la protéine T. Aussi, il ne nous a pas été possible d'obtenir des peptides suffisamment purifiés recouvrant la région C-terminale. Seul un peptide nous apporte des informations complémentaires par rapport aux résultats précédents. Il s'agit du peptide C-3 dont la séquence déterminée par la dégradation d'Edman est la suivante : (Arg)₃-Tyr (Figure 20).

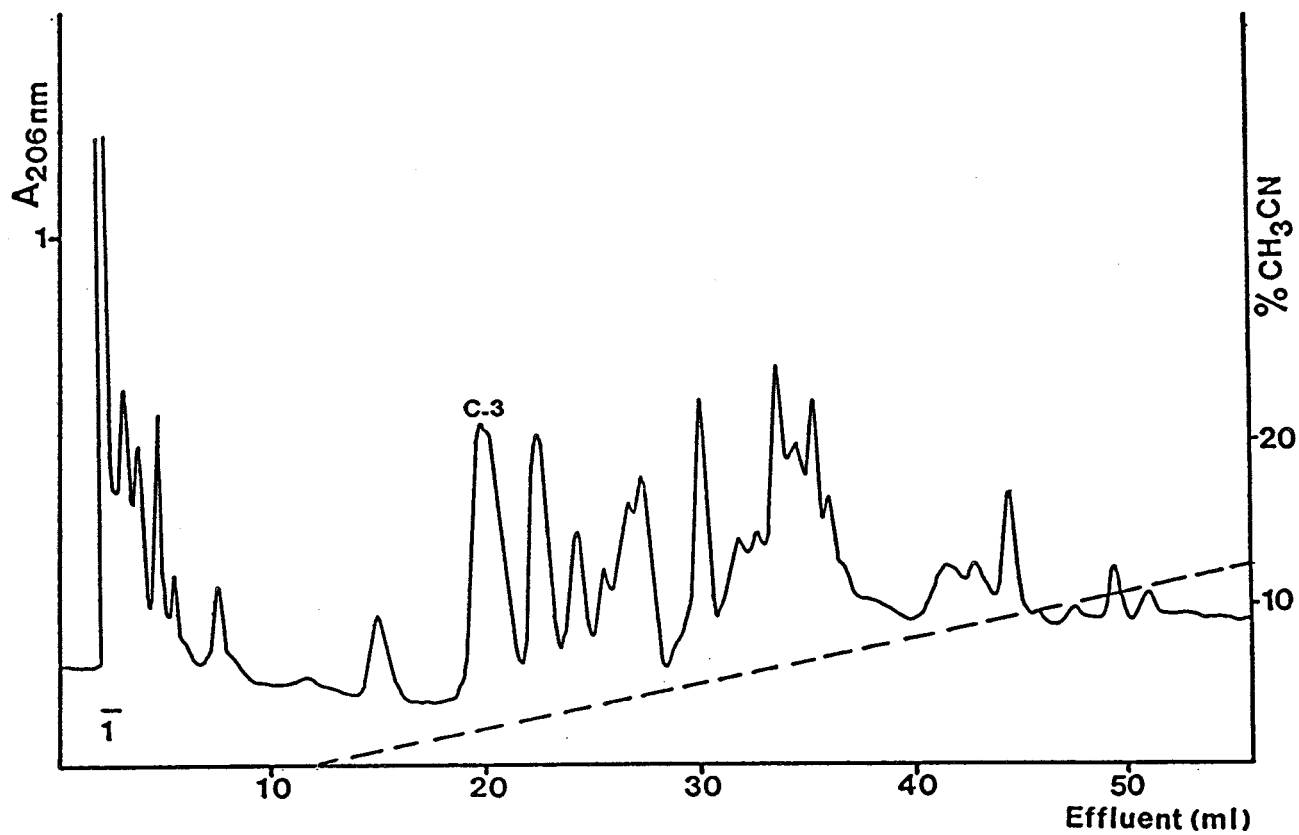


FIGURE 22: Séparation des peptides obtenus après action de la chymotrypsine sur la protéine T. La séparation est réalisée sur une colonne C₁₈ µBondapak dans les conditions décrites dans la légende de la figure 18. Le peptide (Arg)₃Tyr provient de la fraction C-3.

II - 1-4 ETUDES STRUCTURALES DES VARIANTS T1 ET T2

II - 1-4-1 Séparation des variants de la protéine T

La protéine T isolée des extraits HCl 0.4 N ou GuCl 5 M est une fraction hétérogène, composée non seulement de variants structuraux différents mais aussi de formes différemment phosphorylées.

Les premières études structurales réalisées sur la protéine T totale ont montré la complexité des hydrolysats et la difficulté d'isoler des peptides purs. Il nous a paru nécessaire, pour continuer notre étude, de travailler sur les variants séparés.

La séparation des variants de la protéine T a été obtenue après de nombreux essais sur différentes colonnes et dans différentes conditions. Les résultats les plus probants ont été obtenus par CLHP avec une colonne de μ Bondapak C18 analytique éluée par un gradient d'acétonitrile de 0 à 50 % dans le TFA 0,1 % avec introduction d'un palier (Figure 23). Dans ces conditions, les variants T1 et T2 sont élués dans les fractions 4 et 6 respectivement (Tableau III).

II - 1-4-2 Dégradation automatique des variants séparés

Chaque variant de la protéine T a été soumis à la dégradation automatique d'Edman. Les résultats obtenus confirment et complètent ceux acquis précédemment sur la protéine T totale. Ils sont rassemblés dans les figures 29 A et B, page 73.

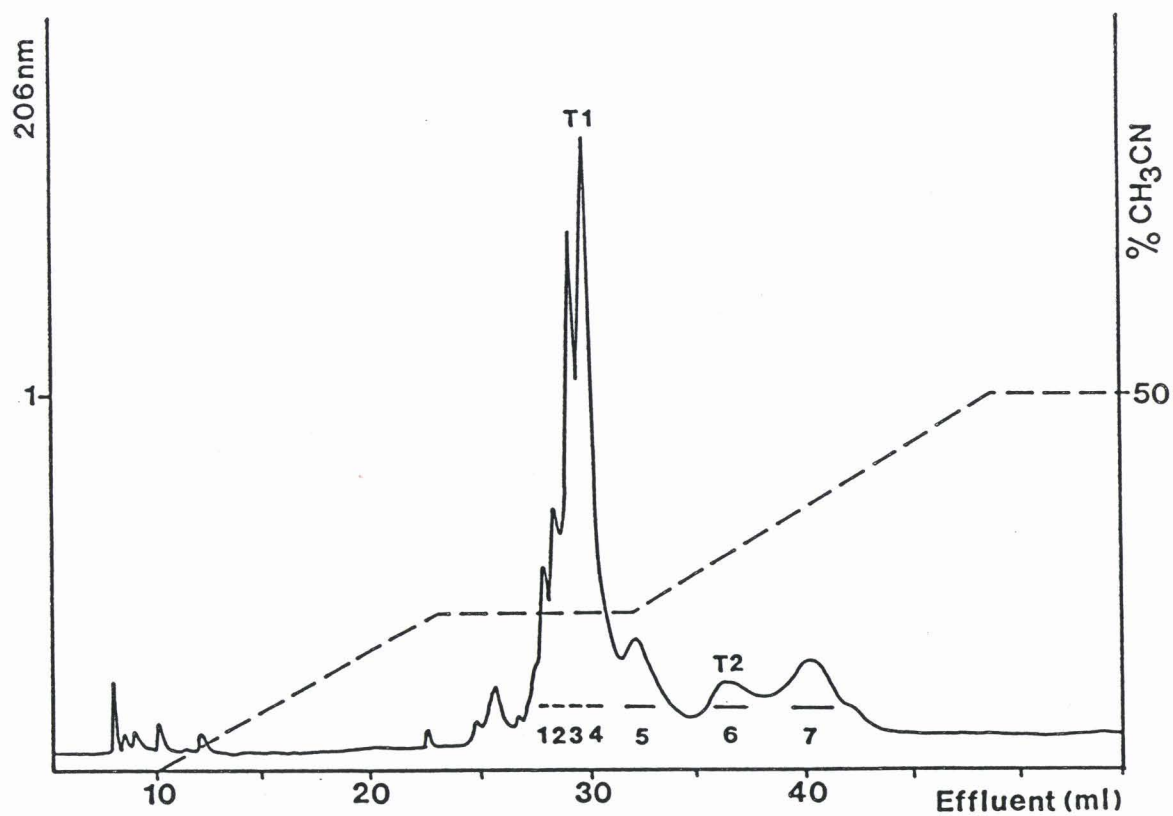


FIGURE 23: Séparation des variants de la protéine T sur colonne C₁₈ μ Bondapak. Les conditions de chromatographie sont décrites dans la légende de la figure 18. Les fractions 4 et 6 correspondent aux variants purs T1 et T2. Les autres fractions sont des mélanges de variants sous différentes formes phosphorylées.

TABLEAU III

Compositions en acides aminés des variants T1 et T2 obtenus après hydrolyse totale acide de 24 heures.

	Variant T1	Variant T2
Asp	1.3 (1)	2.0 (2)
Thr		0.8 (1)
Ser	6.3 (9)	5.6 (8)
Glu	2.4 (2)	1.5 (1)
Pro	1.0 (1)	1.0 (1)
Gly	2.4 (2)	2.2 (2)
Ala	3.3 (3)	2.4 (2)
Val	1.0 (1)	1.9 (2)
Met	1.2 (3)	0.8 (2)
Leu	3.1 (3)	3.8 (4)
Tyr	4.6 (5)	3.4 (5)
Lys	3.9 (4)	3.9 (4)
Arg	44.0 (44)	43.0 (43)
Total	78	77

Aucune correction n'a été faite pour les pertes occasionnées lors de l'hydrolyse. Les chiffres entre parenthèses correspondent aux nombres de résidus déduits de la séquence.

Ces résultats ont permis d'attribuer les substitutions en positions 7, 12 et 16 déjà observées à chacun des deux variants et d'en identifier une quatrième en position 18 où nous trouvons une méthionine dans le variant T1 et une leucine dans le variant T2. L'ambiguïté concernant le doublet Ser-Ser observé aux positions 34-35 dans la séquence réalisée sur le mélange des variants a pu être levée puisque chaque variant séquencé contient une seule sérine à ce niveau, le variant T1 en position 35 et le variant T2 en position 34 (Figure 20).

II - 1-4-3 Stratégie de coupure des variants

Les résultats décrits ci-dessus ayant permis d'établir sans ambiguïté la séquence N-terminale (résidus 1 à 50), le but de notre travail sur les variants séparés est de déterminer la séquence C-terminale riche en arginine de ces deux protéines.

Le choix des coupures enzymatiques est assez restreint puisque cette région de la protéine ne contient que quatre acides aminés constitutifs : Ser, Pro, Tyr, Arg. La stratégie utilisée a été inspirée de celle employée pour la protamine Sp (MARTIN-PONTHIEU *et al.*, 1990) qui présente aussi cette caractéristique :

- coupure par la chymotrypsine dans des conditions limitées pour réduire le nombre de peptides.
- coupure par la protéase d'*Astacus fluviatilis* pour obtenir des recouvrements.

II - 1-4-4 Hydrolyses enzymatiques des variants T1 et T2

* *Chymotrypsine*

Les diagrammes d'élution des hydrolysats chymotrypsiques des variants T1 et T2 sont présentés dans la figure 24. Les courbes sont très comparables et les analyses sur gel de polyacrylamide des fractions majeures ne permettent pas de déceler des différences notables entre les deux variants (Figure 25).

Les compositions en acides aminés des peptides chymotrypsiques des variants T1 et T2 sont présentées dans le tableau IV. Certains de ces peptides ont pu être replacés dans la séquence connue de la protéine uniquement d'après leur composition en acides aminés. Il s'agit des peptides C-1 et C-2 provenant respectivement des fractions 6 et 4 du diagramme d'élution (Figure 24).

- Un peptide majeur, le peptide C-2 est observé dans les fractions 4 et 5 de chaque hydrolysat. Les compositions en acides aminés des peptides C-2 des variants T1 et T2 montrent quelques différences au niveau des résidus d'acide aspartique et d'acide glutamique, ainsi que des résidus de leucine et de méthionine. Ceci montre nettement qu'ils proviennent de la zone amino-terminale de la protéine et qu'ils recouvrent les régions 16-43 du variant T1 et 16-47 du variant T2.

- Un peptide majeur est présent dans la fraction 6. Là encore, les compositions en acides aminés des peptides provenant des 2 variants T1 et T2 sont très caractéristiques. On observe des différences au niveau des résidus de thréonine, sérine, acide glutamique, alanine, valine, méthionine et de leucine. On peut donc affirmer qu'ils recouvrent les régions 1 à 18 du variant T1 et 1 à 15 du variant T2.

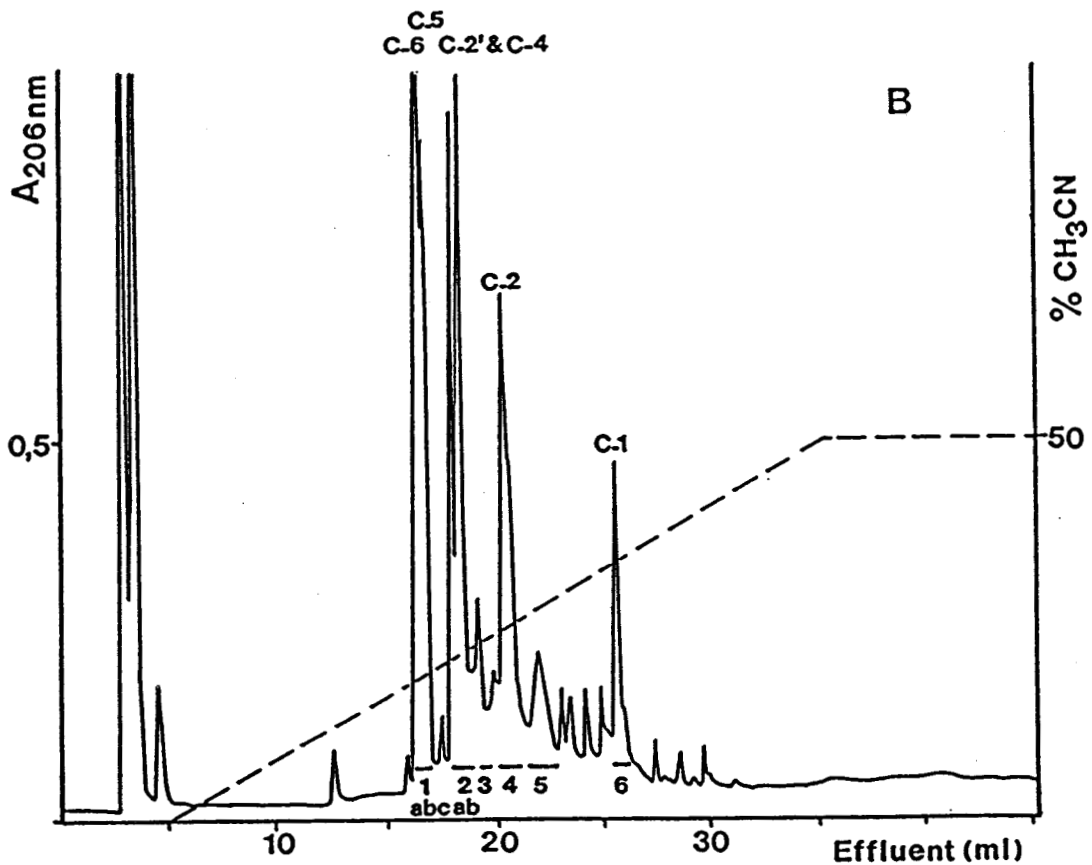
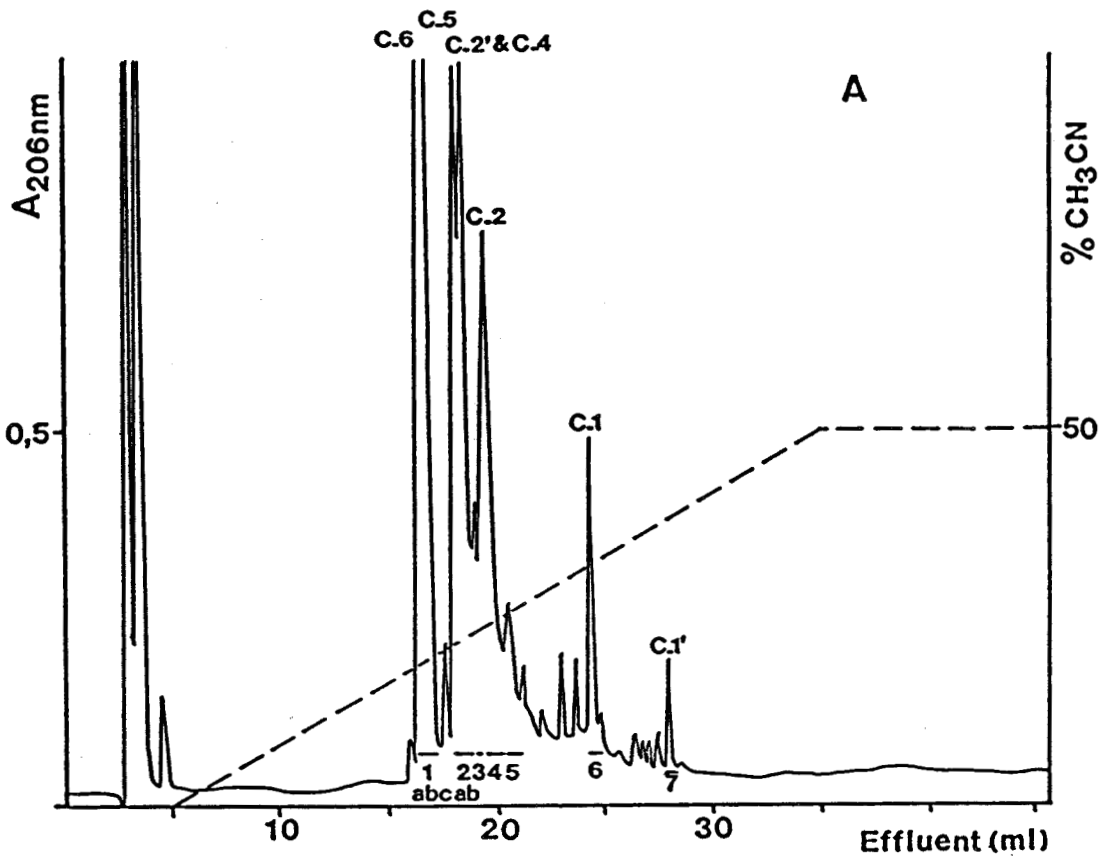


FIGURE 24: Séparation sur colonne C_{18} des peptides obtenus après action de la chymotrypsine à pH 5,0 sur les protéines T1 (A) et T2 (B). Les conditions de la chromatographie sont décrites dans la légende de la figure 18. Les peptides sont symbolisés par la lettre C et numérotés en fonction de leur position dans la séquence des protéines. L'analyse de ces peptides a aussi été effectuée par électrophorèse sur gel de polyacrylamide (cf. Fig. 25).

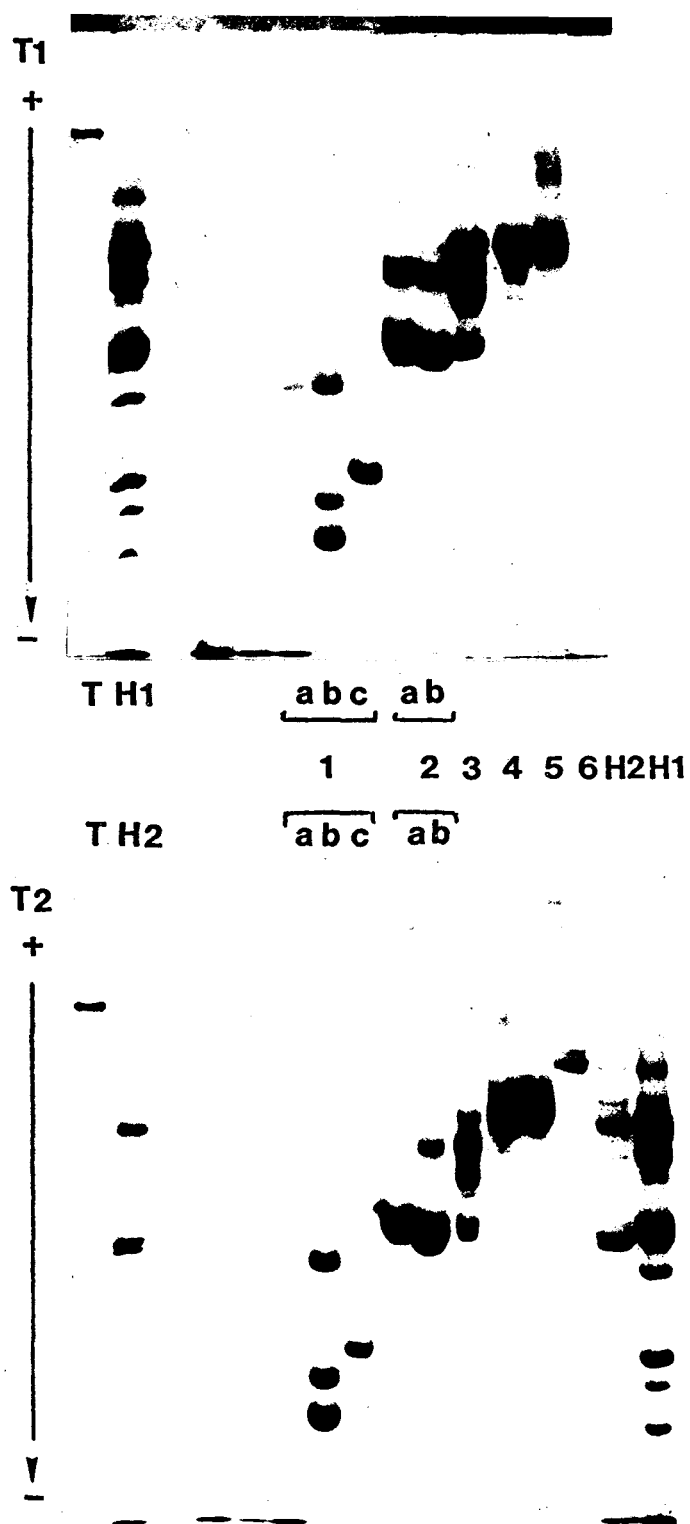


FIGURE 25: Analyse électrophorétique des peptides obtenus après action de la chymotrypsine à pH 5,0 sur les protéines T1 et T2. L'électrophorèse est réalisée sur un gel à 22 % de polyacrylamide, en milieu: acide acétique 0,9 M/urée 2,5 M. Après une pré-électrophorèse de 1h30, la migration des échantillons est réalisée pendant 2h sous un ampérage constant de 22 mA. Première piste: protéine T intacte, seconde piste: hydrolysat total de la protéine T1 en H1, T2 en H2. Les pistes 1a → 6 correspondent aux fractions de la figure 24.

TABLEAU IV-B

Compositions en acides aminés des peptides chymotrypsiques (C) après hydrolyse de 24 heures du variant T2.

	C-1 1-15	C-2 16-47	C-2' 19-43	C-4 48-62	C-5 63-68	C-6 69-77
Acide aspartique	1.2 (1)	0.9 (1)				
Thréonine	1.0 (1)					
Sérine	1.5 (1)	3.4 (4)	3.2 (4)	0.8 (1)	0.8 (1)	1.0 (1)
Acide glutamique	1.2 (1)					
Proline		1.0 (1)	0.9 (1)			
Glycine		2.1 (2)	2.2 (2)			
Alanine	1.8 (2)					
Valine	2.0 (2)					
Méthionine	1.7 (2)					
Leucine	2.3 (2)	2.2 (2)				
Tyrosine		1.4 (2)	0.8 (1)	1.2 (1)	1.0 (1)	1.0 (1)
Lysine	2.7 (3)	1.1 (1)	1.0 (1)			
Arginine		19.0 (19)	16.0 (16)	13.2 (13)	3.9 (4)	7.0 (7)
Total	15	32	25	15	6	9
N° des fractions	6	4	2a,2b	2a,2b	1c	1a

Les valeurs sont exprimées en moles/100 moles. Aucune correction n'a été faite pour les pertes occasionnées lors de l'hydrolyse. Les chiffres entre parenthèses correspondent aux nombres de résidus déduits des séquences.

- Les peptides présents dans les fractions 2a et 2b de chaque hydrolysats semblent homologues à la fois d'après leur migration électrophorétique et leur composition en acides aminés. Celle-ci montre nettement qu'ils proviennent de la région C-terminale de la protéine. Leur séquence a été déterminée par dégradation automatique d'Edman (Figure 26). Les fractions 2a et 2b du variant T1 renferment le même peptide majeur C-4 provenant de la région C-terminale (résidus 48-63) et le même peptide mineur C-2' provenant de la région N-terminale (résidus 19-43).

De la même façon, les fractions 2a et 2b du variant T2 renferment un peptide majeur C-4 (résidus 48 à 62) dont la séquence est cependant légèrement différente. Par contre, le peptide mineur C-2' recouvre la même zone (résidus 19-43).

* *Protéase d'Astacus fluviatilis*

Les diagrammes d'élution des peptides obtenus par hydrolyse des protéines T1 et T2 avec la protéase d'*Astacus fluviatilis* sont présentés dans la figure 27. Ils possèdent certaines similitudes mais aussi beaucoup de différences qui se manifestent dans le profil électrophorétique des hydrolysats totaux (Figure 27) particulièrement au niveau des peptides de mobilité électrophorétique élevée.

En ce qui concerne ces hydrolysats, nous ne présenterons que les peptides qui ont été nécessaires pour établir la séquence complète des variants T1 et T2, c'est-à-dire uniquement les peptides de la région C-terminale qui permettent d'aligner les peptides chymotrypsiques. La composition de ces peptides est présentée dans le tableau V.

