J Mol Evol (1995) 40:663–670

JOURNAL OF MOLECULAR EVOLUTION

Sequence Analysis and Structural Features of the Largest Known Protamine Isolated from the Sperm of the Archaeogastropod *Monodonta turbinata*

M. Daban,[1] A. Martinage,[2] M. Kouach,[2] M. Chiva,[1,3] J.A. Subirana,[1] P. Sautière[2]

[1] Departament d'Enginyería Química, ETSEIB, Universitat Politècnica de Catalunya, 08028, Barcelona, Spain
[2] URA, CNRS 1309, Chimie des Biomolécules, Institut Pasteur, 59019 Lille, France
[3] Departament de Ciències Fisiològiques Humanes i de la Nutrició, Universitat de Barcelona, 08028, Spain

Received: 4 April 1994 / Revised: 8 August 1994 / Accepted: 14 August 1994

Abstract. Protamine of the archaeogastropod mollusc *Monodonta turbinata* has been isolated and characterized. With a mass of 13,476 Da, it is the largest known protamine. Amino acid sequence of this protamine (106 residues) was established from data provided by automated sequence analysis and mass spectrometry of the protein and of its fragments. The primary structure of the NH_2-terminal region exhibits repetitive sequence motifs "Basic-Ser" (mainly R-S) and both central and COOH-terminal regions are composed by arginine clusters. The amino acid sequence of *Monodonta turbinata* protamine shows structural similarities with other protamines from invertebrates and from birds and mammals.

Key words: Protamine — Sperm basic proteins — Archaeogastropod — Mollusc

Introduction

Protamines (or sperm-specific proteins) are very basic molecules found in the nuclei of spermatozoa in almost all animal species (Bloch 1969). These proteins display an enormous diversity (see reviews by Poccia 1986; Kasinsky 1989; Oliva and Dixon 1991), and despite the fact that they have been studied for more than a century (Miescher 1874), their structural and evolutionary relationships are understood only in limited cases (Oliva and Dixon 1991; Retief et al. 1993; Retief and Dixon 1993). They share the same general function: electrostatic neutralization and interaction with DNA with subsequent nuclear condensation. This fact imposes general constraints over their amino acid composition and primary structure (sequence), which have been solved in many different ways during evolution, as is shown by the great interspecific variability of these proteins. Compare for instance the protamines from vertebrates (Kasinsky et al. 1985; Chiva et al. 1989; Oliva and Dixon 1991), bivalve molluscs (Subirana et al. 1973; Ausió 1988), and cephalopod molluscs (Martin-Ponthieu et al. 1991; Wouters-Tyrou et al. 1991), or other invertebrates (Saperas et al. 1992; Chiva et al. 1992). Such variability makes this protein "family" an interesting model of protein evolution. However, it is necessary, when studying these proteins, to take into account the taxonomic context of the animal species in which they appear since the terminology "protamine" does not indicate—in an evolutionary sense—a real protein family, but instead a group of proteins with the same localization and function, and with a high percentage of lysine and/or arginine (Subirana 1983).

The phylum Mollusca offers a great evolutionary interest because of its biological diversity. In this phylum, sperm nuclear basic proteins ("protamines") have been extensively investigated (Subirana et al. 1973; Ausió 1988; Daban et al. 1990; Martin-Ponthieu et al. 1991; Wouters-Tyrou et al. 1991; Daban et al. 1991a,b; Mogensen et al. 1991). The great diversity of these proteins—histones, protaminelike proteins (PL proteins), or true protamines—even within the same species, has been shown through analytical electrophoresis and amino acid

Correspondence to: P. Sautière

composition. At this time, sequence studies have been only performed on sperm nuclear basic proteins of a cephalopod (*Sepia officinalis*) (Martin-Ponthieu et al. 1991; Wouters-Tyrou et al. 1991; Schindler et al. 1991) and a bivalve (*Mytilus*) (Ausio and McParland 1989; Carlos et al. 1993a,b; Ruiz-Lara et al. 1993). The partial amino acid sequence of the protaminelike protein EM-1 from the bivalve mollusc *Ensis minor* has been also reported (Giancotti et al. 1992). These studies demonstrate that the PL proteins from external fertilizing bivalves apparently have little similarity with protamines from cephalopods which are more evolved molluscs with internal fertilization.

In this paper, we present the sequence analysis and structural features of the *Monodonta turbinata protamine. Monodonta turbinata* belongs to Archaeogastropoda, a primitive group of gastropod molluscs with a primitive type of reproduction (external fertilization) (Daban et al. 1990, 1991b; Chiva et al. 1991).

Methods

Monodonta turbinata (Archaeogastropoda, Trochacea) was collected in May and June from the Mediterranean coast of Catalunya (Spain).

Endoproteinase Lys-C was from Boehringer. Thermolysin was from Merck. *Astacus fluviatilis* proteinase was from Serva. Carboxypeptidase B treated with iPr2*P* F was purchased from Sigma. Acetonitrile for reverse-phase HPLC was obtained from Carlo Erba. All reagents and solvents for gas-phase sequencing were from Applied Biosystems. All other reagents were of the highest purity available.

