Characterization of the sperm chromatin structure in *Mytilus californianus* and *Spisula solidissima*

by

Fan Zhang

B.Sc, Anhui University, 1991

A Thesis Submitted in Partial Fulfilment of the
Requirements for the Degree of

MASTER OF SCIENCE

in the Department of Biochemistry and Microbiology

We accept this thesis as conforming
to the required standard

[redacted]

Dr. Juan Ausio, Supervisor (Department of Biochemistry and Microbiology)

[redacted]

Dr. Santosh Misra, Departmental Member (Department of Biochemistry and Microbiology)

[redacted]

Dr. David B. Levin, Outside Member (Department of Biology)

[redacted]

Dr. Robert Burke, External Examiner (Department of Biology)

Supervisor: Dr. Juan Ausio

ABSTRACT

The aim of this thesis was to establish the chromatin structure and organization in the sperm of two bivalve mollusks: *Mytilus californianus* and *Spisula solidissima* . Nuclease digestion and crosslinking studies showed that the sperm chromatin of *Mytilus californianus* has a organization consisting of small groups of oligonucleosomes interspersed between long stretches of nucleoprotamine filaments. The nucleoprotamine complexes result from the association of PL-II, PL-III and PL-IV with DNA. The spatial arrangement of different PL proteins within these complexes is discussed. In the case of *Spisula*, it was found that its major protamine-like component (PL-I) contains at least two cysteines located between tryptophan and tyrosine residues of the globular region of the molecule. It was also found that the PL-I exhibits a head to tail arrangement in the sperm chromatin of *Spisula*.

Examiners: [redacted]

Dr. Juan Ausio, Supervisor (Department of Biochemistry and Microbiology)

[redacted]

Dr. Santosh Misra, Departmental Member (Department of Biochemistry and Microbiology)

[redacted]

Dr. David B. Levin, Outside Member (Department of Biology)

[redacted]

Dr. Robert Burke, External Examiner (Department of Biology)

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

A (Ala):	Alanine
Asp:	Aspartic acid
ATP:	Adenosine triphosphate
CD:	Circular dichroism
B-P:	Basic-phosphate
C (Cys):	Cysteine
Da:	Dalton
DMF:	Dimethylformamide
DNA:	Deoxyribonucleic acid
DTBSP:	Dithiobis (succinimidyl propionate)
DTT:	Dithiothreitol
EDTA:	Ethylenediamine-tetraacetic acid
Glu:	Glutamine
HPLC:	High performance liquid chromatography
Ile:	Isoleucine
K (Lys):	Lysine
Leu:	Leucine
Met:	Methionine
MNase:	Micrococcal nuclease
mRNA:	Messenger ribonucleic acid
MW:	Molecular weight
NBS:	N-bromosuccinimide
PAGE:	Polyacrylamide gel electrophoresis
PCA:	Perchloric acid
Phe:	Phenylalanine
PL:	Protamine-like
PMSF:	Phenylmethylsulfonyl fluoride
qBBr:	Quanternary salt of monobromotrimethylammoniobimane
R (Arg):	Arginine
RNA:	Ribonucleic acid
S (Ser):	Serine

SDS:	Sodium dodecyl sulphate
T (Thr):	Threonine
TE:	Tris-EDTA
TEMED:	N, N, N', N'-tetramethylethylenediamine
TFA:	Trifluoroacetic acid
Tricine:	N-tris (hydroxymethyl)methyl glycine
Tris-HCl:	Tris (hydroxymethyl) aminomethane-HCl
UA:	Urea-acetic acid
W (Trp):	Tryptophan
Y (Tyr):	Tyrosine

ACKNOWLEDGEMENTS

I am grateful to Dr. Santosh Misra, Dr. David B. Levin, the members of my supervisory committee, for their advice, expertise and positive input over this period of study.

I am indebted to my colleagues, LeAnn Howe, Phil, Nieve, Jason, Dustin, Jon. I thank them for their help in routine work, special instructions, and the English language.

I acknowledge the enjoyment of my studies at the University of Victoria in the Department of Biochemistry and Microbiology. I thank the departmental staff (secretaries and technicians) for their assistance.

I was financially supported by University Fellowship.

The most special thanks of all goes to Dr. Juan Ausio, my supervisor. Every progress I achieved during the last three years has been related to his guidance and understanding.
I am grateful.

INTRODUCTION

Spermatogenesis is the cellular differentiation process that is responsible for the production of male gametes. Sperm development occurs in three main phases. The first phase is a mitotic stage of stem cell proliferation. The second phase is a meiotic stage in which diploid spermatocyte reduces to haploid spermatids. The final stage is spermiogenesis in which round spermatids undergo morphogenesis to obtain the final particular mature sperm shape. Mature sperm cells, or spermatozoa of most animal species are elongated cells consisting typically of a head, middle region and tail (Fig 1). The dense nucleus forms the majority of the head. The mature sperm head also contains acrosome which is a Golgi-derived organelle whose contents are important for the penetration of the egg membrane. The middle region is populated by mitochondria.

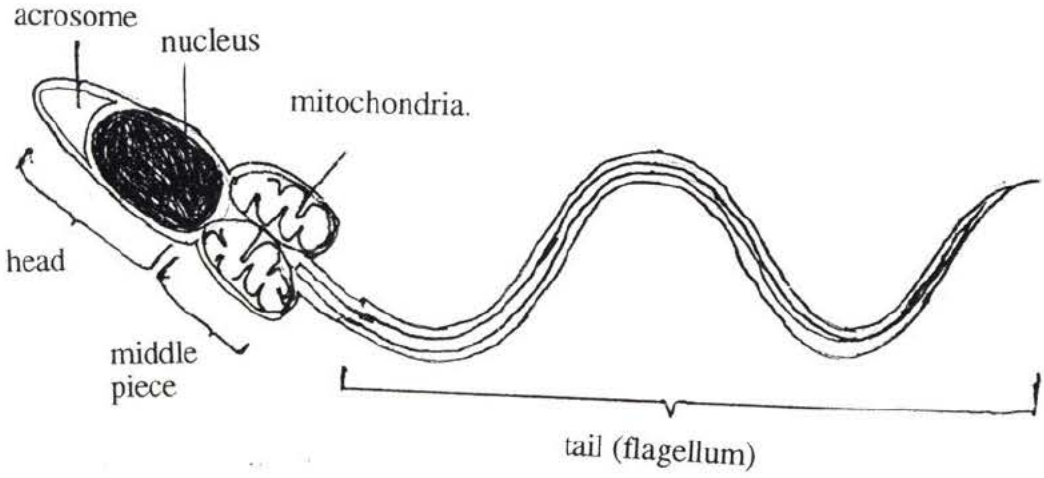


Figure 1. Typical structures of an animal sperm cell.

During spermiogenesis, the extent of nuclear condensation is generally correlated with the replacement of somatic histones by protamines or protamine-like proteins in different species. In some species, such as trout, nearly all somatic histones are replaced. In mammals and some mollusks 10% to 20% of the DNA retains somatic histones, whereas, in winter flounder somatic histones form 75% of the total protein content in mature sperm nuclei. In several species, including frog, sea urchin, sea cucumber, horseshoe crab and goldfish, sperm nuclei contain only histones; no protamines are present (1). From the compositional point of view, these nuclear sperm-specific proteins can be arranged in a discrete number of basic types: histone, protamine and protamine-like (PL) consisting of proteins with an intermediate composition between histones and protamines.

these different species is complicated compared to their somatic chromatin structure in that these proteins within the chromatin exhibit a broad spectrum of both compositional and structural variability between different taxonomic groups (2, 3, 4). Research on protamine-like proteins has been carried out mainly in mollusks. This taxonomic group is of special interest because it contains organisms with nuclear sperm-specific proteins that can be considered representative of each of the three protein types. The focus of this work was on the structure and function of the sperm nuclear proteins in the chromatin condensation during spermiogenesis of two invertebrate species of Bivalve Mollusks, *Spisula solidissima* and *Mytilus californianus*.

I. Protamines

1. Classification of protamines

During spermiogenesis the somatic histones of chromatin are replaced by highly basic proteins, which condense and compact DNA within the sperm nucleus. In most species, these proteins are known as protamines. Different animals have diverse protamines in their mature sperm. According to the amino acid composition, protamines can be divided into three major types (2):

1. *Salmon* type, or monoprotamines contain predominantly arginine as in the case of clupeine (herring), salmine (salmon), iridine (trout) and thynine (tuna) (5).

2. Mammalian type protamines (also called stable protamines or basic keratins) are arginine rich, highly basic and contain cysteine. They were first identified in mammalian sperm, such as mouse, boar, ram, stallion and human, and has also been seen in insect sperm.

3. *Mytilus* type protamines (intermediate protamine) are intermediate type proteins with characteristics of monoprotamines and somatic histones in that they contain both lysine and arginine. Intermediate protamines have been found in bivalve mollusks, as well as in echinoderm, amphibians and fish.

2. Composition of protamines

Protamines of the *Salmon* type contain a very high content of arginine (60-80%) and serine (10-15%). *Salmon* type protamines have a very low content or lack of acidic and hydrophobic amino acids as well as a general lack of lysine and cysteine.

In mammalian type protamines, the arginine content is 50-60%. Unlike *Salmon* type protamines, cysteine is present and can form inter- or intramolecular disulfide linkages and hence stabilize the sperm chromatin in the mammalian sperm nucleus. This novel

feature is rare or completely absent in the other types of protamines.This protamine contains little or no lysine.

In bivalve mollusks (*Mytilus* type), the major nuclear proteins contain similar amounts of arginine and lysine yet either amino acid may predominate depending upon the species. The total content of these two amino acids can amount to 40-50%; thus, the protamines in this group have a lower charge compared to protamines of the *Salmon* and mammalian types. In addition, there exist a large amount of serine, approximately 20%, present in most species (6).

All these different types of protamines share two common features: lysine and arginine content accounts for 45-80% of all the amino acids. Serine and threonine account for 10-25%.

Despite some common features, the *Salmon*, mammalian and Mytilus types of protamines have significant differences. Marine snail protamines contain a reduced amount of lysine (5-10%), tuna fish protamines have a high valine content, and squid protamines contain a high tyrosine percentage. Despite this variation within individual species, the acidic (Glu, Asp) and bulky hydrophobic residues (Met, Ile, Leu, Phe) rarely occur in protamines. Tyrosine, however, occurs in the most evolved species (squid, fowl, mammals), whereas, histidine usually is absent, except in some fish species and mammals (3).

3. Molecular size

The molecular size of protamines varies significantly. The *Salmon* type protamines are very small, only 27 to 34 amino acids long compared to 50 amino acids in mammalian type protamines. In *Mytilus* type protamines, a small size is not a general characteristic and some extent up to 300 amino acid residues, as in the case in *Spisula solidissima* (7).

4. Protamine microheterogeneity

In fish the protamines contain a mixture of proteins with a closely related composition (8). Mammalian, ram and rat sperm nuclei contain only one protamine type at maturity. Mouse sperm nuclei contain two protamine types and human sperm have three (1). In Mollusks, protamines show microheterogeneity in different species (see section II of introduction).

5. Sequence of protamine

Protamine sequences of vertebrates (bird, mammal, amphibian and bony fish) and

invertebrate (different species of mollusc) have been compared (9) (Fig 2). The protamine sequences of species of various mollusks indicated that in both *Mondonta turbinata* (archaeogastropod mollusc) and *Mytilus edulis* (bivalve mollusc), alternating basic-serine residues were present in the amino-terminal region. Seven alternating R-S and one K-S existed in residues 10-24 of the former and four R-S and two K-S in the latter species. Serine and threonine residues were putatively phosphorylated. The alternating basic-phosphorylatable residue was called the B-P residue. However, the cephalopod *Sepia officinalis* lacks the characteristic alternating B-P region. In the central region and carboxy-terminal portion of protamines from *M. turbinata*, arginine was the main basic residue and they were found to exist in clusters. The arginine clusters were found in the carboxy-terminal region of *M. edulis* and throughout the entire protamine of *S. officinalis*. In addition, the protamine sequence in *M. turbinata* had notable features. In the central region (residues 28-67), arginine clusters alternated with triplets of other residues. These triplets were SAS, TAS, SVS, and SRS. This central region did not contain any lysine residues. The carboxy-terminal of the molecule (residues 68-106) also contained arginine clusters, but they alternated with heterogeneous groups of amino acids. The separation between clusters was made by a single amino acid (T, A) or by heterogeneous groups of four, five, or eight residues, some of which contained isolated arginine and lysine residues. Consequently, lysine appeared in the carboxy-terminal of the protamine but did not disrupt the arginine clusters. The central region of the protamine from the bivalve *M. edulis* was less basic and had a distribution of basic amino acids which was similar to that found in the carboxy-terminal part of histone H1. In histone H1, the uniform charge distribution and lack of long clusters of basic amino acids prevents aggregation when histone H1 interacts with DNA so that chromatin is available for transcription (10).

a)
10 30
* * * *
ARAV RRRR A BSRSRSRKSRSRSR SAKR SAS RRR SRS AG RRRRRR TAS
50 70 90
* * * * *
RRRRR SAS RRR SVS RRRRRR SAKSRG RRRR GRKV RRRR VKRAGRKG RRR
* 106
T RRRRRR A RR

b)
10 30 50
* * * * *
PSPT RR S RKSRSKSRSRSR SASASPGKAAKRARSKTPRRG KKR ARSPSKKA
70 89
* * *
RRR SRSTKKTAA KRRKR SSSP KKRR SAG KRR VRA KKKK

c)
10 30
50
* * * *
RRRRRR S RRRRRR S RRR SRSPY RRR Y RRRRRRRRRR S RRRR Y RRRR SYS
50 57
*
RRR Y RRRR

d)
10 30 50
* * * * *
ARYR RSRTRSRS PRS RRRRRR SG RRR SP RRRRR YGSA RR S RR SVGG RRRR
* 57
RRRRRRR YYGS

e)
10 30 50
* * * * *
ARYR CC RSKSRS RC RRRRRR C RRRRRR CC RRRRRR CC RRRR SYTIRCKR Y

f)
10 30 35
* * *
PP RRKR VSSAP RRRRR TY RR TTA HKH QDRPVHRGH

g)
10 30 32
* * *
P RRRR SSSRPV RRRRR PRVS RRRRRR GG RRRR

Figure 2. Amino acid sequences belonging to some sperm-specific proteins: a. *Monodonta turbinata* protamine (archaeogastropod mollusc); b. Protein PL-I from *Mytilus edulis* (bivalve mollusc); c. Protamine from *Sepia officinalis* (cephalopod mollusc); d. *Gallus domesticus* (bird) protamine; e. P1-protamine from mouse (*Mus sp*) (mammal); f. Protamine from *Bufo japonicus* (amphibia); g. Typical protamine (iridine 2b) from *Onchorhynchus mykiss* (bony fish) (adapted from Daban, M. et al, 9).

When the mollusc protamines were compared with those of vertebrates, it was found that protamines of birds (61 amino acid residues) and P1 protamines of mammals (50 residues) contained alternating RS (also KT or RT) in the amino-terminal region (figure 2, d and e). In mammals, the amino-terminal region of the P1 protamines, which included the alternating residues SRSR, was conserved in evolution. It may have functional importance during the displacement of histones by protamine in spermiogenesis where serine or threonine residues in the protamines are phosphorylated when displacing nucleosomal histones and dephosphorylated after finishing the displacement. This forms a more compact and transcriptionally inactive chromatin (11). These B-P groups did not appear in protamines from amphibian or fish. Protamines from birds and mammals developed from genes related to those of bony fish protamines (12). This fact implied that the B-P groups in the amino-terminal part of these molecules had an evolutionary appearance independent from that of the same alternating B-P groups found in mollusc *M. turbinata* protamine.

From all this, overall conclusions can be drawn: Some large protamines (archaeogastropods, birds, mammals) display a similar molecular organization. Namely, an amino-terminal domain containing alternating B-P residues and the rest of the molecule occupied with arginine clusters. Protamines may have up to three clearly distinguishable domains: (a) a domain with clusters of basic residues, found in all protamines; (b) an amino-terminal domain with an alternating B-P region; and (c) a region with a sequence related to the carboxy-terminal part of histone H1, found thus far only in the central region of *Mytilus* protamine. In invertebrates, some mollusc protamine such as that in *Sepia officinalis* also contains *Salmon* type protamine (arginine rich monoprotamine), unlike other intermediate *Mytilus* type protamine.

6. Protamine displacement

During spermiogenesis, histones are replaced by protamines. When protamine messenger RNA (mRNA) reaches the cytoplasm, the amino acid residues are translated by diribosomes (disomes) which are present in high concentration during the period of rapid protamine synthesis. These disomes incorporate arginine rapidly but lysine very slowly which suggests that they are principally engaged in protamine synthesis (8). The newly synthesed protamines are phosphorylated by adenosine triphosphate (ATP) in cytoplasm. This reaction is catalyzed by protamine kinase. Modification of protamines by phosphorylation of their serine residues would reduce the net charge of protamine considerably and hence it would alter its interaction with DNA. The phosphorylated

protamines then penetrate through the nuclear envelope and accumulate in the nucleus. Prior to replacement, histones become highly acetylated. Presumably, this hyperacetylation greatly reduces the electrostatic attraction between the histones and DNA and facilitates histone removal. At the same time the protamines are dephosphorylated. Removal of phosphate groups from the protamines greatly increases their positive charge and attraction between the protamines and DNA. Histone replacement then proceeds, probably simply by competition for the phosphate groups of DNA by the more strongly charged protamines.

The process of histone replacement by protamines has been studied in depth in trout. Protamines appear in the mid-spermatid cells and are deposited by a mechanism that involves phosphorylation of the nascent protamine. The first phosphate is incorporated onto protamines instantly upon their entry into the nucleus. Two other phosphate groups are incorporated within the next 10 hours. Over the following 5 to 10 days, a slow process of dephosphorylation occurs (13). This long dephosphorylation period corresponds to the displacement of the somatic histones from chromatin and the extreme condensation of nucleoprotamine, typical of the mature sperm nucleus (14). The function of protamine phosphorylation is unknown, but it may modulate the protamine-DNA association. The introduction of up to 6 or 8 negative charges on the basic protamine molecule could decrease ionic contacts between the phosphoprotamine and DNA, allowing several binding conformations to be searched in the competition for histone-DNA interactions. Size and net charge of protamine species would be a factor in their efficiency of nucleosome disassembly. The size effect has been proven *in vitro* by comparison of the relatively inefficient, small salmine and iridine species with the larger and more efficient rooster (galline), squid, and cuttlefish protamines (15). Replacement of histones by protamines starts at certain locations and progressively spreads throughout the chromatin structure. Protamine binding seems to be cooperative, to both chromatin and naked DNA. Once histone displacement occurs, DNA may be packaged into a highly condensed state via the gradual dephosphorylation of protamine, which reinstates the full basic charge.

7. Function of protamines

Protamines may be present in mature sperm instead of histones to protect DNA from degradation and keep its structure compact for the transfer of the intact genetic material to the egg cells. The genetic information of DNA is not operational when protamine binds to DNA.

The basic hypotheses about the biological role of protamines in the sperm nucleus includes: the condensation of the nucleus and streamlining of the sperm (16); the

condensation and protection of the genetic message (17); and the generation of an imprinted/blank state of the male genome preceding initiation of embryogenesis in the *fertilized ovum (17).*

The first hypothesis states that the function of the histone to protamine transition is to condense and streamline the sperm head in order to reduce the weight and volume of nuclear material to be transported by the sperm cell (16). Bloch (2) argued that the sperm protein type was of little consequence as far as condensation *per se* was concerned. For example, the heads of the sperm of *Rana pipiens* (frog) and of *Loligo opalescens* (squid) were both lozenge shaped and highly condensed. However, the protein in frog sperm was a histone and in squid the protein was a monoprotamine. Therefore, there was no relationship in different species between different aspects of sperm function and the type of sperm nuclear protein. Bloch (2) suggested that the variability of the sperm nuclear proteins reflected an evolutionary insignificance to a relatively unimportant protein in an inactive nucleus. However, Kasinsky (18) showed that an evolutionary tendency becomes obvious between the type of sperm basic protein (degree of basicity) and the biology of fertilization (sperm motility, condensation and streamlining). For example, the species having internal fertilization contained sperm proteins with higher condensing ability, from which the sperm acquired a competent hydrodynamic shape. The shape permits it to swim through the viscous surroundings of the female tract, such as in the case of reptiles, birds, mammals, insects. In these cases monoprotamines, cysteine containing protamines and intermediate protamines are dominant. In contrast, the nuclear sperm proteins of species having external fertilization exhibit a lower condensing ability as is the case in frogs, horseshoe crabs, bonyfish and some mollusks. Somatic histones and intermediate protamines are present in these animals. This proposed relationship between sperm basic protein type and the biology of fertilization is consistent with the hypothetical evolutionary pathway put forward by Subirana et. al (19): somatic histone ---> histone-like ---> protamine-like (e.g., *Mytilus*) ---> arginine protamines ---> tyrosine and cysteine protamines. Although condensation can reduce volume and drag, nuclear shapes in sperm from certain organisms, such as in rat, seem to be poorly designed for swimming. It was argued that for many species the shaping of sperm nuclei was guided by processes independent of the DNA condensation process. The shaping of sperm nuclei and the pattern of chromatin condensation may be determined by interactions between DNA and a nuclear matrix (20). The nuclear matrix could direct the morphogenesis of sperm nuclei by organizing DNA into a specific pattern which will determine the ultimate shape of chromatin when condensed nonspecifically by basic proteins.

The second hypothesis states that the highly basic proteins of sperm can protect the genome from environmental mutagens. As sperm proteins become more basic, a more efficient packing of DNA is achieved. As DNA is packed tighter, the probability of it being damaged decreases. Water, enzymes and bacteria have a smaller chance of reaching it (21). The need for efficient protection of the DNA in spermatozoa is obvious, not only as a way of preserving the gene content, but also because these cells have to be transported and often stored for considerable lengths of time in the absence of DNA repair mechanisms. The appearance of cysteine bridges in mammals could reinforce this protection idea. However, as it was pointed out by Risley (17), the only effective mechanism for the protection of DNA was the ability to repair damage. Yet DNA repair mechanisms were disabled in late spermatids and sperm and only functions after fertilization (17). In addition, protamines were themselves targets of alkylating agents in germ cells and thus may enhance the sensitivity of sperm to certain types of mutagenesis (22).

The imprinted/blank hypothesis favoured by Risley (17), suggested that the compaction in sperm produced a male gamete with a genetic blank slate in which the egg can inscribe the molecular information essential for the direction of embryonic development. By this mechanism, the egg would gain immediate control of the male genome, enhancing the prospects for a coordinated regulation of expression of the embryonic genome. Both the imprinted and blank states are consistent. The imprinting could be accomplished either by introducing DNA modifications or new structures in some parts of the genome, or by removing existing DNA modifications or structure. For example, protamines can compete with DNA regulatory binding proteins which are important for spermatogenesis but not essential during embryogenesis, by displacing them from the cis-acting sequences.

The question of the function(s) of protamines in the sperm nucleus is by no means solved. It provides a challenge in the understanding of the origin and evolution of the protamine genes. Specifically, how their post-meiotic expression might be controlled and why their structures are so variable in phylogeny.

8. DNA-protamine interaction

Much of the research in DNA-protamine interactions has focused on the mode of binding of protamines to the double-helical DNA. X-ray diffraction, spectroscopic techniques, electron microscopy and chemical modification have all been employed to elucidate the complex interactions. Knowledge of the physical structure of the nucleoprotein complexes includes evidence that nucleoprotamine complexes DNA

molecules *in vivo* and tend to arrange themselves in a parallel fashion in the B form. Also, the structure of the protamine in the DNA-protamine complexes was shown to partially extend, intermediate between the α-helix and fully extended ß-structure (23).