Les peptides A-1 retrouvés dans chacun des deux variants sont d'une importance capitale. En effet, ils contiennent 3

T1 fraction 2a	Peptide C-2'	1 Lys-Gly-Gly-Arg-Arg-Arg-Arg-Arg-Arg-Ser-Arg-Arg-Arg-Arg-Arg-Arg- ...
	Peptide C-4	Arg-Arg-Arg-Arg-Arg-Arg-Arg-Arg-Arg-Arg-Ser-Arg-Arg-Arg-Arg- Tyr
T1 fraction 2b	Peptide C-2'	Lys-Gly-Gly-Arg-Arg-Arg-Arg-Arg-Arg-Ser-Arg- X -Arg-Arg-Arg-Arg- ...
	Peptide C-4	Arg-Arg-Arg-Arg-Arg-Arg-Arg-Arg-Arg-Arg-Ser-Arg-Arg-Arg-Arg- Tyr
T2 fraction 2a	Peptide C-2'	Lys-Gly-Gly-Arg-Arg-Arg-Arg-Arg-Arg-Ser-Arg-Arg- ...
	Peptide C-4	Arg-Arg-Arg-Arg-Arg-Arg-Arg-Arg-Ser-Arg-Arg-Arg-Arg-Arg-Tyr
T2 fraction 2b	Peptide C-2'	Lys-Gly-Gly-Arg-Arg-Arg-Arg-Arg-Arg-Ser-Arg-Arg-Arg-Arg-Arg-Ser- ...
	Peptide C-4	Arg-Arg-Arg-Arg-Arg-Arg-Arg-Arg-Ser-Arg-Arg-Arg-Arg-Arg-Tyr

FIGURE 26: Structure primaire des peptides C-2' et C-4 contenus dans les fractions 2a et 2b de l'hydriysat chymotrypsique de chaque variant T1 et T2 (cf. FIG. 24 et 25)

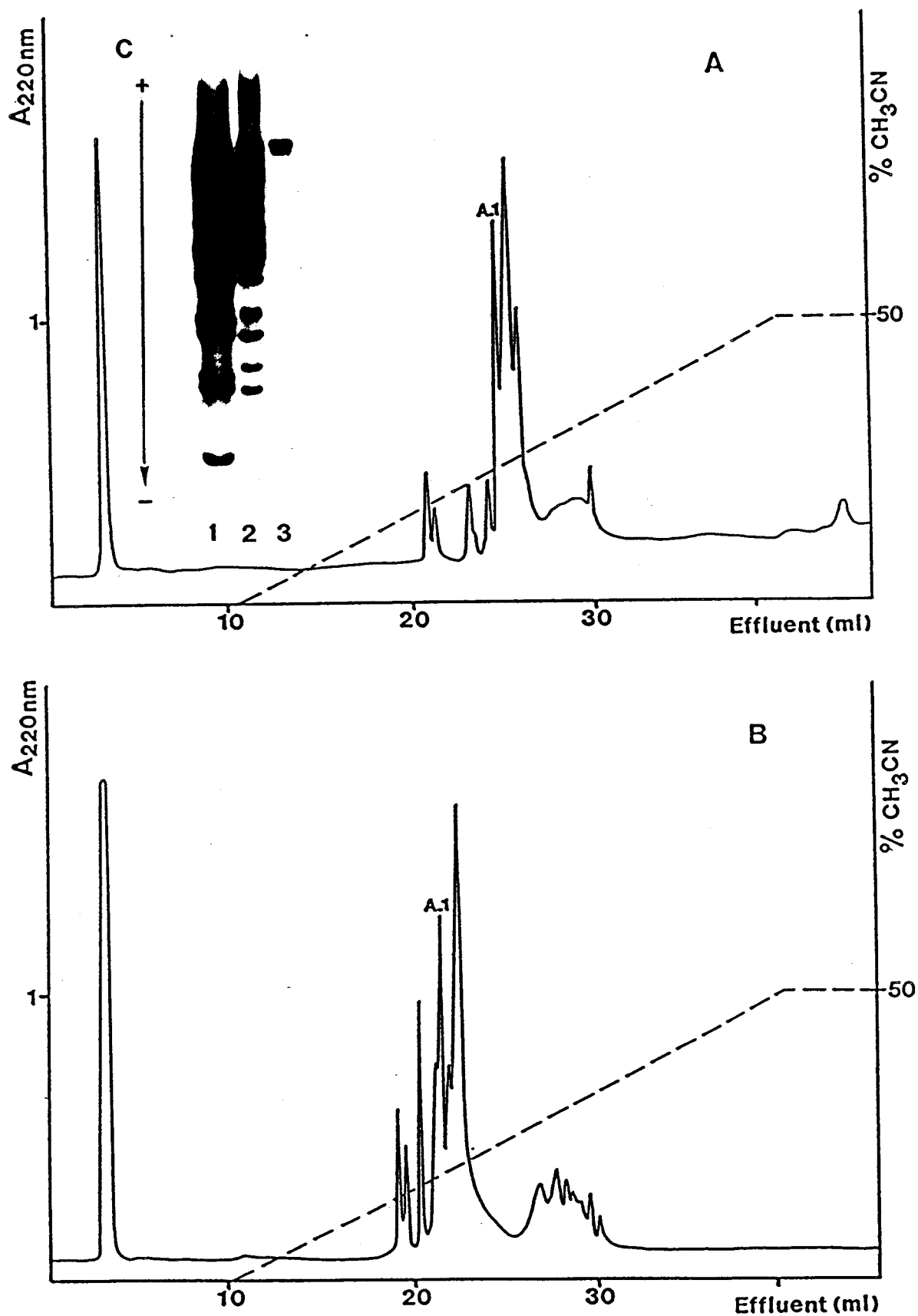


FIGURE 27: Séparation sur colonne C_{18} μ Bondapak des peptides obtenus par action de la protéase d'*Astacus fluviatilis* sur la protéine T1 (A) et T2 (B). Les conditions de la chromatographie sont indiquées dans la légende de la figure 18. C: Analyse électrophorétique des hydrolysats de la protéine T1 (piste 1) et T2 (piste 2) par action de la protéase d'*Astacus fluviatilis*. Les protéines T1 et T2 intactes sont présentées dans la piste 3. Les conditions électrophorétiques sont décrites dans la légende de la figure 25.

TABLEAU V

Composition en acides aminés des peptides A-1 obtenus à partir de l'hydrolyse par la protéase d'*Astacus fluviatilis* sur les variants T1 et T2

	Variant T1 A-1 58-78	Variant T2 A-1 56-77
Sérine	2.9 (3)	2.8 (3)
Tyrosine	2.6 (3)	2.8 (3)
Arginine	15.0 (15)	16.0 (16)
Total	21	22

Les valeurs sont exprimées en moles/100 moles. Les chiffres entre parenthèses correspondent aux nombres de résidus déduits de la séquence des peptides.

des 5 résidus de tyrosine présents dans chacun des variants. Ils ont tous deux été soumis à la dégradation automatique d'Edman. Les résultats de séquence sont présentés dans la figure 28.

Comme les deux premiers résidus de tyrosine ont déjà été placés dans la séquence des 50 premiers résidus, ce peptide ne peut se placer qu'à l'extrémité C-terminale de la molécule. Il contient aussi trois résidus de sérine, ce qui vient confirmer sa position C-terminale puisque les cinq ou six autres résidus de sérine (selon le variant) ont déjà été placés dans les 50 premiers résidus. Les séquences de ces deux peptides montrent que le peptide issu du variant T1 possède un résidu d'arginine de moins que celui du variant T2. Ces résultats sont confirmés par les masses des peptides A-1 déterminées par spectrométrie : pour A-1(T1), 3192.2 Da et pour A-1(T2), 3348.8 Da, ce qui correspond à la masse calculée pour ces peptides si l'on considère qu'ils possèdent chacun un groupement phosphate.

II - 1-4-5 Séquences totales des protéines T1 et T2

Les séquences en acides aminés des protéines T1 et T2 ont été établies par séquençage des protéines intactes et de leurs fragments obtenus par clivage enzymatique. Elles sont présentées dans les figures 29 A et B.

* Variant T1

- Le séquençage de la protéine intacte a permis d'identifier les 35 premiers résidus de la molécule.

- Les peptides C-2 et C-2' qui recouvrent cette région N-terminale permettent de prolonger la séquence jusqu'au résidu de tyrosine en position 43.

- Nous pouvons d'après les données acquises sur la protéine T totale compléter cette séquence jusqu'au résidu d'arginine en position 50.

T1 A-1 Ser-Arg-Arg-Arg-Arg-Tyr-Arg-Arg-Arg-Arg-Ser-Tyr-Ser-Arg-Arg-Arg-Tyr-Arg-Arg-Arg-Arg 1 5 10 15 20

T2 A-1 Ser-Arg-Arg-Arg-Arg-Tyr-Arg-Arg-Arg-Arg-Ser-Tyr-Ser-Arg-Arg-Arg-Tyr-Arg-Arg-Arg-Arg 1 5 10 15 20

FIGURE 28: Séquences déterminées par la dégradation d'Edman des peptides A-1 obtenus après hydrolyse des variants T1 et T2 par la protéase d'*Astacus fluviatilis*.

- Le peptide C-4 ne peut être placé qu'après un résidu de tyrosine. Il est recouvert lui-même dans sa région C-terminale par le peptide A-1 qui renferme les trois derniers résidus de tyrosine de la protéine. Nous pouvons donc placer ce peptide C-4 après la tyrosine en position 47.

- Les peptides C-5 et C-6 peuvent être alignés grâce au peptide A-1 et viennent donc confirmer la séquence C-terminale.

La protéine T1 est une molécule de 78 résidus. La masse déterminée par spectrométrie (10 872 Da) est différente de la masse calculée d'après la séquence en acides aminés (10 632.5 Da). Toutefois, cela peut s'expliquer par le fait que la protéine est phosphorylée ; la différence (239.5 Da) correspondrait à la présence de trois groupements phosphate.

** Variant T2*

- Le séquençage de la protéine intacte a permis d'identifier les 42 premiers résidus de la molécule.

- Les peptides C-2 et C-2' qui recouvrent cette région N-terminale permettent de prolonger la séquence jusqu'au résidu 47.

- Pour la région C-terminale de la protéine, le raisonnement est le même que pour la variant T1. Le peptide A-1 recouvre la région C-terminale du peptide C-4 et on peut placer les 30 derniers résidus de la protéine. Cette séquence C-terminale ne peut se situer qu'après le résidu de tyrosine en position 47, en tenant compte du nombre total de résidus de la protéine, et du nombre de résidus de tyrosine et de sérine.

- Les peptides C-5 et C-6 viennent s'aligner grâce au peptide A-1.

La protéine T2 est une protéine de 77 résidus, elle contient donc un résidu d'arginine de moins que la protéine T1. La masse calculée d'après la séquence en acides aminés est de 10 486.3 Da; celle-ci n'a pas pu être confirmée par spectrométrie de masse.

La comparaison des séquences des protéines T1 et T2 est présentée dans la figure 29 C.

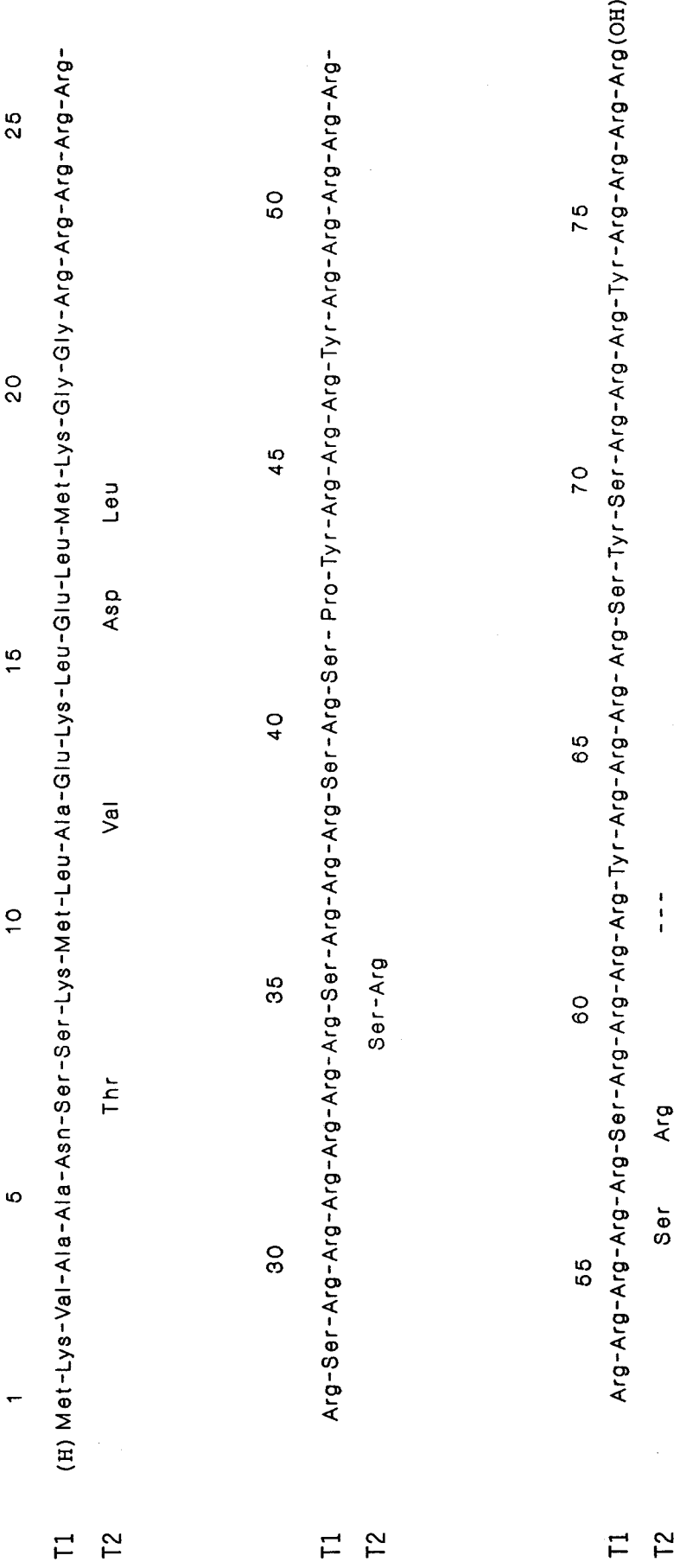


FIGURE 29 C: Comparaison des structures primaires des protéines T1 et T2. Pour la protéine T2, seules les acides aminés différents ont été indiqués. Les deux protéines ont été alignées de façon à avoir un maximum d'homologie, il a été nécessaire de faire apparaître une délétion représentée par (---).

La protéine T spécifique des spermatides de seiche est donc en fait, un mélange de deux variants structuraux majeurs T1 et T2 de 78 et 77 résidus respectivement, et de deux variants structuraux mineurs T1a et T2a de 77 et 76 résidus respectivement. Ces variants mineurs sont dépourvus de méthionine en position N-terminale. Les proportions relatives de ces variants ont pu être établies à partir des données de la dégradation d'Edman des variants et de leurs peptides N-terminaux :

T1 : 60,6 % ; T1a : 6,1 %
T2 : 31,7 % ; T2a : 1,6 %

** Comparaison de la protéine T avec d'autres protéines spécifiques des spermatides*

Lorsque l'on compare la protéine T de seiche à d'autres protéines spécifiques des spermatides (roussette, souris, rat, homme), on peut dégager les caractéristiques récapitulées dans le Tableau VI.

La protéine T de seiche a une taille similaire à celle des protéines spécifiques des spermatides de roussette (environ 80 résidus) et inférieure à celle des protéines de type p-P2 (Pré-protamine P2) et TP2 des mammifères (plus de 100 résidus)

Elle possède une proportion de résidus basiques plus proche de celle observée dans les protéines de type p-P2 que dans les protéines TP1 et TP2 de mammifères. Celles-ci sont plus riches en lysine, comme les protéines S1 et S2 de roussette.

La protéine T de seiche ne renferme ni cystéine, ni histidine ce qui la différencie des autres protéines spécifiques des spermatides. En effet, la présence de cystéine n'est observée que dans les protéines de type P2 et TP2 de mammifères, alors que l'histidine est présente dans toutes les autres espèces citées.

TABLEAU VI

Comparaison de la protéine T de seiche avec d'autres protéines spécifiques des spermatides.

	Seiche	Roussette		Souris (Rat)			Homme
	T	S1	S2	TP1	TP2	pP2	HPI1
Nombre de résidus	78	80	87	54	117	106	101
% Arginine	58,7	33,1	15,8	20,4	14,2	42,7	35,3
% Lysine	5	9,6	20	18,5	9,9	3,5	1,6
Rapport Lysine/Arginine	0,08	0,29	1,26	0,90	0,70	0,08	0,04
% Cystéine	0	0	0	0	4,4	6,8	6,1
% Histidine	0	3,4	2,2	5,5	8,3	12,2	10,9
Nombre et position des acides aminés dicarboxyliques	2 N-t	4 C-t	4 C-t	2 C-t	1 C-t 1 N-t	5 N-t	8 N-t

N-t = extrémité N-terminale ; C-t = extrémité C-terminale.

La protéine intermédiaire de seiche ne renferme que deux résidus acides dans la région N-terminale (en positions 13 et 16). Chez la roussette, mais surtout chez les mammifères dans les protéines de type p-P2, on trouve des centres acides plus importants.

On peut noter aussi une asymétrie plus marquée de la protéine T de seiche par rapport aux autres espèces. En effet, la région N-terminale de cette protéine (résidus 1-21) ne possède aucun résidu d'arginine. Alors que dans les protéines de type p-P2 et TP2 de mammifères, les résidus d'arginine sont répartis tout au long de la séquence, même si des multiplets d'arginine sont situés, comme pour la seiche, dans la région C-terminale. Par contre, dans les protéines S1 et S2 de roussette, les centres basiques sont plus nombreux dans la région N-terminale.

En conclusion, la protéine spécifique des spermatides de la seiche est tout à fait particulière. Si l'on considère l'ensemble des caractères rassemblés dans le Tableau VI, elle ne peut se rattacher à aucun des autres types de protéines intermédiaires. Par certains critères (rapport Lys/Arg identique, distribution des résidus d'arginine), elle se rapproche des protéines de type p-P2 ; par d'autres (taille, absence de cystéine), elle se rapproche de la protéine S1. Par contre, elle présente très peu de similitudes avec les protéines de type TP2.

II - 1-5 PREDICTIONS DE STRUCTURE SECONDAIRE DE LA PROTEINE T

De nombreuses expériences de dénaturation et de renaturation des protéines (ANFINSEN *et al.*, 1972 ; IKAI *et al.*, 1973) ont montré que beaucoup de protéines sont capables de se replier dans leur conformation native après élimination de l'agent dénaturant, parfois même assez rapidement (GAREL et BALDWIN, 1973). Ces expériences indiquent que l'information nécessaire au repliement d'une protéine est contenue principalement dans sa séquence polypeptidique.

Différentes méthodes de prédictions de structures secondaires ont été développées ces dernières années. Elles permettent d'appréhender la structure secondaire d'une protéine. Cependant la structure tertiaire ne reste abordable que par les études de diffraction des rayons X à la condition préliminaire d'obtenir des cristaux de protéines (qui doivent en outre être de bonne qualité).

Afin d'estimer la structure secondaire probable de la protéine T, nous avons appliqué différentes méthodes de prédictions de structures. Ces méthodes peuvent être réparties en trois classes : les méthodes mathématiques, les méthodes de comparaison de structures et les méthodes graphiques.

* *Méthodes mathématiques*

Les méthodes mathématiques, de loin les plus nombreuses, attribuent une valeur de statistique à chaque acide aminé qui correspond à la probabilité pour que cet acide aminé soit dans une conformation précise. Ces méthodes tiennent compte des acides aminés voisins et calculent la probabilité de chacun des états conformationnels pour l'acide aminé considéré. Le dépouillement des résultats est parfois délicat car il arrive

que les valeurs des probabilités des différents états soient très proches les unes des autres pour un même acide aminé. Les méthodes de cette classe que nous avons appliquées sont : les méthodes de CHOU et FASMAN (1978), de GARNIER *et al.* (1978), de GIBRAT *et al.* (1987) et de ROBSON (1971).

* *Méthodes de comparaison de structure*

La méthode de comparaison de structure que nous avons utilisée est la méthode de LEVIN *et al.* (1986). Le principe de cette méthode consiste à dire que de courts segments de protéines de structures primaires identiques ont des structures secondaires identiques. Il y a donc un découpage de la protéine à prédire en heptapeptides successifs décalés d'un résidu à chaque fois et comparaison de leur structure avec des structures contenues dans une banque de données. La banque de données renferme des protéines (62) dont la structure tridimensionnelle a été établie par cristallographie.

* *Méthodes graphiques*

LIM (1974) est à l'origine des méthodes graphiques de prédiction de structures secondaires. L'hypothèse de travail étant que lors de sa synthèse au niveau du ribosome le polypeptide se trouve sous forme d'une hélice α et qu'elle reste sous cette forme si les éléments nécessaires à sa stabilisation sont présents. Cette méthode purement graphique indique les régions susceptibles de se trouver en hélice α et en feuillet β selon la forme des amas hydrophobes d'où son nom de méthode HCA (Hydrophobic Cluster Analysis) donné par GABORIAUD *et al.* (1987). A notre connaissance, c'est la seule méthode qui tienne compte des interactions des acides aminés dans l'espace.

Etant donnée la nature de la protéine T, cette méthode n'a pu fournir d'information que pour la région 1-19. En effet, la séquence après le résidu 19 ne renferme pas d'acides aminés hydrophobes en quantité suffisante pour former des amas.

Il faut cependant noter que prises séparément ces méthodes ont une fiabilité qui ne dépasse que rarement 60%, il est donc indispensable de ne pas se contenter d'une seule méthode. La synthèse des résultats obtenus avec les différentes méthodes est présentée dans la figure 30. Seules les structures prédites par au moins trois des méthodes ont été prises en considération.

Les différentes méthodes sont en accord pour indiquer que la protéine serait essentiellement constituée de structure aperiodique. Par contre, elles sont en désaccord en ce qui concerne la structure de la région 1-19 de la protéine et il ne nous pas été possible de conclure. De plus, les probabilités de structure pour la région 19-fin sont bien souvent très faibles et voisines pour les différents états (la plus forte probabilité semble indiquer des coudes β successifs). Aussi, nous n'avons fait apparaître de structure que lorsque les valeurs de probabilité étaient supérieures à ce qui pouvait être considéré comme "un bruit de fond".

La protéine T renfermerait un fragment d'hélice α (séquence 10-18), un court fragment de feuillet β au voisinage de la tyrosine 47 et quatre coudes β au voisinage des positions 20, 39, 43 et 68.

Il est à noter que les glycines en positions 20 et 21 seraient impliquées dans un coude β juste après l'hélice α et que cette organisation semble identique à celle du site de coupure des précurseurs d'hormones (RHOLAM *et al.*, 1986).

II - 1-6 COMPARAISON DES SEQUENCES DES PROTEINES T ET Sp

II - 1-6-1 Comparaison des cartes peptidiques

La comparaison des compositions en acides aminés des deux protéines spécifiques de la spermiogenèse chez la seiche (protéines T et Sp) (WOUTERS-TYROU *et al.*, 1988) montre qu'elles renferment le même nombre de résidus de proline, de tyrosine, d'arginine et un nombre de résidus de sérine peu différent. Cela suggère qu'il existe entre les protéines T et Sp une analogie structurale importante.

Les premiers résultats de séquence obtenus sur la protéine T totale, nous avaient permis de mettre en évidence une remarquable similitude structurale entre la région 22-39 de la protéine T et la séquence N-terminale de la protamine Sp (résidus 1 à 18). De plus, le peptide C-terminal de la protéine T (résidus 22 à la fin) obtenu par action de l'endoprotéase Lys-C sur la protéine totale a une composition en acides aminés très proche de celle de la protamine Sp (Voir tableau I).

Dans le but de vérifier l'homologie structurale entre les deux protéines, nous avons d'abord réalisé une étude comparative des peptides obtenus par action de la chymotrypsine à pH 5.0 et de l'élastase, d'une part sur la protéine Sp isolée du spermatozoïde de la seiche (WOUTERS-TYROU *et al.*, 1988) et d'autre part sur le peptide C-terminal de la protéine T obtenu par coupure par l'endoprotéase Lys-C de la protéine déphosphorylée.

Les deux types d'hydrolysats ont été effectués dans les mêmes conditions pour chaque protéine (tampon, pH, température, rapport E/S), puis ils ont été déposés sur gel de polyacrylamide (concentration 22 %) en milieu acide-urée. Les résultats obtenus sont présentés dans la figure 31.

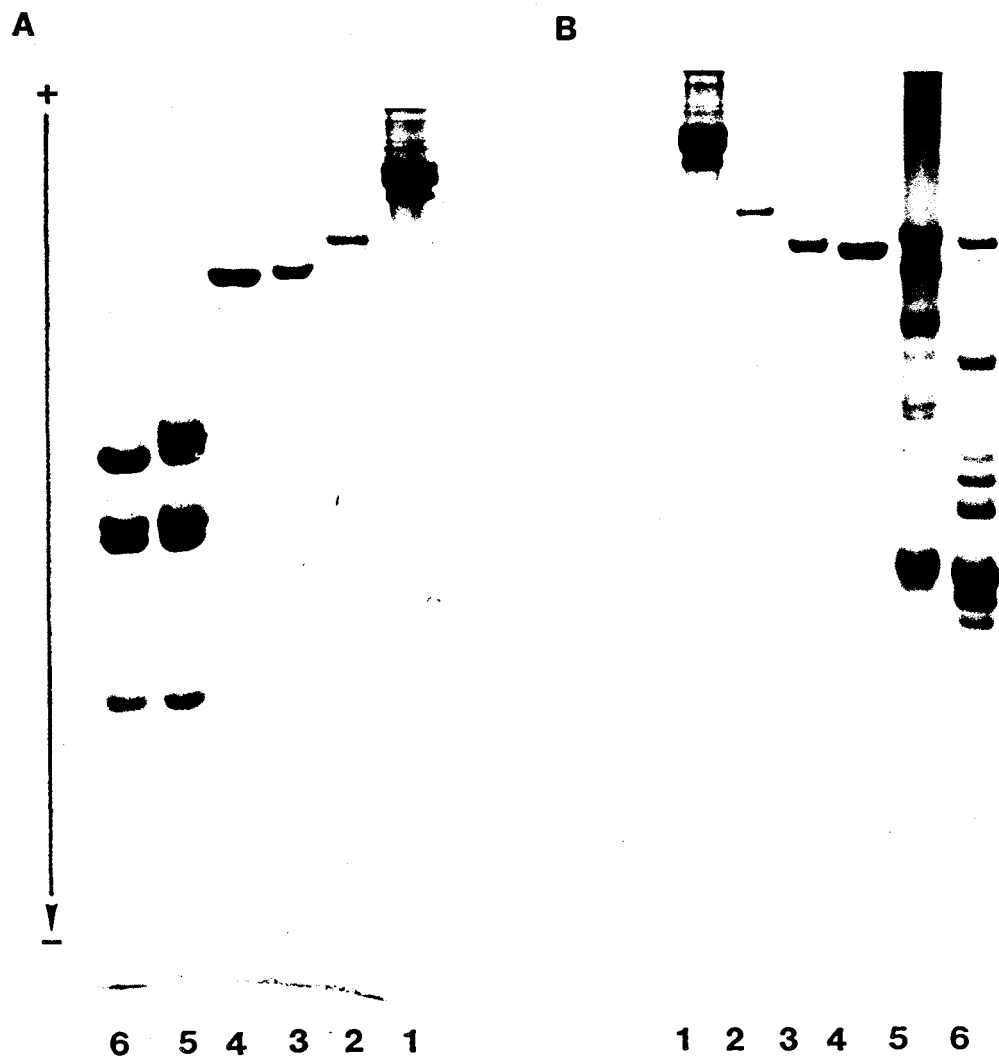


FIGURE 31 - Cartes peptidiques de la protamine Sp et du peptide K4 (Figures 20 et 21) de la protéine T de seiche.

A : après coupure par l'élastase.

B : après coupure par la chymotrypsine à pH 5.

1 : témoin d'histones totales de thymus de veau.

2 : témoin de protéine T.

3 : peptide C-terminal obtenu par coupure par l'endoprotéinase Lys-C de la protéine T déphosphorylée (K4).

4 : témoin de protamine Sp de sperme.

5 : hydrolysats du peptide C-terminal de la protéine T.

6 : hydrolysats de la protamine Sp.

Les conditions électrophorétiques sont décrites dans la légende de la figure 25.

- La séparation des peptides chymotrypsiques est très semblable dans les deux cas. Seul le peptide le plus lent présente une migration légèrement différente qui s'explique par la présence de deux résidus de glycine supplémentaires dans le peptide issu de la protéine T.

- La comparaison des peptides obtenus par action de l'élastase est moins spectaculaire. Cependant la majorité des bandes est commune aux deux hydrolysats, seule l'intensité est différente. Cela pourrait s'expliquer par une hydrolyse incomplète de la protéine et par le nombre plus important de coupures partielles.

En conclusion, l'analyse des deux hydrolysats montre un certain lien de parenté entre les protéines, celui-ci est très net pour l'hydrolysat chymotrypsique, un peu moins pour la coupure par l'élastase.