Protamine Isolation and Purification. The sperm was obtained from mature gonads and the nuclei were purified as described previously (Chiva et al. 1990). Nuclei were successively extracted with 35% acetic acid and 0.25 M HCl. The protamine fraction was recovered from the HCl extract by precipitation with 6 vol of cold acetone. Purification of the protamine was achieved by ion-exchange chromatography on carboxymethylcellulose (CM-52 Whatman) equilibrated in 50 mM sodium acetate buffer, pH 6.0, containing 0.2 M NaCl and eluted with a linear gradient of NaCl in the same buffer.

Electrophoretic Analysis. Purity of the protamine was assessed by polyacrylamide slab gel electrophoresis at pH 3.2 in the presence of 6.25 M urea using a 17% acrylamide concentration (Panyim and Chalkley 1969).

Amino Acid Analyses. Protamine and peptide samples were hydrolyzed in vacuo in 6 M HCl at 110°C for 24 h. Amino acid analyses were performed on a Beckman 6300 amino acid analyzer.

Carboxy-Terminal Analysis. Protamine (1 nmol) dissolved in 0.2 M ammonium bicarbonate pH 8.0 was digested for 3 h at 37°C with carboxypeptidase B treated with iPr2*P* F using an enzyme-to-substrate ratio of 1:25 (by weight).

Enzymatic Hydrolyses. The protamine (about 180 nmol) was dissolved in 0.5 ml of 0.1 M ammonium bicarbonate pH 8.5 and hydrolyzed with endoproteinase Lys-C for 2 h at 37°C using an enzyme-to-substrate ratio of 1:100 (by weight). Hydrolysis was stopped by lowering the pH at 3.0 with formic acid.

The protamine (about 60 nmol) was dissolved in 0.7 ml of 0.1 M ammonium bicarbonate pH 8.0 and hydrolyzed with thermolysin for 4 h at 40°C using an enzyme-to-substrate ratio of 1:100 (by weight). Hydrolysis was stopped by lowering the pH at 3.0 with formic acid.

The protamine (about 60 nmol) was dissolved in 0.6 ml of 0.1 M ammonium bicarbonate pH 8.0 and hydrolyzed with endoproteinase from *Astacus fluviatilis* for 2 h at 30°C, using an enzyme-to-substrate ratio of 1:50 (by weight). Hydrolysis was stopped by lowering the pH at 3.0 with formic acid.

Separation of Peptides. Peptides generated from enzymatic hydrolyses of *Monodonta* protamine were separated by reverse-phase HPLC on C18 Superspher endcapped column (Merck) (250 × 4 mm) using a gradient of acetonitrile in 0.1% trifluoroacetic acid.

Nomenclature of Peptides. Peptides obtained by cleavage of the protamine with endoproteinase Lys-C, thermolysin, and *Astacus fluviatilis* proteinase were designated by K, Th, and A, respectively, and numbered according to their position in the sequence of the protamine.

Sequence Analysis. The protamine and its fragments were submitted to automated Edman degradation on a gas-phase sequencer Applied Biosystems 470A using the 03 RPTH program slightly modified to ensure a better extraction of the 2-anilino-5-thiazolinone of arginine (03C Arg program). Phenylthiohydantoin derivatives of amino acids were identified on line as described in Sautière et al. (1988).

Electrospray Mass Spectrometry. Electrospray mass spectrometry was performed on a VGBio-Q quadrupole mass spectrometer with a mass range of 3,000 Da. The electrospray ion source was operating at atmospheric pressure. Calibration was performed using charged ions from poly(ethyleneglycol)800, which was introduced separately.

The electrospray was emitted at 3,000 V. The extraction cone voltage (Ve) was adjusted to 150 V. The samples were first dissolved in water containing 1% acetic acid and then an equivalent volume of methanol was added. The concentration used was between 20 and 30 pmol/μl. The sample solutions (2–10 μl) were introduced into the ion source at a flow rate of 2 μl/min.

Fast Atom Bombardment Mass Spectrometry (FAB-MS). Positive FAB mass spectrometry was carried out on a concept II HH (Kratos Analytical, Manchester, U.K.) four-sector tandem mass spectrometer, which consists of two double-focusing forward-geometry instruments joined back to back (E.B.E.B.). The spectrometer was equipped with a commercial Kratos FAB source, an ion Tech B11 NF saddle-field fast atom gun (Ion Tech, Teddington, U.K.), and a Kratos DS-90 data system. A beam of Xe atoms of 8-keV impact energy and equivalent to 1-mA emission current was employed to ionize peptides dissolved in matrix (glycerol/water/trifluoroacetic acid in ratio 10/88/2). The FAB-produced ions were accelerated through a potential of 8 kV and mass-selected by using MS-I (ESA-I and the magnetic sector) at a resolution of about 1,500 (full width at 5% height). Cesium iodide was used as standard compound for mass calibration of samples. The peptides were dissolved in deionized water at a concentration of 1 nmol/μl. One microliter of peptide solution was deposited on a stainless-steel target and 1 μl of matrix was added.

Results

Sperm nuclei of the archaeogastropod *Monodonta turbinata* contain one protamine which accounts for 92% of nuclear protein complement, together with 8% of a sperm-specific histone H2B (Colom and Subirana 1981).

