the FEBS Journal

A unique vertebrate histone H1-related protamine-like protein results in an unusual sperm chromatin organization

Núria Saperas[1], Manel Chiva[2], M. Teresa Casas[1], J. Lourdes Campos[1], José M. Eirín-López[3,4], Lindsay J. Frehlick[4], Cèlia Prieto[1], Juan A. Subirana[1] and Juan Ausió[4]

1 Departament d'Enginyeria Química, ETSEIB, Universitat Politècnica de Catalunya, Barcelona, Spain
2 Departament de Ciències Fisiològiques II, Facultat de Medicina, Universitat de Barcelona, L'Hospitalet de Llobregat, Spain
3 Departamento de Biología Celular y Molecular, Universidade da Coruña, Spain
4 Department of Biochemistry and Microbiology, University of Victoria, British Columbia, Canada

Keywords
chromatin; electron microscopy; histone H1; protamine-like protein; X-ray diffraction

Correspondence
J. Ausió, Department of Biochemistry and Microbiology, University of Victoria, Petch Building, Room 220, Victoria, BC, Canada V8W 3P6
Fax: +1 250 721 8855
Tel: +1 250 721 8863
E-mail: jausio@uvic.ca

(Received 13 July 2006, revised 10 August 2006, accepted 11 August 2006)

doi:10.1111/j.1742-4658.2006.05461.x

Protamine-like proteins constitute a group of sperm nuclear basic proteins that have been shown to be related to somatic linker histones (histone H1 family). Like protamines, they usually replace the chromatin somatic histone complement during spermiogenesis; hence their name. Several of these proteins have been characterized to date in invertebrate organisms, but information about their occurrence and characterization in vertebrates is still lacking. In this sense, the genus *Mullus* is unique, as it is the only known vertebrate that has its sperm chromatin organized by virtually only protamine-like proteins. We show that the sperm chromatin of this organism is organized by two type I protamine-like proteins (PL-I), and we characterize the major protamine-like component of the fish *Mullus surmuletus* (striped red mullet). The native chromatin structure resulting from the association of these proteins with DNA was studied by micrococcal nuclease digestion as well as electron microscopy and X-ray diffraction. It is shown that the PL-I proteins organize chromatin in parallel DNA bundles of different thickness in a quite distinct arrangement that is reminiscent of the chromatin organization of those organisms that contain protamines (but not histones) in their sperm.

During spermatogenesis, chromatin undergoes one of the most dramatic rearrangement transitions involving chromatin remodeling in the eukaryotic cell. In this process, the somatic histones from the stem cells are replaced by highly specialized sperm nuclear basic proteins (SNBPs), which not only alter the tertiary structure of chromatin, but also remove the somatic histone epigenetic component resulting from the histone post-translational modifications [1] and histone variants [2,3].

SNBPs can be divided into three major groups or types: protamine (P) type, histone (H) type and protamine-like (PL) type [4,5]. The P type consists of usually small proteins ($4000 \leq M_r \leq 10\,000$) that are very rich in arginine and/or cysteine [6]. These proteins are widely distributed. Representative examples can be found in both vertebrate [7,8] and invertebrate organisms [9,10], where they replace the somatic histones of the stem cells during spermiogenesis [8]. The H type comprises a highly evolutionarily conserved group of chromosomal proteins that are closely linked to the main histone constituents of somatic chromatin [11]. These proteins are present in the chromatin of the mature sperm of invertebrate and vertebrate organisms

Abbreviations
AU, acetic acid/urea; AUT, acetic acid/urea/Triton; BS, bootstrap; CP, confidence probability; E, enzyme; EM, electron microscopy; H, histone; IBT, interior branch test; IDP, intrinsically disordered protein; P, protamine; PCA, perchloric acid; PL, protamine-like; RD, replication dependent; RI, replication independent; S, substrate; SNBP, sperm nuclear basic protein.