II - 1-6-2 Comparaison des séquences

La comparaison des séquences en acides aminés de la protéine T et de la protamine Sp vient renforcer l'idée de "filiation" entre les deux protéines. En effet, deux variants Sp1 et Sp2 ont pu être identifiés (MARTIN-PONTHIEU *et al.*, soumis pour publication). Ils ont une séquence en acides aminés semblable à la séquence C-terminale des variants T1 et T2 (Figures 32 A et 32 B).

Ceci permet de penser que la protéine T1 serait le précurseur potentiel de la protamine Sp1, alors que la protéine T2 serait le précurseur potentiel de la protamine Sp2.

T2	1	Met-Lys-Val-Ala-Ala-Asn-Thr-Ser-Lys-Met-Leu-Val-Glu-Lys-Leu-Asp-Leu-Leu-Lys-Gly-Gly-	10	15	20
T2	25	Arg-Arg-Arg-Arg-Ser-Arg-Arg-Arg-Ser-Arg-Arg-Arg-Ser-Arg-Arg-Ser-Arg-Ser-Pro-	30	35	40
Sp2	1	Arg-Arg-Arg-Arg-Ser-Arg-Arg-Arg-Ser-Arg-Arg-Arg-Ser-Arg-Ser-Arg-Ser-Pro-	5	10	15
					20
T2	45	Tyr-Arg-Arg-Arg-Tyr-Arg-Arg-Arg-Arg-Arg-Ser-Arg-Ser-Arg-Arg-Arg-Tyr-Arg-	50	55	60
Sp2	25	Tyr-Arg-Arg-Tyr-Arg-Arg-Arg-Arg-Arg-Ser-Arg-Arg-Arg-Arg-Tyr-Arg-	30	35	40
T2	65	Arg-Arg-Arg-Ser-Tyr-Ser-Arg-Arg-Arg-Tyr-Arg-Arg-Arg-Arg	70		
Sp2	45	Arg-Arg-Arg-Ser-Tyr-Ser-Arg-Arg-Arg-Tyr-Arg-Arg-Arg-Arg	50	55	

FIGURE 32 B: Comparaison de la séquence de la protéine T2 et de la protamine Sp2.
La séquence de la protamine Sp2 est de MARTIN-PONTHIEU *et al.* (1990).

Le mécanisme réel de cette "transition" protéique déjà décrit chez les protamines de type P2 des mammifères est encore inconnu. Plusieurs possibilités existent :

1. Il s'agirait d'un phénomène de protéolyse limitée, comme cela existe lors de la maturation des hormones polypeptidiques (RHOLAM *et al.*, 1986). Dans ce cas, la coupure protéolytique aurait lieu dans un site caractérisé par la présence d'acides aminés basiques adjacents à la glycine, et localisé dans un coude β . Les résultats de prédiction de structure secondaire font effectivement apparaître un coude β potentiel dans la séquence MKGGR (résidus 18 à 22).

Ce site serait alors identique au site de coupure des protéines pVI et pVII de l'*adenovirus* de type 2 $\begin{matrix} M \\ L \end{matrix} XGG \downarrow X$ (WEBSTER *et al.*, 1989).

Il reste toutefois à isoler la ou les protéases qui seraient responsables de cette protéolyse spécifique.

2. L'apparition successive des protéines T puis Sp résulte d'un mécanisme de régulation de l'expression des gènes.

Afin d'apporter des éléments de réponse en faveur de l'une ou l'autre de ces hypothèses de mécanisme de transition, nous avons entrepris l'étude de l'expression des acides nucléiques du testicule de seiche et notamment la caractérisation des ARN messagers présentée dans la deuxième partie du chapitre.

II - 2 APPROCHE GÉNÉTIQUE DE LA TRANSITION PROTÉINE T → PROTAMINE SP

Au niveau génétique, nous ne pouvions exclure la possibilité de gènes différents pour les protéines T et Sp. D'autre part, la présence d'un gène unique pour ces deux protéines n'exclut pas :

- l'existence de plusieurs ARNm directement dérivés de ce gène
- ou même l'existence d'un seul ARNm initial qui, par maturation différentielle par épissage, aboutit à des ARNm différents.

Pour entreprendre cette étude, il a été nécessaire d'introduire les techniques de biologie moléculaire au sein de notre laboratoire.

La stratégie que nous avons utilisée a consisté à extraire les ARN messagers, à identifier et à enrichir les fractions d'ARN codant pour la protéine T, puis à obtenir à partir de ces fractions, une banque d'ADN complémentaire. Cette banque d'ADNc devant servir de base à des travaux ultérieurs dont l'objectif final est l'étude des gènes des protéines T et/ou Sp, et de leurs régulations.

Nous avons pu effectuer les premières étapes de ce travail jusqu'à l'élaboration de la banque d'ADN complémentaire, de son criblage et du séquençage d'un ADNc correspondant à une partie de la protéine T2.

II - 2-1 EXTRACTION ET PURIFICATION DES ARN MESSAGERS

L'isolement des ARN messagers se fait en deux étapes, la première consiste à extraire les ARN cytoplasmiques à partir du testicule, la seconde consiste à isoler les ARNm sur colonne d'oligo (dT)-cellulose.

Au cours de l'extraction des ARN cytoplasmiques, l'intégrité des noyaux a été contrôlée par microscopie afin de vérifier l'absence de toute contamination de ces ARN par le matériel nucléaire.

Par la technique utilisée, nous avons obtenu 16 mg d'ARN totaux à partir de 7 g de testicule. Puis les ARN totaux ont été séparés en deux fractions :

- la fraction des ARN dépourvus de queue poly A (ARNs de transfert, ribosomiaux et certains messagers)
- la fraction des ARN messagers dotés d'une queue poly A. Cette dernière fraction a été obtenue après deux passages successifs sur colonne d'oligo (dT)-cellulose.

Les ARN messagers poly A+ représentent environ 1,2 % des ARN cytoplasmiques totaux. Cette valeur est en accord avec celle obtenue à partir du testicule de roussette par BERLOT-PICARD *et al.* (1986).

II - 2-2 TRADUCTION *IN VITRO* DES ARN CYTOPLASMIQUES

L'analyse des différentes fractions d'ARNm a été réalisée par traduction en système de lysat de réticulocyte de lapin. Après traduction, nous ajoutons dans les produits et dans les témoins de traduction un mélange de protéines entraînées

afin de limiter les pertes au cours du traitement de l'échantillon tel qu'il est décrit dans l'appendice technique.

ELSEVIER (1982) et YELICK *et al.* (1987) ont montré que l'addition de ces protéines entraînées ne provoquait aucun artefact de migration électrophorétique sur gel de polyacrylamide.

Pour vérifier l'efficacité du système de traduction, nous utilisons un témoin d'ARNm de "B-M-V" (Virus de la mosaïque du brome). En outre un témoin de traduction sans adjonction d'ARN, nous permet d'obtenir la traduction des ARNm endogènes du lysat qui ne sont pas totalement éliminés par la nucléase micrococciale. Ce dernier témoin permet d'identifier après électrophorèse sur gels de polyacrylamide (gels en milieu acide-urée ou en présence de SDS) certaines bandes contaminantes liées au système de traduction notamment au niveau des globines (pour exemple, se reporter à la figure 35).

L'analyse électrophorétique des protéines acido-solubles traduites à partir des ARN totaux (Figure 33) révèle des protéines dont la migration électrophorétique est identique à celle de la protéine T et à celle de la protéine Sp. Ces deux protéines représentent respectivement 14 et 12 % de la radioactivité incorporée dans les protéines acido-solubles.

Les traductions effectuées à partir des ARN poly A+ nous permettent de conclure que les ARNm de la protéine T de seiche possèdent une queue poly A+ comme les messagers des protéines intermédiaires de la spermiogenèse caractérisées chez le rat et la souris.

Il faut remarquer que la protamine Sp ne renferme pas de méthionine (WOUTERS-TYROU *et al.*, 1988). La présence d'une bande radioactive à ce niveau ne peut donc se concevoir que selon deux hypothèses :

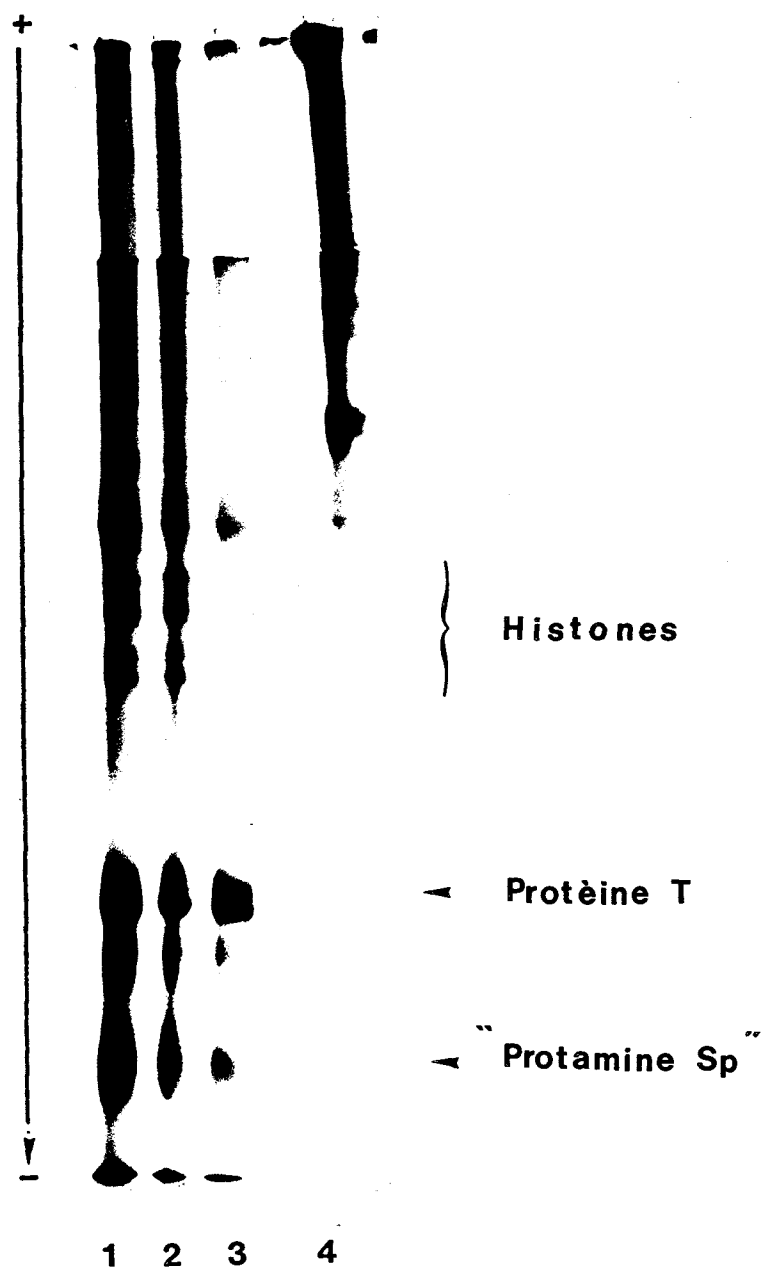


FIGURE 33: Analyse électrophorétique sur gel de polyacrylamide à 14% en milieu acide/urée des produits de traduction acido-solubles des ARN totaux (pistes 1 et 2) et des ARN poly A⁺ (piste 3), et des ARN du virus de la mosaïque du brome (piste 4). Après coloration et séchage, le gel est mis en autoradiographie.

- Il s'agit d'une autre protéine de migration électrophorétique identique à celle de la protamine Sp.
- la méthionine initiatrice n'est pas excisée dans le système de traduction utilisé, comme cela a été observé chez la truite (LOUIE et DIXON, 1972).

La présence d'une méthionine initiatrice non excisée semble être confirmée par une importante diminution de l'intensité de cette bande radio-marquée après action de l'aminopeptidase M pendant un temps court, directement sur le gel. Un enregistrement densitométrique effectué sur l'autoradiographie a permis de montrer que les surfaces des pics de la protéine T et de la "protéine Sp" diminuaient dans des proportions identiques. Il est évident que cette manipulation n'apporte rien sur la nature même de la protéine suspectée comme étant la protamine Sp.

II - 2-3 ENRICHISSEMENT DES ARNm DE LA PROTEINE T

L'ultracentrifugation sur gradient de saccharose des ARNm poly A⁺ a été utilisée pour enrichir une fraction en ARN codant pour la protéine T. Les meilleurs résultats ont été obtenus sur gradient de 10 à 40 % de saccharose. Les gradients ont été étalonnés avec des protéines de coefficients de sédimentation connus. Ces protéines sont localisées par leur activité enzymatique et nous avons contrôlé que la présence des ARNm dans le gradient ne modifiait pas cet étalonnage.

L'enregistrement densitométrique à 258 nm d'un gradient type est présenté dans la figure 34.

Grâce à la reproductibilité des résultats obtenus, les fractions de plusieurs gradients peuvent être rassemblées.

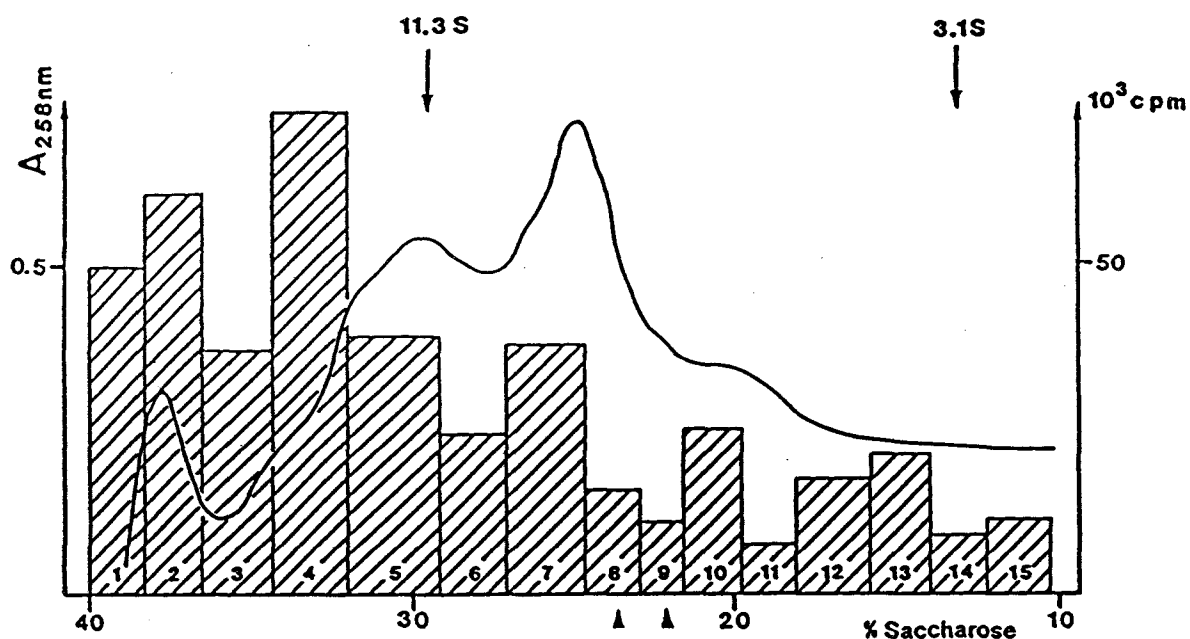


FIGURE 34: Enrichissement des ARN poly A⁺ par ultra-centrifugation sur gradient de saccharose 10 à 40 %. 150 µg d'ARNm sont déposés sur le gradient et centrifugés à 40 000 rpm (SW41 Beckman, 14h). Le gradient est prélevé par le bas du tube et l'absorbance à 258 nm est suivie en continu (—). 3 µg d'ARNm de chaque fraction sont traduits en lysat de réticulocyte de lapin et la radioactivité incorporée est mesurée (▨). Des protéines de constantes de sédimentation connues ont servi à étalonner les gradients (catalase: 11,3 S et phosphoglycérate kinase: 3,1 S). Les fractions collectées sont numérotées de 1 à 15 depuis le bas du tube.

Les produits de traduction obtenus en système de lysat de réticulocyte de lapin des différentes fractions de gradients ont été analysés par électrophorèse en gel de polyacrylamide (Figure 35). Le gel de polyacrylamide en milieu acide-urée montre la répartition des ARNm de la protéine T dans le gradient tandis que le gel de polyacrylamide en milieu SDS indique la répartition des ARNm des autres protéines non acido-solubles.

Cette technique nous a permis d'obtenir des fractions enrichies en ARNm de la protéine T (fractions 8 et 9 de la figure 35-A) dans lesquelles on a également une forte diminution des ARNm codant pour les protéines non acido-solubles (fractions 8 et 9, figure 35-B) Ces ARNm auraient une constante de sédimentation d'environ 6 à 7S.

Nous pouvons remarquer la présence de deux variants de protéine T tout au long du gradient de saccharose. Ce phénomène d'étalement des ARNm dans un gradient de saccharose a déjà été signalé pour les ARNm de protamines de truite, dans des conditions similaires (GEDAMU et DIXON, 1976).

D'autre part, nous devons noter la présence de deux protéines coenrichies avec la protéine T dont l'une a une migration électrophorétique semblable à celle de la protamine Sp. La seconde protéine est mineure en intensité et présente une migration électrophorétique intermédiaire entre celle de la protéine T et de la protamine Sp. Une protéine de migration électrophorétique identique a été caractérisée dans un extrait HCl 0.4 N de spermatozoïdes de seiche. Cette protéine renferme 22 % d'arginine, 11 % d'acide aspartique et de sérine et environ 4 % de cystéine (Tableau VII).

La chute importante de la quantité de radioactivité incorporée dans ces fractions (pistes 8 et 9 de la figure 34) peut certainement être attribuée à la faible teneur en méthionine des protéines traduites.

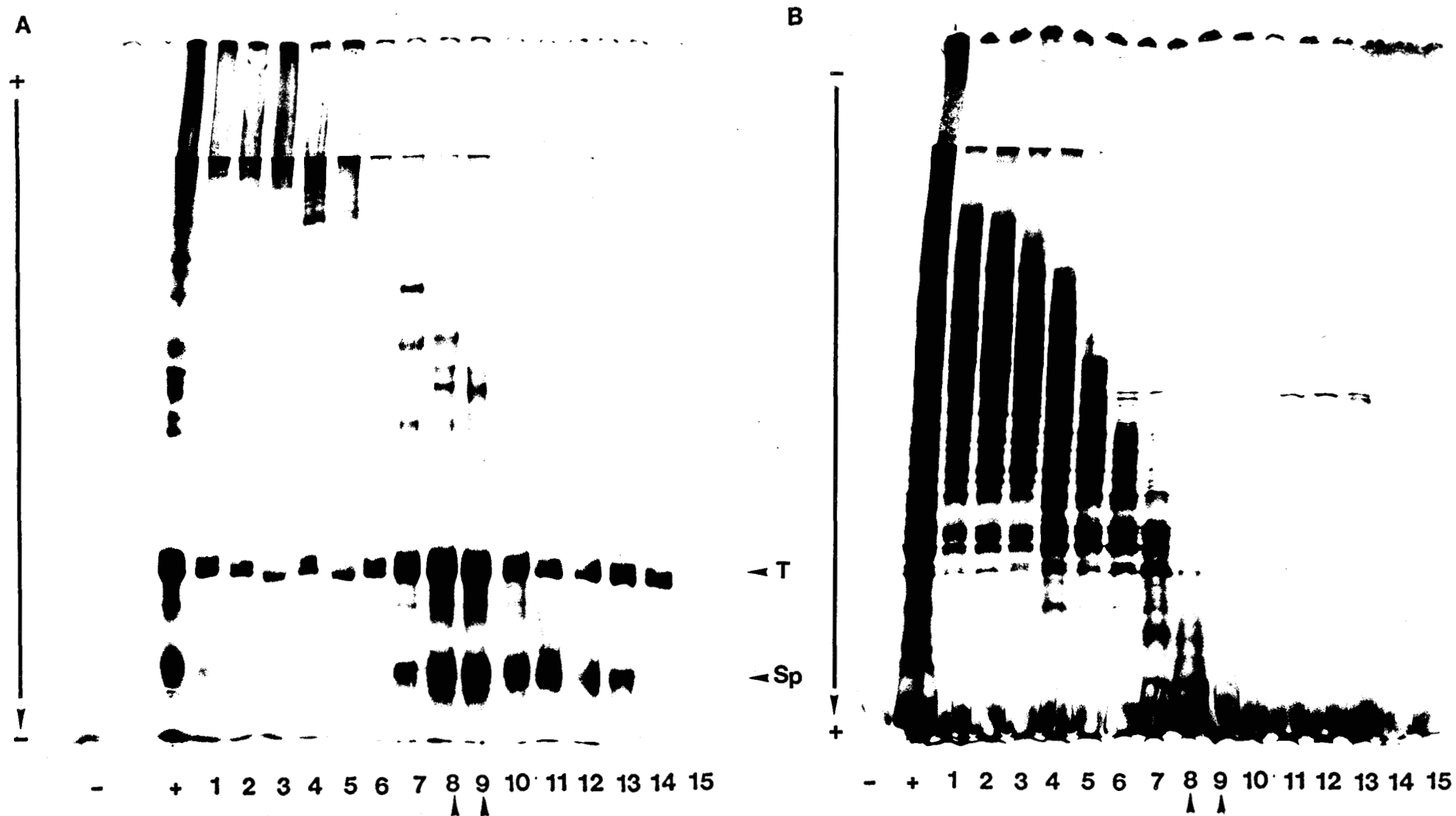


FIGURE 35: Analyses électrophorétiques sur gels de polyacrylamide des produits de traduction des ARNm séparés sur gradient de saccharose. A: gel de polyacrylamide à 14 % en milieu acide/urée.

B: gel de polyacrylamide à 12 % en présence de SDS.

Les numéros des échantillons correspondent aux numéros des fractions collectées dans le gradient de saccharose. -: témoin négatif des ARNm endogènes du lysat de réticuloocyte de lapin. +: témoin de traduction des ARN poly A⁺. Après coloration et séchage les gels sont mis en autoradiographie.

TABLEAU VII

Composition en acides aminés de la protéine de migration électrophorétique intermédiaire entre les protéines T et Sp (BOUTILLON, 1987)

Asx	10.6 (6)*
Thr	3.3 (2)
Ser	11.0 (6)
Glx	4.5 (2)
Pro	3.7 (2)
Gly	6.8 (4)
Ala	3.9 (2)
Cys	3.9 (2)
Val	3.3 (2)
Met	1.6 (1)
Ile	3.3 (2)
Leu	2.8 (2)
Tyr	3.5 (2)
Phe	4.4 (2)
His	2.6 (1)
Lys	9.4 (5)
Arg	22.0 (12)
Trp	n.d.
Total	(55)

Les résultats sont exprimés en moles/100 moles d'acides aminés.

Les valeurs pour la Thréonine et la Sérine sont déduites de l'extrapolation linéaire au temps zéro des valeurs obtenues après 24 et 72 heures d'hydrolyse.

* Les nombres entre parenthèses représentent le nombre de résidus réel établi en estimant la masse molaire de la protéine d'après sa migration électrophorétique en gel de polyacrylamide par rapport à des protéines connues.

Ces fractions enrichies ont été utilisées pour l'élaboration de la banque d'ADNc.

II - 2-4 BANQUES D'ADNc

Le choix du vecteur de clonage utilisé pour créer les banques d'ADNc a été guidé par les moyens d'investigation mis à notre disposition pour le criblage des banques. En effet, la nature de la protéine étudiée et notamment son caractère très basique rend extrêmement difficile l'obtention d'anticorps spécifique de la protéine T.

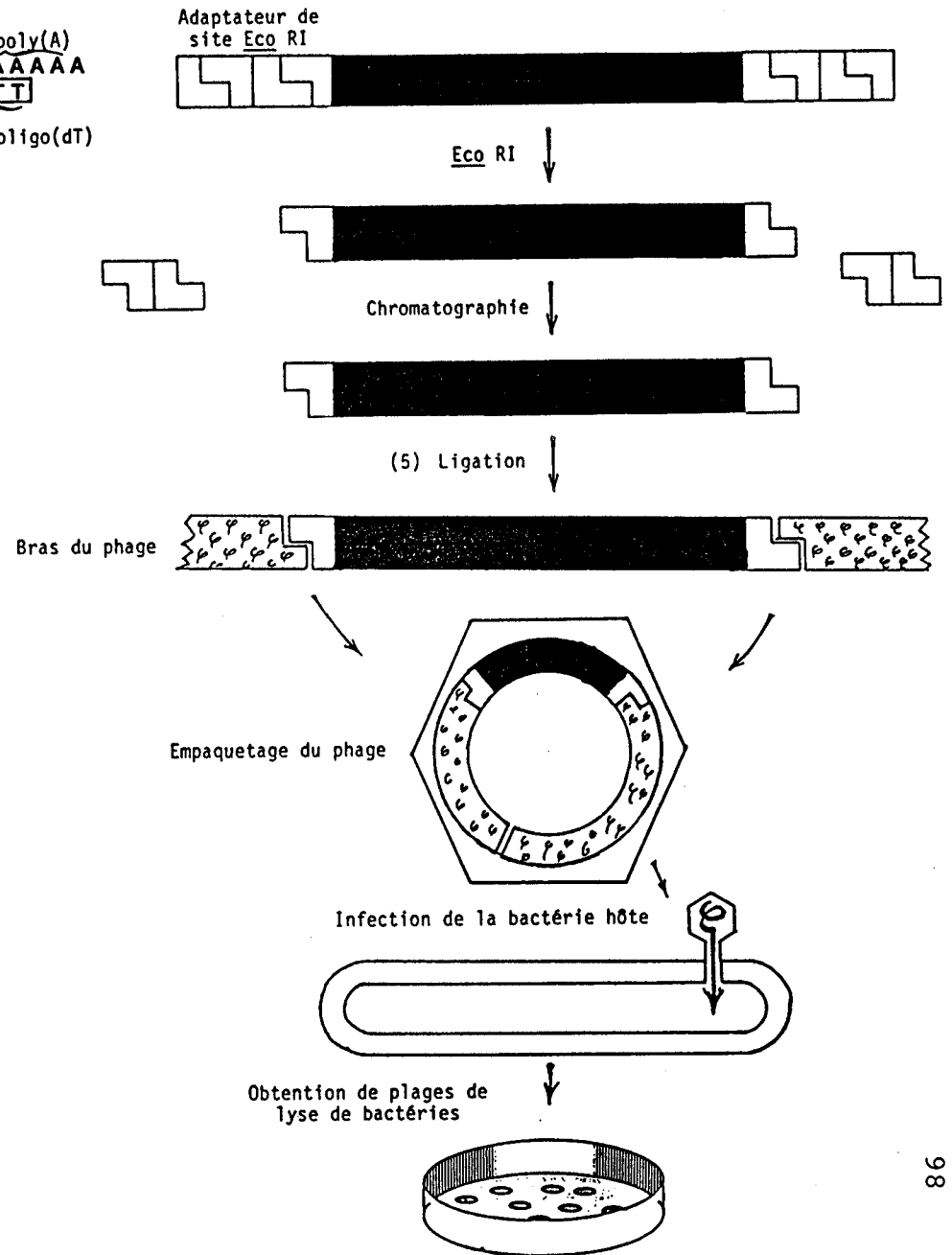
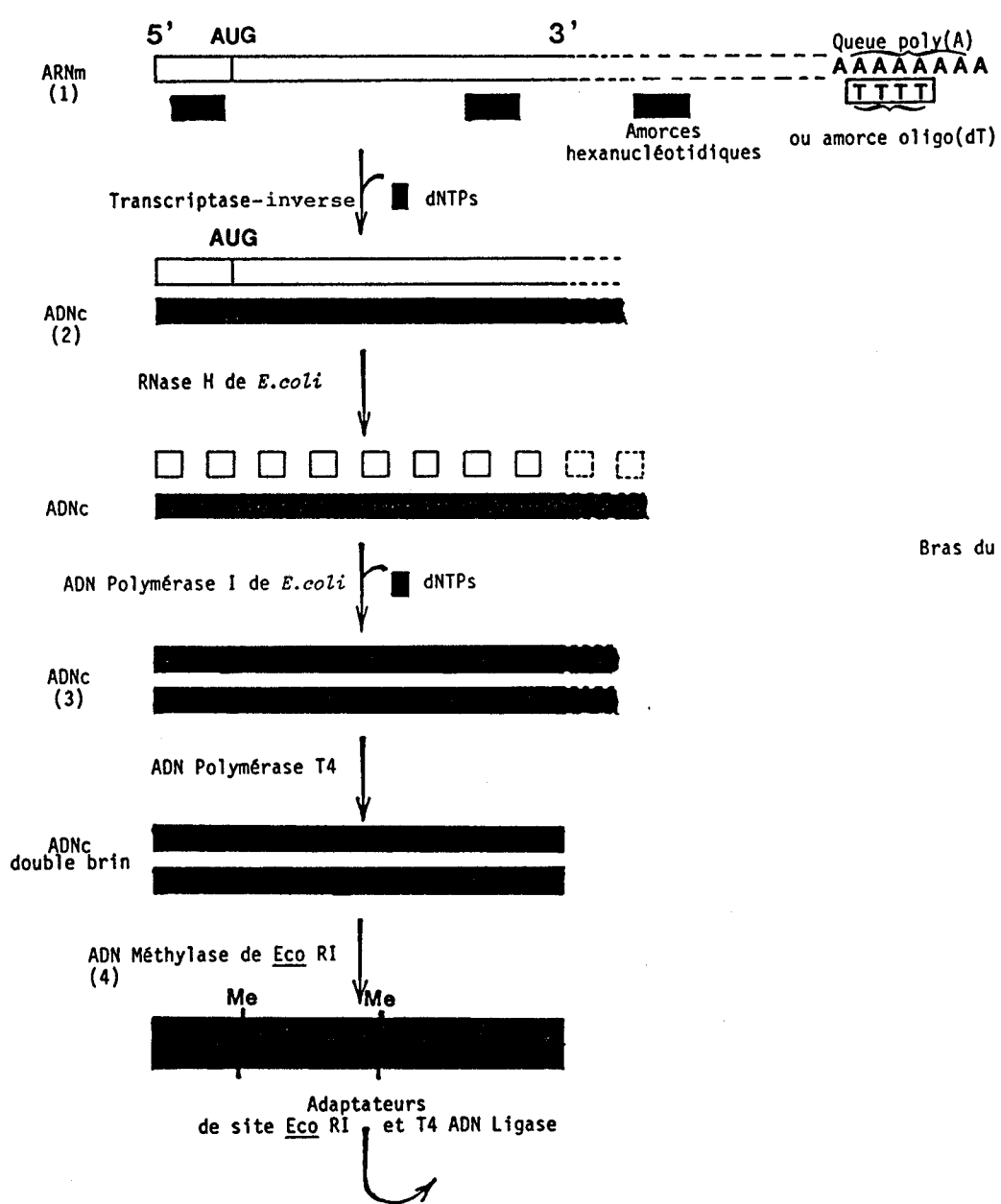
Aussi, dans un premier temps, nous avons utilisé le phage λ gt 10 et le plasmide pUC 13 comme vecteurs de clonage, afin de pouvoir utiliser les techniques d'hybridation de sondes oligonucléotidiques.