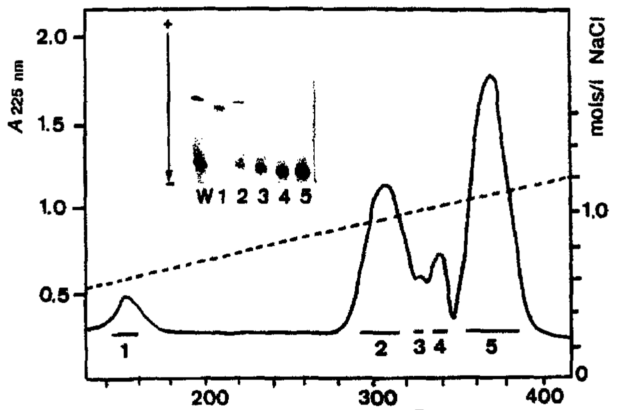

Fig. 1. Fractionation by ion-exchange chromatography on carboxymethylcellulose of the 0.25 M HCl extract of *Monodonta turbinata* sperm nuclei. Column was equilibrated with 50 mM sodium acetate pH 6.0 containing 0.2 M sodium chloride. Elution was performed with a linear gradient of sodium chloride in the same buffer. The **insert** shows the polyacrylamide gel electrophoresis of fractions 1–5 and of the whole extract (*W*). Electrophoresis was performed on slab gel (160 × 180 × 0.75 mm) at pH 3.2, in 0.9 M acetic acid/6.25 M urea, using a 17% acrylamide concentration, for 2.5 h at 22 mA.

This protamine was obtained in pure form after ion-exchange chromatography on carboxymethylcellulose (Fig. 1). It was eluted with 1.1 M NaCl in fraction 5. Its amino acid composition is given in Table 1.

The high amount of arginine (57%) and the small number of constitutive amino acids (Daban et al. 1991b) are characteristic of a true protamine. *Monodonta* protamine clearly differs from other known mollusc protamines in its basic amino acid content. In the bivalve, *Mytilus californianus,* the major protamine (ø1), has 22.5% lysine and 27% arginine, whereas the three variants of the minor protein Ø3 contain 50% of lysine (Ausió and McParland 1989). Lysine is present in a small amount (6%) in protamine of the gastropod *Monodonta turbinata* (this work) and it is absent in protamine of the cephalopod *Sepia officinalis* (Martin-Ponthieu et al. 1991). Furthermore, the protamine of another gastropod, the neogastropod *Murex brandaris,* contains equivalent amounts of lysine and arginine (37%) (Unpublished results).

Monodonta protamine (106 residues) has a molecular mass of 13,476.8 Da calculated from the sequence, which is in perfect accordance with that determined by electrospray mass spectrometry, 13,475.7 ± 1.9 Da (Fig. 2). The minor series of peaks yielded a mass of 13,573 ± 2.9 Da. The difference (98 Da) between these two masses can be attributed to an adduct of phosphate.

The amino-terminal sequence of *Monodonta* protamine was determined up to residue 38 after automated Edman degradation of the intact protein. The digestion of the protamine with carboxypeptidase B for 1 h released two arginine residues (Fig. 3).

The remainder of the sequence was unambiguously established by sequencing peptides generated from cleavage of the protein with endoproteinase Lys-C, thermolysin, and *Astacus fluviatilis* endoproteinase. The amino acid compositions of the peptides used to elucidate the complete sequence of *Mondonta* protamine are presented in Tables 1 and 2.

From the cleavage of protamine with endoproteinase Lys-C, six peptides (K1 to K6) of the seven expected were obtained (Table 1 and Fig. 3). The C-terminal peptide (residues 93–106) was not recovered from the hydrolysate. The peptide K3 (residues 27–70) was further hydrolyzed with thermolysin in 0.1 ammonium bicarbonate, pH 8.0, for 4 h at 40°C using an enzyme-to-substrate ratio of 0.1 μg/nmol. The five peptides obtained (K3-Th1 to K3-Th5) altogether correspond to the sequence 29-70 (Table 2 and Fig. 3). The mass of peptide K3 (5,619.5 Da) calculated from the sequence data is fully consistent with the mass measured by electrospray mass spectrometry (5,620 ± 1.3 Da) (Fig. 4).

Cleavage of *Monodonta* protamine with thermolysin yielded two useful peptides—Th1 and Th2. Peptide Th1 (residues 60–80) overlaps peptides K3 and K4. Peptide Th2 (residues 89–103) overlaps peptide K6 and allows extension of the sequence of the protamine up to residue 103 (Table 2 and Fig. 3). Finally, when the protamine was hydrolyzed with *Astacus fluviatilis* endoproteinase, two peptides were found very useful in establishing its complete sequence. Peptide A1 (residues 78–88) overlaps peptides K4, K5, and K6. Peptide A2 (residues 97–106) corresponds to the carboxy-terminal sequence of the protamine (Table 2 and Fig. 3). The mass of this peptide (1,439.7 Da) calculated from the sequence data is in full agreement with the mass measured by FAB-MS (1,439.7 Da) (Fig. 5).