Taken together, in the nucleoprotamine molecule, B-form DNA and partly extended protamine combine in some way so that the basic groups in the protamine bind to the phosphate groups of DNA (23).

Based on these findings, several models relating the physical structure of nucleprotamine have been proposed.

A model based upon the *Salmon* type protamines which contain several polyarginine clusters (Fig 3, Salmine) was proposed by Balhorn (24). Polyarginine segments bind to the phosphodiester backbone of DNA in an extended conformation. As adjacent arginine residues alternately bind to the opposite DNA strand, the phosphodiester chains of DNA are linked along the entire length of DNA. As a result of this linking, the DNA would be prevented from making the normal B-to-A or B-to-C conformation changes that are observed when the hydration of free DNA is decreased, transitions that do not appear to be allowed in DNA-protamine or DNA-polyarginine complexes. Subsequently, the polyarginine segment binds in the minor groove of DNA. Interspersed groups of non-arginine amino acids loop out away from the minor groove which allows the short polyarginine segments (4-6 residue clusters) to be located adjacent to one another (Fig 4), therefore, neutralizing every phosphate residue in one turn of DNA.

A second model was based upon Mammalian protamines which contain polyarginine clusters flanked by nonbasic amino acid residues located in the C-terminal and N-terminal "tail" fragments (Fig 3, bull). In addition, all mammalian protamines contain numerous cysteine residues which allow them to generate disulfide cross-links between adjacent protamine molecules. This makes nucleoprotamine complexes difficult to dissociate from DNA.

The mammalian protamine model suggests a similar pattern to the *Salmon* model of DNA binding. However, in mammalian or higher vertebrates, the presence of C- and N-terminal peptide "tails" refines sperm chromatin packaging. The N- and C-terminal peptide "tails" of mammalian protamines expose individual protamines to be cross-linked around DNA through the formation of inter- and intraprotamine disulfide bridges. The central polyarginine segment binds in the minor groove of DNA, similar to the *Salmon* type, while the N- and C-terminal tails protrude into the major groove of adjacent molecules of DNA, thus minimizing the volume required for packing (24).

```
Salmine AI   - - - - - - - - - - - - - - -  - PRRRR - - - - - SSRR PVRRRR - - - - - RPRV SRRRRRR - -GGRRRR - - - - - - - - -
Galline      ARYR - - - -SRGRSRS  RRTRRRRSPRSGRRR SPRRRRSRRRRRYGS ARRSRRSGGVRRRRY GSRRRRRRY
Bull         ARYR-CCLTHSGSRC  RRRRRRR - - - - CRR RRRR - - - - - - - - FG RRRRRR - - -V - - - CY  - TVIRCTRQ
```

Figure 3. Amino acid sequence homologies of vertebrate protamines (adapted from Subirana, J.A., 3).

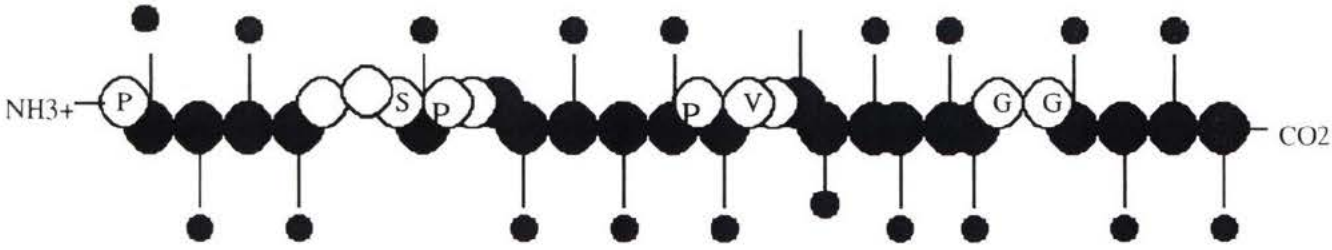


Figure 4. Binding of salmon protamine to the minor groove of DNA. Big solid circles reprresent arginine residues of protamine. Small solid circles represent phosphates in one turn of a DNA helix. Three peptide segments, S-S-S-R-P-V, P-R-V-S and G-G, loop up above the groove and allow the four polyarginine segments in salmine to move into position adjacent to one another and bind to and neutralize every phosphate residue in one turn of DNA (adapted from Balhorn, R., 24).

A third model proposed by Warrant and Kim (25) was based upon crystallographical data on the interactions between protamines and transfer RNA (tRNA) from yeast. From their analysis they concluded that all protamines have a structure composed of three or four α-helical domains connected by two or three flexible joints (Fig 6). The α-helical domains were more likely to lie on the major groove of DNA (Fig 5) which was consistent with a model by Subirana (3). Two or three arginines of each α-helix domain neutralize and form hydrogen bonds to two or three phosphates across the major groove. The remaining arginines function similarly with the phosphates of neighbouring double helix or helices. Thus, each domain is binding or cross-linking two or more double helices together, thereby causing a heavy condensation of DNA double helices (25).

The advantages of this protamine model are that the α-helical domains have structural rigidity sufficient to stabilize DNA double helix-double helix cross-linking or "condensation". In addition, the joints between the α-helical domains provide flexibility in the arrangement of the domains, thus two DNA double helices in different arrangements can be brought together.

The inability to detect a fundamental unit of condensed sperm chromatin which contains nucleoprotamines as compared to the relatively well-characterized structure of the nucleosome of somatic chromatin which contains histones is more likely due to the extremely high variability of the amino acid sequence of protamines. This has made the determination of conserved structure of nucleoprotamine difficult.

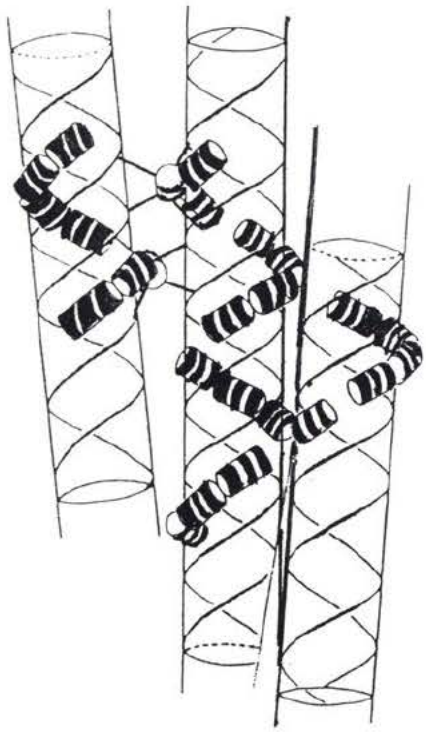


Figure 5. Schematic drawing to show parts of nucleoprotamine structure. Two types of possible cross-linking are shown. Between the left and middle DNA double helices, there are two α–helical segments with arginine sides chain shown as straight lines. Each segment forms two contacts across a major groove of one DNA double helix and two across a minor groove of the other DNA double helix. Between the middle and the right DNA double helices, three α-helical segments from one protamine molecule contact the major groove of the middle DNA double helix and the fourth segment contacts the major groove of the right DNA double helix (adapted from Warrant, R.W., Kim, S.H., 25).

Salmine Al	P++++SSS+	PV+++++	P+VS++++++	GG++++
Iridine	P++++SSS+	PV+++++	P++VS++++++	GG++++
Thynnin	P++++EAS+	PV+++++	Y++STAA+++++	VV++++
Clupine	A++++SSS+	PI++++	P+++TT++++	AG++++
	HELIX A →	HELIX B →	HELIX C →	HELIX D

Figure 6. α-helical domains of some protamines. All the known sequences of protamines (salmine, iridine, thynnin, clupine) are arranged into three or four α-helical domains connected by two or three flexible joints (adapted from Warrant, R.W., Kim, S.H., 25).

II.The Sperm Nuclear Basic Proteins of Bivalve Mollusks

During spermatogenesis in invertebrates, some phylogenetic groups such as sea urchins (26) maintain their histones. Mollusks, on the other hand, have their histones replaced by proteins of an intermediate composition between somatic histones and typical protamines. The protein composition of the sperm chromatin of the invertebrates is extremely heterogeneous (18).

In the case of bivalve mollusks, the major basic protein constituents of their sperm chromatin can be classified into four basic groups based on their electrophoretic mobility (27): PL-I, PL-II, PL-III and PL-IV (PL=protamine-like). Based on this classification, the Bivalve mollusks were grouped into five major categories: Pectinidae (group O), Veneridae (group I), Cardiidae (group II), Tellinidae (group III) and Mytilidae (group IV) (Table I, II and Fig 7).

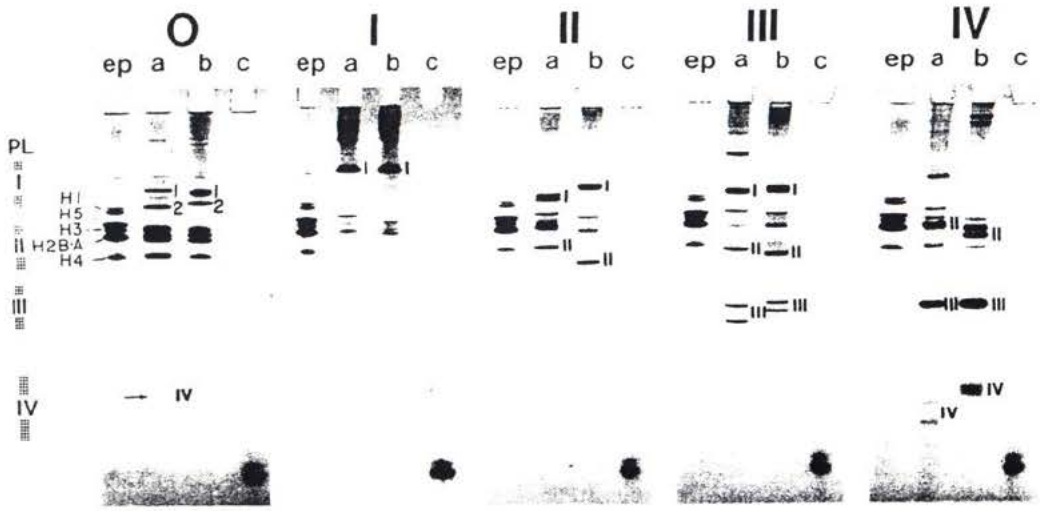


Figure 7. Whole basic proteins from the sperm nuclei of some bivalve mollusks. The roman numbers correspond to the different protamine groups. O, a: *Pecten maximus;* O, b: *Crassostrea gigas*; I, a: *Callista chione*; I, b: *Spisula solidissima*; II, a: *Cardium edule*; II, b: *Ensis ensis*; III, a: *Macoma nasuta*; III, b: *Donax trunculus*; IV, a: *Lithophaga lithophaga*; IV, b: *Mytilus edulis*; ep=chicken erythrocyte histones, c=clupein (protamine) (adapted from Ausio, J., 27).

Table I. Classification of the bivalve mollusks according to their protamine-like composition (*)

Group	Protamine-like composition
Pectinidae-O	(Mostly histones)
Veneridae-I	PL-I
Cardiidae-II	PL-I, PL-II
Tellinidae-III	PL-I, PL-II, PL-III
Mytilidae-IV	PL-II, PL-III, PL-IV

*. From Ausio (1986) (4).

Table II. Bivalve Mollusks classified according to Kuhn-Schnyder (28), Coan-Carlton (29) or according to their protamine-like composition (Ausio, 4, 30).

Subclass	Order	Family	Genera	Species	Protamine-like group
Pteriomorphia	Pteroida	Ostreidae	*Aequipecten*	*irradiaus*	O
Pteriomorphia	Pteroida	Pectinidae	*Crassostrea*	*gigas*	O
			Pecten	*maximus*	O
Anomalodesmata	Pholadomyoida	Lyonsidae	*Agriodesma*	*saxicola*	I
Anomalodesmata	Pholadomyoida	Lyonsidae	*Mytilimeria*	*nutalli*	I
Heterodonta	Veneroida	Veneridae	*Callista*	*chione*	I
Heterodonta	Veneroida	Mactridae	*Spisula*	*solidissima*	I
		Mactridae	*Tresus*	*capax*	I
		Veneridae	*Protothaca*	*staminea*	I
		Veneridae	*Tapes*	*japonica*	I
		Veneridae	*Saxidomus*	*giganteus*	I
Heterodonta	Veneroida	Cardiidae	*Cerastoderma*	*edule*	II
	Veneroida	Cardidae	*Clinocardium*	*nuttallii*	II
	Veneroida	Solenidae	*Ensis*	*ensis*	II
	Veneroida	Solenidae	*Ensis*	*minor*	III
Heterodonta	Veneroida	Tellinidae	*Macoma*	*nasuta*	III
Heterodonta	Veneroida	Donacidae	*Donax*	*trucalus*	III
Isofilibranchia	Mytiloida	Mytilidae	*Mytilus*	*edulis*	IV
		Mytilidae	*Mytilus*	*californianus*	IV
Isofilibranchia	Mytiloida	Mytilidae	*Lithophaga*	*lithophaga*	IV
	Myoida	Myidae	*Mya*	*arenaria*	II
	Myoida	Hiatellidae	*Panope*	*generosa*	I

Pectinidae (group O)

The group Pectinidae (O) includes those bivalve mollusks whose sperm protein consists only of histones which are similar to somatic histones (H2A, H2B, H3, H4). Unlike somatic histones, they contain a sperm specific histone H1 (arginine-rich) with a lower electrophoretic mobility. This sperm-specific H1 usually shows microheterogeneity or polymorphism. The components (H1-1 and H1-2) can often be distinguished (Fig 7, O, a,b). In some cases a minor protein with high mobility also appears to be present [Fig 7, O, a, (called S protein in *Swiftopecten swifti* and *Glycymeris yessoensis*)].

During spermatogenesis in species belonging to this O group, the somatic histones are not displaced by protamine-like proteins. Instead, they are replaced by a new set of specific germinal histones which maintain a nucleosomal chromatin structure similar to that in somatic cells (31). However, in the latter, the linker DNA of the nucleosome is significantly elongated. And the S-protein, like histone H1, is shown to be located on the linker DNA of a nucleosome. The location of the S protein and histone H1 in sperm chromatin suggests that the S protein may, like histone H1, may be involved in chromatin condensation. It has been postulated that the organization of the linker part of a nucleosome, which is dependent on the length of its DNA and the properties of the bound proteins, is important for the higher order structure organization of chromatin in this group.

The other groups of Bivalve mollusks all contain some highly specialized protamine-like proteins (PL) which show an amino acid composition intermediate between histones and protamines. These groups have different combination of PL components (PL-I, II, III, IV) which often coexist with a small percentage of somatic histones in the sperm nucleus. In most cases, the PL proteins account for 70-80% of the total basic nuclear sperm proteins.

Veneridae (group I)

In group Veneridae (I), there is a major PL-I component which exhibits a very low electrophoretic mobility (Fig7, I, a,b). It coexists with somatic histones at a ratio of 3:1 (w/w). The composition and structure of PL-I is similar to histone H1 (as discussed in part III). The organization of the sperm chromatin in this I group consists of a segregated structural mixture in which short oligonucleosome clusters are scattered within long PL-I nucleoprotamine filaments (32). *Spisula solidissima* can be considered as one of the representative species of this group.

Cardiidae (group II)

In the cardiidae group, the number of sperm chromosomal proteins increases to two: PL-I with a relatively low electrophoretic mobility, and PL-II with a mobility slightly higher than histone H4 (Fig 7, II, a,b). Both proteins coexist with the complete typical set of somatic like histones. The structural arrangement of nucleosome with PL-I, PL-II is not clear yet.

Tellinidae (group III)

The tellinidae group includes PL-I, PL-II as in Cardiidae, as well as PL-III, which has a higher electrophoretic mobility than any of the histones. PL-III represents approximately 20% of total sperm proteins, exhibits a microheterogeneity and/or polymorphism (Fig 7, III, a,b) similar to that observed in protamines. The sperm chromatin is organized by these protamine-like proteins PL-I, PL-II and PL-III. The mechanism by which the PL proteins interact with each other and with DNA to form their structural organization is not known. Like other groups of Bivalvia mollusks, mature sperm of the group III organisms still retain some fraction of the somatic-type histone complement. The significance of this remnant histone fraction is still not clear.

Mytilidae (group IV)

The mytilidae group, PL-I appears to be absent and replaced by PL-IV, which has the highest electrophoretic mobility of all PL proteins. The sperm nucleus of the group IV organisms consists of approximately 80% PL proteins (including PL-II, PL-III and PL-IV) and 20-25% histones. Interestingly, the chromatin structure and organization resulting from the association of these proteins with DNA is similar to that found in the group Veneridae. Short oligonucleosomal stretches interspersed in long bundles of PL-DNA nucleofilaments (33). DNA retains its B conformation in these nucleofilament complexes. PL-II, PL-III and PL-IV are most likely the proteins which are responsible for the fibrilar DNA organization (27).

Despite the above mentioned differences existing between all these groups, common features exist. Except for *Mytilus* PL-II and PL-IV, all the other PL proteins have similar amino acid compositions: 20-30% arginine, 20-30% lysine, 30-40% serine+alanine. Tryptophan is absent in all except for *Spisula* PL-I, which contains one residue per molecule. Protamine-like proteins coexist with histones to varying degrees. However, the full complement of histones is always present. This is important in that PL proteins do not

completely replace any individual histone during spermatogenesis. On the other hand, it may be an indirect evidence that the DNA regions associated with these histones in the sperm, must have a folded nucleosomal structure similar to that described for the chromatin organization in somatic cells (4).

The only group with exceptions to these common features is the group Pectinidae (group O). Most of the organisms of the O group lack PL proteins. Instead, they contain somatic-like histones, except for the histone H1 variants which seem to be different from the somatic counterparts. In several instances, such as *Swiftopecten,* a small amount of a high electrophoretic mobility PL is present. Although, they still retain a typical somatic-like nucleosomal organization. Also in Mytilidae (IV), the PL composition is the most complex and the histone level is the lowest in all the groups. The reason of this difference is unclear (4).

The nature of the interaction between protamine-like proteins and DNA *in vivo* is unclear. Nevertheless, *in vitro* studies have shown that the nuclear protamine-like protein complex is very similar to typical nucleoprotamine structures (34).

III. Nuclear proteins and chromatin organization during spermatogenesis in *Spisula solidissima*

During spermatogenesis in *Spisula solidissima* [a member of the Veneridae (I) group (27)], spermatocytes and spermatids undergo an important biochemical turnover which involves the synthesis of PL-I. As a result, 75% of the histones which were originally present in the spermatogonial cells are displaced by this sperm-specific component. This protein component is called protamine-like protein (PL) because it is only present in spermatic cells, and has an amino acid composition similar to protamines. The mature sperm also contains a small amount of the complete histone set, which retains the same relative stoichiometry as in somatic chromatin. The protamine-like component (PL-I) represents 75-77% of the total sperm nuclear protein, while the histone complement accounts for 23-25%.

The histones from the core histone complement have amino acid compositions very similar to those of somatic histones from calf thymus (34). This is not surprising because core histones have been extremely conserved throughout evolution. Only the amino acid composition of H2B exhibits some differences with a slightly lower lysine content and higher arginine. It was believed that these differences might reflect the spermatic nature of this histone (35).

The linker histone of Spisula's sperm has similar physical properties as somatic

histone H1. For example, they are extractable in 5% PCA, they have a molecular weight of 20,700 (about 200 amino acids), the proteins lack of tryptophan, and they have a similar amino acid composition. Also, the pattern of trypsin digestion shows that sperm-specific histone H1 has a trypsin-resistant core, like the somatic histone H1. A unique feature in this histone H1 is the presence of 2 cysteines located in the globular region (35). In addition, H1 is extractable from sperm chromatin in salt at 0.6M NaCl.

PL-I

Spisula's sperm PL-I contains approximately 300 amino acids residues (MW 33,500) and is significantly larger than fish protamines (30-50 amino acids). PL-I however has some similarities to histone H1. The composition of PL-I is rich in lysine (24.8%) and alanine (14.2%) and aromatic amino acids as tyrosine (0.3%) and phenylalanine (0.3%; one per molecule). However, PL-I has high arginine (23.1%) and serine (21.7%) content and a low proline composition. In addition, PL-I contains one tryptophan, which is a unique property for a chromosomal protein. Tryptophan has never been reported previously in either histones or protamines (7).

Other features of the PL-I protein is that it has a tendency to aggregate at high ionic strength or high pHs, and is highly charged. It was initially thought that the aggregation was the result of hydrophobic interactions between specific regions of this protein probably involving alanine residues and the neutralization of some of the lysine residues at high pH (7).

Structure of PL-I

Trypsin digestion studies show that *Spisula* sperm PL-I has a trypsin-resistant core which is 75 amino acids long (MW 7,615 Da). This trypsin resistant peptide migrates slightly faster than the somatic histone H1 core in urea/acetic acid gels. The only tryptophan residue of PL-I is found within the core region, as was established by fluorescence spectroscopy. The amino acid sequence of PL-I core is (36):

Lys-Gly-Ser-Ser-Gly-Met-Met-Ser-Met-Val-Ala-Ala-Ala-Ile-Ala-Ala-Asn-Arg-Thr-Lys-Lys-Gly-Ala-Ser-Ala-Gln-Ala-Ile-Arg-Lys-Tyr-Val-Ala-Ala-His-Ser-Ser-Leu-Lys-Gly-Ala-Val-Leu-Asn-Phe-Arg-Leu-Arg-Arg-Ala-Leu-Ala-Ala-Gly-Leu-Lys-Ser-Gly-Ala-Leu-Ala-His-Pro-Lys-Gly-Ser-Ala-Gly-Trp-Val-Leu-Val-Pro-Lys-Lys

The amino acid composition of the PL-I core is significantly different from that of

the consensus sequence from the corresponding region of histone H1 family proteins. PL-I contains three aromatic amino acid (Phe, Tyr, Trp) residues (7). The core region is preceded by an N-terminal 'tail' of 150 residues and followed by a C-terminal 'tail' of 75 residues (36).

CD spectra shows that core region has twice as much α-helix as the total PL-I protein. Similarly, the anti parallel ß-sheet composition of the core is double that of total PL-I and three times that of the 'tails'. No parallel ß-sheet was detected in PL-I. The flanking 'tails' contain more ß-turns and random coil conformation than the core region (36).

Measurement of the diffusion coefficient by analytical ultracentrifuge indicates that at the tertiary level, the core has a compact globular conformation whereas the whole PL-I molecule is quite asymmetric and the N- and C-terminal tails are highly extended in solution (36).

Chromatin organization in the sperm of *Spisula*

PL-I coexists with histones at a ratio of 4:1 (w/w) in the sperm of *Spisula* suggesting that PL-I must have an important role in the organization of chromatin. Micrococcal nuclease digestion and chromatin salt fractionation showed that *Spisula* sperm chromatin had a heterogeneous structure. The protein/DNA ratio 0.83 (w/w) was slightly higher than that of protamine containing sperm nuclei (0.65-0.67) and lower than that of typical somatic chromatin (1.0). An early model was proposed by Ausio and Van Holde (32) was designed to account for this chromatin heterogeneity. Accordingly, short oligonucleosomal chromatin regions were interspersed within very long stretches of PL-I containing regions (32). Therefore, during the late stages of spermatogenesis in *Spisula*, DNA was packed in two different organizations. Most of the DNA (80%) remained associated with PL-I, whereas the rest was packed in nucleosomes. Yet, it is unclear why some nucleosomes remain in the mature sperm of *Spisula* as well as other *bivalve mollusks*. It has been shown that the presence of nucleosomal histones is not adventitious (32). It appears possible that the nucleosomal DNA fraction may have a critical role in the fate of this chromatin fraction during the early development stages that follow oocyte fertilization (32). Despite all that is known, the details of how PL-I protein binds to DNA are still unclear.