Les étapes de synthèse et de clonage des ADNc sont rappelées dans la figure 36.

II - 2-4-1 Synthèse de l'ADNc

Les rendements obtenus pour la synthèse des deux brins d'ADN complémentaires ont été calculés à partir de l'incorporation d' $[\alpha^{32}\text{P}]d\text{CTP}$. Ces rendements sont de l'ordre de 20 % et de 92 % pour la synthèse du premier et du second brin respectivement.

L'analyse électrophorétique sur gel de polyacrylamide non dénaturant montre que ces fragments d'ADN ont des tailles majoritairement comprises entre 250 et 650 paires de bases (Figure 37).



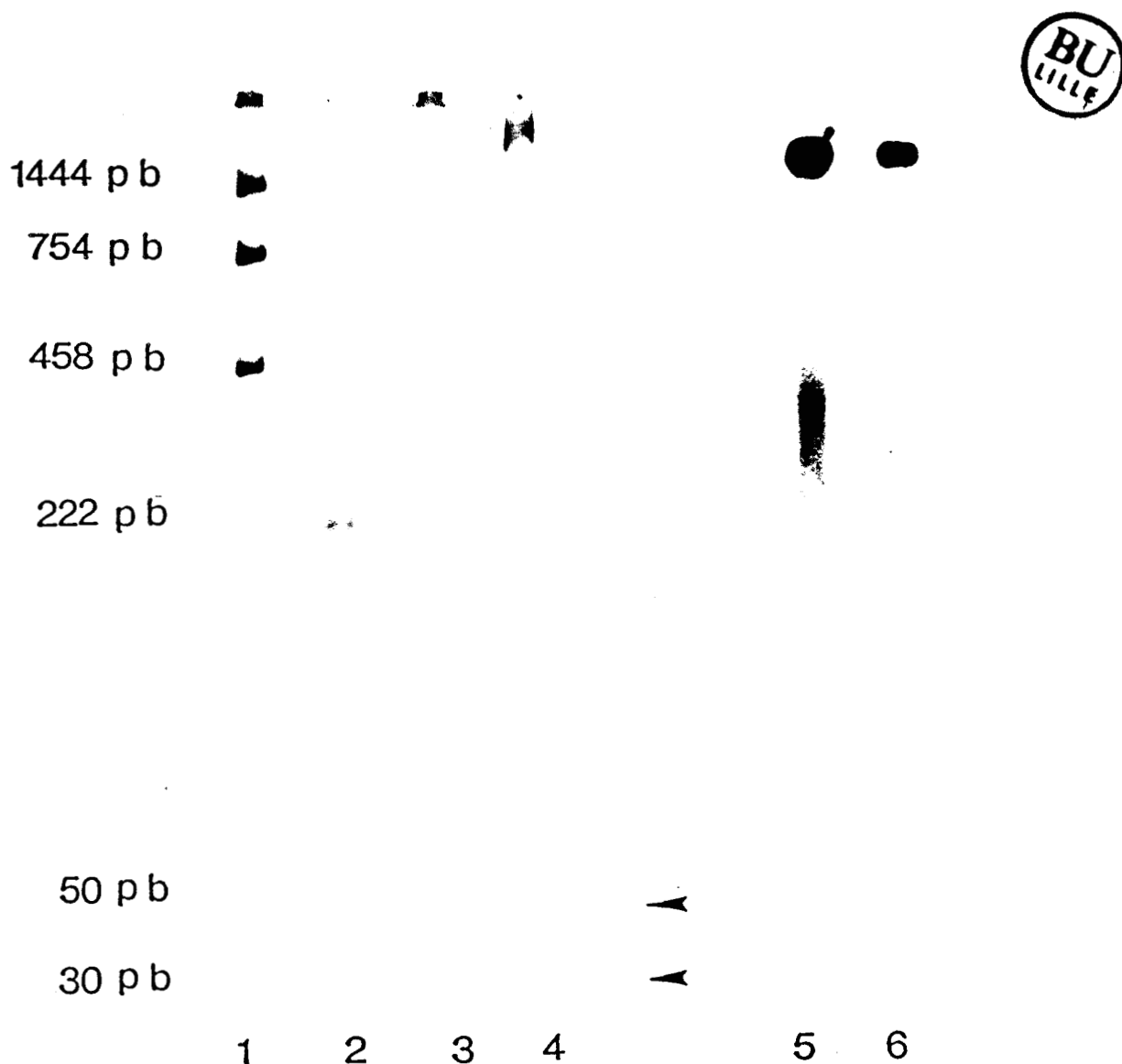


FIGURE 37: Diagramme électrophorétique en gel de polyacrylamide à 5% en conditions non dénaturantes des ADNc synthétisés. Les témoins (pistes 1 à 4) ont été obtenus après coupure du plasmide pUC13 par différentes enzymes, puis marquage par [α - 32 P] dCTP par action du fragment Klenow de l'ADN polymérase I.

- 1: témoin de taille moléculaire obtenu avec *Taq* I.
- 2: témoin de taille moléculaire obtenu avec *Ava* II.
- 3: témoin de taille moléculaire obtenu avec *Hind* III et *Bam* HI.
- 4: témoin de taille moléculaire obtenu avec *Hind* III et *Eco* RI.
- 5: ADNc synthétisé avec l'amorce oligo-(dT).
- 6: ADNc synthétisé avec les amorces hexanucléotidiques au hasard.

Ce résultat concorde bien avec les tailles connues d'ADNc correspondant à des protamines, à leurs précurseurs ou à des protéines intermédiaires.

II - 2-4-2 Clonage

Le phage λ gt 10 accepte des fragments d'ADNc dans l'unique site de coupure par EcoRI localisé sur le gène répresseur "CI". Ce gène est indispensable à l'établissement d'un cycle lysogénique lors de la transformation des bactéries hôtes. La souche parentale de λ gt 10 a une voie lysogénique préférentielle car le gène "CI" est actif. Elle donne des plages de lyse troubles sur des boîtes de Pétri lorsqu'elle infecte des souches d'*E.coli* L87. L'infection, réalisée par des bactériophages recombinants "CI", produit des plages au centre clair. Cette caractéristique permet de reconnaître les bactériophages recombinants. Enfin, lorsque les souches bactériennes telles que NM514, possédant la mutation "hfl" (high frequency lysogeny), sont infectées, seuls les bactériophages recombinants forment des plages de lyse sur la boîte.

Le plasmide pUC 13 renferme un polylinker contenant un site unique EcoRI situé aux environs de l'extrémité 5' du gène Lac Z codant pour la β -galactosidase. Ce site est utilisé pour insérer les fragments d'ADN complémentaires. L'utilisation du pUC 13 permet une sélection des clones positifs par la couleur grâce à l'inactivation de la β -galactosidase.

Le rendement de clonage des ADNc en phages λ gt 10 est de 2 à 4 10^6 phages recombinants par μ g d'ADNc. Pour le clonage des ADNc en plasmide pUC 13 le rendement par μ g d'ADNc est de 10^4 plasmides contenant un insert pour 10^6 plasmides.

II - 2-4-3 Criblage des banques d'ADNc

Dans un premier temps, nous avons utilisé des oligonucléotides de synthèse (de 18 et 30 bases) correspondant à la séquence amino-terminale de la protéine T. Cette séquence est celle qui présente le nombre de combinaisons nucléotidiques le plus faible. Cependant la dégénérescence du code génétique pour les acides aminés concernés est telle que le nombre de combinaisons possibles est encore considérable (128 pour l'oligonucléotide de 18 bases, plus de 16.000 pour celui de 30 bases). L'utilisation de l'inosine a permis de réduire à 32 le nombre d'oligonucléotides à synthétiser en mélange (Figure 38).

1ère sonde

- Séquence protéique : Met-Lys-Val-Ala-Ala-Asn
- Séquence nucléotidique : 3' TAC-TTT-CAC-CGA-CGA-TTA 5'
choisie
 C G T T G
 C C
 G G

2ème sonde

- Séquence protéique : Met-Lys-Val-Ala-Ala-Asn-Ser-Ser-Lys-Met
- Séquence nucléotidique : 3' TAC-TTT-CAC-CGI-CGI-TTA-AGI-AGI-TTT-TAC 5'
choisie
 C G G C
 A
 T

FIGURE 38 - Séquences nucléotidiques des sondes de 18 et 30 bases.

Ces deux sondes utilisées dans les conditions usuelles d'hybridation (MANIATIS *et al.*, 1982), en faisant varier la température, la composition des tampons d'hybridation et de lavage, n'ont pas fourni de résultat. Nous avons donc décidé d'utiliser la technique d'hybridation ARN/ADN qui est résumée dans la figure 39.

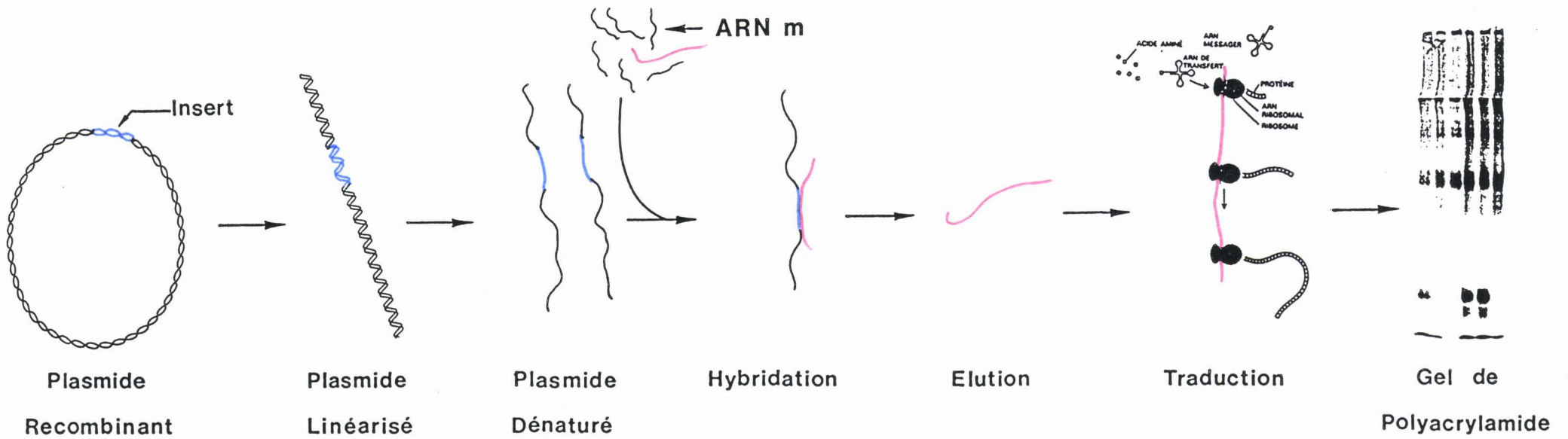


FIGURE 39: Principe de la sélection des clones par hybridation ADNc/ARNm. Les ARNm sont mis en hybridation sur l'ADNc inséré dans le plasmide. L'ARNm fixé est récupéré et traduit en lysat de réticulocyte de lapin. Les produits de traduction sont ensuite analysés sur gel de polyacrylamide en différentes conditions.

Le principe consiste à fixer l'ADN des plasmides contenant les ADNc sur des membranes de nitrocellulose, puis à utiliser l'ADN immobilisé pour hybrider des ARN messagers spécifiques. L'ARN poly A⁺ hybridé sert de marqueur de l'insert présent dans le plasmide. Après hybridation avec les ADN plasmidiques recombinants, les ARNm hybridés sont récupérés et traduits *in vitro* en lysat de réticulocyte de lapin. L'analyse sur gel de polyacrylamide permet d'identifier les clones s'hybridant préférentiellement avec les ARNm de la protéine T. Nous avons contrôlé l'absence d'hybridation des ARNm poly A⁺ sur les plasmides non recombinants.

Les différents plasmides recombinants ont été criblés par groupes de cinq. Cette technique a permis de sélectionner trois groupes de cinq clones dont l'ADN s'hybride avec les ARNm codant pour la protéine T et pour les deux autres protéines coenrichies au cours du gradient de sucrose. Aucune protéine non acido-soluble n'est détectée sur le gel de polyacrylamide en milieu SDS.

Chacun de ces quinze clones a été contrôlé séparément et nous avons ainsi pu sélectionner trois clones positifs : les clones 31, 39 et 48 (Figure 40). Ces trois clones ont une taille d'environ 100 nucléotides, taille trop petite pour correspondre à la totalité de l'ARNm de la protéine T. Les études de structure primaire de ces inserts (Figure 41) ont montré qu'il s'agissait en fait d'un seul et unique clone renfermant un signal de polyadénylation, ce qui le situe dans la région 3' non codante de l'ARNm. Nous devons noter que ces inserts possèdent des structures secondaires importantes, ce qui perturbe leur séquençage et entraîne des compressions dans le gel.

L'insert du clone 48 a été libéré du plasmide par action de l'enzyme Eco RI et marqué par [α ³²P] dCTP afin d'être utilisé comme sonde.

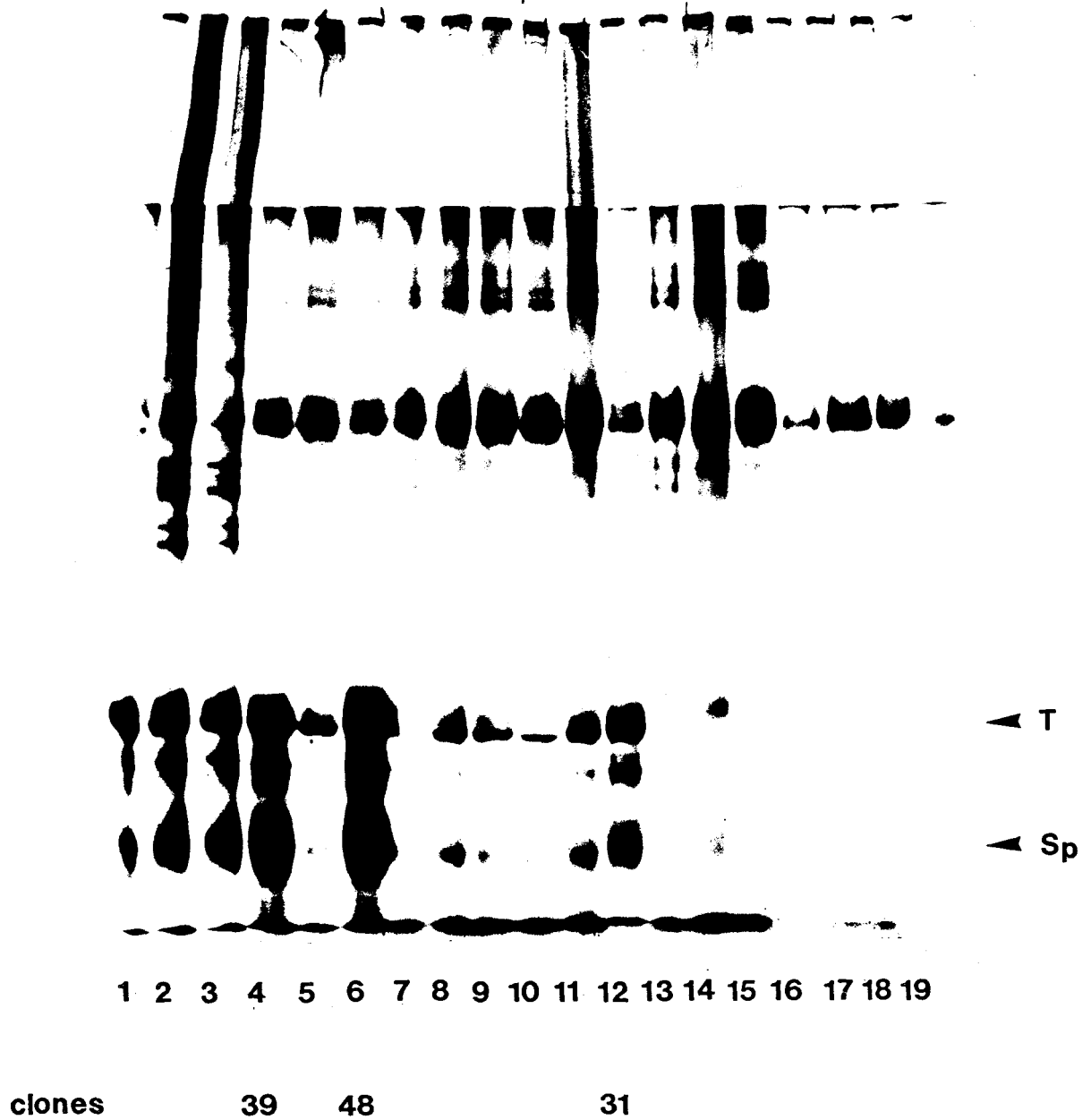


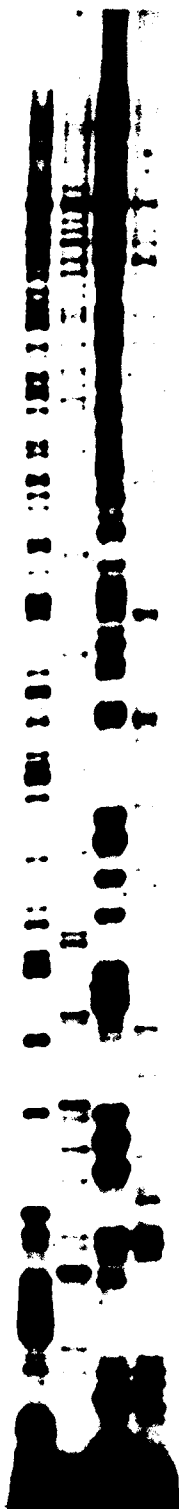
FIGURE 40: Etape finale de sélection des clones par la technique d'hybridation ADNc/ARNm.

Les produits de traduction acido-solubles des ARNm fixés sur les clones sont analysés sur gel de polyacrylamide à 14 % en milieu acide/urée.

1: témoin de traduction avec les ARN poly A⁺ de testicule de seiche. 2 et 3: témoins de traduction avec les ARN totaux. 4 à 18: produits de traductions obtenus avec les ARNm hybridés sur différents clones.

19: produits de traductions après hybridation sur le plasmide pUC13 ne contenant pas d'insert.

A



A G C T

B

1 CAATATCTTTTCATAAGGAAGGGCTAAGGAA

30 CTGTCCTTATGGCGGGGCCTACAAAAGT

60 CAGCAGCAATAAACTTTTTTCGAAAGAT

90 TAAAAAAAAA

FIGURE 41 - Séquence du clone 48.

- A. Analyse sur gel de polyacrylamide à 8 %, urée 7 M des produits de séquençage du clone 48 en plasmide pUC 13, obtenus d'après la méthode de SANGER et COULSON et marqués au [$\alpha^{32}\text{P}$] dCTP.
- B. Séquence nucléotidique : le signal de polyadénylation est souligné par---

- Dans un premier temps, cette sonde (appelée sonde 48) a été utilisée pour déterminer la taille de l'ARN susceptible de coder pour la protéine T. Les ARN totaux, les ARN poly A⁺ et la fraction n°9 enrichie sur gradient de saccharose ont été déposés sur gel d'agarose à 1,4 % en conditions dénaturantes. Après l'électrophorèse, ils ont été transférés sur membrane de nitrocellulose. La sonde 48 a permis de caractériser une bande d'ARN dont la taille est comprise entre 700 et 800 paires de bases (Figure 42).

- Dans un second temps, la sonde 48 a été utilisée pour cribler la banque de phages λ gt10. Une première étude de la banque de phages a permis d'identifier de très nombreux clones positifs. Sur 10⁴ phages testés, plus de 150 clones se sont hybridés avec la sonde. Ceci peut s'expliquer principalement par le fait que nous avons fabriqué la banque avec les fractions enrichies en ARNm codant pour les protéines T et/ou Sp. Il n'est pas possible d'exclure que le clone 48, qui contient une queue poly A⁺ de 10 nucléotides, ne s'hybride pas avec d'autres clones correspondant à des ARN poly A⁺.

Sur 50 clones, nous avons déterminé la taille des inserts par une méthode rapide (SERGHINI *et al.*, 1989) et retenu ceux dont la taille était comprise entre 300 et 800 paires de bases. Chaque clone ainsi sélectionné a de nouveau été hybridé avec la sonde 48 afin d'éliminer toute possibilité de "faux positif" (Figure 43).

Nous nous sommes tout d'abord intéressé au clone le plus grand : le clone 5/14 (14ème clone de la 5ème boîte). L'insert de ce clone a une taille de 760 paires de bases. Ceci correspond à la taille probable de l'ARN messager de la protéine T telle que nous l'avons déterminée au cours de la manipulation précédente.

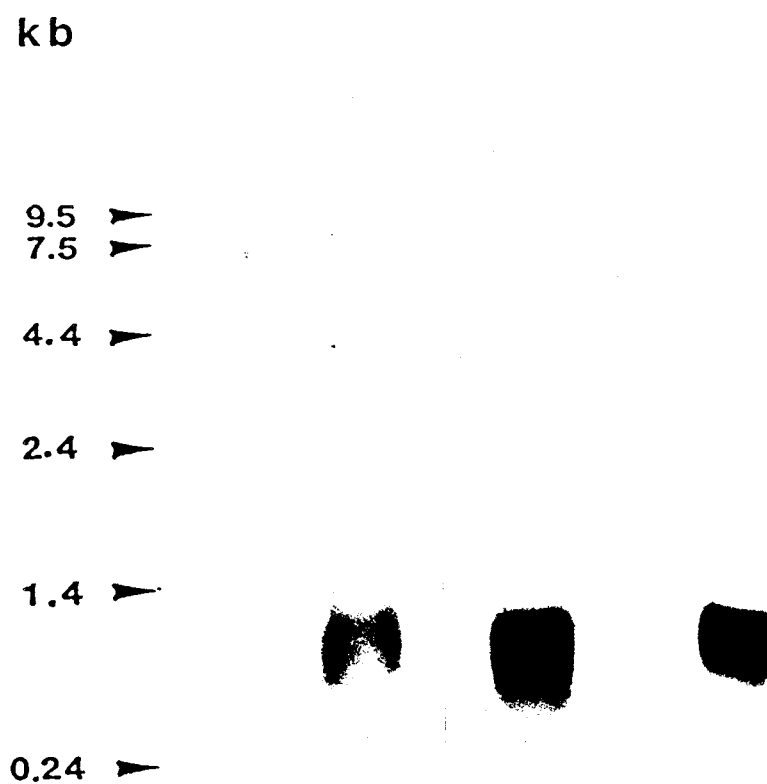


FIGURE 42: Utilisation de la sonde 48 afin de déterminer la taille de l'ARNm codant pour la protéine T, après électrophorèse et transfert sur membrane de nitrocellulose.

- 1: ARN poly A⁺ enrichis sur gradient de saccharose (fraction 9).
- 2: ARN totaux.
- 3: ARN poly A⁺.



FIGURE 43: Sélection de clones de la banque de phages λ gt10 en utilisant comme sonde l'insert 48 marqué au ^{32}P .
A: Détection de clones positifs parmi d'autres clones.
B: Contrôle d'un des clones précédemment sélectionné après amplification.

II - 2-4-4 Sous clonage de l'insert 5/14

L'insert 5/14 a été sous cloné dans le site Eco RI du vecteur plasmidique pUC 13. Après amplification dans la souche bactérienne JM 109, l'ADN plasmidique a été extrait et la carte de restriction de l'insert établie (Figure 44). L'insert renferme des sites de coupure pour les endonucléases de restriction Acc I, Ava I, Dra I, Nar I et Rsa I.

Parallèlement, nous avons effectué une nouvelle hybridation ARN/ADN selon le schéma décrit précédemment et analysé les produits de traduction par électrophorèse. Nous avons pu ainsi contrôler que les ARNm hybridés avec ce clone sont ceux codant pour la protéine T et les protéines coenrichies avec la protéine T.

Afin d'en établir la séquence, nous avons sous cloné l'insert 5/14 dans le site Eco RI du phage M13 mp18, M13 mp19 (Figure 45).