Basic amino acids of *Monodonta* protamine, especially arginine, are distributed along the polypeptide chain, mainly located in sequences Arg-Ser in the amino-terminal part of the protein and scattered in clusters of four, five, or six residues in the C-terminal part of the protein. Furthermore, three clusters of arginine residues are in palindromic sequences:

46 54
Ala–Ser-Arg-Arg-Arg-Arg-Arg-Ser-Ala

61 68
Ser-Arg-Arg-Arg-Arg-Arg-Arg-Ser

79 88
Arg-Lys-Val-Arg-Arg-Arg-Arg-Val-Lys-Arg

Serine residues are located in the amino-terminal part (residues 1–71) of the protamine and are always adjacent to an arginine residue.

Discussion

Two different regions within the primary structure of *M. turbinata* protamine may be recognized:

Table 1. Amino acid composition of *Monodonta turbinata* protamine P2 and of peptides generated by cleavage of the protein with endoproteinase Lys-C (K)[a]

Amino acid	P2 (mol/mol)	K1, 1–17	K2, 18–26	K3, 27–70	K4, 71–80	K5, 81–87	K6, 88–92
Thr	1.6 (2)			1.0 (1)			
Ser[b]	13.2 (18)	2.5 (3)	3.4 (4)	6.9 (10)	0.6 (1)		0.6 (0)
Gly	4.1 (5)	1.0 (1)	0.7 (0)	1.2 (1)	1.7 (2)		1.5 (1)
Ala	9.5 (10)	1.7 (3)	1.2 (1)	4.1 (4)			1.1 (1)
Val	3.5 (4)	0.7 (1)		1.1 (1)		1.6 (2)	
Lys	5.9 (6)	1.0 (1)	1.0 (1)	1.0 (1)	1.0 (1)	1.0 (1)	1.0 (1)
Arg	61.9 (61)	9.8 (9)	3.0 (3)	28.2 (26)	5.8 (6)	4.4 (4)	2.0 (2)
Total residues	106	17	9	44	10	7	5
Mass[c]	13,476.8			5,619.5			
Mass[d]	13,475.7 ± 1.9	n.d.	n.d.	5,620.3 ± 1.3	n.d.	n.d.	n.d.

[a] Values in parentheses are the number of residues/molecule of protein or peptide derived from the sequence. n.d., not determined
[b] Uncorrected values for hydrolytic losses
[c] Calculated masses in Da
[d] Masses in Da, measured by electrospray mass spectrometry (ESMS)

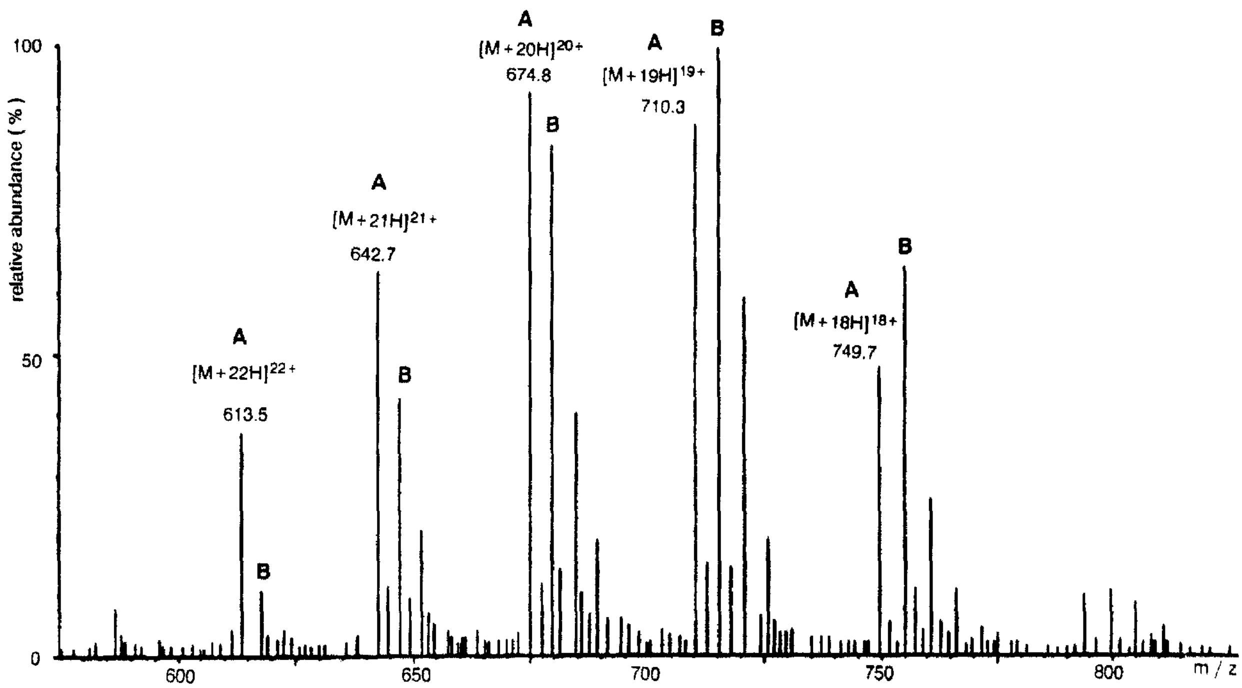

Fig. 2. Electrospray mass spectrum of *Monodonta turbinata* protamine. The major series (**A**) of multicharged ions with 18–22 charges yields a mass of 13475.7 ± 1.9 Da. Minor series (**B**) of peaks yields a mass of 13573.3 ± 2.9 Da. The 98-Da difference can be attributed to phosphate adduct.