Relationship between PL-I and histones H1-H5:

PL-I from the sperm of *Spisula* is similar to histone H1/H5. The size of PL-I (300

amino acids) is similar to H1 (214 amino acids) and H5 (195 amino acids) when compared to the small fish protamines (about 30-35 amino acids) (8). Secondly, the lysine and alanine content of PL-I is similar to that of the proteins of the histone H1 family. Thirdly, PL-I has a similar structural organization: globular core region flanked by highly charged, less structural tails. Finally, the size of PL-I C-terminal tail is similar to histone H1 and H5: PL-I 85 amino acids, H1 110, H5 90. Yet, the N-terminal tail's length is quite different between PL-I (140 amino acids) and H1 (40 amino acids), H5 (22 amino acids) (37).

The amino acid composition of PL-I is more similar to histone H5 than to histone H1. The overall lysine and arginine content (48%) of PL-I is closer to chicken erythrocyte histone H5 (34%) than to calf thymus histone H1 (29%) (37). The PL-I core region has 39% amino acid sequence similarity to chicken erythrocyte histone H5 core compared to 21% for calf thymus histone H1 core (36). However, the arginine and serine content of PL-I from *Spisula* is much higher than that of calf thymus histone H1 and chicken erythrocyte histone H5 (7). Two aromatic amino acids (Phe and Tyr) are present in *Spisula* PL-I as well as in histone H1 and H5. PL-I has a tryptophan which is neither present in histone H1 nor H5. PL-I from *Spisula* has the same percentage of methionine as histone H5 while molecule of the histone H1 family lacks this amino acid (7).

In terms of secondary structure, PL-I from *Spisula* and histone H1 and H5 all contain a trypsin-resistant core. PL-I from *Spisula* contains 35% anti parallel ß-sheet in its trypsin-resistant core, neither the canonical H1 histone nor the more specialized histone H5 shows any indication of any ß-structure in their trypsin-resistant cores. Instead, the α-helix content for the core regions of histones H1 and H5 is 30% compared to 10% observed for the core of protamine-like protein from *Spisula solidissima* (36).

Birds and mollusks are different branches of the animal evolution tree (deuterostome and protostome). There is no direct line of descent from mollusks to chickens (36). This paradox makes it difficult to understand the possible evolutionary relationship which could exist between *Spisula* PL-I and chicken erythrocytes histone H5.

Taken together, experimental evidence and sequence similarity suggest PL-I is a member of the histone H1 family.

IV. Nuclear proteins and chromatin organization during spermatogenesis of *Mytilus californianus*

The mature sperm of *Mytilus californianus* contains three basic nuclear proteins, PL-II, PL-III and PL-IV, which coexist with a complete set of histones (10-15%). This

histone complement is relatively low compared to other bivalve mollusks (32). All these PL proteins exhibit some extent of microheterogeneity. PL-II has at least two protein components (38), with molecular weights 15,601 Da and 15,865 Da. The PL-II protein was initially considered to be a sperm-specific histone H2B because its amino acid composition and physical properties were most similar to that of histone H2B. Also PL-II had a similar electrophoretic mobility in urea/acetic acid gels to H2B (39). But the solubility of PL-II protein in 5% PCA and its high-lysine content later revealed a similarity to histone H1. In addition, PL-II similarity to histone H1 and to H5, was supported by immunological cross-reactivity studies (40). Also, trypsin digestion of PL-II generates a homogeneous trypsin-resistant peptide which has a totally different electrophoretic behaviour than the mixture of peptides retained upon trypsin digestion of H2B (40).

The amino acid sequence of the trypsin-resistant PL-II core has the same consensus sequence as that of the histone H1 core. There is a substantial similarity sequence between the sequence of the PL-II core and that of the trypsin-resistant core of the spermatic H1 histones from other marine invertebrates.

The secondary structure of PL-II is similar to that of Spisula PL-I (41) which also belongs to the histone H1 family of proteins. The amount of α-helix, β-sheet and β-turn domains in PL-II are similar to those of histone H5 (41). Histone H5 is a cell differentiation-specific histone H1 variant which is found in the nucleated erythrocytes from birds.

The tertiary structure of PL-II consists of a compact globular core flanked by less structured extended N- and C-terminal domains ('tails'). This is similar to the members of the histone H1 family. Therefore, PL-II can be considered to be a highly specialized member of histone H1 family.

In bivalve mollusks, sperm-specific histone H1-like proteins are members of the PL-I class of proteins. Despite the fact that PL-II has a higher electrophoretic mobility than that of most PL-I, its structural features are very similar to those of PL-I, and therefore should be included within this protein group. The presence of a highly specialized histone of the H1-like protein in the sperm of the bivalve mollusks seems to be a general feature shared by all members of this taxonomic group (4).

Other PL proteins exist and showing varying degrees of microheterogeneity. PL-III has a higher electrophoretic mobility than PL-II. Although this protein exhibits an apparent electrophoretic homogeneity, it has been shown to have microheterogeneity (38). Very little is known about the structural features of PL-III. PL-IV consists of at least three subcomponents which are present in constant stoichiometry (42). The reason for this

microheterogeneity is not the result of chemical posttranslational modification such as phosphorylation (43), but the result of different amino acid composition. The average molecular weight of these subcomponents is approximately 5 kDa. The sequence of some of the PL-IV fraction is already available and shows alternating disposition of alanine and lysine. These poly (Lys-Ala-Ala-Lys) or poly (Lys-Ala-Ala) motifs exist in an α-helix conformation (42). PL-IV is closely related to the C-terminal region of histone H1. Both PL-IV and histone H1 are lysine-rich proteins. PL-IV has 43.2% identity to C-terminal end of rabbit thymus histone H1 in a 44 amino acid region (44).

It was hypothesized that it could be possible for PL-II to operate in conjunction with PL-IV. This hypothesis was based on the fact that both PL-II and PL-IV were soluble in 0.6M NaCl, and could be coextracted under these conditions (40). The PL-II protein accounted for 20% and PL-IV for 6% of total sperm nuclear protein which when taken together with the respective molecular weight (PL-II 15,600, PL-IV 6,500), suggested that PL-II and PL-IV were present in equal stoichiometric amounts in the sperm nuclei. PL-IV was compositionally related to C-terminal region of histone H1 and PL-II was related to histone H1 (45).

It has been recently shown that in *Mytilus trossulus*, PL-II and PL-IV are the result of post-translational cleavage of a common protein precursor (46). Therefore, PL-II together with PL-IV represent the complete histone H1-like related protein (PL-I) from the sperm of *Mytilus*. The post-translational cleavage of PL-I has an important implication for the evolution of protamines. Excision of C-terminal domain of the PL-I (PL-II + PL-IV) precursor could be viewed as one of the many different mechanism involved in this evolutionary process (46).

The coexistence of histones and PL proteins in the sperm of *Mytilus* suggests that the sperm chromatin organization of this organism has peculiar features, and that PL proteins are responsible for novel forms of chromatin condensation and packing. However the detailed structural feature of this chromatin organization are very poorly understood at present.

The objective of this thesis is to study the interactions of protamine-like protein of Mollusks with DNA in the two species and to study how the chromatin conformation that arise from this interactions. To this purpose, we have chosen two species *Mytilus californianus* and *Spisula solidissima*, because one contain only one PL-I protein, the other contain three PL (PL-II, PL-III and PL-IV).

MATERIALS AND METHODS

Organisms

Samples of the surf clam, *Spisula* (*Mactra*) *solidissima* were obtained from the Department of Marine Resources from the Marine Biological Laboratory (Woods Hole, MA). The ripeness of the gonads was checked periodically. The gonadal development started by February-March and reached its maximum by June-August, depending on the year. Only these individuals collected at the stage of maximal gonad ripeness were used in the present work.

Mytilus californianus was collected locally. The gonad is ripe almost all year round (mature sperms coexist with immature sperms).

Sperm collection and preparation of nuclei (from *Mytilus californianus*):

After opening carefully the shell, the gills were completely removed and a small incision was made in the gonadal tissue. The sperm, which was spontaneously released, was then collected with the aid of Pasteur pipette and was resuspended in buffer A (0.15M NaCl, 10 mM Tris-HCl, pH 7.5, 0.2 mM phenylmethylsulfonyl fluoride, PMSF) on ice. The sperm suspension was filtered through eight folds of cheese-cloth and was centrifuged at 2000xg in a JA-20 rotor at 4°C for 10 min. The pellet was resuspended gently in buffer B (buffer A + 0.5% Triton X-100, w/w) with a glass rod. After incubation for 10 min on ice, this suspension was centrifuged at 2000xg for 10 min in a JA-20 rotor at 4°C. The resulting pellet was resuspended in buffer A and centrifuged at 2000xg for 10 min at 4°C. The pellet was resuspended in 1 vol of buffer A and 1 vol of cold 80% glycerol and was stored at -80°C.

To quantify the DNA, an aliquot of the nuclei suspension was diluted 400-fold with distilled water. After vigorous vortexing, it was adjusted to 0.5% sodium dodecyl sulfate (SDS). The absorbance (A_{260}) was then measured against an appropriate blank. The extinction coefficient of DNA is 20 at 260 nm.

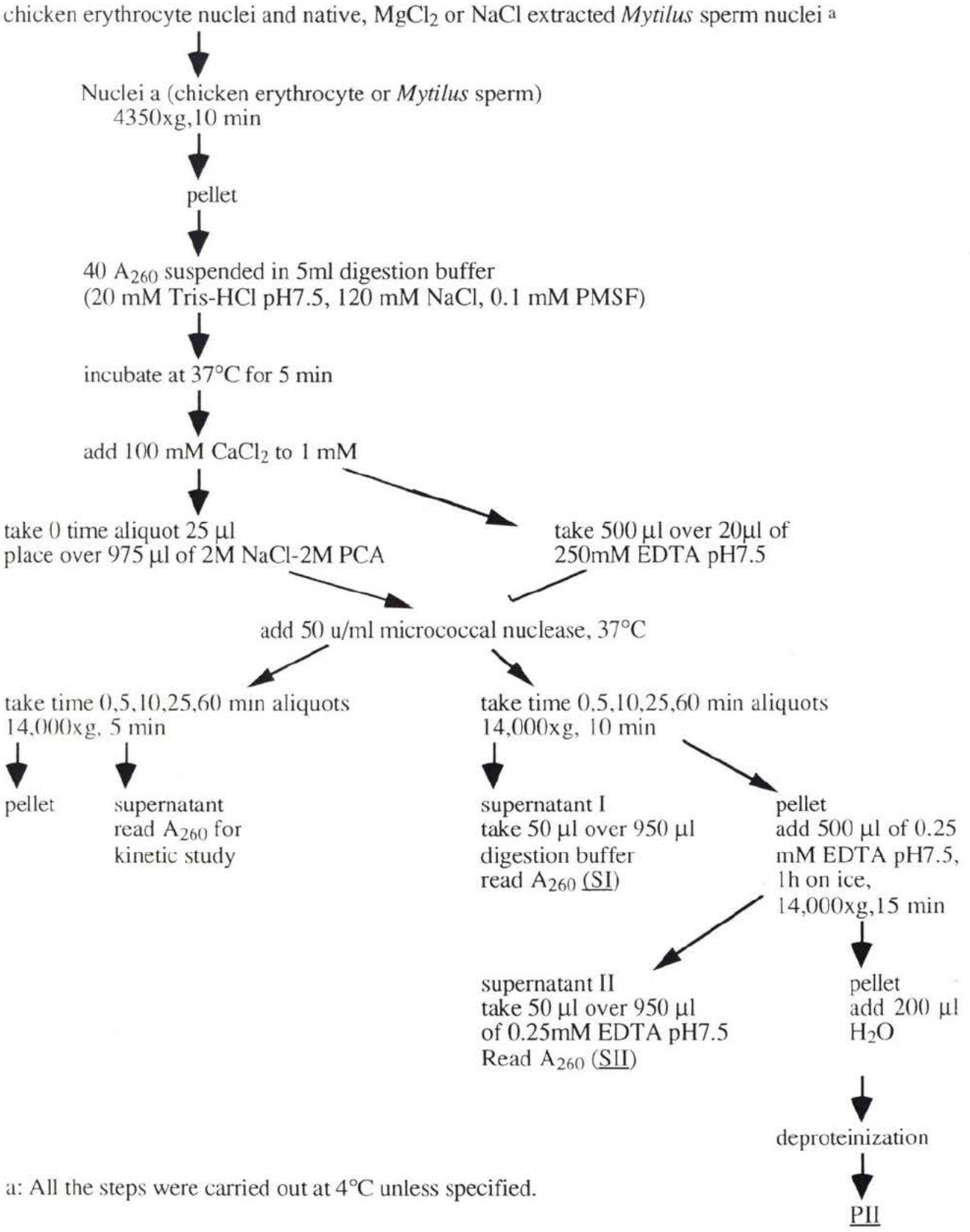
Extraction of nuclear proteins

The nuclear suspension (in glycerol) was spun down at 3000xg for 10 min in a JA-20 rotor at 4°C and the resulting pellet was resuspended in 0.4N HCl and was further homogenized in a Dounce homogenizer with 8-10 strokes. It was then incubated at 37°C for 5 min. The suspension thus obtained was then centrifuged at 14,000xg for 10 min at 4°C. The acid extract was further precipitated with 6 vols of acetone at -20°C for 20 min

and the protein precipitate was dried in a speed vac concentrator.

Chromatin digestion with Micrococcal Nuclease

Digestions were performed in digestion buffer C (20 mM Tris-HCl, 120 mM NaCl, 0.1 mM PMSF 1mM $CaCl_2$, pH7.5). The nuclear solution in this buffer at an 40 absorbance at 260 nm wavelength (A_{260}) was prewarmed to 37°C for 10 min. Digestions were carried out at 37°C with 50 units micrococcal nuclease (Worthington)/ml chromatin solution. The kinetics of DNA digestion was determined through a time course digestion. Following the addition of nuclease, the amount of A_{260} absorbing material soluble in 1M NaCl-1M $HClO_4$ (PCA) was determined at each digestion time point. The absorbtion was corrected for hyperchromicity by dividing the absorbance by a factor of 1.67. Once the right time of digestion was established, the nuclear suspension was digested and the reaction was stopped by addition of EDTA to a final concentration of 20 mM on ice. This was then centrifuged at 14,000xg at 4°C and a supernatant (SI) and a pellet (PI) were obtained. The PI pellet was resuspended and hypotonically lysed in 0.25 mM EDTA pH7.5 for 1 hr on ice, and was centrifuged at 14,000xg for 15 min at 4°C to obtain supernatant II (SII) and pellet II (PII). PII was resuspended in water for protein and DNA analysis. The percentage of chromatin present in each fraction was determined from the absorbance at 260nm of each sample. This was corrected for turbidity according to the method of Subirana (47). The absorbance at 260nm of the total chromatin before digestion was taken to be 100%. The whole process is illustrated as following.



a: All the steps were carried out at 4°C unless specified.

Protein Analysis

The proteins from the SI, SII and PII aliquots were extracted by adding HCl to a final concentration of 0.4N, followed by incubation at 37°C for 5 min. The solution was then centrifuged at 14,000xg at 4°C for 10 min. Supernatants were precipitated by adding 6 vols of acetone at -20°C for 20 min. The cloudy suspension was then centrifuged for 20 min as above. The protein precipitate was dried and analyzed by urea/acetic acid polyacrylamide gel electrophoresis.

DNA Analysis

Deproteinization of SI, SII and PII was performed by adding 1 mg of proteinase K to 10 mg of DNA (as determined by A_{260}, and assuming 1μg DNA≈1μg histone), and incubating at 56°C for 30 min. After incubation, the solution was extracted with phenol/chloroform/isoamyl alcohol (25:24:1). The upper aqueous phase was transferred to an Eppendorf tube. In the case of PII, this pellet was extracted once more with 1-2 vol phenol containing 0.8% SDS to solubilize the nuclear membranes. The procedure was repeated until the interphase was free of any precipitated proteins. The aqueous phase was then ethanol precipitated. The pellet was washed with 2 vols of cold 70% ethanol and air dried. The pellet was finally resuspended in TE buffer (Tris-EDTA pH8.0) and 1/5 vol of 5x DNA sample buffer (0.25% bromophenol blue, 30% glycerol, 0.02% SDS) and loaded onto a 1% agarose gel.

Selective $MgCl_2$ extraction of *Mytilus californianus* nuclear proteins

Nuclear proteins from the sperm of *Mytilus californianus* were suspended in 20 mM Tris-HCl pH7.5 at 0.6mg/ml with different $MgCl_2$ concentrations (40, 80, 160, 200 mM). The mixtures were incubated at 4°C for 30 min and the extracted proteins were solubilized in 0.4N HCl and precipitated as described above and saved for protein gel eletrophoresis.

Selective extraction of the protamines from the sperm of *Mytilus californianus* with 140 mM $MgCl_2$ and 0.6M NaCl

Mytilus sperm nuclei, which had been kept frozen at -60°C in 40% glycerol were thawed and spun down at 3500xg in an JA-20 rotor at 4°C for 10 min. The pellet was resuspended in 20 mM Tris/HCl pH7.5 buffer and the final DNA concentration was adjusted to 0.6mg/ml. Then 1 vol of 20 mM Tris/HCl pH7.5, 280 mM $MgCl_2$ (or 1.2M

NaCl) was added to this and the suspension was pipetted up and down very gently (to prevent lysis) for 30 min at 4°C to let $MgCl_2$ or NaCl extract proteins from *Mytilus* sperm chromatin. The suspension was centrifuged at 3500xg. The supernatant obtained was saved for protein gel electrophoresis. The pellet was resuspended in digestion buffer for micrococcal nuclease digestion (see Chromatin digestion) and protein/DNA analysis. Usually, the chromatin lysed when resuspended in digestion buffer after $MgCl_2$ extraction. In order to prevent this, 0.0125% Nonidet P-40 was added to digestion buffer.

Cysteine location determination of *Spisula* PL-I

1. Reduction of disulfide linkage

Spisula PL-I S-S linkages were reduced under denaturing conditions as described by Kuehl (47). In brief, the proteins, at a concentration of 6 mg/ml in 8 M urea, 25 mM Tris-HCl pH7.5, were reduced in the presence of 4% ß-mercaptoethanol for 2 h at room temperature.

2. 5% polyacrylamide-2.5M urea-acetic acid gel electrophoresis

Both native *Spisula* PL-I and disulfide bond reduced PL-I were loaded onto this 5% polyacrylamide (acrylamide/bisacrylamide=30:2)-2.5M urea-acetic acid gel (Fig 18A and B in P.57).

3. Thiol labelling of *Spisula* PL-I with bromobimane and NBS cleavage

To determine the location of cysteine in the PL-I protein from *Spisula*, 50µg of PL-I protein (in 10µl of water) was suspended in 20mM Na_2PO_4, 10 mM Tricine-NaOH pH7.0 and incubated in 1µl of ß-mercaptoethanol for 1 h at room temperature in order to cleave the disulfide bridges that could be present in the native protein. The solution was acidified by bringing it to 0.4 N HCl and the protein was precipitated with cold acetone as described before. The protein pellet was then dissolved in 10µl of 20 mM Tricine pH7.0, to which a 20-fold excess of fluorescent bromobimane qBBr (quaternary salt of monobromotrimethylammoniobimane, Molecular Probes Inc.) was added and the solution was incubated in the dark for 20 min at room temperature. The protein was reprecipitated with cold 0.4 N HCl-acetone. The protein pellet was then dissolved in 12.5µl of 8 M urea, 5% acetic acid with or without NBS (protein :NBS=1:13, w/w), and incubated for 15 min in the dark at room temperature. Then the mixture was immediately loaded onto a 15% 2.5 M urea-acetic acid PAGE (Fig 18C in P.57). Pyronine was not added to the sample buffer in order to prevent the fluorescence disturbance from pyronine. The labelled peptide bands

were visualized and photographed with an ultraviolet transilluminator using a Kodak wratten gelatin filter No. 15 (49). All the steps were carried out in the dark, including electrophoresis.

Purification and fractionation of cysteine-containing peptide of *Spisula* PL-I

1. Cysteine pyridylethylation

The protein PL-I from *Spisula* was pyridylethylated to precisely quantify the number of cysteines using amino acid analysis. The pyridylethylation procedure used is described as follows: the protein (1 nM) was dissolved in 44 µl of 6.8 M urea, 60 mM Tris-HCl, 1.25 mM EDTA pH7.6, and 2.3% ß-mercaptoethanol. The solution was incubated for 3 h at room temperature in the dark to reduce the disulfide bonds of cysteines which could be present in the protein. Then, 8µl of 4-vinylpyridine were added and incubated for 2 h at room temperature (30). The sample was then immediately desalted using a 10 x 1.5 cm Sephadex G-25M desalting column PD-10 (Pharmacia). The protein elutes first (in tube 4, 5, 6 as determined by absorbance at 230 nm), the salt (urea, Tris) comes out last in tubes 7 to 12 (as measured by refractory index), the elution profile in shown in figure 8.

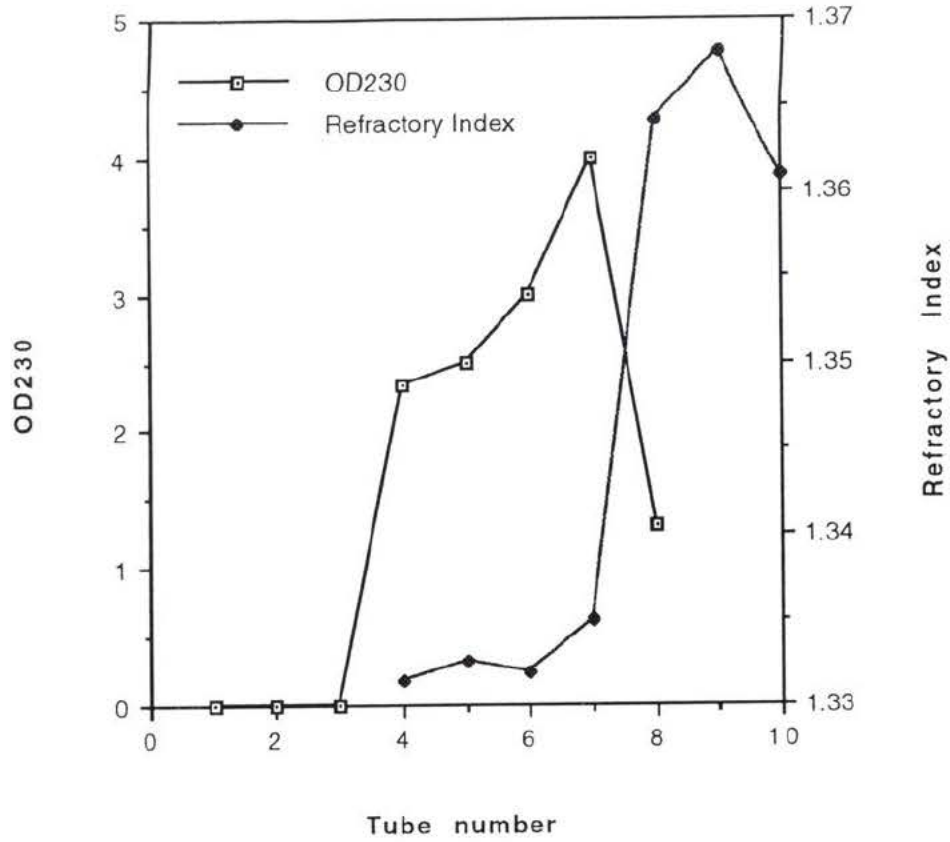


Figure 8. Elution profile of pyridylethylated protein through desalting column, eluted with water.