II - 2-4-5 Séquençage de l'insert 5/14

Le clone 5/14 a été séquencé sous forme d'ADN double brin, ce qui a permis de retrouver la séquence de la sonde 48, avec les mêmes problèmes de compression liés aux structures secondaires. Aussi, avons nous appliqué la technique de séquençage en ADN simple brin en utilisant l'ADN polymérase Tag qui fonctionne à haute température, et en incorporant de la 7-déaza guanidine afin de réduire les structures secondaires.

L'analyse de la séquence a montré qu'en fait, le clone 5/14 ne correspond pas à un seul ARNm, mais qu'il est constitué de deux ARNm placés tête-bêche (Figure 44). Cela implique

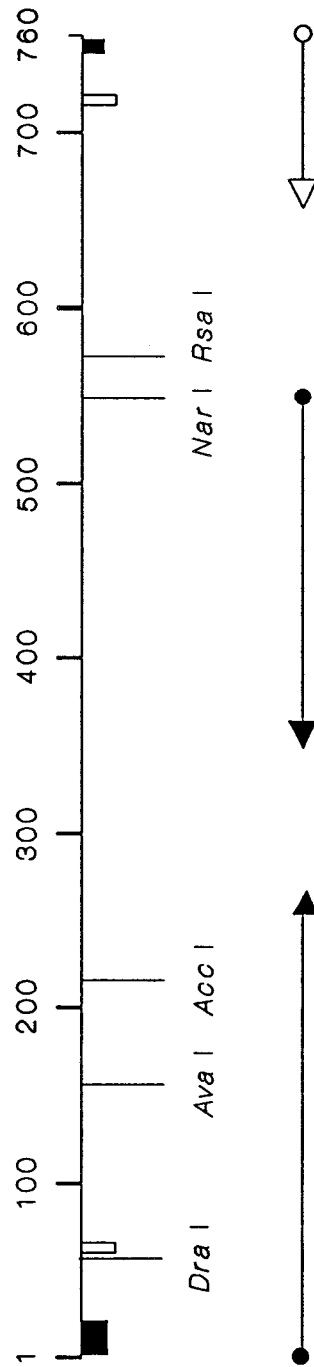


FIGURE 44: Carte de restriction et stratégie de séquence de séquence du clone 5/14. Les flèches indiquent la direction de séquençage et le nombre de nucléotides identifiés. Les ■ indiquent des séquences poly A et □ indiquent les signaux de polyadénylation. Les séquences réalisées en simple brin sont symbolisées par des flèches pleines et celle en double brin par une flèche creuse.

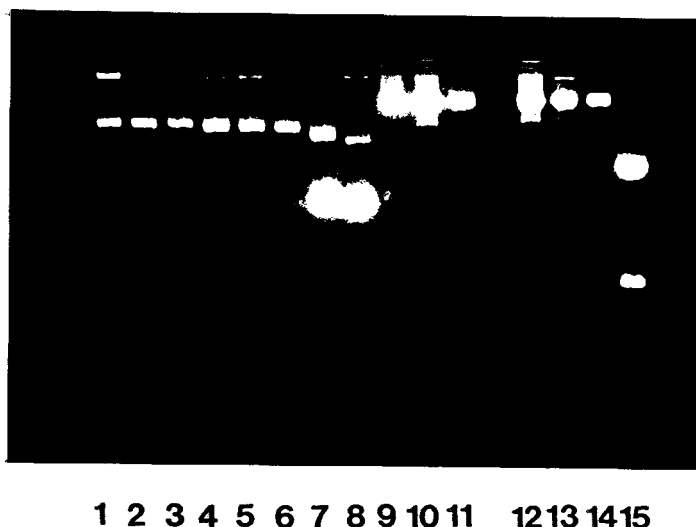


FIGURE 45: Sous-clonage de l'insert 5/14 en phages M13mp18 et M13mp19. L'électrophorèse est réalisée sur un gel d'agarose à 1,8 % en TBE 1X sous un courant constant de 90 volts pendant 2 h.

- 1 et 2: M13mp18 de l'insert 5/14.
- 3: M13mp19 de l'insert 5/14.
- 4 et 5: témoin de ligation en phage M13mp18 avec un insert témoin de 500 pb.
- 6: témoin de ligation en phage M13mp19 avec un insert témoin de 500 pb.
- 7 et 8: plasmide 5/14 et plasmide obtenu avec le témoin de 500 pb en plasmide pUC13 après la première étape de sous-clonage.
- 9, 10 et 11: mêmes échantillons que 4, 5 et 6 après libération des inserts par Eco RI.
- 12, 13 et 14: mêmes échantillons que 1, 2 et 3 après libération des inserts par Eco RI.
- 15: libération de l'insert 5/14 par Eco RI à partir du plasmide à la première étape de sous-clonage.

malheureusement que ce clone ne peut pas couvrir la séquence complète de la protéine T. Il a donc été nécessaire de rechercher un insert intéressant parmi les autres clones.

II - 2-4-6 Etude du clone 5/20 correspondant à la protéine T2

Quatre clones dont les inserts ont des tailles comprises entre 320 et 630 paires de bases ont été sous clonés à la fois en plasmide pUC 13 et en phage M13 (Figure 46). Jusqu'à présent, seule la séquence de l'insert du clone 5/20 a pu être réalisée. Les résultats sont présentés dans la figure 47.

La recherche des phases ouvertes de lecture montre une phase ouverte dont la séquence correspond sans ambiguïté aux résidus 17 à 77 de la protéine T2 déterminée par la dégradation d'Edman. Cette séquence représente les 184 premiers nucléotides de l'insert et se termine par le codon de terminaison TAA.

Sur les 43 résidus d'arginine contenus dans la protéine T2, l'utilisation des codons s'effectue selon les pourcentages suivants :

- AGG : 30,3 % (13 résidus)
- AGA : 18,6 % (8 résidus)
- CGC : 16,3 % (7 résidus)
- CGA : 13,9 % (6 résidus)
- CGG : 11,6 % (5 résidus)
- CGT : 9,3 % (4 résidus)

Contrairement à ce qui se passe chez certaines espèces (BERLOT-PICARD *et al.*, 1986 ; OLIVA et DIXON, 1989), nous pouvons dire que tous les codons de l'arginine sont utilisés dans

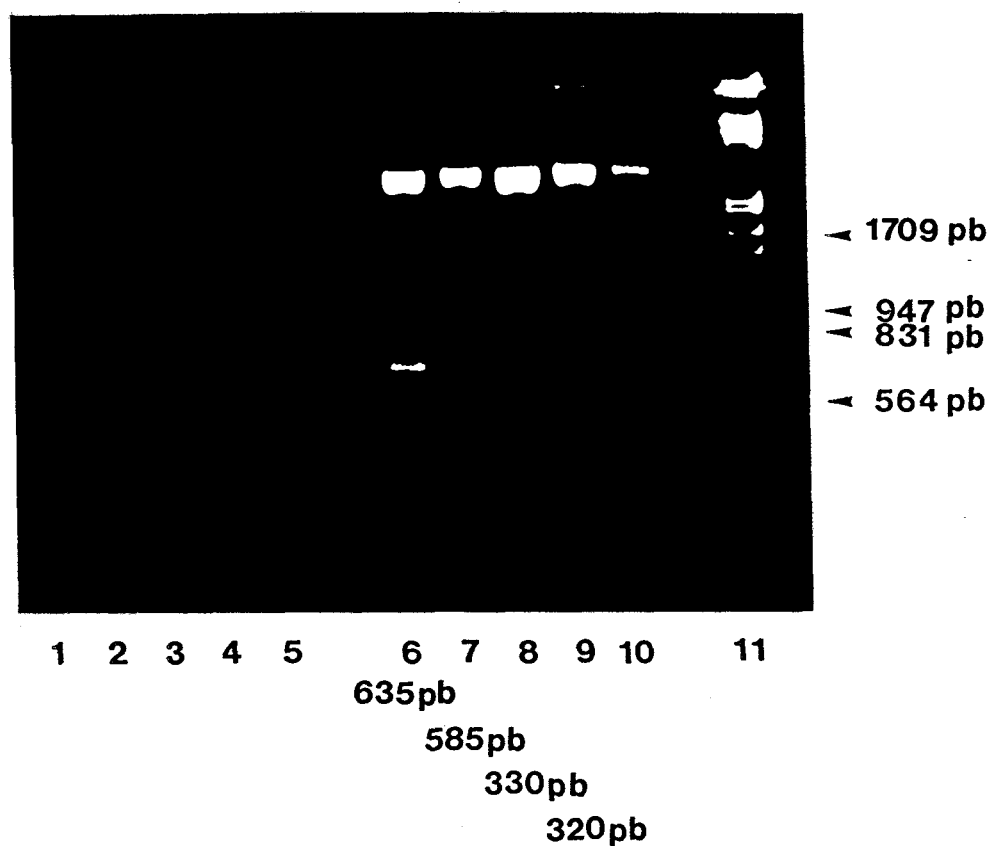


FIGURE 46: Détermination de la taille des inserts sous-clonés en plasmide pUC13. La migration électrophorétique a été réalisée sur un gel d'agarose à 2 % en TBE 1X sous un courant constant de 30 volts pendant 16 h.

- 1: plasmide 8/5.
- 2: plasmide 5/5.
- 3: plasmide 5/20.
- 4: plasmide 5/3.
- 5: plasmide pUC13.
- 6 à 10: mêmes plasmides après libération des inserts par *EcoRI*.
- 11: témoins de tailles moléculaires (Boehringer)

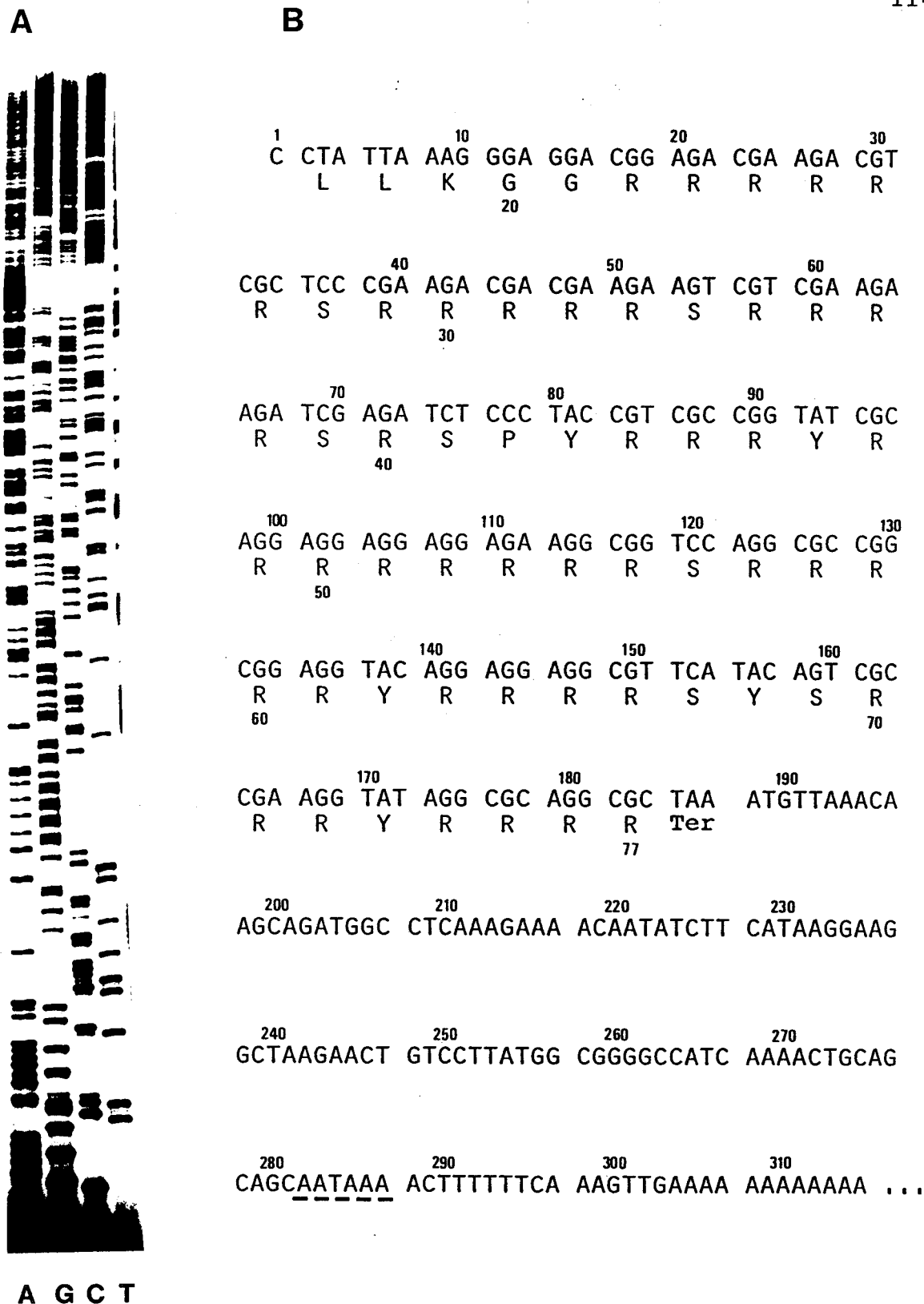


FIGURE 47 - Séquence du clone 5/20.

- A. Analyse sur gel de polyacrylamide à 8 %, urée 7 M des produits de séquençage du clone 5/20 en phages M13mp18 obtenus avec le kit Taq-Track™ Sequencing System (Promega Biotec) et marqués à 1' [α^{35} S] dATP.
- B. Séquence nucléotidique et séquence déduite en acides aminés d'une partie de la protéine T2. Ter représente le signal de terminaison. Le signal de polyadénylation est souligné par---

la protéine T2. Cette caractéristique est partagée avec les gènes des protamines de mammifères (KLEENE *et al.*, 1986 ; YELICK *et al.*, 1987). Dans les gènes de protamines, c'est en général le codon CGC qui est le plus utilisé, suivi par le codon AGG (LEE *et al.*, 1986). Dans le cas de la protéine T2 de seiche, le codon le plus utilisé par contre est le codon AGG.

L'insert du clone 5/20 contient également la région 3' non codante qui représente 118 nucléotides (queue poly A non comprise). Cette région est de taille similaire à celles trouvées pour les autres protamines et protéines intermédiaires. Dans cette région 3' nous retrouvons la séquence du clone 48 qui nous a servi de sonde lors du criblage de la banque de phages. En effet, la séquence de 219 à 303 nucléotides (queue poly A non comprise) est très comparable à celle du clone 48, à quatre délétions et une interversion près. Nous pouvons expliquer cela de plusieurs façons :

- soit il existe des différences dans la nature des clones étudiés. En effet, il est possible que les séquences en 3' des gènes de protéines T1 et/ou Sp soient différentes de celle de la protéine T2.

- soit la différence est liée à l'enzyme utilisée lors du séquençage. En effet, la Taq polymérase choisie lors du séquençage du clone 5/20 agit à haute température et l'incorporation de 7-deaza guanidine donnent de meilleurs résultats par rapport à ceux obtenus avec la transcriptase inverse.

D'autre part, dans cette région 3' le seul signal caractéristique présent est le signal de polyadénylation AATAAA, 16 nucléotides en amont de la queue poly A. L'utilisation de la banque PC "gene" (A. BAIROCH, Genofit) nous a permis de comparer la région 3' du gène de la protéine T2 de seiche avec celle

d'autres protamines et protéines intermédiaires ; l'analyse de ces résultats ne montre aucune analogie. De plus, la séquence de régulation potentielle GCCACCTG présente dans les gènes de protamines de mammifères n'est pas retrouvée, chez la seiche (voir paragraphe I - 3-1-2 page 32).

II - 2-5 DISCUSSION

Au cours de ce travail de biologie moléculaire, nous avons été confrontée à de nombreuses difficultés techniques liées à la particularité du modèle étudié.

La plupart des études de biologie moléculaire réalisées sur les protéines nucléaires basiques ont été effectuées jusqu'à présent sur des animaux d'élevage. Ainsi le travail a été facilité, par la possibilité d'entreprendre des études *in vivo*, et par l'obtention de tissus frais (études sur des populations cellulaires enrichies). Dans le cas du modèle étudié, l'utilisation de matériel congelé limite les moyens d'investigation.

De plus, dans la majorité des travaux, le criblage des banques d'ADNc fabriquées a été facilité par la présence de séquences consensus spécifiques (Figure 10) pouvant être utilisées pour la synthèse de sondes oligonucléotidiques (KRAWETZ *et al.*, 1986) et de sondes ADNc (LEE *et al.*, 1987 ; DOMENJOURD *et al.*, 1988).

Dans le cas de la seiche, ces techniques n'ont pas pu être utilisées car les gènes des protéines intermédiaires et des protamines des Invertébrés ont été très peu étudiés et il n'est pas possible actuellement de dégager des séquences consensus.

De plus, comme nous l'avons vu dans la séquence du clone 5/20, le codon le plus fréquent pour l'arginine (codon AGG) n'est pas le codon habituellement le plus utilisé dans les gènes de protamines. Aussi, nous avons dû utiliser une technique de criblage plus délicate - les hybridations ADN-ARN - qui ont déjà été utilisées avec succès pour isoler des gènes de protéines intermédiaires ou de protamines (HEIDARAN et KISTLER, 1987 ; BERLOT-PICARD *et al.*, 1986, 1987). Cette technique nous a permis d'obtenir le clone 48 qui, à son tour, a été utilisé comme sonde. Nous avons ainsi sélectionné de nombreux clones positifs dans la banque de phages. La séquence du clone 5/20 montre sans ambiguïté la valeur de la sonde 48.

Nos résultats sur l'analyse des produits de traduction semblent en contradiction avec l'hypothèse d'un mécanisme de protéolyse de la protéine T en protéine Sp (présence d'une protéine de migration électrophorétique identique à celle de la protéine Sp dans les produits de traduction). L'analyse des nombreux clones positifs restants permettra peut-être d'éclaircir ce point. Elle permettra sans aucun doute de compléter la séquence nucléotidique de la protéine T2 et de trouver un ADNc correspondant à la protéine T1.

CONCLUSION ET PERSPECTIVES

La spermatogenèse chez la seiche est donc caractérisée par une double transition protéique. Les histones présentes aux premiers stades de la spermatogenèse sont remplacées au stade spermatides rondes par la protéine T de type intermédiaire. Celle-ci est à son tour remplacée au stade spermatides allongées par une protamine typique, la protéine Sp. Cette protéine subsiste dans le spermatozoïde où elle représente alors la protéine basique majeure associée à l'ADN.

- Les structures primaires des variants T1 et T2 et de leurs sous-fractions respectives T1a et T2a qui ont perdu le résidu de méthionine en position N-terminale ont été établies.

La protéine intermédiaire de seiche ne se rattache précisément à aucune autre protéine intermédiaire actuellement connue.

L'alignement optimum des séquences des protéines T et Sp ainsi que leur apparition successive au cours de la spermiogenèse indiquent que les variants de protéine Sp pourraient provenir des variants de protéine T par coupure protéolytique. La présence de précurseurs protéiques des protamines a déjà été évoquée chez les protamines de type P2 uniquement de l'homme et de la souris (les protamines de type P1 de ces espèces n'ont pas cette caractéristique). Il faut toutefois rappeler que le mécanisme de coupure protéolytique n'a jamais été éclairci et que les sites de coupure potentiels diffèrent chez l'homme, la souris et la seiche. Cependant les méthodes de prédiction de structures secondaires indiquent que le site de coupure dans la protéine T ($\begin{matrix} \text{Met} \\ \text{Leu} \end{matrix}$ -Lys-Gly-Gly \downarrow -Arg) serait dans une conformation en coude β favorable à une coupure protéolytique (RHOLAM *et al.*, 1986). De plus, ce site de coupure est un site de coupure connu dans les protéines VI et VII de l'adénovirus de type 2 (WEBSTER *et al.*, 1989).

- Nous avons enrichi fortement les ARNm codant pour la protéine T. Cet enrichissement s'accompagne de celui des ARNm codant pour d'autres protéines acido-solubles parmi lesquelles une protéine a une migration électrophorétique identique à celle de la protamine Sp, et une autre, une migration électrophorétique identique à une protéine riche en cystéine et arginine qui a été caractérisée dans notre laboratoire (BOUTILLON, 1987).

- Nous avons fabriqué une banque d'ADNc plasmidique et une banque d'ADNc en phages λ gt 10.

- Nous avons obtenu dans la banque plasmidique une sonde valable permettant d'identifier les clones codant pour la protéine T et ainsi pu sélectionner un nombre important de clones positifs dans la banque de phages λ gt 10.

- Nous avons déterminé la séquence d'un clone dont la partie codante correspond à la séquence 17-77 de la protéine T2.

- Nous savons que dans la banque de phages λ gt 10 se trouvent les sondes nécessaires pour trouver les gènes de la protéine T et, éventuellement, de la protamine Sp dans le génome de la seiche (clone 5/20, par exemple).

Les différentes possibilités qui s'ouvrent pour éclaircir le mécanisme de transition sont :

- du côté protéique : utiliser la protéine ou un peptide de synthèse pour rechercher une ou plusieurs enzymes présentes dans le testicule, capable(s) de donner naissance à la protéine Sp.

- du côté acide nucléique :
 - . Rechercher dans la banque d'ADNc, des inserts correspondant aux variants de la protamine et de la protéine T.
 - . Dans le cas où une enzyme serait caractérisée, utiliser la mutagenèse dirigée sur l'ADNc de la protéine T pour en étudier la spécificité.
 - . Construire une banque d'ADN génomique et utiliser les inserts intéressants de la banque d'ADNc comme sonde, afin de trouver les gènes complets de la protéine T et, éventuellement de la protamine Sp.
 - . Enfin, la banque d'ADN génomique permettra :
 - * une connaissance de la structure des gènes, des protéines intermédiaires de céphalopodes c'est-à-dire de situer les éventuels introns, les facteurs de régulations et les structures conformationnelles de l'ADN.
 - * une étude des mécanismes de régulation des gènes.

Cette étude pourra ensuite être étendue à d'autres céphalopodes.

APPENDICE TECHNIQUE

I - RÉCOLTE DU MATÉRIEL BIOLOGIQUE

Les zones de pêche sont situées dans la Manche d'une part au large de Wimereux (Station Marine de Wimereux, Pas-de-Calais) et d'autre part au large de Luc-sur-Mer (Station Marine de Luc-sur-Mer, Calvados). La pêche est réalisée en fonction des conditions climatiques et météorologiques sur la période fin avril - début juin au début de la période de reproduction des seiches.

Les animaux sont séparés selon leur sexe et conservés en bassin alimenté en eau de mer. On récolte séparément sur chaque mâle, le testicule, les spermatophores et l'épididyme. Les tissus sont aussitôt plongés dans l'azote liquide et sont conservés ainsi jusqu'à leur stockage en congélateur à -80°C.

II - ÉTUDE BIOCHIMIQUE DE LA PROTÉINE T SPÉCIFIQUE DES SPERMATIDES

II - 1 ISOLEMENT ET PURIFICATION DES PROTEINES BASIQUES

II 1-1 Isolement des protéines basiques

La préparation des protéines se fait à partir de noyaux de testicules selon un procédé décrit par SAUTIERE *et al.* (1981). Toutes les manipulations sont effectuées à +4°C en chambre froide et toutes les solutions préparatives sont additionnées de diisopropylfluorophosphate pour inhiber l'activité des sérine-protéases.

** Isolement et purification des noyaux*

- Les testicules (35 g) sont émincés et homogénéisés dans 200 ml d'une solution de saccharose 0.25 M, CaCl_2 0.003 M, Tris 0.05 M, pH 7.4 (tampon A) à l'aide d'un homogénéiseur de type Waring Blendor (basse vitesse, 2 minutes).

L'homogénat est filtré sur quatre épaisseurs de gaze hydrophile et centrifugé à 2500 xg pendant 20 minutes à 4°C pour éliminer les débris cellulaires. Les noyaux sont lavés deux fois dans 200 ml de tampon A.

- Les noyaux lavés sont remis en suspension dans 200 ml de tampon A mais contenant cette fois du saccharose 2,2 M (tampon B) et centrifugés à 47.000 xg pendant 60 minutes.

** Préparation de la chromatine et extraction des protéines basiques*

- La préparation de la chromatine est fondée sur son insolubilité en milieu isotonique NaCl 0.14 M, Tris 0.005 M, pH 7.1 (tampon C).

Les noyaux purifiés sont homogénéisés dans 200 ml de tampon C dans un Waring Blendor (vitesse rapide, 2 minutes). La chromatine insoluble est récupérée par centrifugation à 2500 xg pendant 20 minutes. Elle est lavée deux fois dans 200 ml du même tampon, puis trois fois dans 50 ml de NaCl 0.35 M pour éliminer les protéines non histones.

- Les protéines basiques sont alors extraites de la chromatine par deux extractions successives (2 fois 18 heures) avec 50 ml d'HCl 0.4 M. Les surnageants contenant les protéines sont rassemblés, dialysés extensivement contre de l'eau distillée à 4°C, dans du tube à dialyse Spectrapore n°3 (limite d'exclusion < 3500) et lyophilisés. Le culot de chromatine résiduelle est extrait pendant 8 h à +4°C par 100 ml d'une

solution de chlorhydrate de guanidine 5 M contenant du 2-mercaptoéthanol 0.1 M. Le surnageant obtenu après centrifugation est dialysé dans les conditions décrites précédemment et lyophilisé.

II - 1-2 Purification de la protéine T

** La protéine T*

La protéine T est purifiée à partir des extraits lyophilisés HCl 0.4 M ou GuCl 5 M de la chromatine du testicule, par chromatographie liquide haute pression.

- L'extrait HCl 0.4 M est dissous dans l'HCl 0.01 M et déposé sur une colonne de C8 Ultrapore (Beckman) équilibrée en milieu TFA 0.1 %. L'élution est réalisée par un gradient d'acétonitrile de 0 à 50 % dans le TFA 0.1 %.

- L'extrait GuCl 5 M est fractionné sur une colonne C18 μ Bondapak (Waters), équilibrée et éluée dans les mêmes conditions.

** Les variants structuraux*

- Les variants structuraux de la protéine T ont été séparés à l'aide d'un gradient discontinu d'acétonitrile dans le TFA à 0,1 %, sur une colonne de C18 μ Bondapak (0 % 6 min. - 0 à 21 % 13 min. - 21 % 10 min. - 21 à 50 % 17 min.)

II - 2 ANALYSES ELECTROPHORETIQUES DES PROTEINES

II - 2-1 Electrophorèse sur gel de polyacrylamide en milieu acide-urée

Les analyses électrophorétiques des protéines et des peptides acido-solubles sont réalisées sur gel de polyacrylamide en milieu acide-urée, dans les conditions présentées dans le tableau ci-dessous.

Pourcentage de polyacrylamide :	Nature des échantillons		
	Protéines	Variants protéiques	Peptides
- des gels de séparation	14 %	17 %	22 %
- des gels de concentration	7 %	7 %	5 %
Rapport en acrylamide/bisacrylamide	38/1	38/1	20/1
Concentrations :			
- en urée	6.25 M	6.25M	2.5 M
- en acide acétique	5,4 %	5,4 %	4,5 %

Le tampon de migration est composé d'acide acétique 0,9 M à pH 2.7. Le gel est soumis à une pré-électrophorèse de deux heures sous courant constant de 22 mA. Les échantillons sont mis en solution la veille de la migration en tampon HCl 0,01 M, Urée 8 M, 2-mercaptoéthanol 0,5 M. L'électrophorèse est réalisée à 22 mA sous température contrôlée (10°C).