- The amino-terminal region (residues 1–27), in which alternating basic-serine residues are found. In this region (as in the rest of the molecule), arginine is the main basic residue and seven alternating R-S and one K-S are found among residues 10–24.
- The rest of the molecule (residues 28–106) in which most of the arginine residues are found in clusters. In the central region (residues 28–67), arginine clusters alternate with triplets of other residues. These triplets are SAS, TAS, SVS, and SRS. This central region does not contain any lysine residue. The COOH-terminal part of the molecule (residues 68–106) also has arginine clusters, but they alternate with heterogeneous groups of amino acid residues. The separation between clusters can be made by a single amino acid residue (T, A) or by heterogeneous groups of four, five, or eight residues, some of which contain isolated arginine and lysine residues. Consequently, lysine again appears in the COOH-terminal part of the protamine, but it never disrupts arginine clusters.

The basic charge (arginines + lysines) increases from 53.6% in the NH_2-terminal region to 71.8% in the COOH-terminal region. It should be noted that in the COOH-terminal region, basic amino acid residues are never found separated by more than two neutral residues. The relative amount in phosphorylatable residues (serines + threonines) decreases from the NH_2-terminal

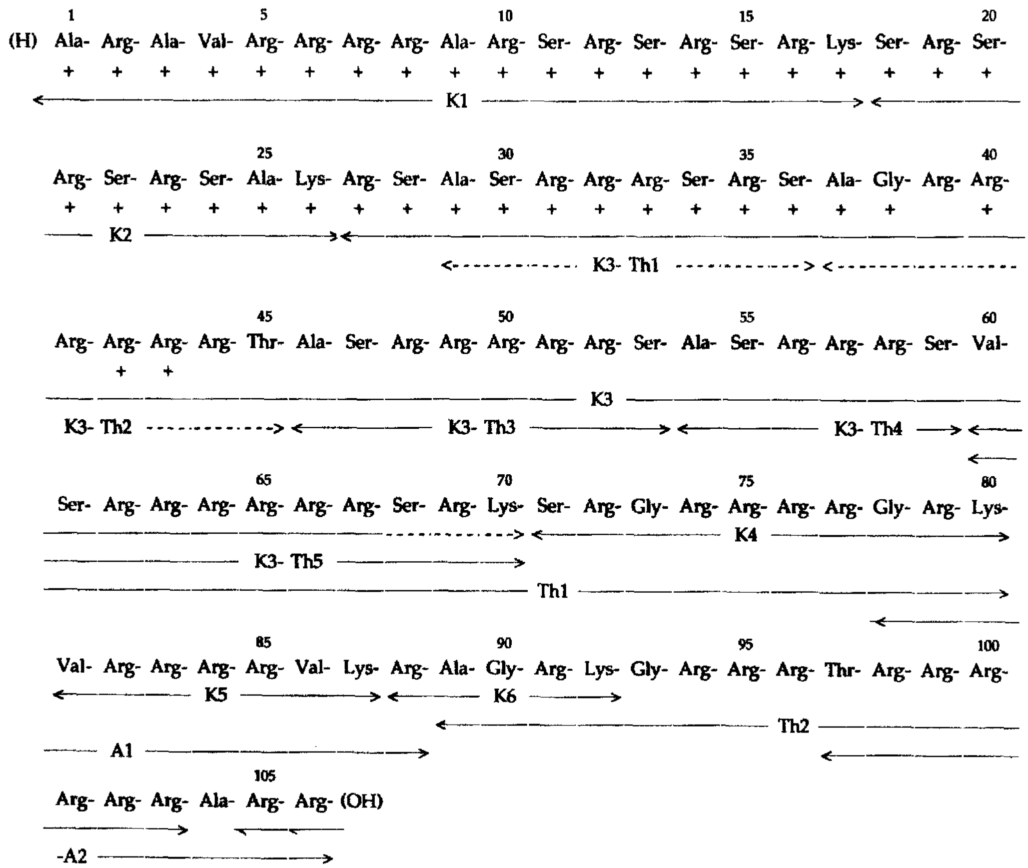

Fig. 3. Amino acid sequence of *Monodonta turbinata* protamine. Methods used for the determination of the sequence are indicated as follows: +++, automated Edman degradation of the protamine; *K-*, endoproteinase Lys-C peptides; *Th-*, thermolysin peptides; *A-*, peptides obtained by hydrolysis with *Astacus fluviatilis* proteinase; ⇌, carboxypeptidase B digestion. Peptides delimited with a *dotted line* were not sequenced.