2. NBS cleavage of pyridylethylated *Spisula* PL-I

The desalted protein was lyophilized overnight. 30μgr of this pyridylethylated *Spisula*'s PL-I protein were cleaved with NBS in 50% acetic acid (protein:NBS=2:1, w/w). The reaction was allowed to proceed for 15 min at room temperature in the dark and was immediately dried in dark until acetic acid had evaporated.

3. Ion-exchange chromatography

Following the above steps, the dried peptides were dissolved in 800 μl of HPLC water and immediately injected into a CM Sephadex C-25 ion-exchange column to separate the peptides generated by the NBS cleavage. The elution buffer was 20 mM Na_2PO_4 at pH7.0 and the peptides were eluted with a salt gradient from 0 to 2M NaCl. The cysteine containing peptides were separated into two major fractions (Fig 19 in P.60), peptides 1,

2, 3 were in fraction 7c, 7d, 7e (peptide 1, 2 contain the cysteines), peptides 4, 5 were in fraction 6d (peptide 4 also contained cysteine). Peptide 4 was chosen to carry out the amino acid sequencing because it is shorter than peptide 1 and 2. Therefore peptide 4 needed to be further purified from fraction 6d.

4. Reverse phase high pressure liquid chromatography (HPLC)

After the fraction 6d was purified by ion-exchange column, it was dialysed at 4°C against 1 L of water with a 2 hour and a 3 hour changes, using Spectra/Por 3 dialysis tubing and lyophilized. Then it was loaded onto a (25x0.46 cm) 300-A C18 Vydac column at 1.0 ml/min flow rate in the presence of 0.1% trifluoroacetic acid (TFA) using an acetonitrile gradient for the purpose of further fractionation and also to remove the salt from previous ion-exchange column. The peptide from the eluted peaks was vacuum dried. Peptide 4 eluted at fraction 7b in this column (Fig 20 in P.61) seems very pure by the criterion of 15% PAGE.

5. Amino Acid Sequencing

Amino acid analysis were obtained using an Applied Biosystems model 420A derivatizer-analyzer system. The hydrolysis was carried out in gas-phase 6N HCl and 1% phenol under an argon atmosphere at 165°C, for 1 h, 2 h and 4 h. In order to quantify cysteine, the protein samples were pyridylethylated before hydrolysis as described above.

Cross-linking of *Spisula* PL-Iwith dithiobis (succinimidyl propionate) (DTBSP)

The nuclei were centrifuged at 4000xg in a JA-20 Beckman rotor for 10 min. The pellet was resuspended in 500µl of 20 mM EDTA, 100 mM Tris-HCl pH8.0, 0.2 mM PMSF to obtain a DNA concentration of about 0.3mg/ml (A_{260}=6). Next, 500µl of 575 mM DTT (dithiothreitol) was added to the mixture and incubated for 2 h at room temperature in order to reduce any of the native disulfide bridges which in the case of *Spisula* could arise from the cysteines present in PL-I. The mixture was then centrifuged at 4000xg for 10 min and the resulting pellet was resuspended in 40 mM Tris-HCl pH8.0, 100 mM sodium Iodoacetate and incubated for 1 h in the dark at room temperature in order to chelate the -SH groups of the cysteines in PL-I. After the incubation, the nuclear suspension was centrifuged at 4000xg for 10 min. The nuclear pellet was washed 2-3 times with cross-linking buffer (0.34 M sucrose, 5 mM triethanolamine, 3 mM $MgCl_2$, 0.2 mM PMSF pH7.5) in order to completely remove the iodoacetate. The resulting pellets were resuspended in 500µl of cross-linking buffer to an A_{260} of approximately 6. Cross-linker

dithiobis (succinimidyl propionate, 50 mg/ml in dimethylformamide, DMF) was then added to a final concentration of 0.2 mg/ml (referred here as to 1x). The amount of cross linking reagent was increased or decreased accordingly. The reaction was allowed to proceed for 10 min at room temperature, and was quenched to a final concentration of 50 mM glycine (50). The mixture was then centrifuged at 1500xg for 10 min. The pellet was dissolved in 50µl of 6M guanidium chloride, 50 mM Tris-HCl pH7.5. The nucleoprotein complexes were precipitated with 5 vols of cold ethanol at -20°C for 10 min. The ethanol precipitates were centrifuged at 14,000xg for 10 min at 4°C and the nuclear proteins were extracted with 100µl of 0.4 N HCl at 37°C for 5 min. These HCl extracts were precipitated with 5 vols of acetone at -20°C for 10 min. The acetone precipitate was collected by centrifugation at 14,000xg for 20 min, and the pellet was air dried.

This cross-linked proteins were first run on 2.5 M urea-acetic acid 5% (acrylamide/bisacrylamide=30:2) polyacrylamide gel (Fig 22A on P.64) with the dimension of 10 cm x 15 cm x 0.5 cm at 120V for about 1 hour. A strip from the first dimension gel was incubated in NBS (N-bromosuccinimide) solution which contained 50 mg/ml NBS, 8 M urea, 5% acetic acid and it was shaken for 15 min at room temperature. Incubation with NBS was carried out for the purpose of studying the organization of *Spisula* PL-I polymer that might arise from the PL-I-PL-I interactions. After incubation, the gel strip was placed on top of a 15% (acrylamide/bisacrylamide=30:0.2) polyacrylamide-2.5 M urea-acetic acid gel of the dimension 10 cm x 15 cm x 1 cm and run at 120 V for 75 min (Fig 22B on P.64).

Cross-linking of *Mytilus* with dithiobis (succinimidyl propionate) (DTBSP)

In order to understand the arrangement of PL-II, PL-III and PL-IV of *Mytilus* sperm, cross-linking study with dithiobis as cross-linker was done the same way as in the case of *Spisula* (see the above mentioned method). In the case of *Mytilus*, the cross-linked proteins were first run on a 16% (acrylamide/bisacrylamide=60:0.4) polyacrylamide-6M urea neutral gel at 120V for 2 h. A strip from the first dimension gel was incubated in O'Farrel solution [10% glycerol (w/v), 5% ß-mercaptoethanol, 2.3% SDS, 62.5mM Tris-HCl pH6.8 (51)] for 5 min to reduce the S-S linkages of the cross-linked proteins. After incubation, the gel strip was placed on top of a 15% (30:0.2) polyacrylamide-2.5 M urea-acetic acid gel and run in the second dimension (Fig 16 in P.52). The gels were then silver stained as described by Oakley (52).

Gel Electrophoresis

1. Urea-Acetic Acid polyacrylamide gel electrophoresis (UA-PAGE)

15% or 5% polyacrylamide (2.5 M urea) 5% acetic acid polyacrylamide gels consists of 15% (acrylamide/bisacrylamide=30:0.2) or 5% (acrylamide/bisacrylamide=30:2) polyacrylamide, 5% acetic acid, 2.5M urea, 0.77mg/ml thiourea 6.6% H_2O_2. The running buffer was 5% acetic acid. Sample buffer (2 fold) consisted of 8 M urea, 5% acetic acid and 0.1% pyronine-Y which was used as a tracking dye. The minislab gels 10 cm x8 cm x 0.77 cm were routinely electrophoresed at 120 V until the dye ran to 0.5 cm above the bottom of the gels.

2. Neutral PAGE

Sixteen percent polyacrylamide-6 M urea neutral gel were prepared by mixing 2.13 ml of acrylamide/bisacrylamide (60:0.4) solution, 4.8 ml of 0.3 M glycylglycine pH7.8, 10 M urea, 1.07 ml water, 50 µl TEMED and 8 µl 10% ammonium persulfate. The running buffer of the upper reservoir contained 30 mM imidazole-HCl pH7.2. The running buffer of the lower reservoir consisted of 30 mM glycylglycine, 10 mM imidazole pH7.2-7.4. The sample buffer (2x) consisted of 30 mM imidazole-HCl pH7.2, 20% sucrose and pyronine-Y. Minislab gels of this type were routinely run at 120V for 2 h.

3. SDS Gel Electrophoresis

SDS-polyacrylamide gel electrophoresis was carried out as described by Laemmli (53). The following stock solutions were used: solution A (30% acrylamide, 0.8% bisacrylamide), solution B (1.5M Tris-HCl pH8.8), solution C (10% SDS), solution D (0.5M Tris-HCl pH6.8). The 15% separating gel was prepared by mixing 7.5 ml A, 3.75 ml B, 0.15 ml C, 3.48 ml water, 6.75 µl TEMED and 84µl of 10% ammonium persulfate. The stacking was prepared by mixing 400µl A, 500µl of D, 1.06 ml water, 20µl of C, 2µl TEMED and 20µl of 10% ammonium persulfate. The sample buffer (2x) contained 0.0625 M Tris-HCl pH6.8, 2% SDS, 10% glycerol, 5% ß-mercaptoethanol and 0.001% bromophenol blue. The running buffer was 0.05 M Tris, 0.38 M glycine, 0.1% SDS. The running conditions was at 150V at 4%C for 40min for 10 cm x 8 cm x 0.77 cm minislab gels. The samples were boiled in hot water for 2-3 min before being loaded on to the gel.

All protein gels were fixed and stained in 25% isopropanol, 10% acetic acid, 0.2% Coomassie blue R250, and destained in 10% acetic acid, 10% isopropanol.

4. DNA gel electrophoresis

1% Agarose horizontal gels were prepared in 40 mM Tris-acetic acid pH8.0, 1.0 mM EDTA. Running buffer was 40 mM Tris-acetic acid pH8.0, 1.0 mM EDTA. 1/5 vol of

5x DNA sample buffer (0.25% bromophenol blue, 30% glycerol, 0.02% SDS) was added to one volume of DNA samples. 10 cm x 7 cm x 0.75 cm gel were ran at 100V for 1 h at room temperature. Agarose gels were stained with 0.5mg/ml ethidium bromide for 30 min (54). The DNA bands were visualized and photographed using an ultraviolet transilluminator. The sizes of the DNA fragments were estimated by comparison with the Lambda BstE11 fragments run as a reference.

Measurement of the length of the DNA fragment and determination of the nucleosome repeat:

The length of the fragments were determined by interpolation in a linear plot of electrophoretic mobility versus DNA size obtained using the *BstE11* restriction fragments of phase Lambda DNA. A plot of the migration distance versus the logarithm of the lengths of the restriction fragments gave a straight line. The sizes of the DNA fragments obtained by micrococcal nuclease digestion of chromatin were determined from the distances migrated by the midpoint of each fragment band relative to the standard line determined from the restriction fragments (55).

RESULTS

I. Characterization of the chromatin structure from *Mytilus californianus*

During spermatogenesis in *Mytilus* the somatic histone complement of the stem cells is partially replaced by protamine-like proteins PL-II, PL-III and PL-IV. Electron microscopy studies indicate that the chromatin in the mature sperm head of *Mytilus edulis,* which belongs to the same category as *Mytilus californianus*, has a coarse structure with granules of about 40-70 nm in diameter (57). X-ray diffraction indicates that no nucleosome-like structures are present in these spermatozoa, but rather, DNA is organized as parallel bundles of molecules (56). The relationship between these parallel bundles and the coarse granular structures observed in mature sperm remains unknown.

It is presently known that PL-II from *Mytilus californianus* is a highly sperm specific H1 histone, and that PL-IV is the C-terminal end of this histone H1 which has been post-translationally cleaved. In contrast, there is less information on the secondary and tertiary structure of PL-III. About 10-15% of histones coexist with these protamine-like proteins. The interaction of PL-II, III and IV proteins with DNA in the chromatin of the sperm of *Mytilus* is not clear. In order to determine how PL-II, III and IV are arranged onto the DNA in this chromatin, some of the basic nuclear PL proteins from sperm chromatin of *Mytilus californianus* were selectively extracted with $MgCl_2$. This allowed micrococcal nuclease to preferentially digest the salt extracted chromatin at the exposed sites in the DNA where protein had been removed by extraction.

1. 140 mM $MgCl_2$ extracts PL-II and PL-IV but not PL-III from *Mytilus* sperm chromatin

$MgCl_2$ concentration was titrated to determine the optimum condition under which the PL protein from *Mytilus* could be selectively extracted. The protein extractions were evaluated by urea/acetic acid gel electrophoresis (UA-PAGE). After extraction with $MgCl_2$, the nuclei were pelleted. Proteins that were selectively solubilized (supernatant) and those remaining in the pelleted nuclei were HCl extracted, acetone precipitated and analyzed. We found that at 140 mM $MgCl_2$, PL-II and PL-IV were completely solubilized in the supernatant (Fig 9 A) whereas PL-III was not extracted until the concentration of $MgCl_2$ was increased to 160 mM (Fig 9).

The proteins that were left in the nuclei pellet after 140 mM $MgCl_2$ extraction were

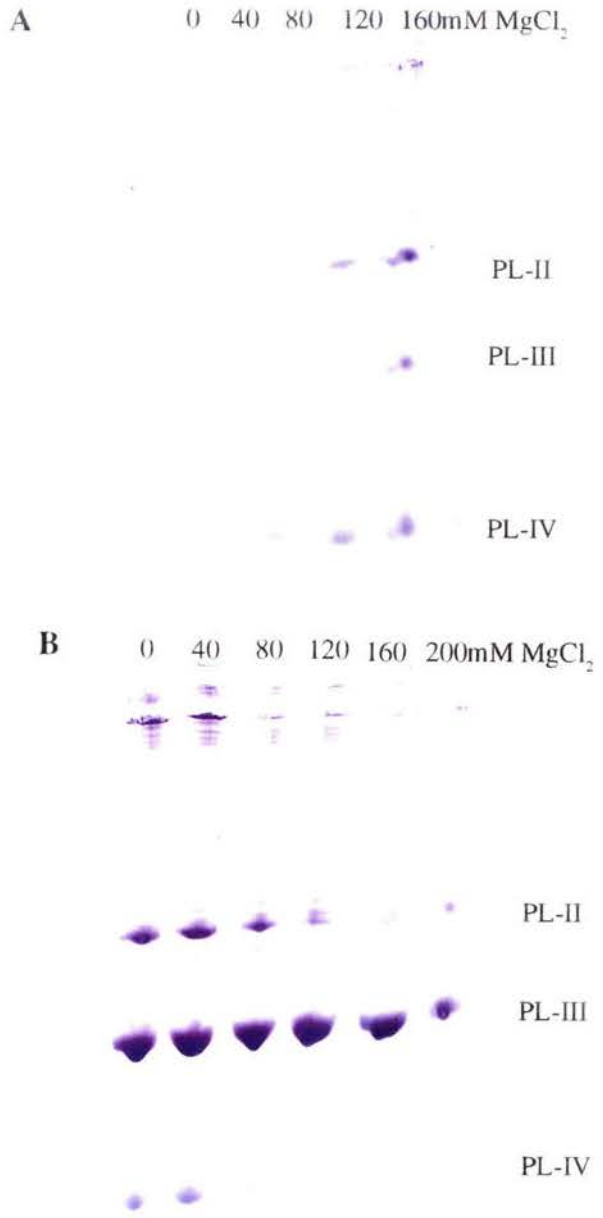


Figure 9. Urea-acetic acid PAGE of the $MgCl_2$ protein extracts from the sperm of *Mytilus californianus*. After extraction with $MgCl_2$, the nuclei were pelleted. The proteins selectively solubilized (supernatant) and those remaining in the nuclei (pellet) were HCl extracted, acetone precipitated and analyzed.
(A) Proteins which were selectively solubilized in the supernatant after $MgCl_2$ extraction.
(B) Correspondence proteins which were not $MgCl_2$ extracted under the same reaction condition.

analyzed by two dimensional PAGE (Fig 10). Most of the PL-II and PL-IV proteins had been extracted under these conditions (Fig 10B) as compared to the amount of PL-II and PL-IV present before $MgCl_2$ extraction (Fig 10A). However, little PL-III was $MgCl_2$-extracted. PL-III did not run in the second dimension because its highly basic charge confers insolubility in the SDS-PAGE system. In addition, two dimensional gel analysis showed that a complete set of histones were also left, running in a region of the gel close to the PL-II protein.

2. Micrococcal nuclease digestion of *Mytilus* sperm chromatin depleted of PL-II and PL-IV

PCA solubility

The sperm nuclei of the bivalve mollusk *Mytilus californianus* contain five histones of the somatic-type in addition to three sperm-specific proteins (PL-II, PL-III and PL-IV). In order to assess the role of these specific proteins in the structural organization of sperm chromatin, the accessibility of this chromatin to MNase digestion was measured. The kinetics of the nuclease digestion were determined by analyzing the percentage of nucleotides soluble in perchloric acid (PCA) over time (Fig 11A). The results shown indicate that MNase is able to digest chicken erythrocyte chromatin at a higher rate than *Mytilus* sperm chromatin. At 85 min, chicken chromatin shows a 30% oligonucleotide solubility in PCA, while that of native *Mytilus* sperm chromatin is only 18%. This indicates that *Mytilus* sperm chromatin is more resistant to nuclease digestion than chicken erythrocyte chromatin. However, when PL-II and PL-IV were $MgCl_2$-extracted from *Mytilus* sperm chromatin, nuclease digestion of the chromatin increased to the same extent as that of chicken chromatin. This suggests that, the PL-II and PL-IV depleted *Mytilus* sperm chromatin and chicken chromatin are digested in a similar fashion and the amount of DNA bound to PL-III in *Mytilus* sperm chromatin must be approximately similar to the amount of DNA protected by nucleosomes in chicken chromatin. 0.6M NaCl can also selectively extract PL-II and PL-IV from *Mytilus* sperm chromatin, but not to the same extent as $MgCl_2$. In this case, more DNA is protected by PL-II and PL-IV, as well as PL-III, in comparison with the $MgCl_2$ extracted sperm chromatin. Therefore, as expected, NaCl extracted sperm chromatin is more resistant to micrococcal nuclease digestion than $MgCl_2$-extracted chromatin, but less resistant than native *Mytilus* sperm chromatin.

After micrococcal nuclease digestion, chromatin was fractionated into three parts (SI, SII and PII) as described in Materials and Methods (Page 27). The protein compositions of the soluble fractions (SI and SII) and the insoluble fraction (PII) were then

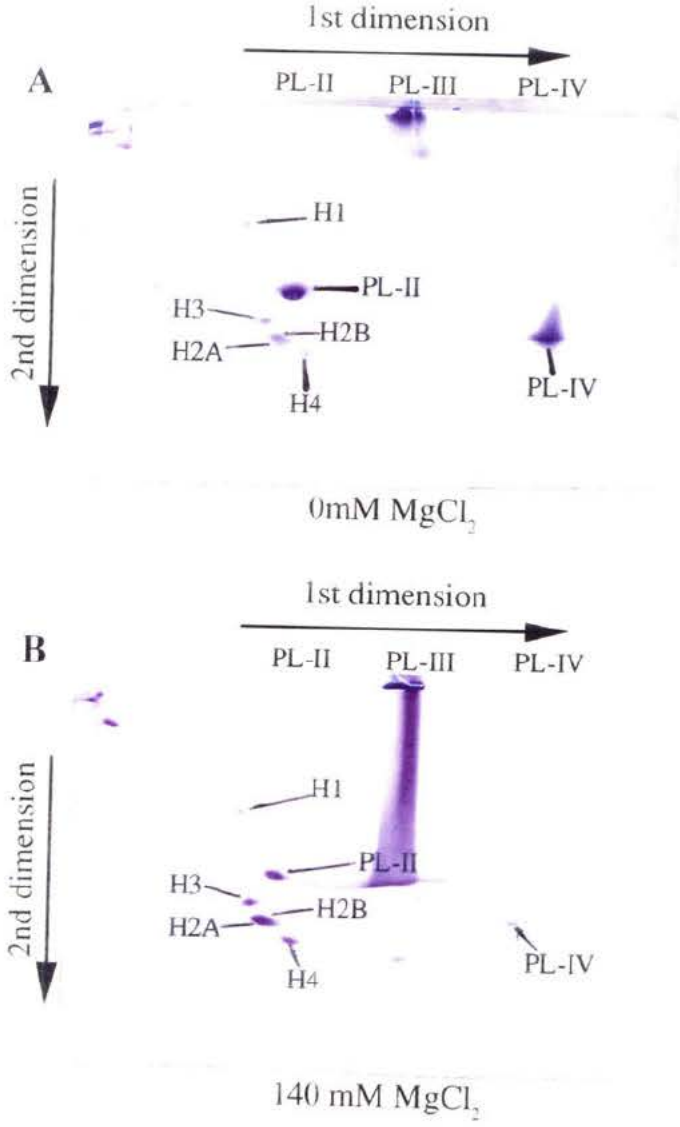
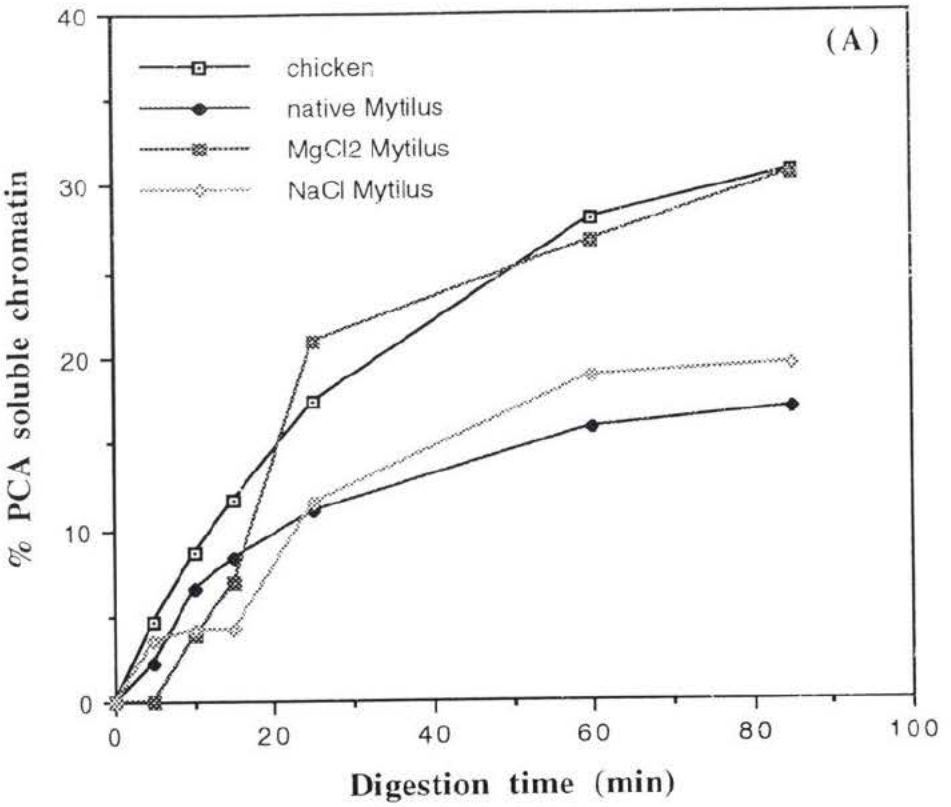
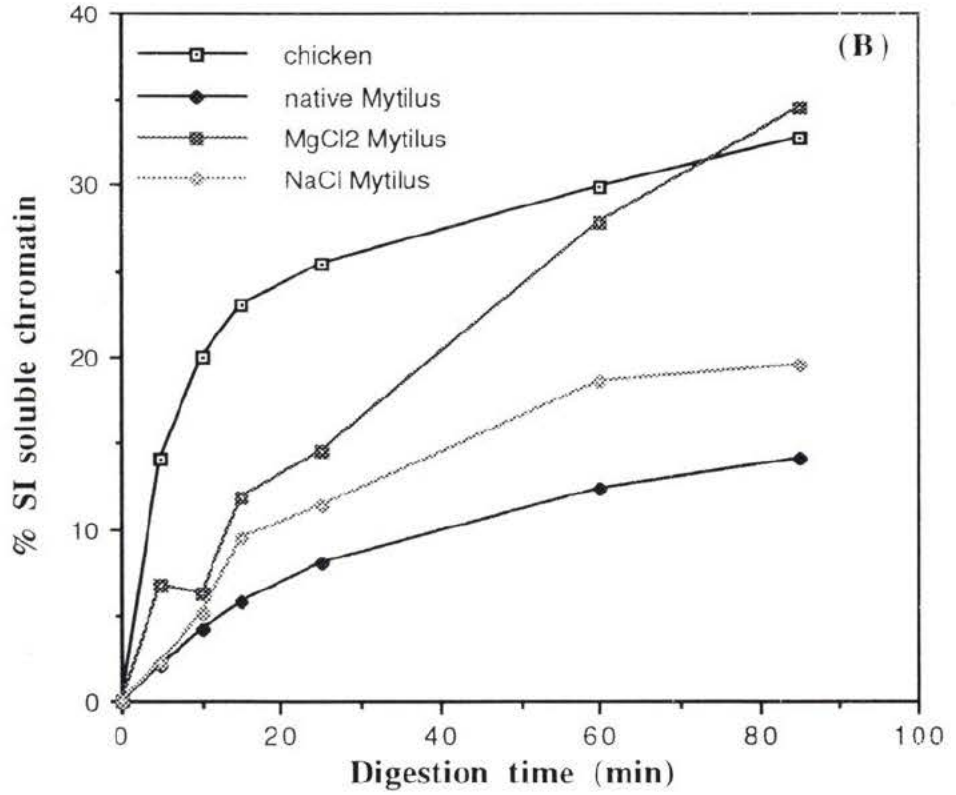


Figure 10. Two-dimensional gel electrophoresis of the proteins from the sperm nuclei of *Mytilus californianus* before $MgCl_2$ extraction (A), and after 140 mM $MgCl_2$ extraction (B). The first dimension was acetic acid-2.5M urea PAGE, and the second dimension was SDS-PAGE. The first dimension gel was incubated in O'Farrel solution for 1 h before the second dimension gel was run.