II - 2-2 Electrophorèse sur gel de polyacrylamide en présence de SDS

Les protéines solubles en présence de SDS sont analysées sur gel de polyacrylamide selon la méthode de tampons discontinus de LAEMMLI (1970) dans un gel homogène de polyacrylamide 18 % (rapport acrylamide/bisacrylamide 37/1) contenant 0,1 % de SDS dans du tampon Tris-HCl 0,375 M à pH 8.8. Le gel de concentration à 5 % de polyacrylamide dans du tampon Tris-HCl 125 mM à pH 6.8, contient 0,1 % de SDS. Chaque échantillon est solubilisé dans 15 μ l de tampon Tris-HCl 62,5 mM, pH 6,8, SDS 3 %, glycérol 10 %, 2-mercaptoéthanol 2,5 % (V/V), bleu de bromophénol 1 %.

La migration dure de 16 à 18 h à 7 mA ou 40 mA pendant 4 h dans le tampon Tris-glycine 16,5 mM et 128 mM respectivement, pH 8,3 contenant 0,1% de SDS.

II - 2-3 Coloration et décoloration des gels

Les gels sont colorés dans une solution de bleu de Coomassie R250 à 0,04 % dans une solution de méthanol 9 % et d'acide picrique 0,08 M, ajusté à pH 7 par de la soude (STEPHANO *et al.*, 1986). La décoloration des gels est effectuée dans une solution d'acide acétique 8 % et de méthanol 45 %.

NB : Les gels en présence de SDS sont lavés dans la solution de décoloration 2 fois 15 minutes avant d'être soumis à la coloration afin d'éliminer le SDS qui réagit avec l'acide picrique.

II - 3 COUPURES ENZYMATIQUES DE LA PROTEINE

II - 3-1 Coupure par l'endoprotéase Lys-C

L'endoprotéase Lys-C est très spécifique. Elle coupe uniquement les liaisons Lys-X. La protéine T a été hydrolysée par l'endoprotéase Lys-C, en tampon bicarbonate d'ammonium 0.1 M de pH 7.8 pendant 16 heures à 37°C, avec un rapport enzyme/substrat de 1/25.

II - 3-2 Coupure par la protéase d'*Astacus fluviatilis*

La protéase d'*Astacus fluviatilis* coupe les liaisons dans lesquelles est engagé le groupe α -aminé des résidus à chaîne latérale courte (Ala, Gly, Ser, Thr). L'hydrolyse est réalisée en tampon bicarbonate d'ammonium 0.1 M, pH 7.8 pendant 20 minutes à 30°C avec un rapport enzyme/substrat de 1/50.

II - 3-3 Coupure par la chymotrypsine

La chymotrypsine coupe les liaisons dans lesquelles sont impliqués les groupements carboxyliques des acides aminés aromatiques. Mais il existe des coupures non spécifiques, des liaisons leucyles et méthionyles. La protéine T a été hydrolysée par la chymotrypsine en tampon bicarbonate d'ammonium 0.1 M, pH 8 à 37°C pendant 2 heures avec un rapport enzyme/substrat de 1/50.

- Le peptide C-terminal obtenu par action de l'endoprotéase Lys-C sur la protéine, de même que les variants séparés, ont été hydrolysés par la chymotrypsine dans des conditions limitées, en tampon acétate d'ammonium 0.1 M, pH 5 à 30°C pendant 4 h avec un rapport enzyme/substrat de 1/200.

II - 3-4 Coupure par l'élastase

L'élastase coupe les liaisons dans lesquelles sont impliqués les groupements carboxyliques des acides aminés hydrophobes non aromatiques. L'hydrolyse est réalisée en tampon bicarbonate d'ammonium 0.1 M, pH 8.0 à 37°C pendant 2 heures avec un rapport enzyme/substrat de 1/100.

Chaque hydrolyse est arrêtée en abaissant le pH par addition d'acide formique pur. L'hydrolysats est alors évaporé à sec et fractionné par chromatographie en phase inverse sur colonne de C18 μ Bondapak (Waters).

La nomenclature des peptides obtenus après hydrolyse est fonction de l'enzyme utilisée :

- K pour l'endoprotéase Lys-C
- C pour la chymotrypsine
- A pour la protéase d'*Astacus fluviatilis*

Les peptides sont numérotés suivant leur position dans la séquence de la protéine.

II - 3-5 Traitement par la phosphatase alcaline

La protéine T est isolée du testicule de seiche sous forme phosphorylée. Elle a donc été soumise à l'action de la phosphatase alcaline afin d'éliminer les groupements phosphates de la protéine et de comparer ensuite les hydrolysats enzymatiques de cette protéine avec les hydrolysats homologues de la protéine Sp (protamine extraite du spermatozoïde de la seiche sous forme non phosphorylée).

L'hydrolyse est réalisée à 37°C pendant 4 heures en tampon Tris-HCl 50 mM, pH 8.0 contenant MgCl₂ 1 mM et ZnCl₂ 0.5 mM avec un rapport enzyme/substrat de 0.75 unités/100 µg de protéine. L'hydrolysate est ensuite refroidi à 0°C et, après addition d'HCl 6 M jusqu'à obtention d'une concentration finale de 0.25 M, la protéine est précipitée par addition de 6 volumes d'acétone froid (-20°C). Après centrifugation, le culot est séché et mis en solution dans l'HCl 0.01 M.

II - 4 ANALYSE DES ACIDES AMINES ET SEQUENCE PROTEIQUE

Les échantillons (200 pmoles) sont hydrolysés sous vide à 110°C pendant 24 h et 72 h par l'HCl 5,6 M. Une goutte de phénol à 1 % est ajoutée pour prévenir la dégradation excessive de la tyrosine. L'analyse des acides aminés est réalisée au laboratoire sur l'analyseur d'acides aminés Beckman 6300.

La dégradation automatique d'Edman des variants de la protéine T et des différents peptides a été réalisée sur le microséquenceur phase gaz Applied Biosystems 470A, en utilisant le programme O3RPTH. Les dérivés phénylthiohydantoïnes des acides aminés ont été séparés sur un analyseur de PTH amino acides Applied Biosystems 120A, couplé au microséquenceur (CHAUVIERE *et al.*, 1989).

Les études de spectrométrie de masse ont été réalisées dans le laboratoire du Dr.A.VAN DORSSELAER au Centre de Neurochimie, Strasbourg.

III - ETUDE DE LA PROTÉINE T ET DE LA PROTAMINE DE SEICHE PAR LA BIOLOGIE MOLÉCULAIRE

Cette partie de l'appendice technique consacrée aux techniques de biologie moléculaire a été volontairement détaillée pour permettre la poursuite et le développement du thème dans le laboratoire.

III - 1 ETUDE DES ARN MESSAGERS DU TESTICULE DE SEICHE

III - 1-1 Extraction des ARN messagers cytoplasmiques

L'isolement des ARN messagers se fait en deux étapes :

- la première consiste à extraire les ARN cytoplasmiques totaux à partir du testicule.
- la deuxième correspond à la purification des ARN messagers poly A⁺ sur colonne d'oligo (dT)-cellulose.

Chaque fois que cela est possible, la vaisselle et les tampons sont traités au diéthylpyrocarbonate pendant la nuit, puis autoclavés pendant deux heures afin d'inhiber les ribonucléases.

** Extraction des ARN cytoplasmiques*

Les ARN cytoplasmiques sont isolés par homogénéisation de 7 g de testicule au Waring-blendor, pendant 30 secondes à vitesse rapide, dans 50 ml de solution A (saccharose 0,25 M, KCl 0,025 M, MgCl₂ 0,005 M, NP 40 0,2 % et Tris-HCl 0,05 M, pH 7,5). L'homogénat est ensuite centrifugé à 12.000 xg pendant

10 minutes. Le surnageant est ajusté en EDTA à 2 mM, SDS à 1 % et NaCl à 0,15 M. Les protéines sont extraites trois fois par un volume de phénol saturé en tampon A sans saccharose ni NP 40, puis deux fois par un volume de chloroforme-alcool isoamylique (24 : 1, v : v), et enfin une fois par un volume d'éther. Les ARN contenus dans la phase aqueuse sont précipités pendant 16 heures à -20°C en présence de 0,1 volume d'acétate de sodium 3 M, pH 5,2, par 2,5 volumes d'éthanol absolu refroidi à -20°C. Après une centrifugation de 20 minutes à 12.000 xg, le culot est lavé à l'éthanol à 70 %, séché sous vide, puis remis en solution dans l'eau.

** Sélection des ARN messagers*

La sélection des ARN poly A⁺ est effectuée selon la méthode de IATROU et DIXON (1977) modifiée comme suit :

Une batterie de colonnes d'oligo(dT)-cellulose (Sigma, 1 g par colonne) couplées à des collecteurs de fractions, est équilibrée en tampon B (Tris-HCl 0,01 M, pH 7,5, KCl 0,5 M), à un débit de 0,3 ml/min. Les échantillons sont déposés dans ce même tampon à raison de 20 mg d'ARN par colonne. L'éluat est recyclé trois fois. Le lavage de la colonne est réalisé en continu par 50 ml de tampon B. Une première élution est faite par 30 ml de tampon C (Tris-HCl 0,01 M, pH 7,5). Les ARNm élués sont recyclés une deuxième fois sur la colonne après addition de KCl 3 M pour rétablir les conditions de dépôt. Ils sont ensuite à nouveau élués, puis précipités par l'éthanol, lavés et remis en solution dans l'eau comme indiqué précédemment pour les ARN totaux. Les colonnes sont régénérées par 10 ml de KOH 0,1 M, lavées en tampon C, puis conservées à 4°C en tampon B.

III - 1-2 Traduction *in vitro* des ARN messagers

Les ARN messagers sont traduits *in vitro* en utilisant comme marqueur radioactif la ^{35}S méthionine (SJ 204, Amersham). Le mélange réactionnel est composé de :

- 21 μl de lysat de réticulocytes de lapin (Promega Biotec, traité à la nucléase micrococcale)
- 3,3 μl de ^{35}S méthionine (50 μCi)
- 1 μl d'acétate de potassium 10 mM
- 1,5 μl d'acides aminés non radioactifs sans méthionine (1 mM chaque)
- 1 μl de RNAsine (1 mg/ml, Promega Biotec)
- 5 μl de solution d'ARN contenant :
 - soit 12 μg d'ARN cytoplasmiques
 - soit 1,2 μg d'ARN poly A⁺

Le temps d'incubation est de 1h30 à 30°C.

* Comptage de la radioactivité incorporée

Une aliquote d'1 μl est prélevée sur chaque échantillon. Ces aliquotes sont diluées par 500 μl d'eau distillée, additionnées de 250 μl d'un mélange d'eau oxygénée à 2 % et de soude 1 M. Après une incubation de 10 minutes à 37°C, les protéines sont précipitées par 1 ml d'acide trichloracétique froid à 25 % pendant 10 minutes dans la glace. Le précipité est ensuite collecté par filtration sous vide sur filtre de verre GF/C Whatman, lavé par 20 ml d'acide trichloracétique à 8% froid, puis par 20 ml d'acétone. Après séchage, les filtres sont comptés en scintillation liquide afin d'estimer l'incorporation de ^{35}S méthionine dans les protéines traduites, ce qui reflète l'efficacité de la traduction.

** Traitement des produits de traduction pour l'analyse électrophorétique*

Après traitement par 1 μ l de ribonucléase A (0,01 mg/ml) pendant trente minutes à 37°C, 40 μ g de protéines entraînées d'un extrait HCl 0,4 M préparé comme en II -1-1, sont ajoutés à chaque échantillon. Les tubes sont centrifugés 30 minutes à 12.000 xg. Le surnageant est mis de côté. Les culots sont extraits par 175 μ l d'HCl 0,4 M sous agitation à 4°C pendant 24 à 48 heures en présence de 1 μ l/ml de PMSF, puis les échantillons sont centrifugés à 12.000 xg pendant 30 minutes à 4°C.

Les culots sont rassemblés avec les premiers surnageants. Avant dépôt sur gel de polyacrylamide en présence de SDS, le pH des échantillons est réajusté à 6,8 avec du tampon Tris 1 M, pH 8 (II - 2). Les extraits HCl sont précipités par 1 ml d'acétone à -20°C pendant 24 heures puis centrifugés à 12.000 xg, 30 minutes. Après séchage, les protéines acidosolubles du culot sont reprises par le tampon de dépôt des gels en milieu acide-urée et analysés comme décrit en II - 2.

Les gels sont ensuite séchés sous vide 1 heure à 70°C, puis mis en autoradiographie par contact avec un film X-OMAT AR5 Kodak (Eastman Kodak) entre deux écrans "renforceurs" pendant 5 à 10 jours.

** Comptage de la radioactivité après analyse sur gel d'électrophorèse*

Des comptages de radioactivité peuvent être obtenus après migration et coloration des gels ou même après autoradiographie des échantillons, en découpant une bande protéique identifiée sur le gel. Cette bande de gel est incubée avec 0,2 ml d'acide perchlorique à 60 % et 0,4 ml d'eau oxygénée à 30 %, pendant une heure à température ambiante. Les comptages sont réalisés en présence de liquide scintillant pour milieux aqueux. Un témoin contenant un morceau de gel non radioactif de la même taille que les autres échantillons est utilisé afin de soustraire le bruit de fond de la méthode.

III - 1-3 Enrichissement d'une classe d'ARN messagers sur gradient de saccharose

900 µg d'ARN poly A⁺, dissous dans 2,4 ml de tampon NaCl 1 M, EDTA 0,01 M, Tris-HCl 0,1 M de pH 7,6 sont chauffés pendant 8 minutes à 65°C, puis refroidis dans la glace. Ils sont déposés sur des gradients de saccharose de 10 à 40 % dans le même tampon, à raison de 150 µg par gradient. La centrifugation est réalisée en rotor SW41 Beckman à 40.000 r.p.m. pendant 14 heures à 20°C. Ensuite les tubes sont percés par le bas, reliés à un densitomètre UV à 258 nm connecté à un enregistreur et des fractions de 300 µl sont collectées. Chaque fraction est précipitée par l'éthanol, lavée et séchée. Grâce aux diagrammes obtenus, les fractions communes de chaque tube du rotor sont rassemblées et analysées dans les conditions de traduction décrites précédemment.

III - 2 SOUCHES BACTERIENNES ET VECTEURS DE CLONAGE

* Souches bactériennes

Trois souches bactériennes ont été utilisées lors de la constitution des banques d'ADNc :

- 2 souches ont servi d'hôte au phage λgt 10
 - la souche *E.coli* L87 : *E.coli*, SupE SupF hsdR r⁻k m⁺k trp r met D Ton A.
 - la souche *E.coli* NM514 : *E.coli*, Hsd R514 (r⁻m⁻) Arg H gal E gal X str A lyc B7 (hfl⁺)
- la souche *E.coli* JM109, qui est un hybride d'*E.coli* K12 et *E.coli* B de génotype end A1 : recA1 endA1 gyrA96 thi hsdR17 supE44 relA1 Δ(lac-proAB) / F' [traD36 proA⁺ proB⁺ lacI^q lacZΔM15] a été choisie lors des clonages en pUC13, en M13mp18, M13mp19.

* *Vecteurs de clonage*

- Le phage lambda gt10 (λ srl λ 1°b527srl λ 3° imm434 (srl 434⁺) srl λ 4°srl λ 5°) est utilisé comme vecteur pour les banques d'ADNc.
- Les phages M13 mp18 et M13 mp19 sont utilisés lors du séquençage en simple brin.
- Le plasmide pUC 13 (Lac Z Amp^r pMB1) a été utilisé lors de la création d'une banque d'ADNc pour le séquençage des inserts en double brin et le sous clonage des ADNc obtenus en phages λ gt 10.

* *Enzymes*

Les endonucléases de restriction, l'ADN ligase du phage T4, l'ADN polymérase de Klenow, l'ADN polymérase I, la Taq ADN polymérase, la T₄ polynucléotide kinase, la désoxyribonucléase libre d'activité ribonucléasique, les ribonucléases, la protéinase K et le lysozyme proviennent de chez Appligène, Biolabs, Boehringer, BRL ou Promega Biotec, et sont utilisées selon les recommandations prescrites par le fabricant.

III - 3 CREATION DE BANQUES D'ADN COMPLEMENTAIRE EN PLASMIDE
ET EN PHAGE

L'ensemble de ces manipulations est réalisé avec les kits Amersham (Réf : RPN 1256 et RPN 1257), ce qui limite la description de ce paragraphe aux principes des manipulations.

III - 3-1 Synthèse d'ADNc

Les ADNc sont synthétisés selon la méthode de GUBLER et HOFFMAN (1983) :

- La transcriptase-inverse est utilisée pour synthétiser le premier brin à partir des ARN messagers enrichis sur gradient de saccharose. Deux types d'amorces (de l'oligo(dT) ou des hexanucléotides au hasard) ont été utilisées pour initier la synthèse. De l' $[\alpha^{32}\text{P}]\text{dCTP}$ a été utilisé pour suivre les rendements réactionnels.
- Le deuxième brin est synthétisé en faisant agir simultanément la ribonucléase H qui digère l'ARNm et libère des amorces afin de permettre à l'ADN polymérase I d'*E.coli* de synthétiser le deuxième brin.

Ensuite l'ADN polymérase T4 élimine les extrémités 3'OH en excès du premier brin.

Deux extractions au phénol-chloroforme, puis une extraction au chloroforme sont effectuées. Les échantillons sont ensuite précipités par l'éthanol pour éliminer les nucléotides libres restants.

III - 3-2 Analyse électrophorétique des ADNc

* *Gel de polyacrylamide*

Les ADNc double brin sont déposés sur gel non dénaturant de polyacrylamide 5 % (acrylamide/bisacrylamide : 29/1) en TBE 1x (Tris-HCl 0.089 M, borate 0.089 M, EDTA 2 mM). Le tampon de dépôt est composé de saccharose à 40% et de bleu de bromophénol à 0.25 %.

* *Marquage des témoins de masse moléculaire*

L'ADN du vecteur pUC 13 a été coupé par les enzymes de restriction suivantes :

- 1 - Hind III - Eco RI
- 2 - Hind III - Bam HI
- 3 - Taq I
- 4 - Ava II

Une extraction phénol-chloroforme, puis une précipitation à l'éthanol sont ensuite réalisées. Les échantillons sont marqués à leurs extrémités 3' par l' $[\alpha^{32}\text{P}]$ dCTP après une incubation d'une heure à 37°C dans les conditions suivantes :

- 0.6 μg d'ADN
- 1 μl de tampon 10x du fragment Klenow de l'ADN polymérase I
- 3 fois 1 μl de dATP, dTTP, dGTP en solution 20 mM.
- 0,8 μl d' $[\alpha^{32}\text{P}]$ dCTP (3000 Ci/mM)
- 1 unité de fragment Klenow de l'ADN polymérase I.

Une fois le marquage effectué, ces échantillons vont servir de témoins de masse moléculaire lors de la migration des ADNc.

III - 3-3 Clonage de l'ADNc en phage λ gt 10 et en plasmide pUC 13

L'ADNc synthétisé est mis en solution dans du tampon TE (Tris 10 mM, EDTA 1 mM, pH 8.0), traité par l'ADN méthylase d'Eco RI, afin de protéger les sites Eco RI potentiels de l'insert. Des fragments d'ADN phosphorylés en 5' renfermant des sites Eco RI sont fixés à l'ADNc par l'ADN ligase du bactériophage T4. Les extrémités cohésives sont libérées par action de l'enzyme Eco RI. L'ADNc à extrémités cohésives est purifié par chromatographie d'exclusion sur colonne de Sephadex G-10 équilibrée et éluée par du tampon TE. Les fractions contenant l'ADNc radiomarqué sont rassemblées et précipitées à l'éthanol.

L'ADNc est ensuite introduit :

- soit dans le site Eco RI du phage λ gt 10. Après ligation, l'ADN phagique est empaqueté *in vitro* et conservé à 4°C, dans 500 μ l de tampon SM (NaCl 0.1 M, MgSO₄ 8 mM, Tris-HCl 50 mM, pH 7.5, gélatine 0.01 %) avec 10 μ l de chloroforme.
- soit dans le site Eco RI du plasmide pUC 13. Après ligation, l'ADN plasmidique est alors introduit dans la bactérie JM109, puis étalé sur boîte de culture comme décrit en III - 4-1.

III - 4 CRIBLAGE DE LA BANQUE D'ADNc PLASMIDIQUE

III - 4-1 Méthode de préparation d'ADN plasmidique

** Obtention de bactéries compétentes*

Un clone de la souche JM109 isolé sur boîte de Pétri est mis en culture une nuit à 37°C dans 20 ml de LB (extrait de levure DIFCO : 5 g/l, peptone DIFCO : 10 g/l, NaCl : 5 g/l). Le lendemain, 1 ml de cette culture est prélevé et inoculé dans 50 ml de LB, mis sous agitation à 37°C pendant environ 3 heures jusqu'à obtention d'une $DO_{550nm} = 0,2$. Les bactéries en phase exponentielle sont centrifugées à 4°C à 5000 xg pendant 10 minutes. Le culot est mis en suspension dans 20 ml de CaCl₂ 0.1 M froid stérile, puis remis à centrifuger dans les mêmes conditions. Le surnageant est ensuite éliminé et 20 ml de CaCl₂ 0.1 M sont ajoutés. Les bactéries en suspension sont alors laissées à 4°C pendant 20 minutes puis centrifugées une dernière fois, avant d'être reprises dans 1 ml de CaCl₂ 1 M et conservées à 4°C.

** Transformation bactérienne*

20 μ l d'ADN (0.5 μ g) sont mis en contact pendant 20 minutes au moins avec 100 μ l de bactéries compétentes. Puis 5 chocs thermiques de 37°C à 4°C sont effectués et la solution est laissée 20 minutes dans la glace. 1 ml de milieu 2XTY (extrait de levure DIFCO : 10 g/l, tryptone DIFCO : 16 g/l, NaCl : 5 g/l) est ensuite ajouté et le tout est mis sous agitation à 37°C pendant une heure.

0.2 ml et 0.8 ml de ce mélange sont étalés sur deux boîtes de Pétri de milieu LB gélosé (Bactoagar 15 g/l DIFCO à rajouter au LB) contenant 50 μ g/ml d'ampicilline, 80 μ g/ml de Bluogal en diméthylformamide (DMFA) et 80 μ g/ml d'IPTG (Isopropyl- β -D-thiogalactopyranoside). Les boîtes sont placées une nuit à 37°C. Les clones de couleur blanche indiquent l'inactivation de la β -galactosidase par l'insertion d'un ADNc. Par contre, lorsque l'enzyme est active, les clones sont de couleur bleue.

** Amplification et purification d'ADN plasmidique*

Les clones sélectionnés sont tout d'abord isolés sur boîte de Pétri à 37°C. Puis ils sont mis en culture dans 20 ml de milieu LB contenant 30 μ g/ml d'ampicilline, sous agitation à 37°C pendant 16 heures. 1 ml de suspension est ajouté à 25 ml de LB-ampicilline et mis en culture pendant 4 heures. 5 ml de cette solution sont alors ajoutés à 300 ml de LB-ampicilline pour une culture d'une nuit.

La culture est ensuite centrifugée 10 minutes à 6000 \times g à 4°C et le culot est repris par 3 ml de solution Tris-HCl 50 mM, pH 8, saccharose 25 %. Les tubes sont placés dans la glace et 1 ml de lysozyme (à 10 mg/ml en Tris-HCl 50 mM, pH 8) est ajouté. Après 10 minutes, on introduit 1 ml d'EDTA 0,2 M, pH 8 et on laisse agir pendant 10 minutes. Puis 4,5 ml d'un mélange de Tris 50 mM, EDTA 25 mM et Triton X-100 à 0,1 % sont ajoutés.

Au bout de 10 minutes, les tubes sont centrifugés à 150.000 xg pendant 45 minutes à 4°C. 9 ml de surnageant sont mélangés à 9 g de CsCl et à 90 µl de bromure d'éthidium à 5 mg/ml. Après 20 heures de centrifugation à 300.000 xg à 16°C, on récupère la bande d'ADN plasmidique. Le bromure d'éthidium est extrait 4 fois par 1 volume d'isopropanol (saturé en CsCl et en H₂O). Les échantillons sont ensuite dialysés contre deux litres de tampon TE, 4 fois 20 minutes puis 1 heure. L'ADN est ensuite conservé à 4°C.

III - 4-2 Sélection des clones positifs par hybridation et traduction d'ARNm

La sélection des clones s'effectue par hybridation d'ARNm sur des ADNc fixés sur des membranes de nitrocellulose. Les ARNm hybridés sont ensuite traduits *in vitro* en système de lysat de réticulocyte de lapin.

Après transformation bactérienne des ADNc intégrés en vecteur pUC 13, tous les clones obtenus sur boîtes sont isolés et amplifiés sur boîtes. Chaque clone est répertorié puis amplifié en milieu liquide afin de conserver une aliquote de chaque culture, à -20°C en présence de 60% de glycérol.

* Préparation de l'ADN

Des mélanges de cinq cultures liquides sont effectués afin de préparer un ensemble de 5 ADN plasmidiques, comme décrit dans le paragraphe III - 4-1. Pour chaque mélange, 60 µg d'ADN plasmidique sont linéarisés en présence de l'enzyme Hind III (une aliquote est déposée sur gel d'agarose pour vérifier la digestion). Puis deux extractions au phénol saturé en TE et une extraction au chloroforme-alcool isoamylique,

sont réalisées. Enfin, l'ADN est précipité à l'éthanol. Après centrifugation (12.000 xg, 15 minutes), rinçage à l'éthanol à 70% et séchage, les échantillons sont repris par 40 μ l d'eau distillée. Une mesure de densité optique est effectuée pour pouvoir appliquer avec exactitude 20 μ g d'ADN sur filtre de nitrocellulose (ou Hybond C, Amersham) à une concentration de 500 μ g/ml. Les échantillons sont préalablement portés à 100°C, 10 minutes, puis refroidis dans la glace. La dénaturation est complétée par traitement à la soude 1 M (1 volume), pendant 20 minutes à température ambiante. La solution est ensuite neutralisée par l'addition d'un demi-volume de tampon NaCl 1 M, sodium citrate 0.3 M, Tris-HCl 0.5 M (pH 8.0) dans l'HCl 1 M.

** Préparation des filtres*

Les échantillons sont déposés par fractions de 5 μ l sur des filtres de nitrocellulose triangulaires de 3 mm² environ pouvant occuper le fond d'un tube Eppendorf. Pour accélérer l'adsorption, les filtres sont placés à l'étuve à 42°C entre chaque dépôt. Lorsqu'ils sont secs, ils sont lavés deux fois pendant 15 minutes par 50 ml de SSC 6X sous agitation, puis séchés pendant 1 heure à température ambiante. La fixation covalente de l'ADN est réalisée par chauffage à 80°C pendant deux heures. Lors de leur première utilisation, ils sont dénaturés pendant une minute à 100°C en présence d'1 ml d'H₂O, puis transférés dans d'autres tubes Eppendorf.

** Hybridation*

Pour chaque filtre, on prépare 500 μ l de solution d'hybridation contenant 50 μ g d'ARN poly A⁺, 65 % (v/v) de formamide désionisée, 20 mM d'acide 1,4 pipérazine diéthanesulfonique

(pH 6.4), 0,2 % SDS, 0,4 M NaCl et 10 µg/ml d'ARN de transfert de foie de Veau (Boehringer). Cette solution est chauffée 10 minutes à 70°C puis incubée avec chaque filtre à 50°C pendant 3 heures.