Table 2. Amino acid composition of peptides generated by cleavage of protamine P2 and of peptide K3 with thermolysin (Th and K3-Th, respectively) and by cleavage of protamine P2 with *Astacus fluviatilis* endoproteinase (A)[a]

Amino acid	K3, 27–70 (mol/mol)	K3-Th1, 29–36	K3-Th2, 37–45	K3-Th3, 46–53	K3-Th4, 54–59	K3-Th5, 60–70	Th1, 60–80	Th2, 89–103	A1, 78–88	A2, 97–106
Thr	1.0 (1)		0.8 (1)					0.9 (1)		0.9 (1)
Ser	6.9 (10)	2.3 (3)		1.7 (2)	1.7 (2)	2.1 (2)	2.5 (3)			
Gly	1.2 (1)		1.0 (1)			0.7 (0)	1.7 (2)	1.6 (2)	1.2 (1)	
Ala	4.1 (4)	1.0 (1)	1.1 (1)	1.3 (1)	1.0 (1)			1.0 (1)		1.0 (1)
Val	1.1 (1)					1.0 (1)	0.9 (1)		2.0 (2)	
Lys	1.0 (1)					1.0 (1)	1.7 (2)	0.8 (1)	2.0 (2)	
Arg	28.2 (26)	4.0 (4)	5.9 (6)	5.0 (5)	3.3 (3)	6.7 (7)	13.7 (13)	10.7 (10)	5.9 (6)	8.0 (8)
Total residues	44	8	9	8	6	11	21	15	11	10
Mass[b]	5619.5									1439.7
Mass[c]	5620.3 ± 1.3	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	1439.7

[a] Values in parentheses are the number of residues/molecule of peptides derived from the sequence. n.d., not determined

[b] Calculated mass in Da

[c] Masses in Da, measured by electrospray mass spectrometry for K3, and by fast atom bombardment mass spectrometry for A2

and central regions (28.6%, 25%, respectively) to the COOH-terminal part (7.7%).

The reason for differentiating these two regions in the molecule also lies with the similarities found within the known sequences of sperm proteins from other species, as shown in Fig. 6. Some of these species are not evolutionarily related to the archaeogastropod mollusc *M. turbinata*. When we compare the sequence of this protamine with that of other molluscs, we find that the major protamine ø1 of the bivalve *M. edulis* (Ruiz-Lara et al.

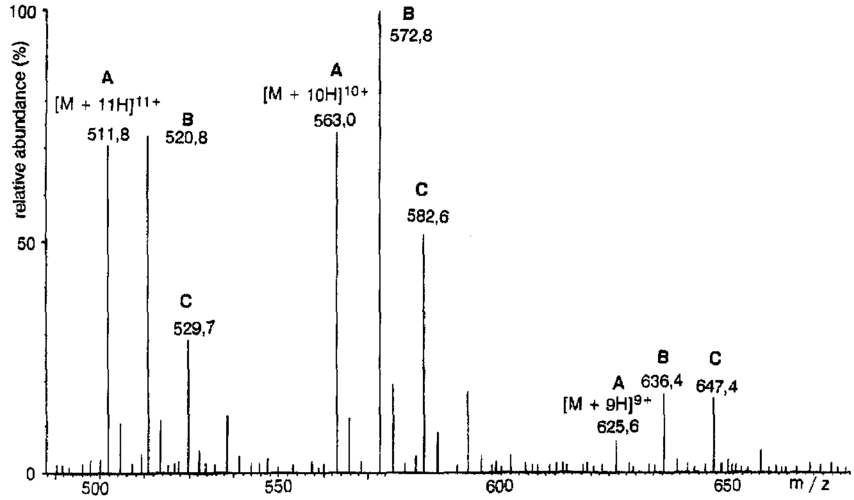

Fig. 4. Electrospray mass spectrum of peptide K3 derived from cleavage of *Monodonta turbinata* protamine with endoproteinase Lys-C. Three series (**A**, **B**, **C**) of multicharged ions with 9–11 charges were detected. The series **A** corresponds to the peptide K3 (5620.3 ± 1.3 Da). The series **B** and **C** correspond to peptide K3 with noncovalently bound phosphate adducts.

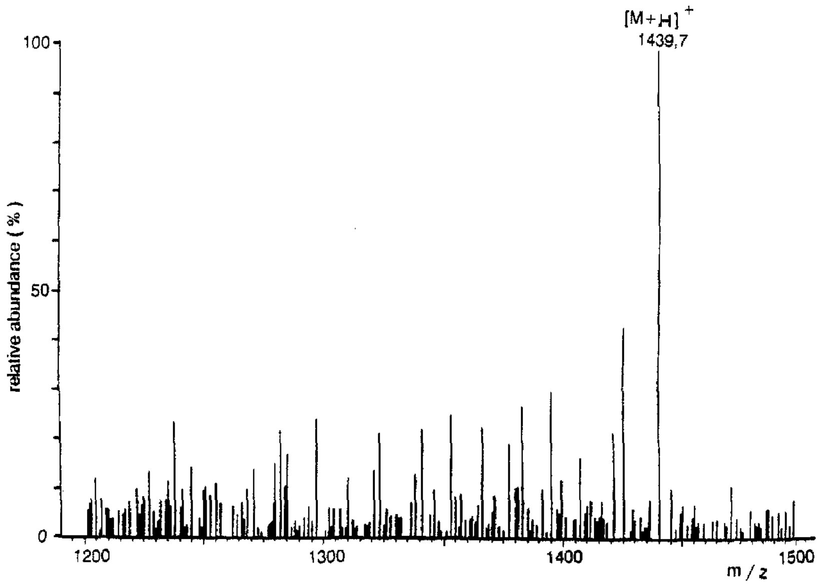

Fig. 5. Fast atom bombardment mass spectrum of peptide A2 derived from cleavage of *Monodonta turbinata* protamine with *Astacus fluviatilis* proteinase.