(A)
chicken
native Mytilus
MgCl2 Mytilus
NaCl Mytilus
% PCA soluble chromatin
Digestion time (min)
0
10
20
30
40
0
20
40
60
80
100


(B)
chicken
native Mytilus
MgCl2 Mytilus
NaCl Mytilus
% SI soluble chromatin
Digestion time (min)
0
10
20
30
40
0
20
40
60
80
100

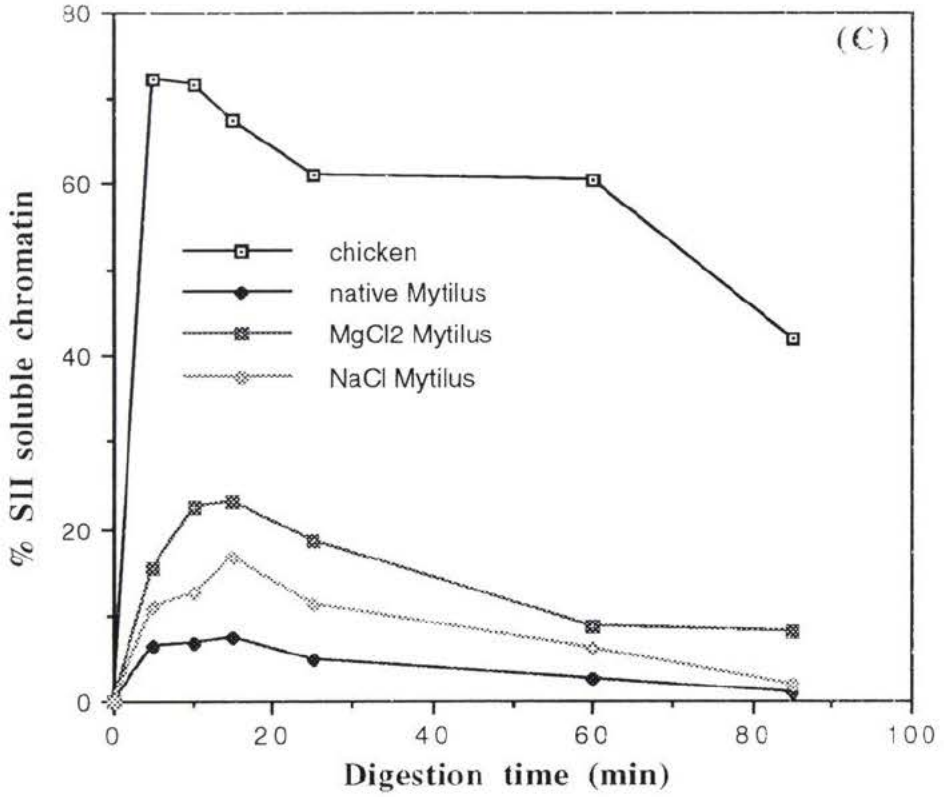


Figure 11. Time course of micrococcal nuclease digestion (50 units/ml) of chicken erythrocyte and *Mytilus* sperm nuclei (at $40A_{260}$) at 37°C.
—□—: buffer A (120 mM NaCl, 20 mM Tris-HCl pH7.5, 1 mM $CaCl_2$, 0.1 mM PMSF), 37°C in chicken erythrocyte nuclei;
—◆—: buffer A, 37°C native *Mytilus* sperm nuclei;
·····◇·····: after extraction of *Mytilus* sperm nuclei with 0.6 M NaCl (in 20 mM Tris-HCl pH7.5, 1 mM $CaCl_2$);
—⊠—: after extraction of *Mytilus* sperm nuclei with 140 mM $MgCl_2$, 1 mM $CaCl_2$ (in buffer A).
A: PCA soluble DNA.
B: Supernatant I (SI) soluble chromatin.
C: Supernatant II (SII) soluble chromatin.
SI is the first supernatant which is solubilized immediately after digestion. Supernatant II (SII) is obtained after lysis against 0.25 mM EDTA pH7.5 of the remaining pellet. PCA-soluble fraction contains the nucleotides released in the digestion (see materials and methods).

analyzed by UA-PAGE (Fig 13). In addition to this, DNA fragments from SI, SII and PII were visualized in 1% agarose gels (Fig 12). Comparison of the DNA and corresponding protein gels of each fraction, when taken together with their respective solubility, may aid to elucidate the chromatin structure and organization in the *Mytilus* sperm nuclei.

Comparison of chromatin solubility, DNA and protein structure in fraction SI between chicken erythrocyte nuclei and *Mytilus* sperm nuclei

In chicken erythrocyte chromatin, fraction SI contains the free oligonucleotides which are also present in the PCA counterpart and it also contains chromatin particles of nucleosome-like sized DNA (100-145bp) and a few oligonucleosomes that are spontaneously released by the digestion with the nuclease and subsequently leak out of the nuclear envelope. In the chicken erythrocyte, it was found that SI solubility increased as a function of time (Fig 11B), but compared to its PCA fraction, there was no noticeable increase in solubility. Chicken erythrocyte chromatin exhibits approximately 30% solubility in the PCA fraction and 30% in fraction SI at 85 min of digestion. This apparent discrepancy in the lack of incremental increase in solubility is attributable to the fact that the amount of DNA released in the form of nucleosome particles is very small. The solubility of *Mytilus* sperm chromatin in the fraction SI (Fig 11B) when digested for 85 minutes is smaller than the solubility of chicken erythrocyte chromatin (30% in SI) digested under the same condition. This can be explained by the highly condensed nature of chromatin in *Mytilus* sperm. It becomes increasingly more difficult for micrococcal nuclease to digest DNA in the *Mytilus* sperm chromatin. After 85 minutes of digestion, the percentage nucleotide solubility of chromatin from $MgCl_2$-extracted *Mytilus* sperm and from chicken erythrocyte in fraction SI are very similar (Fig 11B), suggesting that after PL-II and PL-IV are preferentially extracted, the DNA region protected by PL-III in *Mytilus* sperm chromatin may be similar to that of the nucleosome domain in chicken erythrocyte chromatin. Conversely, NaCl extracted *Mytilus* sperm has a solubility that falls between native *Mytilus* and $MgCl_2$ extracted sperm chromatin because PL-II and PL-IV are not completely removed by NaCl (Fig 14).

Chicken erythrocyte chromatin was used as the control to compare the digestion pattern obtained with native and salt-extracted *Mytilus* sperm chromatin. In fraction SI of chicken erythrocyte chromatin, only mononucleosomes or small oligonucleosomes depleted of H1, with DNA fragment sizes about 145 base pairs or multiples, could be observed (Fig 12A). Linker histones H1/H5 are not present in the nucleosomes of this soluble fraction (Fig 13A) because nucleosomes with linker histones are insoluble in SI. Linker histones

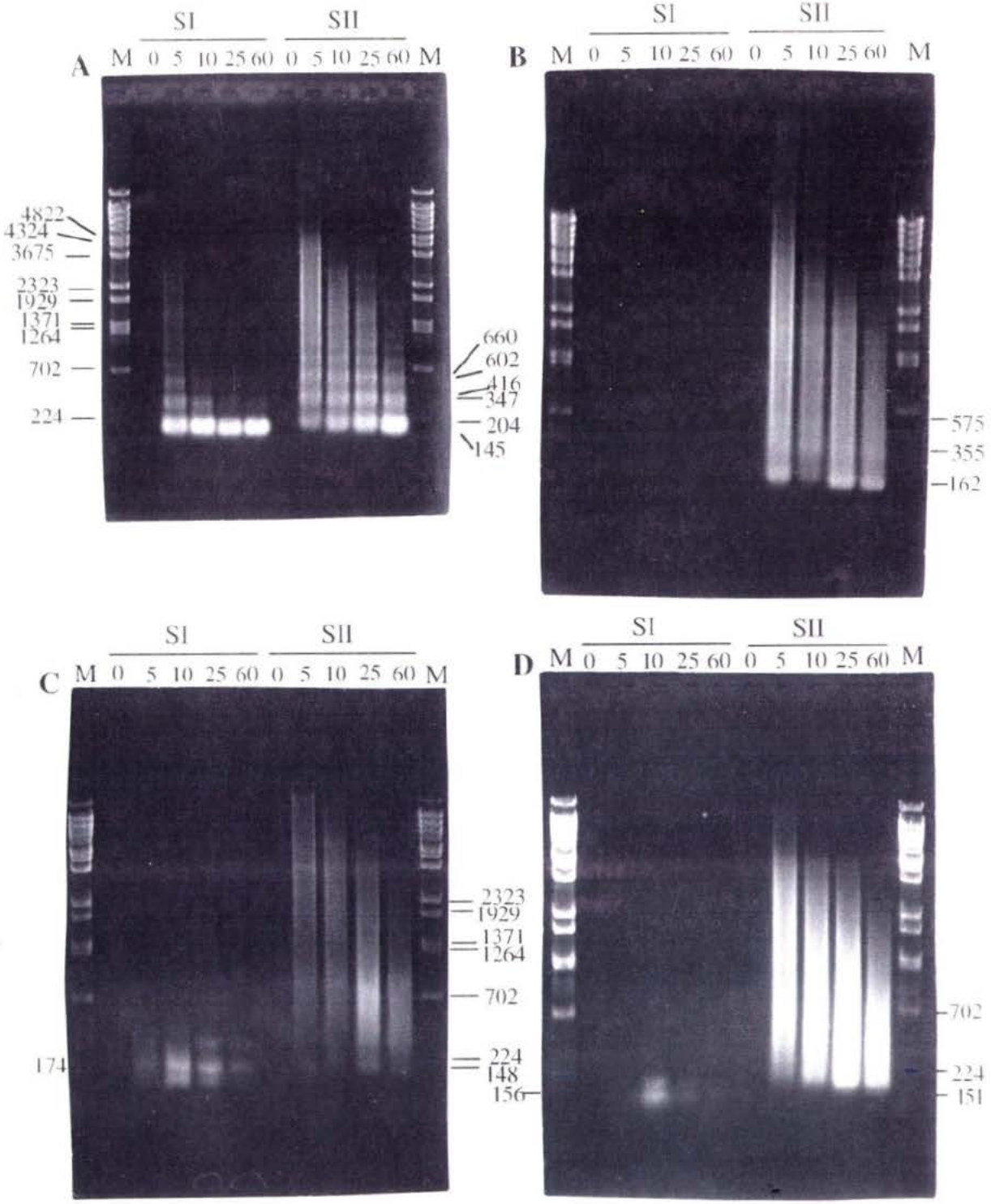


Figure 12. A, B, C, D (description see next page).

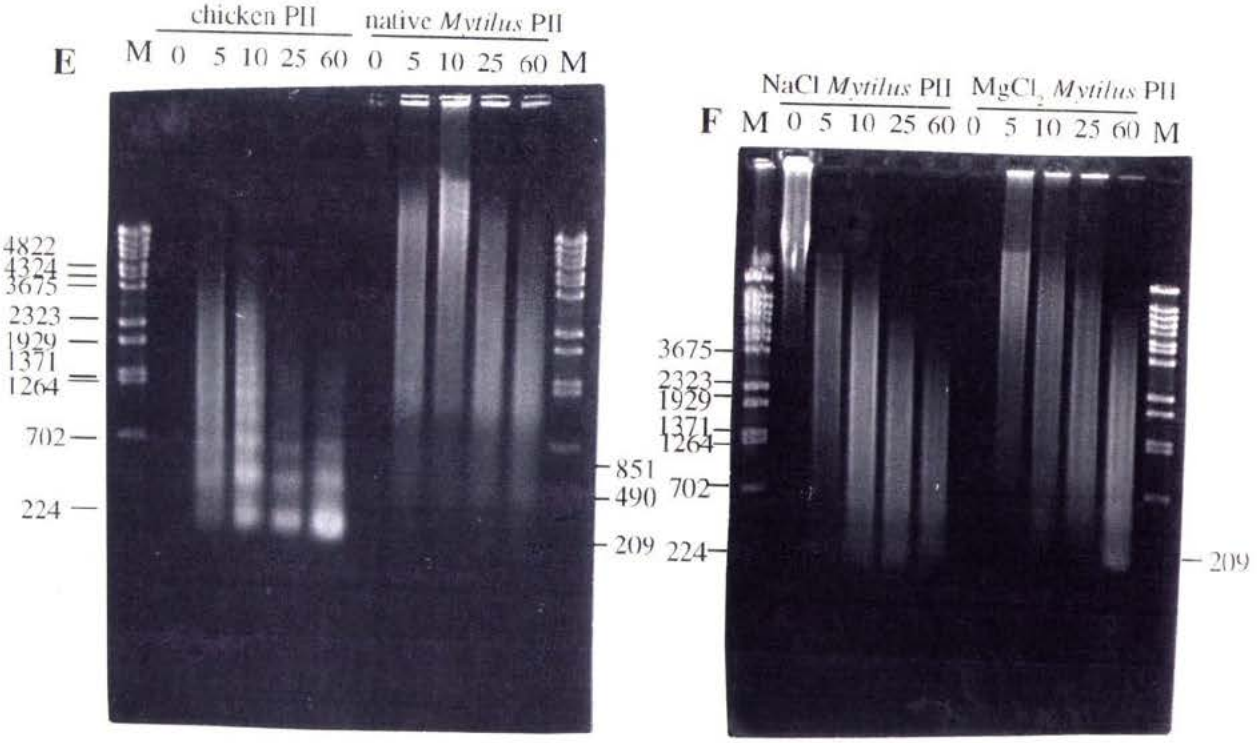


Figure 12. Agarose (1%) slab gel electrophoretic analysis of DNA fragments from the SI, SII and PII fractions after micrococcal nuclease digestion.
A: Chicken erythrocyte chromatin SI, SII (control);
B: *Mytilus* native sperm chromatin (in buffer A) SI, SII;
C: *Mytilus* sperm chromatin extracted with 140 mM $MgCl_2$ SI, SII;
D: *Mytilus* sperm chromatin extracted with 0.6 N NaCl SI, SII;
E: Chicken erythrocyte chromatin PII and *Mytilus* native sperm chromatin PII;
F: 0.6 N NaCl extracted *Mytilus* sperm chromatin PII and 140 mM $MgCl_2$ extracted *Mytilus* sperm chromatin PII.
Lanes 0, 5, 10, 25, 60 min are different digestion times of DNA aliquotes. M is lambda BstE11 fragments.

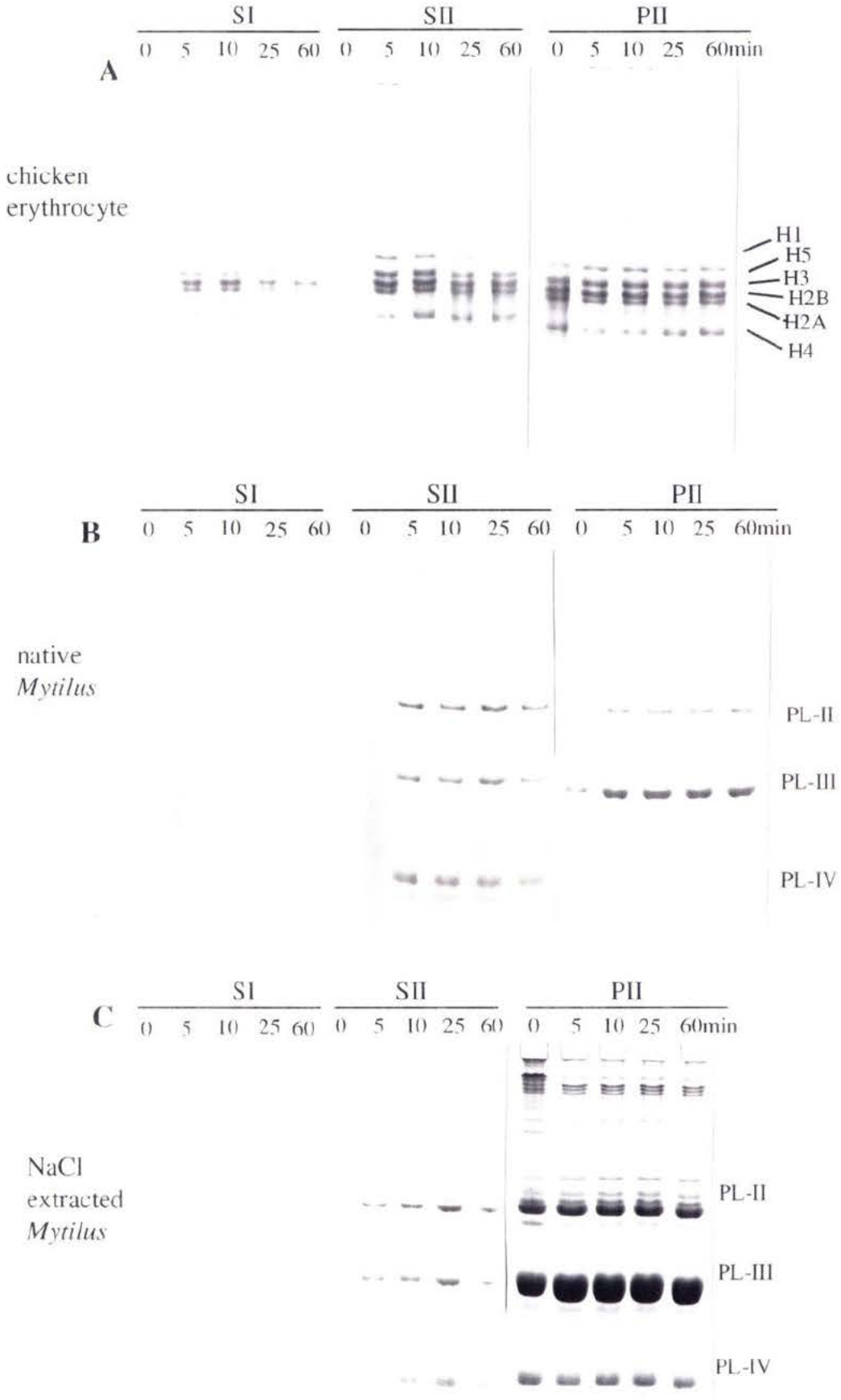

A
SI
SII
PII
0 5 10 25 60 0 5 10 25 60 0 5 10 25 60min
chicken
erythrocyte
H1
H5
H3
H2B
H2A
H4
B
SI
SII
PII
0 5 10 25 60 0 5 10 25 60 0 5 10 25 60min
native
Mytilus
PL-II
PL-III
PL-IV
C
SI
SII
PII
0 5 10 25 60 0 5 10 25 60 0 5 10 25 60min
NaCl
extracted
Mytilus
PL-II
PL-III
PL-IV

Figure 13. 15% urea acetic acid PAGE of nuclear proteins extracted with 0.4 N HCl from the SI, SII and PII fractions after digestion with micrococcal nuclease at the indicated times (in minutes).
A: Chicken erythrocyte chromatin proteins from SI, SII and PII fractions;
B: *Mytilus* native sperm chromatin proteins from SI, SII and PII fractions;
C: 0.6 N NaCl extracted *Mytilus* sperm chromatin proteins from SI, SII and PII fractions.

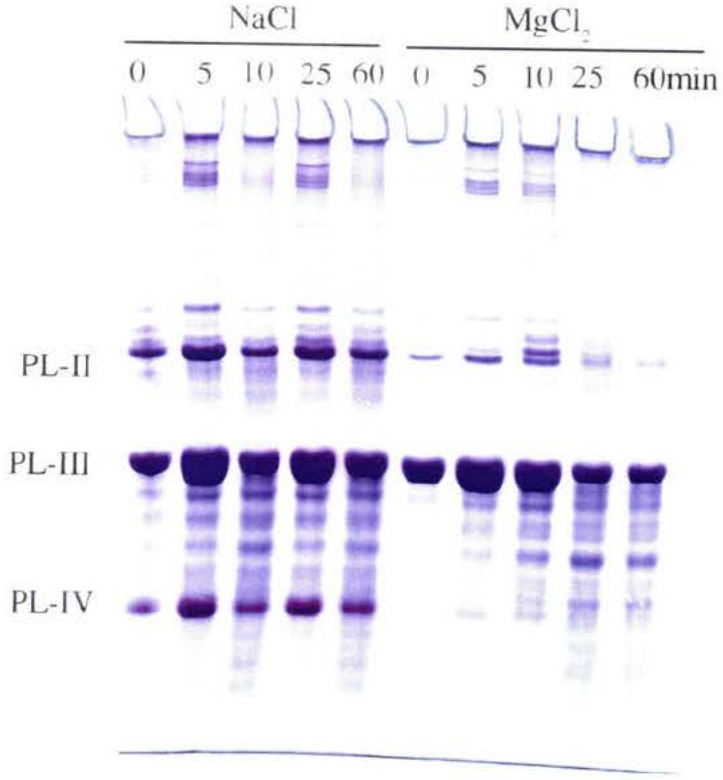


Figure 14. 15% urea acetic acid PAGE of nuclear proteins extracted with 0.4 N HCl from the PII fractions after salt treatment. Left: fraction PII from 0.6N NaCl extracted *Mytilus* sperm chromatin; right: fraction PII from 140mM $MgCl_2$ extracted *Mytilus* sperm chromatin.

neutralize the negative charge of the chromatin DNA which causes the chromatin to precipitate. All the core histones, histone H3, H4, H2A and H2B are, however, present in the protein gel (Fig 13A). When PL-II and PL-IV of *Mytilus* sperm chromatin are depleted with NaCl or $MgCl_2$ and the chromatin digested with micrococcal nuclease, there are some particles released in fraction SI early during the digestion. These particles exhibit DNA fragments of about 156bp and 174bp in the NaCl and $MgCl_2$ SI fractions respectively. These fragment lengths correspond to the nucleosomal DNA size (145bp) from fraction SI of chicken erythrocyte chromatin (Fig 12A), and suggests that nucleosomal arrangement of DNA contributes to the organization of *Mytilus* sperm chromatin.