* *Lavages*

Après ce temps, les filtres sont lavés 12 fois par 1 ml de tampon Tris-HCl 10 mM, pH 7.6, NaCl 0,15 M, EDTA 1 mM et SDS 0,5 %, à 65°C. Les filtres sont agités au Vortex 10 secondes dans la solution de lavage à chaque passage. Puis, on effectue 8 lavages dans les mêmes conditions mais sans SDS.

* *Elution*

Enfin, les filtres sont transférés dans d'autres tubes pour l'éluion qui est réalisée en présence de 300 µl d'H₂O et de 30 µg d'ARN_t par filtre. Le tout est porté à 100°C pendant 1 minute, puis brusquement refroidi dans un bain de carbo-glace et d'éthanol. La décongélation des tubes a lieu dans la glace, puis les filtres sont mis de côté et l'éluat est extrait par 1 volume de phénol-chloroforme. Les ARN sont précipités par 1 ml d'éthanol et 60 µl d'acétate de sodium 2 M pH 5.2, pendant la nuit à -20°C. Les échantillons, après centrifugation et remise en solution dans 5 µl d'eau, sont prêts pour la traduction *in vitro*, comme décrit dans le paragraphe III -1-2.

III - 5 TRANSFERT D'ARNm ET HYBRIDATION

III - 5-1 Marquage d'ADN avec une amorce aléatoire

Nous utilisons le kit Boehringer pour réaliser ces réactions (Cat n°100476) Le brin complémentaire de la sonde est fabriqué à partir de l'extrémité 3'OH d'un hexanucléotide qui sert d'amorce au fragment Klenow de l'ADN polymérase I.

Un fragment de restriction est obtenu à partir d'un gel d'agarose à bas point de fusion. La partie du gel contenant l'ADN à marquer est découpée et transférée dans un tube Eppendorf. Le morceau de gel est pesé afin de pouvoir rajouter un rapport d'eau de 3 ml par gramme de gel, puis l'échantillon est porté à 100°C pendant 7 minutes et des aliquotes de 25 µg sont prélevées.

La réaction est réalisée à partir de 25 mg d'ADN, dénaturé 3 minutes à 100°C, incubé en présence de :

- 0,5 nmole/l de dATP, dGTP, dTTP.
- 2 µl de solution tamponnée d'hexanucléotides.
- 5 µl d' [α³²P] dCTP (50 µCi)
- 1 unité d'ADN polymérase I (fragment Klenow)

Après un minimum de 30 minutes à 37°C, la réaction est arrêtée par addition de 2 µl d'EDTA 0,2M pH 8 ou par chauffage 10 minutes à 65°C.

III - 5-2 Transfert d'ARN et hybridation

Afin d'éviter des problèmes avec les ribonucléases, tout le matériel d'électrophorèse est traité à la soude 1 M pendant 2 heures, puis rincé abondamment avec de l'eau traitée au DEPC (diéthylpyrocarbonate).

Un gel d'agarose à 1,4 % est réalisé en tampon MOPS 1X (Sigma) contenant du DEPC à 0.01 % et de la formamide désionisée préalablement avec de la résine Bio-Rad AG501.

NB : Le tampon MOPS 10X correspond à un mélange d'acide 3-(N-morpholino) propane sulfonique 0,2 M, d'acétate de sodium 50 mM et d'EDTA 10 mM.

Les échantillons d'ARN à déposer sont dénaturés à 68°C pendant 5 minutes en présence de 30 µl de tampon de dénaturation (formamide désionisée 50 % ,v/v ; formaldéhyde 17,8 %,v/v; tampon MOPS 1X) et refroidis dans la glace. Ils sont déposés avec 10 µl de tampon de dépôt (bleu de bromophénol 0,25 % ; xylène cyanol 0,25% et glycérol 30 %). L'électrophorèse est effectuée durant la nuit à 40 volts en tampon MOPS 1X.

- La partie de gel qui contient les témoins est colorée au bromure d'éthidium à 10 mg/ml, pendant 15 minutes à l'obscurité et sous agitation. Puis le gel est décoloré une nuit dans un grand volume d'eau traitée au DEPC. Une photographie polaroïde est prise sous UV.

- La partie du gel qui contient les échantillons à transférer est lavée par 250 ml de SSC 10X (NaCl 1,5 M, citrate de sodium 0,15 M) pendant 15 minutes. Le transfert sur membrane de nitrocellulose est réalisé par capillarité en présence de SSC 20X pendant la nuit. Le lendemain, la membrane est séchée et l'ARN fixé pendant une heure à 80°C.

La membrane est mise ensuite en préhybridation, pendant 4 heures à 42°C en tampon d'hybridation (SSC 6X, formamide désionisée 50 %, SDS 0,5 %, ADN de sperme de hareng 0,2 mg/ml, préalablement soniqué et dénaturé par la chaleur). Par filtre à hybrider, on prépare $\sim 4.10^7$ cpm de sonde marquée comme décrit dans le paragraphe précédent. Cette sonde est préalablement dénaturée en présence de soude 0.02 M par chauffage à 100°C pendant 5 minutes. Puis elle est diluée en tampon d'hybridation et hybridée une nuit à 42°C sous agitation.

Eusuite, le filtre est lavé par 500 ml de tampon SSC 2X et SDS à 1 % pendant 2 minutes sur chaque face sous agitation. Deux autres lavages de 30 minutes à 42°C sous agitation sont réalisés en tampon SSC 0,1X et SDS à 0,5 %. Puis le filtre encore humide est recouvert d'un film plastique et mis en autoradiographie à -70°C. Le film est révélé après 72 heures.

III - 6 CRIBLAGE DE LA BANQUE DE PHAGES

III-6-1 Méthodes de préparation de phages recombinants λ gt 10

** Préparation des bactéries pour l'infection par les phages*

Une colonie bactérienne L87 ou NM514 est inoculée dans 10 ml de LB contenant 0.4 % de maltose. 1 ml est alors prélevé et ajouté à 50 ml de LB contenant 0.4% de maltose pour être mis sous agitation à 37°C, jusqu'à atteindre une DO_{600nm} de 0.5. La culture est alors refroidie dans la glace et centrifugée à 4000 xg 10 minutes à 4°C. Le culot bactérien est resuspendu en 10 mM $MgSO_4$ froid et peut être conservé à 4°C pendant 2 jours au maximum.

** Culture de phages sur boîte de Pétri*

Par cette méthode, le titre de la banque peut être calculé en effectuant une gamme de dilutions à partir des 500 μ l de SM contenant les phages. Une aliquote est alors prélevée et diluée en SM, dans des dilutions successives au 1/10e. 100 μ l de chacune de ces dilutions sont mis en présence de 100 μ l de bactéries préparées pour l'infection (NM514 ou L87) et incubés 20 minutes à 37°C. 4 ml de Top agar liquide à 45°C (Bactotryptone DIFCO 10 g/l, extrait de levures 5 g/l, NaCl 5 g/l, MgSO₄ 2,5 g/l, Bactoagar 10 g/l) sont ajoutés et le mélange est versé rapidement sur une boîte de Pétri (préalablement séchée à 37°C) contenant du LB gélosé. Lorsque le Top agar est pris en masse, les boîtes sont incubées une nuit à 37°C. Le lendemain, le nombre de plages de lyse est compté pour obtenir le titre de la banque.

** Culture rapide de phages en milieu liquide*

Une plage de lyse est mise en présence de 0,5 ml de bactéries "compétentes" et laissée 15 minutes à température ambiante. Puis on ajoute 5 ml de LB contenant 5 mM de CaCl₂. La solution est mise sous agitation pendant 4 heures à 37°C. Après ce temps, les bactéries sont complètement lysées après l'adjonction d'une goutte de chloroforme et une agitation de 5 minutes. Enfin, la culture est centrifugée 10 minutes à 4000 xg et le surnageant qui renferme les phages est conservé à 4°C.

III - 6-2 Hybridation des réplicats de boîtes de phages sur filtre de nitrocellulose

Pour détecter les clones possédant l'insert spécifique du gène recherché, des plages de lyses du phage λ gt 10 sont obtenues sur boîtes comme décrit en III - 6-1.

Les boîtes sont recouvertes de membranes de nitrocellulose BA 85 (Schleicher & Schüll) pendant 30 secondes. Les membranes sont ensuite enlevées et mises à sécher à température ambiante sur papier Whatman 3 MM. Puis, elles sont successivement placées dans trois cuves. La première contient du tampon (NaOH 0.2 M, NaCl 3 M). Le filtre est déposé sur la surface du tampon pendant 2 minutes, puis immergé dans le tampon, 3 minutes. Ensuite, il subit le même traitement dans un tampon Tris-HCl 0.5 M, NaCl 3 M et enfin il est immergé en tampon SSC 2X pendant 5 minutes. Les membranes sont préséchées sur papier Whatman 3 MM pendant 1 heure à température ambiante, puis placées pendant 2 heures dans une étuve à 80°C.

Les membranes sont ensuite lavées en SSC 2X avant d'être mises en sachet plastique contenant du SSC 0.1X et du SDS 0,5 %, pendant 1 heure à 65°C. La solution est ensuite éliminée et une préhybridation est effectuée pendant 2 heures à 42°C en tampon H (SSC 5X, Denhardt's 5X, phosphate de sodium 0,02 M, ADN de sperme de hareng 100 µg/ml, formamide 50 %).

NB : La solution Denhardt's 10X est composée de 2 % de Ficoll 400, 2 % de polyvinylpyrrolidone et 2 % de sérum albumine bovine.

Après ce temps, la solution est éliminée et une hybridation de 20 heures à 42°C est effectuée en tampon H. La sonde radioactive est ajoutée comme décrit en III - 5-2.

Les filtres sont ensuite lavés en SSC 2X, SDS 0,1% afin d'éliminer la solution d'hybridation résiduelle, puis ils sont incubés deux fois pendant 30 minutes en tampon SSC 6X, SDS 0,1 %, à 65°C et enfin 15 minutes en tampon SSC 2X, SDS 0,1 % à 65°C. Les filtres encore humides sont ensuite enveloppés de film plastique et mis en autoradiographie à -70°C.

III - 6-3 Détermination rapide de la taille des inserts

A 5 ml de culture liquide de phages, on ajoute 5 ml d'une solution de SM contenant du PEG (polyéthylène glycol) à 20 % et du NaCl 2 M. Le mélange est laissé 1 heure dans la glace puis centrifugé à 4000 xg, 20 minutes. Le surnageant est complètement éliminé et le culot est repris par 750 µl de LB puis transféré dans un tube Eppendorf contenant 750 µl de résine DEAE-cellulose 52 (Whatman) préalablement équilibrée en LB. Le tube est mélangé doucement par retournement 20 à 30 fois. On le centrifuge alors pendant 5 minutes à 12.000 xg. Le surnageant est récupéré et subit une deuxième fois le même traitement. On ajoute ensuite 32 µl de SDS à 10 % et 13 µl d'une solution de protéinase K à 0.1 mg/ml. Le mélange est incubé, 5 minutes à température ambiante. On ajoute 130 µl d'acétate de potassium 3 M. On incube à 88°C pendant 20 minutes et on refroidit le tube dans la glace pendant 10 minutes. Après une centrifugation de 10 minutes à 12.000 xg, on récupère le surnageant qui est précipité par un volume d'isopropanol froid 10 minutes à -70°C, puis centrifugé 10 minutes à 12.000 xg. Le culot est séché, resuspendu dans 20 µl de T E.

Une digestion enzymatique est alors effectuée pendant 2 heures à 37°C par 6 unités d'enzyme EcoRI afin de libérer l'insert. Pour visualiser ce fragment séparé par électrophorèse sur gel d'agarose à 2 %, en TBE 1X, ce dernier est marqué avec [$\alpha^{32}\text{P}$] dATP par le fragment Klenow de l'ADN polymérase I : on met 1 unité d'enzyme et 1 µCurie de nucléotide pour 10 µl d'échantillon. Le marquage s'effectue à 30°C pendant 30 minutes.

Le gel contenant les échantillons est ensuite fixé à l'acide trichloracétique à 7 %, séché puis autoradiographié.

III - 7 SOUS CLONAGE EN PLASMIDE pUC 13 ET EN PHAGE M13

III - 7-1 Sous clonage de phage λ gt 10 en plasmide pUC 13

** Préparation d'ADN de phage λ gt 10*

Les bactériophages recombinants sélectionnés sont amplifiés sur boîte de Pétri de 90 mm une nuit à 37°C pour obtenir 10⁵ pfu/ml (pfu = plaque forming unit) (il est également possible d'amplifier en milieu liquide). 5 ml de solution SM sont appliqués sur chaque boîte et laissés en contact à 4°C pendant un minimum de 3 heures. La solution est ensuite récupérée et les boîtes rincées par 1 ml de SM. Quelques gouttes de chloroforme sont ajoutées au liquide et le mélange est brièvement agité au Vortex. La solution aqueuse est centrifugée à 4000 xg, 10 minutes à 4°C. Le surnageant est conservé à 4°C avec une goutte de chloroforme.

Des traitements à la ribonucléase A et à la déoxyribonucléase à des concentrations de 1 µg/ml peuvent être effectués afin d'éliminer les constituants bactériens. La solution est ensuite précipitée pendant 1 heure à 4°C, au PEG-NaCl à des concentrations respectives de 10 % et 2,5 M, puis centrifugée 45 minutes à 8000 xg. Le culot de phages obtenu est repris par 4 ml de SM et extrait par le même volume de chloroforme. 0,75 g de chlorure de césium sont ajoutés par ml de surnageant et une ultracentrifugation à l'équilibre est effectuée pendant 16 heures à 300.000 xg.

La bande phagique est récupérée et dialysée 4 fois pendant 20 minutes et une fois pendant 1 heure contre du tampon Tris-HCl 10 mM, pH 7.4, MgCl₂ 10 mM, NaCl 50 mM. L'ADN du phage est obtenu après traitement à la protéinase K (50 µg/ml) en présence

de SDS 0,5 % et d'EDTA 20 mM pendant 1 heure à 65°C et après extraction au phénol-chloroforme. Enfin, l'ADN est précipité à l'éthanol.

** Sous clonage*

L'ADN phagique est digéré par Eco RI afin de libérer l'ADNc. Il est déposé sur gel et, après électrophorèse, l'insert est fixé sur membrane de DEAE-cellulose (NA 45, Schleicher & Schüll). Il est ensuite élué par 400 µl de tampon d'éluion (NaCl 1,5 M, Tris-HCl 20 mM pH 7.5, EDTA 1 mM) porté à 68°C pendant 20 minutes, puis extrait par un volume de phénol-chloroforme. Parallèlement, le plasmide pUC 13 est linéarisé par Eco RI .

Le vecteur et l'insert sont alors mélangés dans un rapport molaire de 1/10 avant d'être précipités à l'éthanol. Après lavage à l'éthanol à 70 %, séchage, remise en solution dans 30 µl d'eau, une microdialyse sur filtre Millipore (0,025 µm, référence VSWP01300) est réalisée pendant 20 minutes. La réaction de ligation est alors effectuée avec l'ADN ligase du phage T4 (BRL), une nuit à 16°C. La solution plasmidique est ensuite diluée 4 fois avec de l'eau pour la transformation en JM109 (paragraphe III - 4-1).

** Minipréparation d'ADN plasmidique*

Nous avons utilisé la méthode décrite par SERGHINI *et al.* (1989). Elle a la particularité d'utiliser une précipitation de l'ADN par deux volumes d'éthanol à 95 % froid, en présence d'acétate d'ammonium 2 M (concentration finale), pendant 15 minutes dans la glace.

III - 7-2 Sous clonage de pUC 13 en phages M13 mp18 et M13 mp19

* *Conditions de sous-clonage*

Elles sont les mêmes que celles décrites précédemment en utilisant cette fois le vecteur phagique M13 mp18 et M13 mp19.

* *Transformation bactérienne*

Le protocole est le même que pour l'ADN plasmidique (décrit en III - 4-1) jusqu'à 20 minutes après les 5 chocs thermiques. L'échantillon est alors divisé en deux (1/5e et 4/5e). Ensuite sont ajoutés dans chaque tube, 35 μ l de bluogal (à 20 mg/ml en DMFA, 35 μ l d'IPTG (à 24 mg/ml), 200 μ l de bactéries JM109 en phase exponentielle et 3 ml de Topagar à 45°C. Leur contenu est mélangé et versé sur des boîtes de Pétri de LB gélosé pré-séchées. Après solidification, les boîtes sont placées à 37°C pendant 16 heures.

* *Préparation d'ADN de phage M13*

Une plage de lyse est prélevée à partir d'une boîte et inoculée dans 1,5 ml d'une culture de JM109 en phase exponentielle diluée au 1/100. Le tout est mis sous agitation à 37°C. Après 6 heures, la solution est mise à centrifuger à 12.000 xg pendant 5 minutes à 4°C. A partir des culots bactériens qui contiennent le forme réplivative de l'ADN phagique, une vérification des inserts sous-clonés est effectuée par la méthode de minipréparation décrite en III - 7-1.

La purification d'ADN de phage M13 en grandes quantités peut alors être réalisée comme suit. Le surnageant (1,5 ml) contenant la forme simple brin est amplifié par infection de

la souche JM109 (5 ml de bactéries en phase exponentielle dans 200 ml de LB) sous agitation, 16 heures à 37°C. Après centrifugation de la culture à 6000 xg, 10 minutes, à 4°C, une préparation d'ADN double brin est réalisée sur le culot comme en III - 4-1. Sur le surnageant, une précipitation d'une heure au PEG 6000 à 4 % et NaCl 0.5 M permet de récupérer après une centrifugation de 30 minutes à 12.000 xg, les phages contenant l'ADN simple brin. Ils sont repris par 400 µl de TE et l'ADN est obtenu après trois extractions au phénol saturé en TE deux extractions au chloroforme-alcool isoamylique, puis deux extractions à l'éther. Après précipitation à l'éthanol, l'ADN est analysé sur gel d'agarose à 0.6 %.

III - 8 SEQUENCAGE DE L'ADN EN DOUBLE ET SIMPLE BRIN

III - 8-1 Principe du séquençage d'ADN par la méthode de SANGER et COULSON (1975)

Cette méthode s'appelle la méthode de terminaison des chaînes car si on incube une amorce oligonucléotidique en présence d'un ADN simple brin, d'ADN polymérase I, des quatre dNTP et d'un ddNTP (le ddTTP par exemple), il y aura arrêt de l'élongation des chaînes, chaque fois qu'un ddTTP sera ajouté à la place d'un dTTP. Si l'on fait 4 incubations différentes de dNTP en présence d'un des quatre ddNTP particulier, pour chaque mélange réactionnel, on obtiendra dans chacun des tubes des fragments qui se termineront par le ddNTP qu'ils contiendront. Il suffit que l'un des quatre dNTP présents soit marqué radioactivement pour visualiser le brin d'ADN néosynthétisé.

III - 8-2 Séquençage à partir de matrice double brin avec le vecteur pUC 13

Cette méthode a été utilisée selon le principe de SANGER et COULSON (1975) avec la transcriptase inverse.

** Gel de séquence*

Il est constitué de 8% d'un mélange d'acrylamide et de bisacrylamide (38%, 2%) en urée 7 M, TBE 1X. La polymérisation se fait en présence de 250 μ l de persulfate à 10% et de 50 μ l de Temed pour 50 ml de gel. L'épaisseur du gel est de 0,3 mm. Le tampon d'électrophorèse est du TBE 1X.

** Préparation d'ADN plasmidique*

1.3 μ g d'ADN dans 8 ml d' H_2O sont dénaturés par 2 μ l de NaOH 2 M pendant 5 minutes, à température ambiante. Puis 7 μ l d'amorce (17 mCi à 2 μ Mol/ μ l) (Boehringer), 3 μ l d'acétate de sodium 3 M pH 5.2 et 75 μ l d'éthanol à 100% sont ajoutés. Les échantillons sont alors placés dans un bain de carboglycérol-éthanol, pendant 20 minutes.

Après une centrifugation (15 minutes, 12.000 xg, 4°C) un rinçage est effectué par 200 μ l d'éthanol à 70%, puis les échantillons sont centrifugés 5 minutes dans les mêmes conditions. Le culot est séché et repris par 11 μ l d' H_2O .

** Réaction de séquençage*

Des mélanges de ddNTP/dNTP sont préalablement effectués comme indiqué dans le tableau ci-après, à partir des solutions mères de ddNTP 5 mM et dNTP 10 mM. Chaque mélange est symbolisé par la lettre du ddNTP correspondant. 1,2 μ l de chaque mélange est placé dans les tubes A, G, C, T.

Composant	mélanges de nucléotides			
	A	G	C	T
ddATP	126 μ M			
ddCTP			1.26 μ M	
ddGTP		12.6 μ M		
ddTTP				12.6 μ M
dATP	60 μ M	250 μ M	250 μ M	250 μ M
dCTP	10 μ M	10 μ M	10 μ M	10 μ M
dTTP	250 μ M	250 μ M	60 μ M	250 μ M
dGTP	250 μ M	250 μ M	250 μ M	60 μ M

Un prémélange est préparé pour toutes les réactions de séquence en conservant les proportions.

- 4 μ l de tampon transcriptase-inverse :

Tris-HCl	300 mM pH 8.3
NaCl	375 mM
MgCl ₂	37,5 mM
DTT	2,5 mM

- 4 μ l de [α ³²P] dCTP (3000 Ci/mmol)

- 0.4 μ l de DTT 50 mM

- 1.6 μ l d'H₂O

- 20 unités de transcriptase-inverse

Les 11 μ l d'ADN sont placés dans 9 μ l de prémélange. Puis 4 fois 4 μ l sont prélevés pour être répartis dans chaque tube: A, G, C, T.

Après 10 minutes à 40°C, sont ajoutés dans chaque tube 1.5 μ l de solution de chasse contenant 0.25 mM de chacun des quatre dNTP. Puis, les tubes sont remis à 44°C pour 8 minutes au bout desquelles on ajoute 4 μ l de tampon d'arrêt (Formamide désionisée en solution dans de l'EDTA 10 mM pH 8, xylène cyanol 0.3% et bleu de bromophénol 0.3%). La solution est dénaturée, 3 minutes à 100°C, puis placée dans la glace. 2 μ l de chaque échantillon sont déposés sur le gel de séquence.

III - 8-3 Séquençage à partir de matrice simple brin avec les vecteurs M13 mp18 et M13 mp19

Nous avons utilisé le kit "Taq Track™ Sequencing Systems" (Promega Biotec). L'enzyme utilisée est la Taq ADN polymérase qui permet une synthèse à de hautes températures. L'amorce utilisée pour la réaction est un 24 mère. Un 7-Deaza dGTP a été substitué au dGTP dans les mélanges de nucléotides A,G,C,T et les nouveaux mélanges de d/ddNTP sont inscrits dans le tableau suivant :

Composant	mélanges de nucléotides			
	A	G	C	T
ddATP	350 µM			
ddGTP		25 µM		
ddCTP			160 µM	
ddTTP				300 µM
dATP	25 µM	250 µM	250 µM	250 µM
7-deaza dGTP	250 µM	25 µM	250 µM	250 µM
dCTP	250 µM	250 µM	250 µM	250 µM
dTTP	250 µM	250 µM	250 µM	25 µM

L'ADN de phage M13 a été préparé comme décrit en III -7-2. La première étape est l'hybridation de l'amorce sur l'ADN simple brin. Elle se fait en incubant 10 minutes à 37°C :

- 0.8 pMole d'ADN simple brin (~ 2 µg d'ADN)
- 1.0 pmole d'amorce (~ 3 ng du 24 mère)
- 5 µl de tampon Taq ADN polymérase 5X
- 2 µl de mélange d'extension (7.5 µM de dGTP, dTTP, dCTP)

La deuxième étape est la réaction d'extension et de marquage où l'on rajoute au tube précédent :

- 2 μl soit 20 μCi de [$\alpha^{35}\text{S}$]dATP (1000 Ci/mmol)
- 1.5 μl de Taq ADN polymérase (5 unités/ μl)

La réaction s'effectue pendant 5 minutes à 37°C.

La troisième étape est la réaction de terminaison qui consiste à prélever des aliquotes de 6 μl du mélange réactionnel précédent pour les répartir dans des tubes A G C T contenant 1 μl des mélanges de d/dNTP appropriés. Les tubes sont alors incubés 5 minutes à 71°C. Après ce temps, la synthèse est arrêtée par 4 μl de solution stop (composition identique à la précédente). Les échantillons sont chauffés 5 minutes à 90°C, puis 3 μl sont déposés sur gel pour chacun d'entre eux.

B I B L I O G R A P H I E

- AIKEN, J.M., ZHAO, H.Z., STATES, J.C. and DIXON, G.H. (1983)
Sequence homologies in the protamine gene family of rainbow trout.
Nucl. Acid. Res. 11, 4907-4922.
- AMMER, H., HENSCHEN, A. and LEE, C.H. (1986)
Isolation and amino acid sequence analysis of human sperm protamines P1 and P2. Occurrence of two forms of protamine P2.
Biol. Chem. Hoppe-Seyler 367, 515-522
- ANDO, T., IWAI, K., ISHII, S., AZEGAMI, M. and NAKAHARA, C. (1962)
The chemical structure of one component of clupeine.
Biochim. Biophys. Acta 56, 628-630.
- ANDO, T. and SUZUKI, K. (1966)
The amino acid sequence of the second component of clupeine.
Biochim. Biophys. Acta 121, 427-429.
- ANDO, T. and SUZUKI, K. (1967)
The amino acid sequence of the third component of clupeine.
Biochim. Biophys. Acta 140, 375-376.
- ANDO, T. and WATANABE, S. (1969)
A new method for fractionation of protamines and the amino acid sequence of salmine and three components of iridine.
J. Protein Res. 1, 221-224.
- ANDO, T., YAMASAKI, M. and SUZUKI, M. (1973)
Isolation, characterization, structure and function.
in "Protamines" Vol.12 (Kleinzeller A., ed.) Springer-Verlag Berlin pp.1-115.
- ANFENSEN, C.B. (1972)
The formation and stabilization of protein structure.
Biochem. J. 128, 737-749.
- AUSIO, J. and SUBIRANA, J.A. (1982)
Nuclear proteins and the organization of chromatin in spermatozoa of *Mytilus edulis*.
Exp. Cell. Res. 141, 39-45.
- BERLOT-PICARD, F., VODJDANI, G. and DOLY, J. (1986)
Nucleotide sequence of a cDNA clone encoding *Scylliorhinus caniculus* protamine Z2.
Eur. J. Biochem. 160, 305-310.
- BLACK, J.A. and DIXON, G.H. (1967)
Evolution of protamine : a further example of partial gene duplication.
Nature 216, 152-154.