1993) also has an alternating region of basic-phosphorylatable residues (B-P), as shown in Fig. 6b. The COOH-terminal part of this protamine (residues 64–89) has clusters of basic amino acids, similar to those found in *Monodonta* (Fig. 6a). On the other hand, the central part of the protein (residues 21–63) is less basic and has a distribution of basic amino acids which is similar to that found in the COOH-terminal part of histone H1 (Subirana 1990). Another mollusc, the cephalopod *S. officinalis* (Martin-Ponthieu et al. 1991), has clusters of arginine throughout the whole protein, as shown in Fig. 6c, and lacks the alternating B-P region.

When we compare the mollusc protamines with those of vertebrates, it is found that protamines of birds (61 amino acid residues) and P1 protamines of mammals (50 residues approximately) also contain alternating RS (also KT or RT) in the NH_2-terminal region, as shown in Fig. 6 (d and e). Particularly in mammals (P1 protamines), the NH_2-terminal region including the alternating residues SRSR, is conserved in evolution, and it may have a functional importance during displacement of histones by protamine in spermiogenesis (Oliva and Dixon 1991, Retief et al. 1993). These B-P groups do not appear in protamines from amphibia or fish, except in dog-fish scylliorhinine Z3 (Kouach et al. 1993) (Fig. 6f and g), which are shorter molecules (less than 40 residues) than those found in birds and mammals. Protamines from birds and mammals have developed from genes related to those of bony fish protamines (reviewed in Oliva and Dixon 1991). This fact implies that the B-P groups in the NH_2-terminal part of these molecules have an evolutionary appearance independent from that of the same alternating B-P groups found in *M. turbinata* protamine.

The rest of the molecule (residues 28–106) has the same organization as that of typical protamines: clusters of arginine separated by small groups of other residues. From a chemical point of view (Subirana 1983), this region has the structure of a true protamine. In fact, it is the longest true protamine known at this time, and its primary structure allows us to establish a great number of

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

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

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

d)          10          *          30          *          50          * 61
ARYR RSRTRSRS PRS RRRRRR SG RRR SP RRRRR YGSA RR S RR SVGG RRRR YGS RRRRRRR Y