Because the histone percentage is low in *Mytilus* sperm chromatin, accounting for less than 20% of the total nuclear sperm-specific proteins, it is expected that only neglible amounts of histones would be detected in the protein gels of the SI fractions from native and NaCl/$MgCl_2$ treated *Mytilus* chromatin (Fig 13 B and C).

Comparison of chromatin solubility, DNA and protein structure in fraction SII between chicken erythrocyte nuclei and *Mytilus* sperm nuclei

Fraction SII from chicken erythrocyte chromatin, which is obtained after hypotonic lysis against 0.25 mM EDTA, contains the larger chromatin fragments that did not 'leak' out of the nuclei into SI. In this fraction (Fig 11C), the chromatin solubility increases sharply at the beginning of the digestion process and then decreases over time. This is because at the beginning of digestion, only a small amount of nucleosomes are digested by the nuclease and leak out of the nuclei (SI), leaving a large amount of chromatin fragments inside the nuclei; as digestion time increases, micrococcal nuclease cleaves more chromatin fragments into oligonucleosomes therefore reducing the bulk chromatin fragments (SII). Chromatin solubility in fraction SII of native *Mytilus* sperm is very low compared to that of chicken erythrocyte chromatin. This is because all of the negatively charged phosphates in the DNA in native *Mytilus* sperm chromatin are neutralized by protamine-like proteins which make this chromatin insoluble and it cannot be detected in fraction SII. In chicken erythrocyte chromatin, however, only one third of the negatively charged phosphates are neutralized by histones, therefore, the chromatin is more soluble and can be detected in fraction SII. $MgCl_2$ extracted *Mytilus* sperm chromatin in fraction SII has a lower solubility than chicken erythrocyte chromatin for a similar reason even though it has the same solubility as chicken erythrocyte chromatin in PCA and the SI fraction. After PL-II and PL-IV have been removed by $MgCl_2$, there is more DNA available for nuclease digestion. This is why the $MgCl_2$-extracted *Mytilus* sperm chromatin has a slightly higher nucleotide

solubility than that of native *Mytilus*.

In fraction SII from chicken erythrocyte chromatin, a full complement of histones are present throughout the entire digestion (Fig 13A). The nucleosomal DNA units observed at 60 min digestion (145bp, 347bp and 602bp, Fig 12A) are smaller than those present at 5 min digestion (204bp, 416bp and 660bp), because linker histones H1/H5 are still present on the nucleosome during the early stages of digestion (Fig 13A). As the digestion progresses, nucleases start to digest linker DNA and the sizes of the nucleosome particles released decrease. In fraction SII of *Mytilus*, three DNA fragment repeats coexist with what appears to be a DNA smear. The sizes of these fragments (162bp, 355bp, 575bp) are comparable to the nucleosomal DNA repeats in fraction SII of chicken (145bp, 347bp, 602bp). The similarity of the repeat pattern in *Mytilus* sperm chromatin and chicken erythrocyte chromatin suggests that the DNA regions protected by nucleosomes in *Mytilus* must have a similar organization as that of chicken erythrocyte chromatin. After extraction of *Mytilus* with NaCl (Fig 12D) or $MgCl_2$ (Fig 12C), the lower part of the DNA smear stops at 151 bp and 148 bp (60 min), respectively. Again these DNA sizes are very similar to the chicken erythrocyte nucleosomal DNA size (145bp).

The DNA smear which is present in *Mytilus* sperm chromatin (Fig 12C and 12D) but not in chicken erythrocyte chromatin (Fig 12A) most likely correspond to a random distribution of the nucleosomes in the sperm chromatin (Fig 15). This random distribution is most likely due to the presence of Pl proteins (Fig 15B) interspersed between nucleosomes.

At the protein level, the stoichiometry of PL-II, PL-III and PL-IV in the SII fraction is almost the same (Fig 13 B, C). This is surprising considering that in native chromatin PL-III is present in a larger amount.

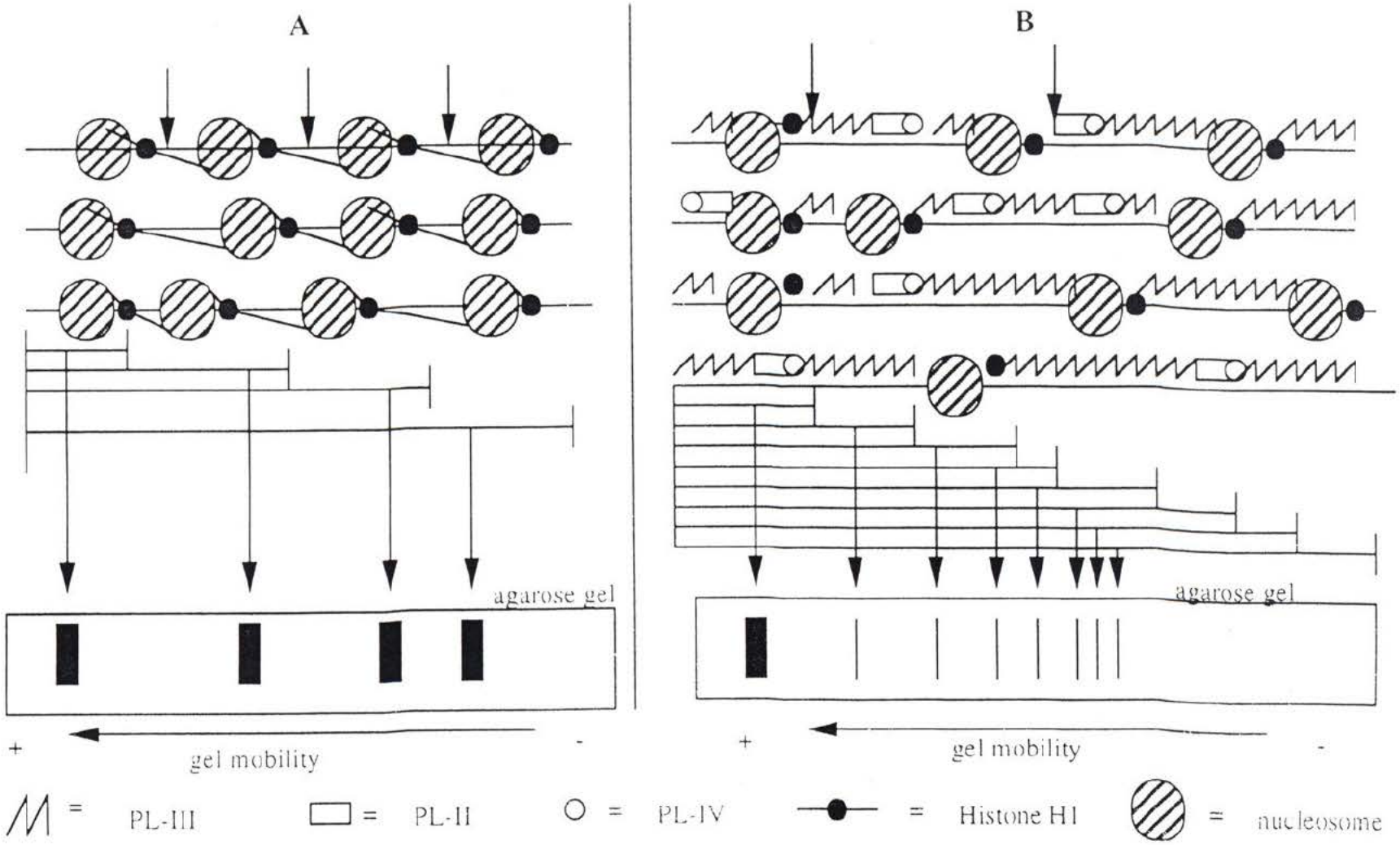


Figure 15. Micrococcal nuclease digestion pattern in A) chicken erythrocyte chromatin and B) *Mytilus/Spisula* sperm chromatin.

Comparison of DNA and protein from fraction PII of chicken erythrocyte nuclei and *Mytilus* sperm nuclei

Fraction PII contains the insoluble, large fragments of chromatin. In *Mytilus*, this DNA fraction again has a smeared appearance (Fig 12 E and F). In fraction PII of $MgCl_2$-extracted *Mytilus* sperm chromatin (Fig 12F), a DNA band of 209bp is additionally observed which has the same size as the nucleosome repeat in fraction PII of chicken erythrocyte chromatin (Fig 12E). The slightly larger size of this band in PII fraction, compared to the nucleosomal unit length in SII (148bp), could be explained by the residual linker histones that remain bound to the nucleosomes in the insoluble chromatin and protect them from further cleavage.

When comparing the ratio of PL proteins in fraction PII of NaCl and $MgCl_2$-extracted *Mytilus* sperm chromatin (Fig 14), it is apparent that there is less PL-II and PL-IV in $MgCl_2$-extracted chromatin than in NaCl-extracted *Mytilus* chromatin.

II. Cross-linking of proteins in the sperm chromatin from *Mytilus californianus*

A cross-linking reagent (DTBSP) was used to determine the spatial organization of PL-II, PL-III, PL-IV and histones in chromatin, as described in Materials and Methods. After cross-linking, the proteins were extracted with 0.4N Hydrochloric acid (HCl). A number of low mobility bands, corresponding to polymerized proteins, were observed (Fig 16A). In order to obtain an interpretable amount of cross-linked product, a series of reactions were performed in which the cross-linker ratio was varied (Fig 16A). As the proportion of cross-linker increased, the amount of mono-PL-II, PL-III and PL-IV decreased and the oligomers increased. Upon reduction of the cross-linked products and electrophoresis in a second dimension, their monomeric components are recovered (Fig 16B) thus allowing to identify the nature of the PL polymers.

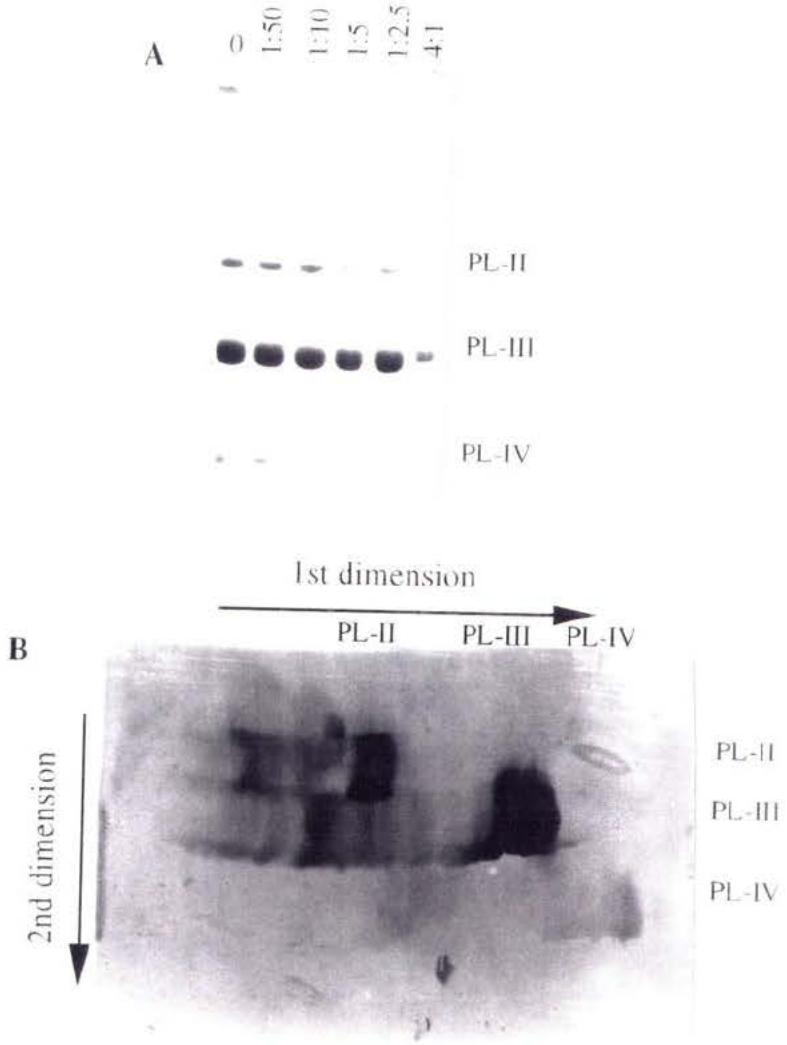


Figure 16. Cross-linking analysis of *Mytilus* PL proteins. (A) 15% urea-acetic acid polyacrylamide gel electrophoresis of crosslinked nuclear proteins from the sperm of *Mytilus californianus*. (B) Silver stained two-dimensional electrophoresis of basic nuclear proteins from sperm of *Mytilus* californianus cross linked with dithiobis (succinimidyl propionate; protein:diothiobis=3:2, protein concentration 0.3mg/ml). The first dimension: neutral 6M urea 16% polyacrylamide gel, and the second dimension: 2.5M urea/acetic acid 15% polyacrylamide gel.

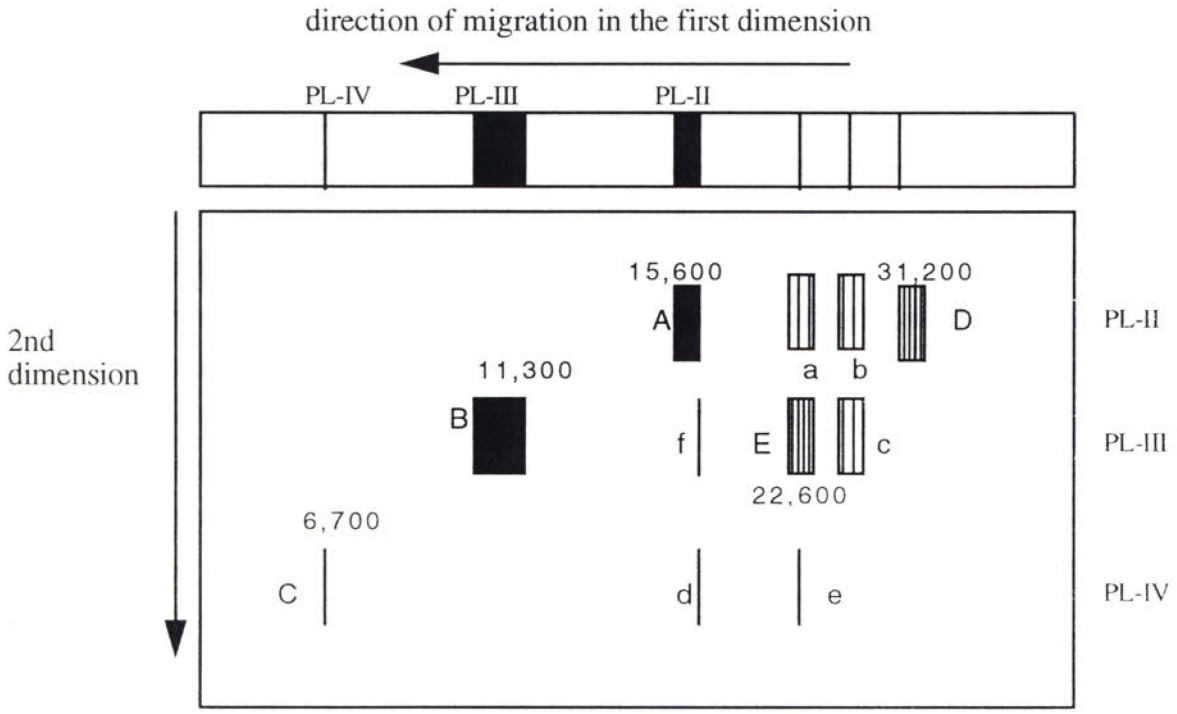


Figure 17. Schematic drawing of crosslinked PL proteins from *Mytilus californianus* after separation by 2D PAGE.

From the schematic drawing (Fig 17) of the second dimension gel (Fig 16B), we can see PL-II monomer (band A; MW=15,600 Da), PL-III monomer (Band B; MW=11,300 Da), and PL-IV monomer (band C; MW=6,700 Da). The PL-III monomer is twice as intense as PL-II monomer, while the PL-IV monomer is faint. Band D is the PL-II dimer with a molecular weight of 31,200 Da, and band E is the PL-III dimer with molecular weight of 22,600 Da.

In order to interpret bands "a", "b", "c", "d", "e" and "f", deduction by elimination is applied. The possibilities for band "a" are a PL-II dimer, a PL-II/PL-III dimer or a PL-II/PL-IV dimer. Firstly, band "a" could not be a PL-II dimer because this structure has already been identified (band D). It could also not be derived from a PL-II/PL-III dimer, because the theoretical molecular weight of PL-II/PL-III is 26,900 Da. This would appear between band D (31,200 Da) and band E (22,600 Da), however, it is instead close to band E. Therefore, band "a" is most likely derived from a PL-II/PL-IV dimer, whose MW is 22,300 Da, and band "e" represents PL-IV. However, based on molecular weights, band "a" and "e" should run a little faster than band E if they are PL-II/PL-IV dimer because its MW (22,300 Da) is a little smaller than the PL-III dimer (22,600 Da). This can be explained by the fact that protein mobility in the UA-PAGE system, depends not only on molecular size, but also on charges. It can be seen that the PL-III dimer is more charged (100 positive charges, Table III) than the PL-II/PL-IV dimer (82 positive charges). Therefore, the PL-III dimer could run at the same place as the PL-II/PL-IV dimer even though its molecular weight is slightly larger.

Band "b" could not be a PL-II dimer. It could also not come from a PL-II/PL-IV dimer because no PL-IV can be detected on the PL-IV lane at the same position as band "b". Therefore, band "b" most likely comes from PL-II/PL-III dimer. In the second dimension, at the same mobility of band "b", band "c" was found in the PL-III lane which supports the belief that a PL-II/PL-III dimer has been formed. The MW of a PL-II/PL-III dimer (26,900 Da) is also consistent with this complex's gel mobility, which falls between the PL-II dimer (band D) and PL-III dimer (band E).

Band "f" is not a PL-III dimer because the gel mobility of the crosslinked complex in the first dimension is very close to a PL-II monomer (band A). It is not a PL-II/PL-III dimer (26,900 Da) for the same reason. It most likely comes from a PL-III/PL-IV dimer based on two facts: 1) there is PL-IV (band d) at the same gel mobility as PL-III (band f); and 2) the MW of PL-III/PL-IV dimer is 18,000 Da. The reason that it runs almost at the same position as PL-II monomer (band A) (15,600 Da) is that PL-III/PL-IV dimer has 105 (50+55) positive charges, much larger than PL-II monomer (27 charges). Therefore, even

though its MW is larger, it still runs faster than the smaller protein with fewer charges.

The results from the above analysis are organized in table III:

Table III: Comparison of PL complex from the two dimensional PAGE (Fig 17).

Band Name	Protein complex name	MW (Da)	charges	Band intensity
A	PL-II monomer	15,600	27	++++
B	PL-III monomer	11,300	50	++++++++
C	PL-IV monomer	6,700	55	+
D	PL-II dimer	31,200	54	+++
E	PL-III dimer	22,600	100	++++
a-e	PL-II/PL-IV dimer	22,300	82	++
b-c	PL-II/PL-III dimer	26,900	77	++
f-d	PL-III/PL-IV dimer	18,000	105	++

No PL-II/PL-III/PL-IV complex nor PL-IV dimer can be identified.

To determine the spacial arrangement of these PL proteins on chromatin, their cross-linking tendencies were measured by comparing the intensity of the various dimers (see Table III). Although the PL-III dimer (band E) is slightly more intense than the PL-II dimer (band D), the PL-III monomer (band B) exhibits twice the intensity of the PL-II monomer (band A). This suggests that the PL-II molecules cross-link with each other better than the PL-III molecules. All of the heterogeneous dimers have similar intensity, but are less intense than the PL-III dimer (band E) suggesting that the cross-linking tendency of all the heterogeneous dimers (PL-II/PL-III, PL-II/PL-IV, PL-III/PL-IV) is similar, but less, than PL-III dimers. This line of reasoning can be expressed as follows:

2PL-II ≥ 2PL-III ≥ PL-II/PL-IV, PL-II/PL-III, PL-III/PL-IV

III. Cysteine characterization for the PL-I protein from *Spisula solidissima* sperm chromatin

1. Two cysteines are present in PL-I protein

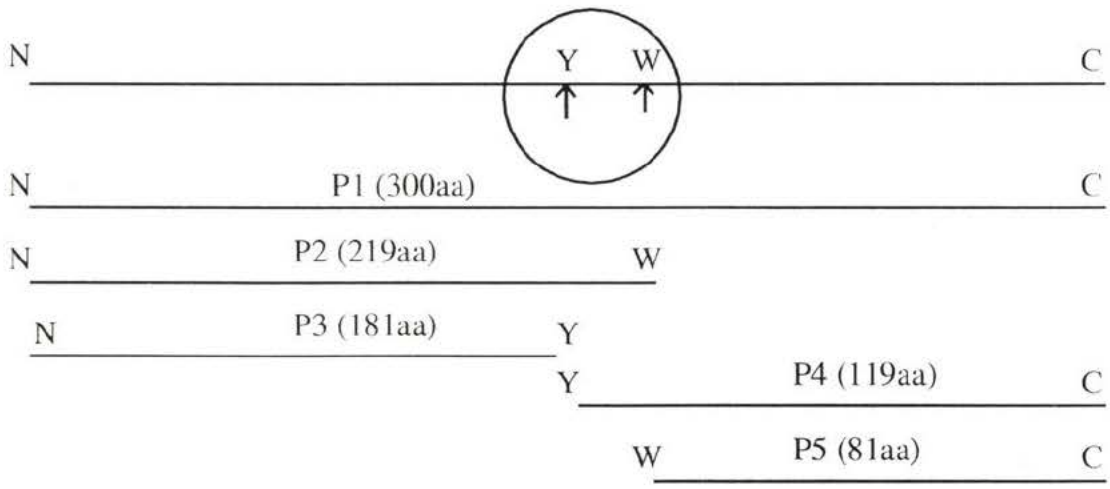
The sperm of *Spisula solidissima* is unique in that it only contains one protamine-like component (PL-I) which as in *Mytilus*, coexists with the complete set of somatic histones (31). This PL-I protein is a member of the histone H1 family. Ausio et al. (35) found that this PL-I protein consists of a trypsin resistant globular core of about 75 amino acids flanked by an N-terminal tail of approximately 150 amino acids and a C-terminal tail

of about 75 amino acids. Previous studies on the amino acid composition of PL-I did not detect the presence of cysteine within this protein (19, 7). There are difficulties identifying cysteine residues by conventional methods. In this work, we have conclusively demonstrated the presence of cysteine in the *Spisula* PL-I protein (Fig 18A). When purified PL-I protein (Fig 18A, lane a) was treated with ß-mercaptoethanol under denaturing conditions (8M urea), the reduction of disulphide bridges to free sulphydryl groups resulted in the production of PL-I monomers (Fig 18A, lane b) from PL-I oliogomers (Fig 18A, lane a). The existence of PL-I oligomers larger than dimers indicates that there are at least 2 cysteines in every PL-I protein monomer.