- BLOCH, D.P. (1969)
A catalog of sperm histones.
Genetics 61, 93-111.
- BOUTILLON, C. (1987)
Phosphorylation des protéines nucléaires basiques au cours de la spermiogenèse de la Seiche.
Mémoire de D.E.A. (Université des Sciences et Techniques de Lille I)
- BOWER, P.A., YELICK, P.C. and HECHT, N.B. (1987)
Both P1 and P2 protamine genes are expressed in mouse, hamster and rat.
Biol.Reprod. 37, 479-488.
- BRETZEL, G. (1972)
Über thynnin, das protamin des thunfisches. Die sequenz der komponente YI.XII. Mitteilung über die struktur der protamine in der untersuchungsreihe von E.WALDSCHMIDT-LEITZ und Mitarbeitern.
Hoppe Seylers Z.Physiol.Chem. 353, 1362-1364.
- BRETZEL, G. (1973)
Über thynnin, das protamin des thunfisches. Die aminosäuresequenz von thynnin Z2. XIV. Mitteilung über die struktur der protamine in der untersuchungsreihe von E.WALDSHMIDT-LEITZ und Mitarbeitern.
Hoppe Seylers Z.Physiol.Chem. 354, 543-549.
- BUCCI, L.R., BROCK, W.A., GOLDKNOPF, I.L. and MEISTRICH, M.L. (1984)
Characterization of high mobility group protein levels during spermatogenesis in the rat.
J.Biol.Chem. 259, 8840-8846.
- BUCCI, L.R., BROCK, W.A. and MEISTRICH, M.L. (1982)
Distribution and synthesis of histone I subfractions during spermatogenesis in the rat.
Exp.Cell.Res. 140, 110-118.
- CALVIN, H.I. (1976)
Comparative analysis of the nuclear basic proteins in the rat, human, guinea pig, mouse, and rabbit spermatozoa.
Biochim.Biophys.Acta 434, 377-389.
- CANDIDO, E.P.M. and DIXON, G.H. (1972)
Trout testis cells. III. Acetylation of histones in different cell types from developing trout testis.
J.Biol.Chem. 247, 5506-5510.
- CHAUVIERE, M., LAINE, B., SAUTIERE, P. and CHEVAILLIER, Ph. (1983)
Purification and characterization of two basic spermatid-specific proteins isolated from the dog-fish *Scylliorhinus caniculus*.
FEBS Lett. 152, 231-235.

- CHAUVIERE, M., MARTINAGE, A., BRIAND, G., SAUTIERE, P. and CHEVAILLIER, Ph. (1987)
Nuclear basic protein transition during sperm differentiation. Amino acid sequence of a spermatid-specific protein from the dog-fish *Scylliorhinus caniculus*.
Eur.J.Biochem. 169, 105-111.
- CHAUVIERE, M., MARTINAGE, A., BRIAND, G., SAUTIERE, P. and CHEVAILLIER, Ph. (1989)
Nuclear basic protein transition during sperm differentiation : primary structure of the spermatid-specific protein S2 from the dog-fish *Scylliorhinus caniculus*.
Eur.J.Biochem. 180, 329-335.
- CHEVAILLIER, Ph. (1983)
Some aspects of chromatin organization in sperm nuclei. in "The sperm cell". Proc. 4th Int.Sym.Spermatology (André, J., ed.) Nijhoff, M. The Hague. p.179-196.
- CHEVAILLIER, Ph., MARTINAGE, A., GUSSE, M. and SAUTIERE, P. (1987)
Amino acid sequence of scylliorhinine Z1 and comparison of primary structure of the protamines of the dog-fish *Scylliorhinus caniculus*.
Biochim.Biophys.Acta 914, 19-27.
- CHOU, P.Y. and FASMAN, G.D. (1978)
Empirical predictions of protein conformation.
Ann.Rev.Biochem. 47, 251-276.
- COELINGH, J.P., MONFOORT, C.H., ROZIJN, T.H., GEVERS-LEUVEN, J.A., SHIPHOF, R., STEYEN-PARVE, E.P., 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.
- COUPPEZ, M., MARTIN-PONTHIEU, A. and SAUTIERE, P. (1987)
Histone H4 from cuttlefish testis is sequentially acetylated.
J.Biol.Chem. 262, 2854-2860.
- DAVIES, P.L. and DIXON, G.H. (1979)
Sequences of large T1 ribonuclease-resistant oligonucleotides from protamine mRNA : the overall architecture of protamine mRNA.
Nucl.Acids.Res. 7, 2323-2345.
- DAVIES, P.L., DIXON, G.H., FERRIER, L.N., GEDAMU, L. and IATROU, K. (1976)
The structure and function of protamine mRNA from developing trout testis.
Prog.Nucleic Acid, Res.Mol.Biol. 19, 135-155.
- DIXON, G.H., AIKEN, J.M., JANKOWSKI, J.M., Mc KENZIE, D.I., MOIR, R. and STATES, J.C. (1986)
Organization and evolution of the protamine genes of salmonid fishes.
in "Chromosomal proteins and gene expression". Vol A 101 (Reeck, R.G., Goodwin, G.A. and Puigdomenech, P, ed.) Plenum Publishing Corporation, pp.287-314.

- DOMENJOURD, L., FRONIA, C., UHDE, F. and ENGEL, W. (1988)
Sequence of human protamine 2 cDNA.
Nucl. Acid. Res. 16, 7733.
- DYNAN, W.S. and TJIAN, R. (1983)
The promoter-specific transcription factor Sp1 binds to
upstream sequences in the SV40 early promoter.
Cell 35, 79-87.
- DYNAN, W.S. and TJIAN, R. (1985)
Control of eukaryotic messenger RNA synthesis by sequence-
specific DNA-binding proteins.
Nature 316, 774-778.
- ELSEVIER, S.M. (1982)
Messenger RNA encoding basic chromosomal proteins of mouse
testis.
Dev. Biol. 90, 1-12.
- FINK, J.S., VERHAVE, M., KASPER, S., TSUKADA, T., MANDEL, G. and
GOODMAN, R.H. (1988)
The CGTCA sequence motif is essential for biological acti-
vity of the vasoactive intestinal peptide gene CAMP-regu-
lated enhancer.
Proc. Natl. Acad. Sci. USA 85, 6662-6666.
- GABORIAUD, C., BISSERY, V., BENCHETRIT, T. and MORNON, J.P. (1987)
Hydrophobic cluster analysis, an efficiency new way to com-
pare and analyse amino acid sequences.
FEBS Lett. 224, 149-155.
- GAREL, J.R. and BALDWIN, R.L. (1973)
Both the fast and slow refolding reactions of ribonuclease
A yield native enzyme.
Proc. Natl. Acad. Sci. USA 70, 3347-3351.
- GARNIER, J., OSGUTHORPE, D.J. and ROBSON, B. (1978)
Analysis of the accuracy and implication of simple methods
for predicting the secondary structure of globular proteins.
J. Mol. Biol. 120, 97-120.
- GEDAMU, L. and DIXON, G.H. (1976a)
Assay of protamine messenger RNA from rainbow trout testis.
J. Biol. Chem. 251, 1446-1454.
- GEDAMU, L. and DIXON, G.H. (1976b)
Purification and properties of biologically active rainbow
trout testis protamine mRNA.
J. Biol. Chem. 251, 1455-1463.

- GEDAMU, L. and DIXON, G.H. (1979)
Heterogeneity of biologically active deadenylated protamine mRNA components isolated from rainbow trout testis.
Nucl. Acid. Res. 6, 3661-3672.
- GEDAMU, L., IATROU, K. and DIXON, G.H. (1977)
Isolation and characterization of trout testis protamine mRNAs lacking poly(A).
Cell 10, 443-451.
- GEDAMU, L., WOSNICK, M.A., CONNOR, W., WATSON, D.C. and DIXON, G.H. (1981)
Molecular analysis of the protamine multi-gene family in rainbow trout testis.
Nucl. Acid. Res. 9, 1463-1482.
- GIBRAT, J.F., GARNIER, J. and ROBSON, B. (1987)
Further developments of protein secondary structure prediction using information theory. New parameters and consideration of residue pairs.
J. Mol. Biol. 198, 425-443.
- GILMOUR, R.S. and DIXON, G.H. (1972)
Biosynthesis of protamine in trout testis. *In vitro* synthesis of protamine in a trout liver cell-free system.
J. Biol. Chem. 247, 4621-4627.
- GREGORY, S.P., DILLON, N.O. and BUTTERWORTH, P.H.W. (1982)
The localization of the 5'-termini of *in vivo* and *in vitro* transcripts of a cloned rainbow trout protamine gene.
Nucl. Acid. Res. 10, 7581-7592.
- GRIMES, S.R., MEISTRICH, M.L., PLATZ, R.D. and HNILICA, L.S. (1977)
Nuclear protein transitions in rat testis spermatids.
Exp. Cell. Res. 110, 31-39.
- GUBLER, U. and HOFFMAN, B.J. (1983)
A simple and very efficient method for generating cDNA libraries.
Gene 25, 263-269.
- GUSSE, M. and CHEVAILLIER, Ph. (1981)
Microelectrophoretic analysis of basic protein changes during spermatogenesis in the dog-fish *Scylliorhinus caniculus* (L).
Exp. Cell. Res. 136, 391-397.
- GUSSE, M., SAUTIERE, P., BELAÏCHE, D., MARTINAGE, A., ROUX, C., DADOUNE, J.P. and CHEVAILLIER, Ph. (1986)
Purification and characterization of nuclear basic proteins of human sperm.
Biochim. Biophys. Acta 884, 124-134.

- HECHT, N.B. (1987)
Gene expression during spermatogenesis.
in "Cell Biology of Testis and Epididymis," vol 513
Annals of the New York Academy of Sciences, pp.90-101.
- HEIDARAN, M.A. and KISTLER, W.S. (1987 a)
Isolation of a cDNA clone for transition protein 1 (TP1),
a major chromosomal protein of mammalian spermatids.
Gene 54, 281-284.
- HEIDARAN, M.A. and KISTLER, W.S. (1987b)
Transcriptional and translational control of the message
for transition protein 1, a major chromosomal protein of
mammalian spermatids.
J. Biol. Chem. 262, 13309-13315.
- IATROU, K., GEDAMU, L. and DIXON, G.H. (1979)
Protamine messenger RNA : partial purification and charac-
terization of a heterogeneous family of polyadenylated
messenger components.
Can. J. Biochem. 57, 945-956.
- IKAI, A., FISH, W.N. and TANFORD, C. (1973)
Kinetics of unfolding and refolding of proteins.
J. Mol. Biol. 73, 165-184.
- JOHNSON, P.A., PESCHON, J.J., YELICK, P.C., PALMITER, R.D. and
HECHT, N.B. (1988)
Sequence homologies in the mouse protamine 1 and 2 genes.
Biochim. Biophys. Acta 950, 45-53.
- KHRAPUNOV, S.N., KADURA, S.N. and DRAGAN, A.I. (1988)
Rearrangements of chromatin structure during spermatoge-
nesis of squid.
Eur. J. Biochem. 175, 609-613.
- KISTLER, W.S., KEIM, P.S. and HEINRIKSON R.L. (1976)
Partial analysis of the basic chromosomal protein of rat
spermatozoa.
Biochim. Biophys. Acta 427, 752-757.
- KISTLER, W.S., NOYES, C., HSU, R. and HEINRIKSON, R.L. (1975)
The amino acid sequence of a testis-specific basic protein
that is associated with spermatogenesis.
J. Biol. Chem. 250, 1847-1853.
- KLEENE, K.C., BORZORGZADEH, A. FLYNN, J.F., YELICK, P.C. and
HECHT, N.B. (1988)
Nucleotide sequence of a cDNA clone encoding mouse tran-
sition protein 1.
Biochim. Biophys. Acta 950, 215-220.

- KLEENE, K.C., DISTEL, R.J. and HECHT, N.B. (1985)
Nucleotide sequence of a cDNA clone encoding mouse protamine 1.
Biochemistry 24, 719-722.
- KLEENE, K.C. and FLYNN, J.F. (1987)
Characterization of a cDNA clone encoding a basic protein, TP2, involved in chromatin condensation during spermiogenesis in the mouse.
J. Biol. Chem. 262, 17272-17277.
- KORNBERG, R.D. (1974)
Chromatin structure : a repeating unit of histones and DNA. Chromatin structure is based on a repeating unit of eight histone molecules and about 200 DNA base pairs.
Science 184, 868-871.
- KOZAK, M. (1986)
Bifunctional messenger RNAs in eukaryotes.
Cell 47, 481-483.
- KRAWETZ, S.A., CONNOR, W. and DIXON, G.H. (1987)
Cloning of bovine P1 protamine cDNA and the evolution of Vertebrate P1 protamines.
DNA 6, 47-57.
- KRAWETZ, S.A., CONNOR, W. and DIXON, G.H. (1988)
Bovine protamine genes contain a single intron. The structure of the two alleles.
J. Biol. Chem. 263, 321-326.
- KRAWETZ, S.A. and DIXON, G.H. (1988)
Sequence similarities of the protamine genes : implications for regulation and evolution.
J. Mol. Evol. 27, 291-297.
- LAEMMLI, U.K. (1970)
Cleavage of structural proteins during the assembly of the head of bacteriophage T4.
Nature 227, 680-685.
- LEE, C.H., BARTELS, I. and ENGEL, W. (1987a)
Haploid expression of a protamine gene during bovine spermatogenesis.
Biol. Chem. Hoppe Seyler 368, 807-811.
- LEE, C.H., HOYER-FENDER, S. and ENGEL, W. (1987b)
The nucleotide sequence of a human protamine 1 cDNA.
Nucl. Acid. Res. 15, 7639.

- LEE, C.H., MANSOURI, A., HECHT, W., HECHT, N.B. and ENGEL, W. (1986)
Nucleotide sequence of a bovine protamine cDNA.
Biol.Chem.Hoppe Seyler 367, 131-135.
- LEVIN, J.M., ROBSON, B. and GARNIER, J. (1986)
An algorithm for secondary structure determination in proteins based on sequence similarity.
FEBS Lett. 205, 303-308.
- LIM, V.I. (1974)
Structural principles of the globular organization of protein chains. A stereochemical theory of globular protein secondary structure.
J.Mol.Biol. 88, 857-872.
- LOUIE, A.J. and DIXON, G.H. (1972a):
Trout testis cells. I. Characterization by deoxyribonucleic acid and protein analysis of cells separated in velocity sedimentation.
J.Biol.Chem. 247, 5490-5497.
- LOUIE, A.J. and DIXON, G.H. (1972b)
Trout testis cells. II. Synthesis and phosphorylation of histones and protamines in different cell types.
J.Biol.Chem. 247, 5498-5505.
- LOUIE, A.J. and DIXON, G.H. (1972c)
Kinetics of enzymatic modification of the protamines and a proposal for their binding to chromatin.
J.Biol.Chem. 247, 7962-7968.
- LUERSSSEN, H., MAIER, W.M., HOYER-FENDER, S. and ENGEL, W. (1989)
The nucleotide sequence of rat transition protein 2 (TP2) cDNA.
Nucl.Acids.Res. 17, 3585.
- MALI, P., SANDBERG, M., VUORIO, E., YELICK, P.C., HECHT, N.B. and PARVINEN, M. (1988)
Localization of protamine 1 mRNA in different stages of the cycle of the rat seminiferous epithelium.
J.Cell.Biol. 107, 407-412.
- MANIATIS, T., FRITSCH, E.F. and SAMBROOK, J. (1982)
in "Molecular cloning : a laboratory manual".
Cold Spring Harbor, New York.
- MARTINAGE, A., ARKHIS, A., ALIM, E., SAUTIERE, P. and CHEVAILLIER, Ph. (1990)
Molecular characterization of nuclear basic protein HPI1, a putative precursor of human sperm protamines HP2 and HP3.
Eur.J.Biochem. (soumis pour publication).

- MARTINAGE, A., GUSSE, M., BELAÏCHE, D., SAUTIERE, P. and CHEVAILLIER, Ph. (1985)
Amino acid sequence of a cysteine-rich, arginine-rich sperm protamine of the dog-fish *Scylliorhinus caniculus*.
Biochim.Biophys.Acta 831, 172-178.
- MARTIN-PONTHIEU, A., WOUTERS-TYROU, D., BELAÏCHE, D., SAUTIERE, P. and VAN DORSSELAER, A. (1990)
Cuttlefish sperm protamines : amino acid sequences of two distinct variants.
Eur.J.Biochem. (soumis pour publication)
- MAZRIMAS, J.A., CORZETT, M., CAMPOS, C. and BALHORN, R. (1986)
A corrected primary sequence for bull protamine.
Biochim.Biophys.Acta 872, 11-15.
- Mc KAY, D.J., RENAUX, B.S. and DIXON, G.H. (1985)
The amino acid sequence of human sperm protamine P1.
Bioscience Reports 5, 383-391.
- Mc KAY, D., RENAUX, B.S. and DIXON, G.H. (1986)
Rainbow trout protamines. Amino acid sequences of six distinct proteins from a single testis.
Eur.J.Biochem. 158, 361-366.
- MEISTRICH, M.L., TROSTLE, P.K. and BROCK, W.A. (1981)
Association of nucleoprotein transitions with chromatin changes during rat spermatogenesis.
in "Bioregulations of reproduction" (G.Jagiello and H.J.Vogel, ed.), Academic Press, New York, pp 151-166.
- MITCHELL, P.J. and TJIAN, R. (1989)
Transcriptional regulation in mammalian cells by sequence-specific DNA binding proteins.
Science 245, 371-378.
- MORI, N., FUNATSU, Y., HIRUTA, K. and GOTO, S. (1985)
Analysis of translational fidelity of ribosomes with protamine messenger RNA as a template.
Biochemistry 24, 1231-1239.
- MUNOZ-GUERRA, S., AZORIN, F., CASSAS, M.T., MARCET, X., MARISTANY, R.A., ROCA, J. and SUBIRANA, J.A. (1982)
Structural organization of sperm chromatin from the fish *Carassius auratus*.
Exp.Cell.Res. 137, 47-53.
- NAGATA, S., MANTEI, N. and WEISMANN, C. (1980)
The structure of one of the eight or more distinct chromosomal genes for human interferon- α .
Nature 287, 401-408.

- OLINS, A.L. and OLINS, D.E. (1974)
Spheroid chromatin units (v bodies).
Science 183, 330-332.
- OLIVA, R. and DIXON, G.H. (1989)
Chicken protamine genes are intronless. The complete genomic sequence and organization of the two loci.
J. Biol. Chem. 264, 12472-12481.
- OUDET, P., GROSS-BELLARD, M. and CHAMBON, P. (1974)
Electron microscopic and biochemical evidence that chromatin structure is a repeating unit.
Cell 4, 281-300.
- POCCIA, D. (1986)
Remodeling of nucleoproteins during gametogenesis, fertilization, and early development.
Int. Review Cytol. 105, 1-65.
- PLATZ, R.D., GRIMES, S.R., MEISTRICH, M.L. and HNILICA, L.S. (1975)
Changes in nuclear proteins of rat testis cells separated by velocity sedimentation.
J. Biol. Chem. 250, 5791-5800.
- RICHARD, A. (1971)
Contribution à l'étude expérimentale de la croissance et de la maturation sexuelle de *Sepia officinalis* L. (Mollusque Céphalopode).
Thèse présentée à la Faculté des Sciences de l'Université de Lille.
- RAO, B.J., BRAHMACHARI, S.K. and SATYANARAYANA, M.R. (1983)
Structural organization of the meiotic prophase chromatin in the rat testis.
J. Biol. Chem. 258, 13478-13485.
- RHOLAM, M., NICOLAS, P. and COHEN, P. (1986)
Precursors for peptide hormones share common secondary structures forming features at the proteolytic processing sites.
FEBS Lett. 207, 1-6.
- ROBSON, B. and PAIN, R.H. (1971)
Analysis of the code relating sequence to conformation in proteins: possible implication for the mechanism of formation of helical regions.
J. Mol. Biol. 58, 237-259.
- ROUSSEAU-PREVOST, R., ENGELHARDT, R.P., ROUSSEAU, J., WOUTERS-TYROU, D. and SAUTIERE, P. (1988)
Nuclear protein transitions in cuttle-fish spermiogenesis: immunocytochemical localization of a protein specific for the spermatid stage.
Gamete Res. 19, 277-290.

- SANGER, F. and COULSON, A.R. (1975)
A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase.
J.Mol.Biol. 94, 441-448.
- SAUTIERE, P., BELAÏCHE, D., MARTINAGE, A. and LOIR, M. (1984)
Primary structure of the ram (*Ovis aries*) protamine.
Eur.J.Biochem. 144, 121-125.
- SAUTIERE, P., BRIAND, G., GUSSE, M. and CHEVAILLIER, Ph. (1981)
Primary structure of a protamine isolated from the sperm nuclei of the dog-fish *Scylliorhinus caniculus*.
Eur.J.Biochem. 119, 251-255.
- SAUTIERE, P., GUSSE, M., BRIAND, G., MARTINAGE, A. and CHEVAILLIER, Ph. (1984)
Primary structure of scylliorhinine S4, a protamine isolated from sperm nuclei of the dog-fish *Scylliorhinus caniculus*.
Biochim.Biophys.Acta 791, 82-86.
- SAUTIERE, P., MARTINAGE, A., BELAÏCHE, D., ARKHIS, A. and CHEVAILLIER, Ph. (1988)
Comparison of the amino acid sequences of human protamines HP2 and HP3 and of intermediate basic nuclear proteins HPS1 and HPS2. Structural evidence that HPS1 and HPS2 are pro-protamines.
J.Biol.Chem. 263, 11059-11062.
- SCHAFFNER, W., KUNZ, G., DAETWYLER, H., TELFORD, J., SMITH, H.D. and BIRNSTIEL, M.L. (1978)
Genes and spacers of cloned sea urchin histone DNA analyzed by sequencing.
Cell 14, 655-671.
- SELLOS, D. and KMIECIK, D. (1985)
Characterization of the histones and protamines from the sperm of the marine worm *Platynereis dumerilii*.
Comp.Biochem.Physiol. 80B, 119-126.
- SERGHINI, M.A., RITZENTHALER, C. and PINCK, L. (1989)
A rapid and efficient "miniprep" for isolation of plasmid DNA.
Nucl.Acids.Res. 17, 3604.
- SINCLAIR, G.D. and DIXON, G.H. (1982)
Purification and characterization of cytoplasmic protamine messenger ribonucleoprotein particles from rainbow trout testis.
Biochemistry 21, 1869-1877.

- SINGH, J. and RAO, S.M.R. (1987)
Interaction of rat testis protein TP, with nucleic acids
in vitro. Fluorescence quenching, UV absorption and thermal
denaturation studies.
J.Biol.Chem. 262, 734-740.
- SPECKERT, W., KENNEDY, B., DAISLEY, St.L. and DAVIES, P. (1983)
Primary structure of protamine from the northern pike *Esox
lucius*.
Eur.J.Biochem. 136, 283-289.
- STATES, J.C., CONNOR, W., WOSNICK, M.A., AIKEN, J.M., GEDAMU, L. and
DIXON, G.H. (1982)
Nucleotide sequence of a protamine component C_{II} gene of
Salmo gairdneri.
Nucl.Acids.Res. 10, 4551-4563.
- STEPHANO, J.L., GOULD, M. and ROJAS-GALICIA, L. (1986)
Advantages of picrate fixation for staining polypeptides in
polyacrylamide gels.
Anal.Biochem. 152, 308-313.
- SUBIRANA, J.A. (1983)
Nuclear proteins in spermatozoa and their interactions
with DNA.
in "The sperm cell". Proc. 4th Int.Sym. Spermatology
(André, J., ed.) Nijhoff, M. The Hague. p 197-213.
- SUBIRANA, J.A., COZCOLLUELA, C., PALAU, J. and UNZETA, M. (1973)
Protamines and other basic proteins from spermatozoa of
molluscs.
Biochim.Biophys.Acta 317, 363-379.
- TANHAUSER, S.M. and HECHT, N.B. (1989)
Nucleotide sequence of the rat protamine 2 gene.
Nucl.Acids.Res. 17, 4395.
- TOBITA, T., TSUTSUMI, H., KATO, A., SUZUKI, H., NOMOTO, M., NAKANO, M.
and ANDO, T. (1983)
Complete amino acid sequence of boar protamine.
Biochim.Biophys.Acta 744, 141-146.
- VON HOLT, C., STRICKLAND, W.N., BRANDT, W.F. and STRICKLAND, M.
(1979)
More histone structures.
FEBS Lett. 100, 201-218.
- WEBSTER, A., RUSSEL, S., TALBOT, P., RUSSELL, W.C. and KEMP, G.D.
(1989)
Characterization of the *Adenovirus* proteinase : substrate
specificity.
J.Gen.Virol. 70, 3225-3234.

- WIGLE, D.T. and DIXON, G.H. (1970)
Transient incorporation of methionine at the N-terminus
of protamine newly synthesized in trout testis cells.
Nature 227, 676-680.
- WOUTERS-TYROU, D., MARTIN-PONTHIEU, A., BRIAND, G., SAUTIERE, P.
and BISERTE, G. (1982)
The amino-acid sequence of histone H2A from cuttlefish
Sepia officinalis.
Eur.J.Biochem. 124, 489-498.
- WOUTERS-TYROU, D., MARTIN-PONTHIEU, A., RICHARD, A. and SAUTIERE, P.
(1988)
Characterization of the nuclear basic proteins specific of
spermiogenesis in cuttlefish *Sepia officinalis*.
Biochim.Biophys.Acta 953, 86-94.
- WOUTERS-TYROU, D., MARTIN-PONTHIEU, A., SAUTIERE, P. and BISERTE, G.
(1981)
Acetylation of histone H4 in chicken erythrocyte and cuttle-
fish testis chromatin.
FEBS Lett. 128, 195-200.
- YELICK, P.C., BALHORN, R., JOHNSON, P.A., CORZETT, M., MAZRIMAS, J.A.,
KLEENE, K. and HECHT, N.B. (1987)
Mouse protamine 2 is synthesized as a precursor whereas
mouse protamine 1 is not.
Mol.Cell.Biol. 7, 2173-2179.
- YULIKOVA, E.P., EVSEENKO, L.K., BARATOVA, L.A., BELYANOVA, L.P.,
RYBIN, V.K. and SILAEV, A.B. (1976)
The primary structure of sturine B, a protamine from cas-
pian sturgeon.
Bioorg.Khim. 2, 1613-1617.
- YULIKOVA, E.P., RYBIN, V.K. and SILAEV, A.B. (1979)
The primary structure of stellin A.
Bioorg.Khim. 5, 5-10.



Résumé

Chez la seiche *Sepia officinalis*, la spermiogenèse est caractérisée par une double transition protéique nucléaire : histones → protéine intermédiaire (protéine T) → protamine (protéine Sp).

La similarité des compositions en acides aminés des protéines T et Sp suggère l'existence d'une filiation entre ces deux protéines. Il peut s'agir soit d'un phénomène de protéolyse post-traductionnel, soit d'un mécanisme de régulation au niveau génique.

1. La structure primaire de la protéine T a été établie. Elle comprend deux variants majeurs T1 et T2 de 78 et 77 résidus, respectivement. Leur séquence est caractérisée par une asymétrie de la molécule, tous les résidus d'arginine étant accumulés dans les 2/3 C-terminaux de la protéine. Cette région basique des protéines T1 et T2 présente une identité structurale complète avec la séquence des protamines Sp1 et Sp2. Le site éventuel de coupure serait localisé au niveau d'un coude β , comme chez les précurseurs d'hormone.

2. L'enrichissement des ARNm codant pour la protéine T s'accompagne de ceux codant pour une protéine de migration électrophorétique semblable à la protamine Sp. Nous avons fabriqué avec ces ARNm enrichis 2 banques d'ADNc, l'une en plasmide pUC 13 et l'autre en phages λ gt 10. Nous avons obtenu, par la méthode "d'hybridation ADNc/ARNm, puis traduction des ARNm", une sonde valable dans la banque plasmidique. Cette sonde, appelée sonde 48, nous a permis de sélectionner un nombre important de clones positifs dans la banque de phages λ gt 10. Parmi ces clones, le clone 5/20 a été séquencé. Il comprend la partie codant pour la protéine T2 (résidus 17 à 77) et la région 3' non codante.

Grâce à ces travaux, l'étude des mécanismes de régulation des gènes codant pour les protéines spécifiques de la spermiogenèse de la seiche va pouvoir se poursuivre et se développer.

Mots clefs : Mollusques Céphalopodes - Spermiogenèse - Protéines spécifiques des spermatozoïdes - Protamines - Précurseur - Structure primaire - Variants structuraux - Séquence nucléotidique.