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
```

Fig. 6. Amino acid sequences belonging to some selected sperm-specific proteins: **a** *Monodonta turbinata* protamine (archaeogastropod mollusc) (this work); **b** protein ø1 from *Mytilus edulis* (bivalve mollusc) (Ruiz-Lara et al. 1993); **c** protamine from *Sepia officinalis* (cephalopod mollusc) (Martin-Ponthieu et al. 1991); **d** *Gallus domesticus* (bird) protamine (Oliva and Dixon 1991); **e** P1-protamine from mouse (*Mus sp*) (mammal) (Oliva and Dixon 1991); **f** protamine from *Bufo japonicus* (amphibia) (Takamune et al. 1991); **g** typical protamine (iridine 2b) from *Onchorhynchus mykiss* (bony fish) (Oliva and Dixon 1991).

amino acid identities with other protamines (for instance, the *Gallus domesticus* protamine is approximately 80% identical to *M. turbinata* when some gaps are introduced in their sequence).

The comparison of *M. turbinata* protamine sequence with other sperm proteins suggests the following conclusions:

1. 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.
2. This comparison shows that 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 COOH-terminal part of histone H1, found thus far only in the central region of *Mytilus* protamine.

Acknowledgments. The authors thank A. Hemez and M.J. Dupire for their technical assistance. This work was supported by grants from the Centre National de la Recherche Scientifique and from the Commission of the European Communities (grant CT91-0619) and in part by a DGICYT grant, PB90-0605.

References

Ausió 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

Ausió J, McParland R (1989) Sequence and characterization of the sperm-specific protein ø3 from *Mytilus californianus.* Eur J Biochem 182:569–576

Bloch DP (1969) A catalog of sperm histones. Genetics (Suppl) 61: 93–111

Carlos S, Jutglar L, Borrell I, Hunt DF, Ausió J (1993a) Sequence and characterization of a sperm-specific histone H1-like protein of *Mytilus californianus.* J Biol Chem 268:185–194

Carlos S, Hunt DF, Rocchini C, Arnott DP, Ausió J (1993b) Post-translational cleavage of a histone H1-like protein in the sperm of *Mytilus.* J Biol Chem 268:195–199

Chiva M, Kulak D, Kasinsky HE (1989) Sperm basic proteins in the turtle *Chrysemis picta:* characterization and evolutionary implications. J Exp Zool 249:329–333

Chiva M, Rosenberg E, Kasinsky HE (1990) Nuclear basic proteins in mature testis of the ascidian tunicate *Styela montereyensis.* J Exp Zool 253:7–19

Chiva M, Daban M, Rosenberg E, Kasinsky HE (1991) Protamines in polyplacophors and gastropods as a model for evolutionary changes in molluscan sperm basic proteins. In: Baccetti B (ed) Comparative spermatology 20 years after. Serono Symp Publications from Raven Press, vol 75, pp 27–30

Chiva M, Kulak D, Rosenberg E, Kasinsky HE (1992) A protamine in a crustacean *Balanus nubilus* (Cirripedia, Thoracica) and its coexistence with acidic proteins in sperm nuclei. Comp Biochem Physiol 102B:935–939

Colom J, Subirana JA (1981) Presence of H2B histone in spermatozoa from marine gastropods. Exp Cell Res 131:462–465

Daban M, Morriconi E, Kasinsky HE, Chiva M (1990) Characterization of the nuclear sperm basic proteins in one archaeogastropod: comparison of protamines between species. Comp Biochem Physiol 96B:123–127

Daban M, Kasinsky HE, Lafargue F, Chiva M (1991a) Nuclear sperm basic proteins (protamines) in chitons (Polyplacophora). Compositional and structural analogies with protamines of other molluscs. Comp Biochem Physiol 98B:437–443

Daban M, Chiva M, Rosenberg E, Kasinsky HE, Subirana JA (1991b) Protamines in prosobranchian gastropods (Mollusca) vary with different modes of reproduction. J Exp Zool 257:265–283

Giancotti V, Buratti E, Santucci A, Neri P, Crane-Robinson C (1992) Molluscan sperm proteins: *Ensis minor.* Biochim Biophys Acta 1119:296–302

Kasinsky HE, Huang SY, Mann M, Roca J, Subirana JA (1985) On the diversity of sperm histones in the vertebrates. IV. Cytochemical and amino acid analysis in Anura. J Exp Zool 234:33–46

Kasinsky HE (1989) Specificity and distribution of sperm basic proteins. In: Histones and other basic nuclear proteins. CRC Press, Boca Raton, FL

Kouach M, Jaquinod M, Belaïche D, Sautière P, van Dorsselaer A, Chevaillier P, Briand G (1993) A corrected primary structure for dog-fish *Scylliorhinus caniculus* protamine Z3. Biochim Biophys Acta 1162:99–104

Martin-Ponthieu A, Wouters-Tyrou D, Belaiche D, Sautiére P, Schindler P, Van Dorsselaer A (1991) Cuttlefish sperm protamines. 1. Amino acid sequences of two distinct variants. Eur J Biochem 195:611–619

Miescher F (1874) Das protamin, eine neue organische base aus den samenfades des rheinlachses. Ber 7:376–379

Mogensen C, Carlos S, Ausió J (1991) Microheterogeneity and interspecific variability of the nuclear sperm proteins from *Mytilus.* FEBS Lett 282:273–276

Oliva R, Dixon GH (1991) Vertebrate protamine genes and the histone-to-protamine replacement reaction. Prog Nucleic Acid Res Mol Biol 40:25–94

Panyim S, Chalkley R (1969) High resolution acrylamide gel electrophoresis of histones. Arch Biochem Biophys 130:337–346

Poccia D (1986) Remodeling of nucleoprotamines during gametogenesis, fertilization, and early development. Int Rev Cytol 105:1–65

Retief JD, Dixon GH (1993) Evolution of pro-protamine P2 genes in primates. Eur J Biochem 214:609–615

Retief JD, Winkfein RJ, Dixon GH, Adroer R, Queralt R, Ballabriga J, Oliva R (1993) Evolution of protamine P1 genes in primates. J Mol Evol 37:426–434

Ruiz-Lara S, Prats E, Casas MT, Cornudella L (1993) Molecular cloning and sequence of a cDNA for the sperm-specific protein ø1 from the mussel *Mytilus edulis.* Nucleic Acids Res 21:2774

Saperas N, Chiva M, Ausió J (1992) Purification and characterization of the protamines and related proteins from the sperm of a tunicate *Styela plicata.* Comp Biochem Physiol 103B:969–974

Sautière P, Martinage A, Belaiche D, Arkhis A, Chevaillier P (1988) Comparison of the amino acid sequences of human protamines HP2 and HP3 and of intermediate basic nuclear proteins HPS1 and HPS2. Structural evidence that HPS1 and HPS2 are pro-protamines. J Biol Chem 263:11059–11062

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

Subirana JA (1983) Nuclear proteins in spermatozoa and their interactions with DNA. In: Andre J (ed) The sperm cell. Martinus Nijhoff, The Hague

Subirana JA (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

Subirana JA, Cozcolluela C, Palau J, Unzeta M (1973) Protamines and other basic proteins from spermatozoa of molluscs. Biochim Biophys Acta 317:364–379

Takamune K, Nishida H, Takai M, Katagiri C (1991) Primary structure of toal sperm protamines and nucleotide sequence of their cDNAs. Eur J Biochem 196:401–406

Wouters-Tyrou D, Chartier-Harlin MC, Martin-Ponthieu A, Boutillon C, Van Dorsselaer A, Sautière P (1991) Cuttlefish spermatid-specific protein T. Molecular characterization of two variants T1 and T2, putative precursors of sperm protamine variants Sp1 and Sp2. J Biol Chem 266:17388–17395