2. Determination of the location of cysteine in the protein PL-I from *Spisula*

In an attempt to locate the cysteine residues within the different structural domains of the PL-I protein, we used a fluorescent tag (bromobimane, qBBr) to label the thiol groups of the cysteines. PL-I protein was first reduced by ß-mercaptoethanol so that the thiol group could be labelled. After labeling, the protein was treated with NBS to cleave it into well-defined fragments. The peptides generated in this way were then analyzed by urea-acetic acid PAGE, on which the fluorescently labelled peptides were visualized and photographed under transmitted ultraviolet light.

As a control, nuclear proteins from the sperm of *Spisula,* which contain both PL-I and histones, were used. In addition to PL-I, cysteine-containing histones H1 and H3 fluoresced (Fig 18 C1). The fact that a labelled PL-I protein band shows up under the transmitted UV light indicates that, like histones H1 and H3, PL-I from *Spisula* also contains cysteine residues (Fig 18 C2). After NBS cleavage, which cleaves at tryptophan and tyrosine residues [NBS/protein ratio of 12.5 (w/w) under 8M urea and 5% acetic acid], PL-I protein was cleaved into five peptides (Fig 18 B3). Of the five NBS-generated peptides, only bands 1, 2, and 4 are labelled with the dye (Fig 18 C3). The labelled bands of the PL-I protein were not as intense as the PL-I control and histone labelled bands (Fig 18 C1) because NBS partially quenched the fluorescence from qBBr. From these experiments, we can conclude that these three cleaved peptide fragments all contain cysteine. The identification of these 3 peptide fragments can be postulated from the sequence data available for the PL-I protein (Ausio, unpublished).



If the 2 cysteines were located between tryptophan (W) and tyrosine (Y), as indicated by the arrows in the globular region then 3 fluorescent peptides would be obtained. This is the only region that all of these three peptides overlap.

3. Purification of the cysteine-containing peptides obtained from NBS cleavage of PL-I

The partial sequence of the globular domain of PL-I has already been published (31). However, cysteine was undetected in that sequence. Therefore, an attempt was made to establish the correct amino acid sequence between the tryptophan and tyrosine residues in the globular core region of PL-I. The PL-I protein was pyridylethylated so that the cysteine residues could be detected by conventional amino acid sequencing. After pyridylethylation, PL-I was cleaved by NBS and the peptides subsequently purified by HPLC to obtain the cysteine-containing peptides. These were then sequenced by conventional amino acid sequencing techniques.

Figure 19 shows that the separation of NBS peptides on an ionic exchange column. Fraction 6d contained the cysteine-containing peptide 4 which was selected for amino acid sequencing. Figure 20 shows the further purification of peptide 4 from fraction 6d by reverse phase HPLC (fractions 7b and 7c).

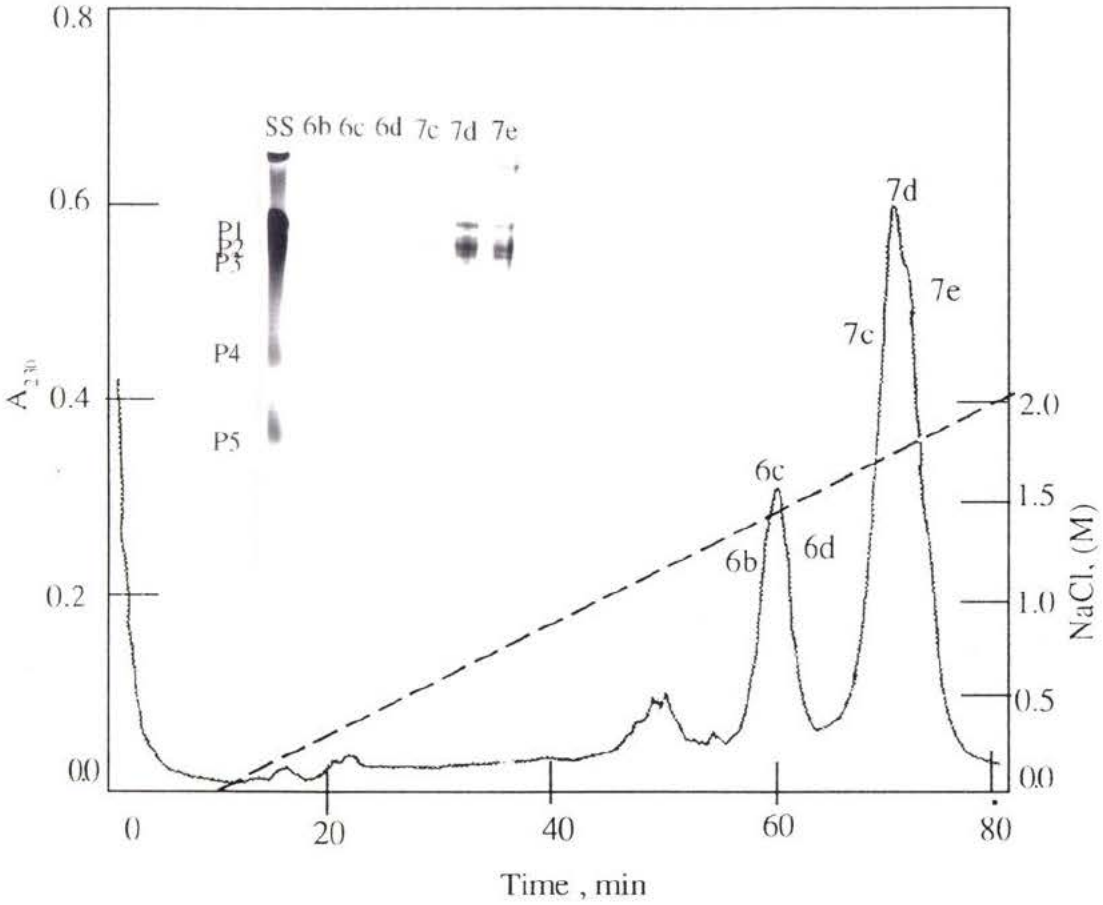


Figure 19. Ionic exchange HPLC on CM300 Sephadex fractionation of pyridylethylated PL-I protein cleaved with NBS. The elution buffer was 20 mM Na_2PO_4 at pH7.0. The peptides were eluted with a salt gradient from 0 to 2M NaCl. Flow rate was 0.5ml/min, the detection wavelength was 230nm. The inset shows the electrophoretic analysis (in 15% polyacrylamide 2.5M urea-5% acetic acid PAGE) of the different fractions along the chromatographic profile.

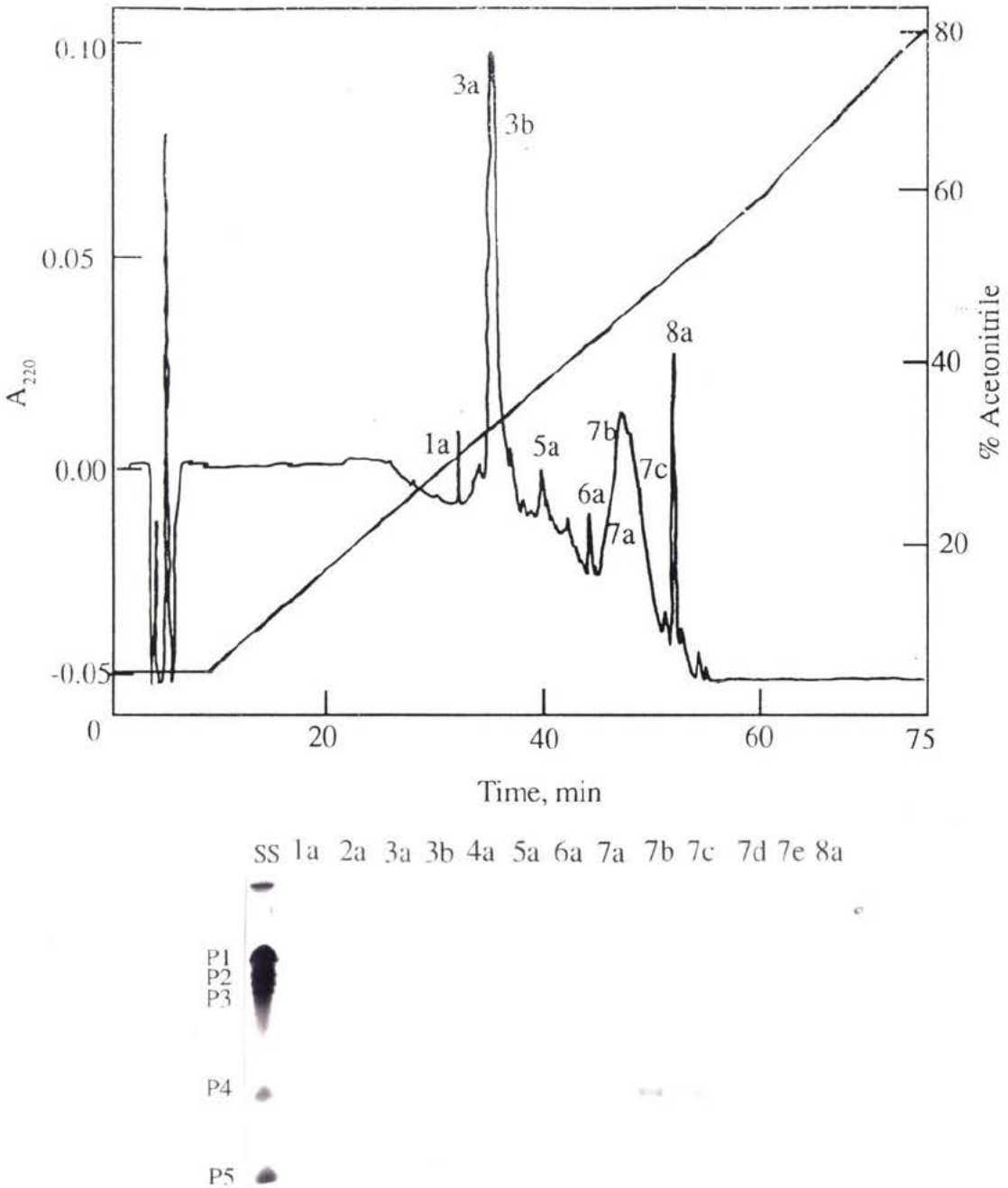


Figure 20. Reversed phase HPLC fractionation of peptide 4 from fraction 6d of the ion-exchange column. The HPLC column was a 25x0.46 cm 300-A C18 Vydac column. Flow rate was 1.0 ml/min. Detection wavelength was 220nm. The inset shows the electrophoretic analysis (in 15% polyacrylamide 2.5M urea-5% acetic acid PAGE) of different fractions along the chromatographic profile.

4. Amino acid sequencing of peptide 4

Purification of the peptide 4 allowed an attempt at determining its sequence by automatic Edman degradation sequencing analysis. However, because the quantity of peptide was too little after purification and the peptide was still not pure, amino acid sequencing was unsuccessful.

IV. Cross-linking PL-I from sperm nuclei of *Spisula solidissima* with DTBSP

According to previous studies on the chromatin structure of the sperm of *S. Solidissima* by Ausio and coworkers (31), the somatic-like histones and the major protamine-like component are segregated into two different domains in which short oligonucleosome stretches are interspersed within very long protamine-containing regions. About 80% of the DNA is complexed only with PL-I protein, which is a member of the histone H1 family (35). However, the spatial arrangement of the PL-I protein in these long nucleofilaments is not known. To determine if the PL-I molecule is spatially arranged head to tail (Fig 21 A), or tail to tail and head to head (B), crosslinking analysis was performed with DTBSP and PL-I. This bifunctional reagent has a maximum cross-linking span of about 11-12Å, and the resulting PL-I dimers were analyzed by two dimensional SDS PAGE (Fig 22B).

The PL-I dimer is formed (Fig 22A, see arrow). This dimer was then cleaved by NBS at tyrosine (Y) and tryptophan (W) residues of the PL-I globular core and then electrophoresed in the second dimension. Five peptides were produced by NBS cleavage (Fig 22B).

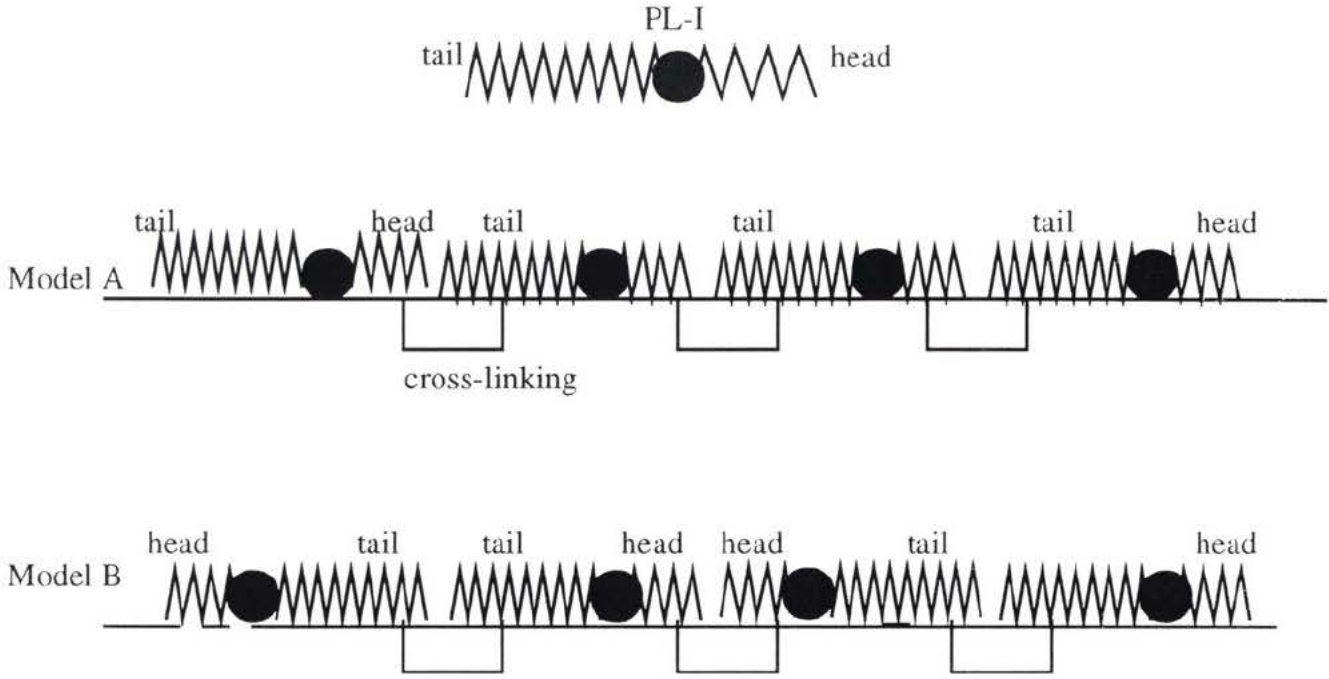


Figure 21. Two models of *Spisula* sperm chromatin organization.
(A) PL-I arranged head to tail.
(B) PL-I arranged tail to tail and head to head.

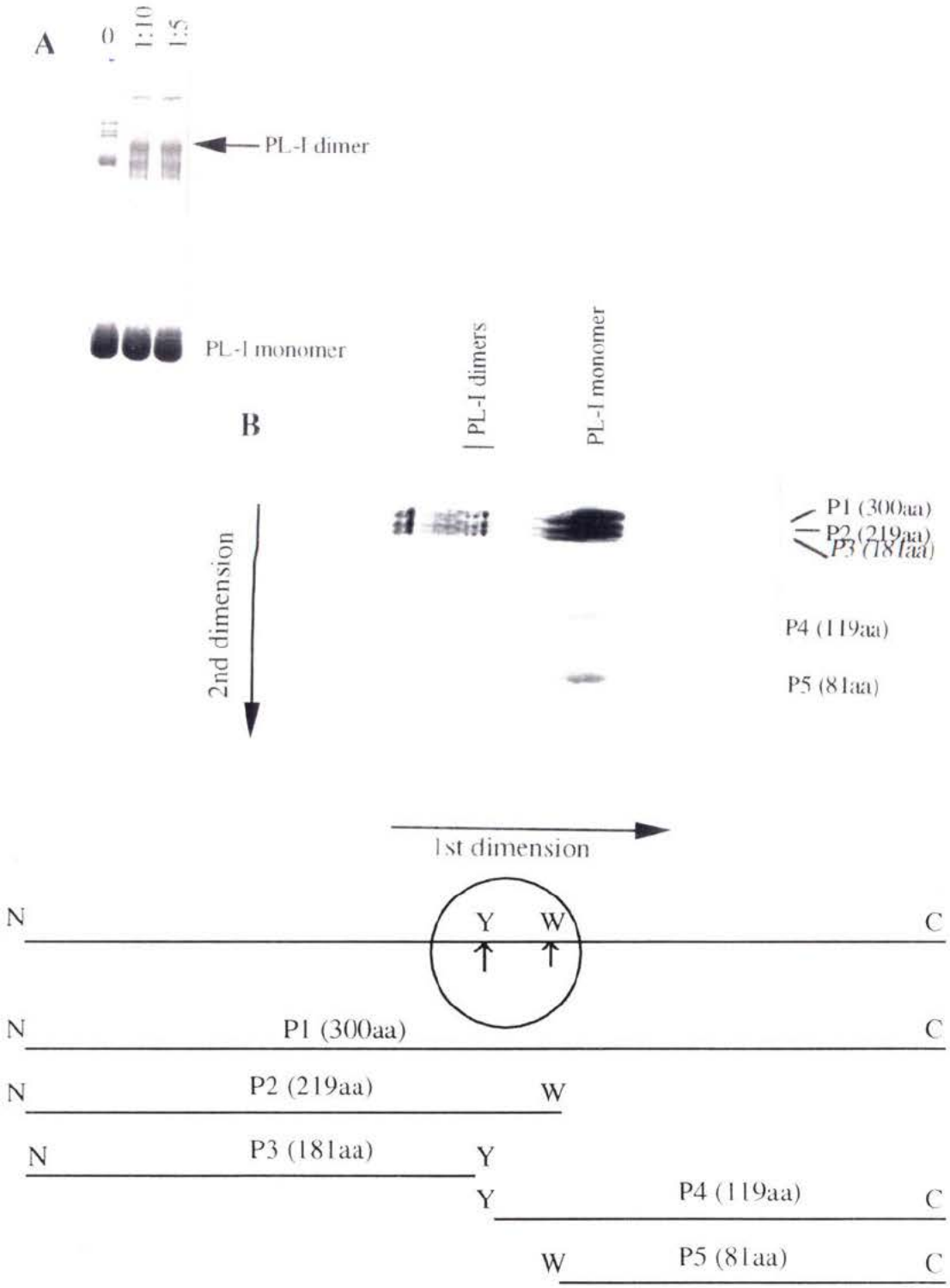


Figure 22. Cross-linking analysis of *Spisula* PL-I protein. (A) 5% polyacrylamide-2.5M urea-5% acetic acid PAGE analysis of crosslinked nuclei with diothiobis (succinimidyl propionate) from the sperm of *Spisula solidissima*. (B) Two dimensional electrophoresis of the proteins cross-linked of a ratio 3/10 of PL-I/diothiobis (w/w, described as 1/5x in the figure) after NBS cleavage.

The presence and absence of the peptides after NBS cleavage of crosslinked PL-I in the proposed models are compared with the actual observation (Table IV). The 2D gel analysis showed that the PL-I dimer was cleaved to the same five peptides as that of the monomer. From the two models proposed, model B is improbable according to the peptide pattern of the cleaved dimer. For model B to be plausible, there should be peptides P2, P3, P4, P5 and crosslinked P2-P2, P5-P5 present, while P1 should not be present; however, P1 was identified, but P2-P2 and P5-P5 were not. Therefore, model A is the accepted case. The dimer is crosslinked from the head of one PL-I molecule to the tail of the next, as shown. The possible cleaved peptides are P2, P3, P4, P5 and crosslinked P5-P2 or P4-P3, which is actually P1. All of these peptides are identified in the gel (Fig 22B). The alternative model is that PL-I is randomly arranged (i.e. it consists of model A and B). If the situation depicted by models A and B occurred with the same frequency, then all the peptides expected from both models should be present. However, this is not the case. If either A or B accounts for more than 50% of the chromatin, then the UA-PAGE would only detect the dominant case. Therefore, our results seem to disprove a random organization of PL-I.

Table IV. Presence and absence of the peptides after NBS cleavage of crosslinked PL-I in the proposed 2 models of *Spisula* sperm chromatin.

Model	Peptides present	Peptides absent
A	P1, P2, P3, P4, P5	
B	P2, P3, P4, P5, 2P2, 2P5	P1
observed	P1, P2, P3, P4, P5	

Therefore, it was concluded that on sperm chromatin filament, PL-I protein has a regular head-to-tail arrangement as presented in model A.

DISCUSSION

I. Structure of the sperm chromatin in *Mytilus californianus*

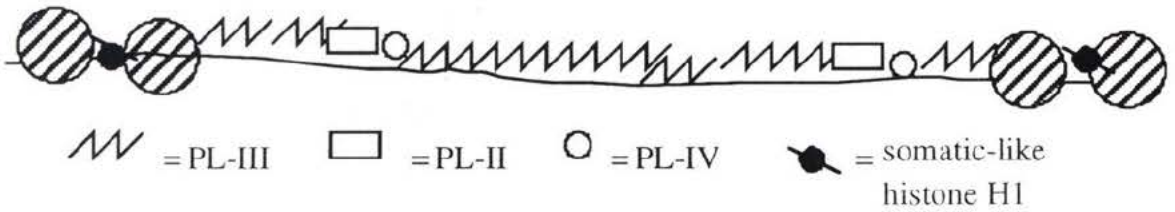
It was demonstrated that the DNA of *Mytilus* sperm is quite resistant to the action of micrococcal nuclease. The susceptibility to MNase is determined by the amount of nucleotides rendered soluble in perchloric acid. The amount of acid soluble DNA produced by digestion of sperm chromatin did not exceed 20% in 85 min compared to 30% for histone-containing chicken erythrocyte chromatin over the same time period. This is most likely due to tighter packing of the chromatin DNA by the sperm-specific proteins PL-II, PL-III and PL-IV. After PL-II and PL-IV have been selectively removed by salt extraction, the micrococcal nuclease susceptibility of fraction SI increases, reaching, after 60 min, levels of digestion similar to those of chicken erythrocyte chromatin. The significance of this is not yet clear.

The appearance of nucleosomal DNA in fraction SI after PL-II and PL-IV removal, and of nucleosomal DNA and multimers in fraction SII before salt extraction similar to that of chicken erythrocyte chromatin, suggest the presence of some nucleosomal structures in the *Mytilus* sperm chromatin. However, this nucleosomal fraction does not represent a very large fraction of the overall sperm chromatin. Indeed, a protein gel of fraction SII of *Mytilus* indicates a low percentage of histones compared to the PL proteins. In addition, in *Mytilus,* the susceptibility of DNA to MNase as determined from the PCA fraction (Fig 11A) and from SI (Fig 11B) are very similar. This suggests that the amount of oligonucleosomes present in fraction SI is also very small. In addition, even though fraction SII from *Mytilus* has a DNA band of similar size to that of nucleosomes in the fraction SII from chicken erythrocyte chromatin (Fig 12A), the majority of DNA in the *Mytilus* SII fraction has a smeared appearance. Furthermore, in the SII fraction (Fig 11C), the chromatin sensitivity to MNase of *Mytilus* sperm is only one fifth of that of chicken erythrocyte. Together all these facts suggest that only small region of DNA in *Mytilus* sperm chromatin is organized into nucleosomes.

In mammalian sperm, chromatin cannot be digested by MNase because all the chromatin is well protected by protamines (21). In chicken erythrocyte nuclei, all the chromatin is digested by nuclease because its chromatin is nucleosomal. The overall susceptibility of DNA to MNase of *Mytilus* sperm chromatin (both SI and SII) under native condition is very low (17%) compared to chicken erythrocyte chromatin solubility (100%) (Fig 11B and C). After PL-II and PL-IV removal, the overall chromatin solubility (both SI and SII) increases only to 18-35% in *Mytilus*. This resistance to digestion when taken

together with the fact that PL-III is the major component of the corresponding PII fraction (Fig 14), suggests that PL-III acts as protamine in *Mytilus* sperm chromatin. In fact PL-III is the major protein component of the PII fraction from both native and PL-II/PL-IV depleted chromatin (Figs 13B PII and 13C PII). In contrast, with mammalian sperm, in *Mytilus* a small portion of the sperm chromatin can be digested suggesting that chromatin in *Mytilus* sperm has a structural organization that is intermediate between protamine-containing vertebrate sperm and histone-containing cells.

All this allows us to put forward the following model:



In this model of *Mytilus* sperm chromatin, PL-III works as a normal protamine and the PL-III region are interspersed by short oligonucleosome regions.

However, the particular arrangement of PL-II and PL-IV remain yet to be established. In order to try to clarify this, cross-linking experiments of the *Mytilus* nuclear proteins were performed. The results from these experiments indicated that PL-II, PL-III and PL-IV are capable of intermolecular crosslinking. This suggests that PL-II, PL-III and PL-IV are in close proximity. PL-II and PL-IV are produced by posttranslational cleavage of a single polypeptide (46) and this is most likely responsible for the close proximity of PL-II and PL-IV detected in the cross-linking studies. PL arrangement in the above model explains the presence of the heterodimers PL-II/PL-III, PL-II/PL-IV, PL-IV/PL-III and PL-III/PL-III homodimer. However it cannot explain the presence of the PL-II/PL-II homodimer. The crosslinking tendency of the PL-II dimer is even greater than the crosslinking tendency of the PL-III dimer. This can be summarized as:
crosslinking tendency: 2PL-II>2PL-III>PL-II/PL-III, PL-II/PL-IV, PL-III/PL-IV (see table III)

It is expected that PL-III would intermolecularly crosslink because it accounts for 75% of the PL proteins, therefore occupying the majority of the chromatin DNA. However, the high crosslinking tendency of the PL-II dimer allows us to putward the following model:

In this model, two PL-II/PL-IV units are grouped together with the PL-II molecules adjacent to one another, and the PL-IV facing PL-III. This model does not exclude the possibility that there might also be single PL-II/PL-IV units dispersed among PL-III molecules. This model is also consistent with the nuclease digestion data explained above.

In conclusion, the studies of nuclease digestion and protein crosslinking in *M. californianus* demonstrated that the organization of sperm chromatin involves a mixture of two structures; small groups of oligonucleosomes interspersed between long stretches of nucleoprotamine filaments which result from the association of PL-II, PL-III, PL-IV with the DNA (Fig 23C and D). Based on this, we propose the following model for the histone to PL transition undergone by *Mytilus* sperm chromatin during spermatogenesis. In the stem cell, at the onset of spermatogenesis, chromatin structure is nucleosomal with linker histone H1 bound to where the DNA enters and exits each nucleosome as in somatic chromatin (Fig 23A). During sperm development, PL-II/PL-IV is most likely initially synthesized as a single molecule and functions in a manner similar to somatic histone H1. Together with PL-III, the PL-II/PL-IV precursor is synthesized in the cytoplasm and enter the nucleus where they begin to replace the nucleosomes (Fig 23B). Consequently core histones and the somatic linker histone H1 are displaced and degraded. When the PL-II/PL-IV precursor finds its site of attachment in the nucleosome that are going to be displaced is when it most likely undergoes posttranslational cleavage into PL-II and PL-IV. By the end of spermatogenesis most of the nucleosomes have been replaced by these PL proteins. A few residual nucleosomes are left in the chromatin which retain the somatic-like linker histone H1 attached to them. The rest of the DNA is occupied by PL-II, PL-III and PL-IV. The spatial arrangement of these PL could be that single PL-II/PL-IV units disperse among PL-III molecules (Fig 23C) or two PL-II/PL-IV units group together with PL-II molecules adjacent to one another, and the PL-IV facing PL-III (Fig 23D).

The model proposed here for *M. californianus* sperm chromatin organization is in complete agreement with the model suggested by Ausio (32) for *Spisula*. Small groups of oligonucleosomes are interspersed in long stretches of nucleoprotamine filaments. However in the case of *Spisula*, these nucleoprotamine filaments consist of a single protein (PL-I) instead of the three PL proteins seen in *Mytilus*.

In *Mytilus* sperm chromatin the complexes of DNA and PL-III may behave as typical nucleoprotamines. In fact *in vitro* studies carried out on DNA-PL-III complexes indicate a packaging of the resulting nucleoprotein complexes into a regular hexagonal system (34) very similar to what has been described for the vertebrate protamine-DNA complexes (3). In contrast, complexes of DNA and PL-IV show a slightly lower degree of

order.

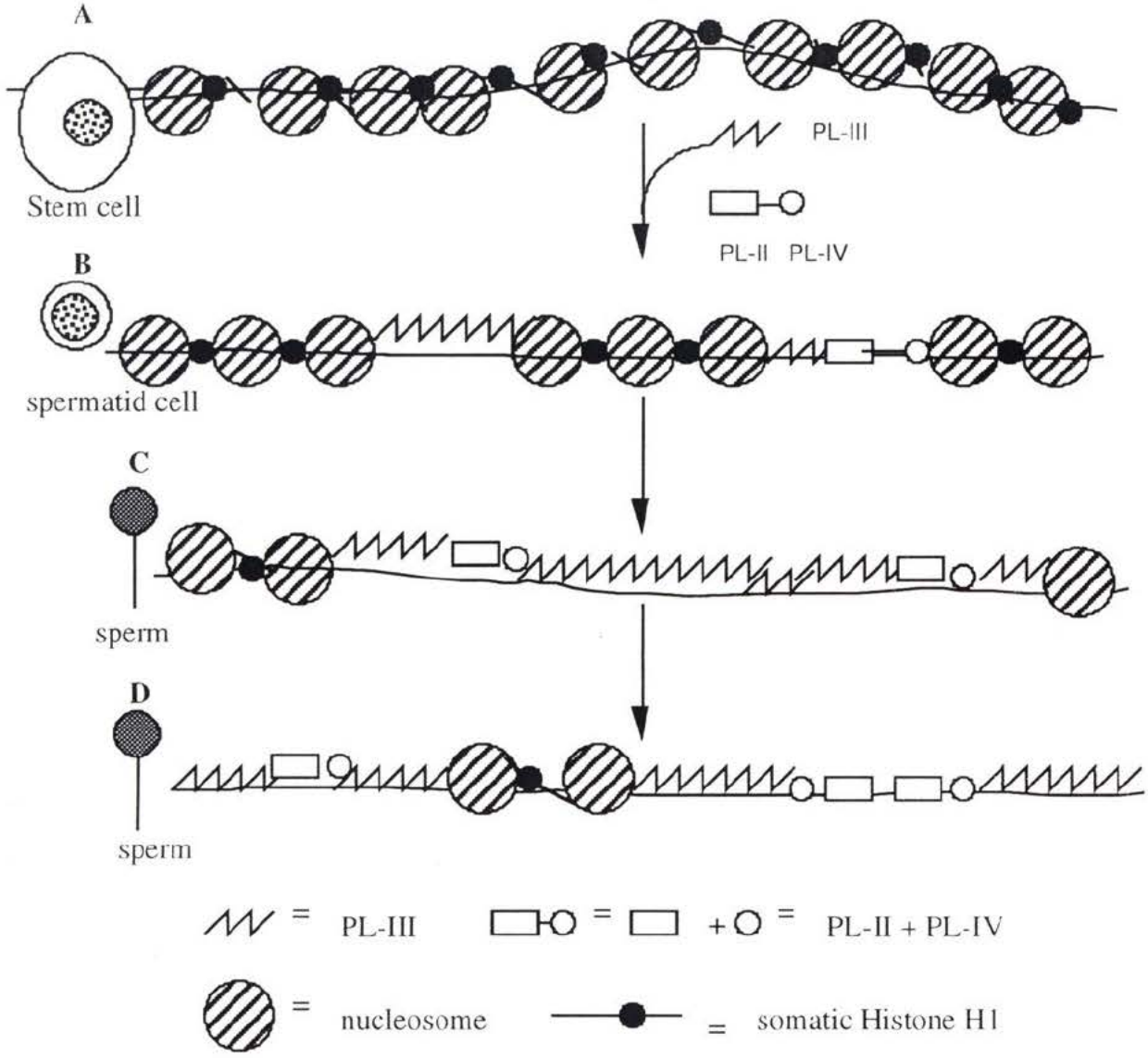


Figure 23. Chromatin structure differentiation during spermatogenesis of *Mytilus californianus*. (A) In the somatic stem cell, chromatin appears as a regular structure. (B) During development from stem cell to spermatid cell PL-III, PL-II and PL-IV begin to replace some of the nucleosomes as well as the somatic histone H1. PL-II and PL-IV exist as a single molecule at this time. (C) During the development from spermatid cell to sperm cell, more nucleosomes are replaced by these PL proteins, the PL-II/PL-IV precursor is split once bound to the DNA and becomes two separate proteins PL-II and PL-IV (D).

II. Sperm chromatin structure of *Spisula solidissima*

In a previous sperm chromatin study of *Spisula solidissima* by Ausio (31), the chromatin structure was described as short oligonucleosome stretches interspersed within very long PL-I-containing regions. However, the spatial arrangement of PL-I on the chromatin structure was not clear at that time. From the crosslinking study in this work, it was found that PL-I is arranged in a regular fashion. The N-terminal region of one PL-I molecule directly follows the C-terminal region of the next PL-I molecule (head to tail arrangement) (Fig 21, model A).

In this work, the presence of cysteine was also detected for the first time in PL-I (Fig 18A). We have shown that there are at least two cysteines in this molecule indicated by the presence of PL-I oligomers larger than dimers. The location of these cysteines was demonstrated to be within the globular region of the PL-I molecule between tryptophan (W) and tyrosine (Y). However, a more precise location could not be determined because of the low yield and impurity of the peptides used for sequencing.

The biological significance of the dual chromatin organization observed in both *M. californianus* and *S. solidissima* sperm is not clear. It is proposed, however, that the histone-chromatin region may play an important role in the genetic activation in the early stages of development following fertilization (58).

REFERENCE

1. **Wolfe, S.L. 1993.** *Molelcular and cellular biology.* Pp.1119-1120. Wadsworth Publishing Company, Belmont, California.

2. **Bloch, D.P. 1969.** A catalog of sperm histones. *Genetics (Suppl.)* **61**: 93-110.

3. **Subirana, J.A. 1983.** Nuclear proteins in spermatozoa and their interactions with DNA. In: Andre J (ed). *The sperm cell.* pp.197-213. Martinus Nijhoff, The Hague.

4. **Ausio, J. 1986.** Structural variability and compositional homology of the protamine-like components of the sperm from the bivalve mollusks. *Comp. Biochem. Physiol.* **85**: 439-449.

5. **Felix, K. 1960.** Protamines. *Adv. Protein Chem.* **15**: 1-53.

6. **Colom, J., J.A. Subirana. 1979.** Protamines and related proteins from spermatozoa of molluscs. *Biochim. Biophys. Acta* **581**:217-227.

7. **Ausio, J., J.A. Subirana. 1982.** A high molecular weight nuclear basic protein from the Bivalve Mollusc *Spisula solidissima. J. Biol. Chem.* **257**: 2802-2805.

8. **Ando, T. 1973.** Protamines--isolation, characterization, structure and function. Springer-Verlag New York.

9. **Daban, M., A. Martinage, M. Kouach, M. Chiva, J.A. Subirana, P. Sautière. 1995**. Sequence analysis and structural features of the largest known protamine isolated from the sperm of the archaeogastropod *Monodonta turbinata. J. Mol. Evol.* **40**: 663-670.

10. **Subirana, J.A. 1990.** Analysis of the charge distribution in the C-terminal region of histone H1 as related to its interaction with DNA. *Biopolymers.* **29**: 1351-1357.

11. **Retief, J.D., R.J. Winkfein, G.H. Dixon, R. Adroer, R. Queralt, J. Ballabriga, R. Oliva. 1993.** Evolution of protamine P1 genes in primates. *J. Mol. Evol.* **37**: 426-434.

12. **Oliva, R., G.H. Dixon. 1991.** Vertebrate protamine genes and the histone-to-protamine replacement reaction. *Prog Nucleic Acid Res. Mol. Biol.* **40**: 25-94.

13. **Louie, A.J., E.P.M. Candido, G.H. Dixon. 1973.** Enzymatic modifications and their possible roles in regulating the binding of basic proteins to DNA and in controlling chromosomal structure. *Cold Spring Harbor Symposia on Quantitative Biology.* Pp. 803-819. Vol. **38**. Cold Spring Harbor Laboratory.

14. **Dixon, G.H., E.P.M. Candido, B.M. Honda, A.J. Louie, A.R. Macleod, M.T. Sung. 1974.** The biological roles of post-synthetic modifications of basic nuclear proteins. *The structure and function of chromatin.* pp. 229-258. Elsevier, N. Holland. Ciba Foundation.

15. **Oliva, R., D. Bazett-Jones, C. Mezquita. 1987.** Factors affecting nucleosome disassembly by protamines *in Vitro*. Histone hyperacetylation and chromatin structure, time dependence, and the size of the sperm nuclear proteins. *J. Biol. Chem.* **262**:17016-25.

16. **Wolfe, S.L. 1981.** Biology of the cell. Pp. 268. 2nd ed., Wadsworth Pub. Co., Belmont, California.

17. **Risley, M.S. 1988.** Chromatin organization in sperm. In: *Chromasomes: Eukaryotic, Prokaryotic and Viral* . Pp. 61-85. K. Adolph, ed. CRC Press, Boca Raton, Florida.

18. **Kansinsky, H.E. 1989.** Specificity and distribution of sperm basic proteins. In: *Histones and other basic nuclear proteins*. Hnilica L., Stein G. And Stein J., eds. CRC Press, Boca Raton, Florida. Scientific publications. pp.25

19. **Subirana, J.A., C. Cozcolluela, J. Palau, M. Unzeta. 1973.** Protamines and other basic proteins from spermatozoa of molluscs. *Biochim. biophys. Acta.* **317**: 364-379.

20. **Risley, M.S., R.A. Eckhardt, M. Mann, H.E. Kasinsky.1982.** Determinants of sperm nuclear shaping in the genus *Xenopus*. *Chromosoma*. **84**, 557-569.

21. **Subirana, J.A. 1975.** On the biological role of basic proteins in spermatozoa and during spermiogenesis. In: *The biology of the male gamete.* Pp. 239-244. Duckett J.C. And Racey P.A., Eds. Academic Press, New York.

22. **Sega, G.A., J.G. Owens. 1978.** Ethylation of DNA and protamine by ethyl methanesulfonate in the germ cells of male mice and the relevancy of these molecular targets to the induction of dominant lethals. *Mutat. Res.* **52**: 87-106.

23. **Bradbury, E.M., W.C. Price, G.R. Wilkinsin. 1962.** (1) Polarized infrared studies on nucleoproteins. I. Nucleoprotamine. *J. Mol. Biol.* **4**: 39-49.

24. **Balhorn, R. 1982.** A model for the structure of chromatin in mammalian sperm. *J. Cell Biol.* **93**, 298-305.

25. **Warrant, R.W., S.H. Kim. 1978.** α-Helix-double helix interaction shown in the structure of a protamine-transfer RNA complex and a nucleoprotamine model. *Nature*. **271**, 130-135.

26. **Zalenskaya, I.A., A.O. Zalensky. 1980.** Basic chromosomal proteins of marine invertebrates ----I. Sperm histones of nine sea urchin species. *Comp. Biochem. Physiol.* **65B**: 369-373.

27. **Ausio, J. 1992.** Presence of a highly specific histone H1-like protein in the chromatin of the sperm of the bivalve mollusks. *Mol. Cell Biochem.* **115**: 163-172.

28. **Kuhn-Schnyder, E., H. Rieber. 1986**. Handbook of Paleozoology. The John Hopkins University Press, Baltimore and London.

29. **Coan, E.V., J.T. Carlton. 1975.** In light's Manual: Intertidal Invertebrates of the Central California Coast: 3rd edn. (Edited by Smith and Carlton). University of California Press, Berkeley.

30. **Ausio, J. 1992.** Purification and biochemical characterization of the nuclear sperm-specific proteins of the Bivalve mollusks *Agriodesma saxicola* and *Mytilimeria nuttalli*. *Ref: Biol. Bull.* **182**: 31-40.

31. **Zalenskaya, I.A., N.A. Odintsova, I.V. Vorb'ev. 1985.** Chromatin from sperm of Bivalvia mollusks. *FEBS Lett.* **188**: 243-247.

32. **Ausio, J., K.E. Van Holde. 1987.** A dual chromatin organization in the sperm of the Bivalve mollusk: Spisula solidissima. *Eur. J. Biochem.* **165**: 363-371.

33. **Zalensky, A.O., Z.V. Avromova. 1984.** Nucleosomal organization of a part of chromatin in mollucks sperm nuclei with a mixed basic protein composition. *Mol. Biol. Rep* **10**: 69-74.

34. **Ausio, J., P. Suau. 1983.** Structural heterogeneity of reconstituted complexs of DNA with typical and intermediate protamines. *Biophys. Chem.* **18**: 257-267.

35. **Ausio, J., K.E. van Holde. 1988.** The histones of the sperm of *Spisula solidissima* include a novel, cysteine-containing H1 histone. *Cell Differentiation.* **23**: 175-190.

36. **Ausio, J., A.Toumadje, K.E. van Holde. 1987.** Structural characterization of the trypsin-resistant core in the nuclear sperm-specific protein from *Spisula solidissima*. *Biochem.* **26**: 975-982.

37. **Libertini, L.J., J. Ausio, K.E. van Holde. 1988.** Highly cooperative binding to DNA by a histone-like, sperm-specific protein from *Spisula solidissima*. *Biopolymers* **27**: 1459-1477.

38. **Mogensen, C., S. Carlos, J. Ausio. 1991.** Microheterogeneity and interspecific variability of the nuclear sperm proteins from *Mytilus*. *FEBS.* **282**: 273-276.

39. **Ausio, J., J.A. Subirana. 1982.** Nuclear proteins and the organization of chromatin in spermatozoa of Mytilus Edulis. *Exp. Cell Res.* **141**: 39-45.

40. **Carlos, S., L. Jutglar, I. Borrell, D.F. Hunt, J. Ausio. 1993.** Sequence and characterization of a sperm-sepcific histone H1-like protein of *Mytilus californianus*. *J. Biol. Chem.* **268**: 185-194.

41. **Jutglar, L., J.I. Borrell, J. Ausio. 1991.** Primary, secondary, and tertiary structure of the core of a histone H1-like protein from the sperm of *Mytilus*. *J. Biol. Chem.* **266**: 8184-8191.

42. **Ausio, J., R. McParland. 1989.** Sequence and characterization of the sperm-specific protein f3 from *Mytilus californianus*. *Eur. J. Biochem.* **182**: 569-576.

43. **Ausio, J. 1988.** An unusual cysteine-containing histone H1-like protein and two protamine-like proteins are the major nuclear proteins of the sperm of the bivalve mollusc *Macoma nasuta*. *J. Biol. Chem.* **263**: 10141-10150.

44. **Bustin, M., R.D. Cole. 1970.** Regions of high and low cathionic charge in a lysine-rich histone. *J. Biol. Chem.* **245**: 1458-1466.

45. **Pheland, J.J., J. Colom, J.A. Subirana, R.D. Cole. 1974.** A lysine-rich protein from spermatozoa of the mollusc *Mytilus edulis*. *J. Biol. Chem.* **249**: 1099-1102.

46. **Carlos, S., D.F. Hunt, C. Rochini, D.P. Arnott, J. Ausio. 1993.** Post-translational cleavage of a histone H1-like protein in the sperm of *Mytilus*. *J. Biol. Chem.* **268**: 195-199.

47. **Subirana, J.A. 1973.** Studies on the thermal denaturation of nucleohistones. *J. Mol. Biol.* **74**: 363-367.

48. **Kuehl, J. 1979.** Synthesis of high mobility group proteins in regenerating rat liver. *J. Biol. Chem.* **254**: 7276-7281.

49. **Kosower, N.S. 1987.** Thiol labeling with bromobimanes. In: *Methods in Enzymology*. **143**: 76-84.

50. **Thomas, J.O. 1989**. Chemical cross-linking of histones. In: *Methods in Enzymology*. Vol **170**: 549-571.Wassarman P.M. Ed., Academic Press Inc., San Diego, California.

51. **Davie, J. R. 1982.** Two-dimensional gel systems for rapid histone analysis for use in minislab polyacrylamide gel electrophoresis. *Anal. Biochem.* **120**: 276-281.

52. **Oakley, B.R. 1980.** A simplified ultrasensitive silver stain for detecting proteins in polyacrylamide gels. *Anal. Biochem.* **105**: 361-363.

53. **Laemmli, U.K. 1970.** Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*. **227**:680-685.

54. **Sambrook, J., E.F. Fritsch, T. Maniatus. 1989.** Molecular Cloning: a laboratory manual. Vol. 3. (second Edition) Cold Spring Harbor, New York.

55. **Morris, N.R. 1976.** A comparison of the structure of chicken erythrocyte and chicken liver chromatin. *Cell*. **9**: 627-632.

56. **Ausio, J., J.A. Subirana. 1982.** Nuclear proteins and the organization of chromatin in spermatozoa of *Mytilus Edulis*. *Exp Cell Res* **141**: 39-45.

57. **Longo, F.J., E.J. Dornfeld. 1967.** The fine structure of spermatid differentiation in the Mussel, *Mytilus Edulis*. *J. Ultrastruct Res.* **20**: 462-469.

58. **Olivares, C., M.L. Vera, S. Ruiz-Lara. 1993.** Coexistence of two chromatin structures in sperm nuclei of the bivalve mollusc *Protothaca thaca*. *Mol. and Cell. Biochem.* **125**: 87-95.

VITA

Surname: Zhang Given Names: Fan

Place of Birth: Qingdao, Shandong, People's Republic of China

Educational Institutions Attended:

University of Victoria	1993-1996
Anhui University (China)	1987-1991

Degrees Awarded:

B.Sc	Anhui University (China)	1991

Honours and Awards:

University of Victoria Fellowship	1993-1994

Publications:

Rocchini, C., F. Zhang, J. Ausio. 1995. Two highly specialized histone H1 proteins are the major chromosomal proteins of the sperm of the sea anemone *Urticina (Tealia) crassicornis*. *Biochem.* 34 (48): 15704-12.

PARTIAL COPYRIGHT LICENSE

Title of Thesis:

Characterization of the sperm chromatin structure in *Mytilus californianus* and *Spisula solidissima*

Author [redacted]

Fan Zhang
June 20, 1996