 © 2006 The Authors Journal compilation © 2006 FEBS

that replace the somatic histones of the stem cells with a somatic-like histone complement by the end of spermiogenesis. Often, they consist of highly differentiated sperm-specific variants, such as in echinoderms [12,13] and other invertebrates [14,15], although they may in other instances retain compositional identity with the somatic counterpart [16]. The PL type is an intermediate type between histones and protamines, and was originally described and characterized in bivalve molluscs [17,18]. It consists of a highly heterogeneous group of SNBPs that are rich in both arginine and lysine and are phylogenetically related to somatic-type linker histones (histone H1 family) [19]. Like the P type, these SNBPs can replace the stem histones to different extents during spermiogenesis [20].

PL-I proteins are histone H1-related PL proteins that have now been quite extensively described in invertebrate organisms from molluscs [17,18,21] to tunicates [22]. In vertebrate organisms, SNBPs of the PL-I type have been described in fish [23] and amphibians [4], but they have not been studied in much detail. In this group of organisms, the sperm chromatin organization resulting from SNBPs of the H type [16,24] and the P type (which prevails in reptiles, birds and mammals) [25,26] have been quite well characterized. However, very little information is available to date on the PL protein-mediated organization of chromatin, and what little information is available has come mainly from invertebrate organisms [20,27]. In this sense, the red mullet (*Mullus*) is the only vertebrate described to date with PL proteins as the only SNBPs organizing its sperm nuclei chromatin. It is therefore interesting to analyze the chromatin structure resulting from the association of these proteins with DNA.

In this work, we show that SNBPs from *Mullus surmuletus* consist of two compositionally related PL-I proteins. We characterize the main PL-I protein component and analyze the chromatin structure resulting from the association of these two proteins with DNA using several biochemical (micrococcal nuclease digestion), structural [electron microscopy (EM) and X-ray diffraction] and phylogenetic approaches.

Results

Characterization of the SNBPs of the fish *M. surmuletus*

Compositional SNBP analysis of the striped red mullet *M. surmuletus* and the red mullet *Mullus barbatus* has already shown that the 0.4 M HCl extracts from the mature sperm nuclei of these species consist mainly of a protein doublet that migrates in the core histone region in acetic acid–urea (AU)/PAGE (Fig. 1A) [28]. As seen in Fig. 1A and in Fig. 1B, which shows acetic acid–urea–Triton (AUT)/PAGE of the same extract, there is no histone complement in the SNBP composition of *M. surmuletus*. The M_r values for the two proteins obtained from a mixture of both of them were 20 300 and 20 354 (Fig. 1C).

By combining cation exchange chromatography (CM-Sephadex C-25) and RP-HPLC, the two major SNBP bands of *M. surmuletus* could be purified to complete purity (results not shown). Table 1 shows the amino acid compositions of both fractions, which are almost identical. A comparative compositional analysis with histone H1 as well as with other PL-I proteins already allows the identification of these two proteins as putative members of the PL-I family of proteins [4,5,29].

The availability of pure fractions allowed us to proceed with the sequencing of these proteins. However, attempts to obtain any sequences from the N-terminal end of the slowly moving band proved completely unsuccessful. The same difficulty was found for the slow band of the closely related species *Mullus barbatus*. This suggested that this band may be acetylated at its N-terminal end. Indeed, N-terminal α-amino acetylation of histone H1 was early ascribed to the difficulty encountered in sequencing some H1 histones by Edman degradation [30], and the occurrence of α-amino acetylation in some histone H1s has been recently confirmed by proteomic analysis [31]. Therefore, we decided to focus most of our sequencing efforts on the main SNBP component, i.e. the fast-moving fraction. Figure 1D shows a summary of the sequencing results obtained.

The N-terminal domain, obtained by direct sequencing of the whole protein, was found to contain a 20 amino acid region highly enriched in basic amino acid residues, with a cluster of five phosphorylatable residues at the beginning of the molecule. This region also contains two SPBB (where B = K or R) motifs. These SPBB motifs are like the ones that are present at the N-terminal and C-terminal tails of the unusually large sperm-specific H1 and H2B histones that are found in echinoderms [12] and are also present in other somatic histones [32]. Sequencing of the five N-terminal residues of the other mullet species, *M. barbatus*, has shown that they are identical in both species.

The C-terminal region is extremely repetitive. Our inability to completely sequence the whole protein using Edman degradation simply reflects the problems arising from this repetition. Similar problems were

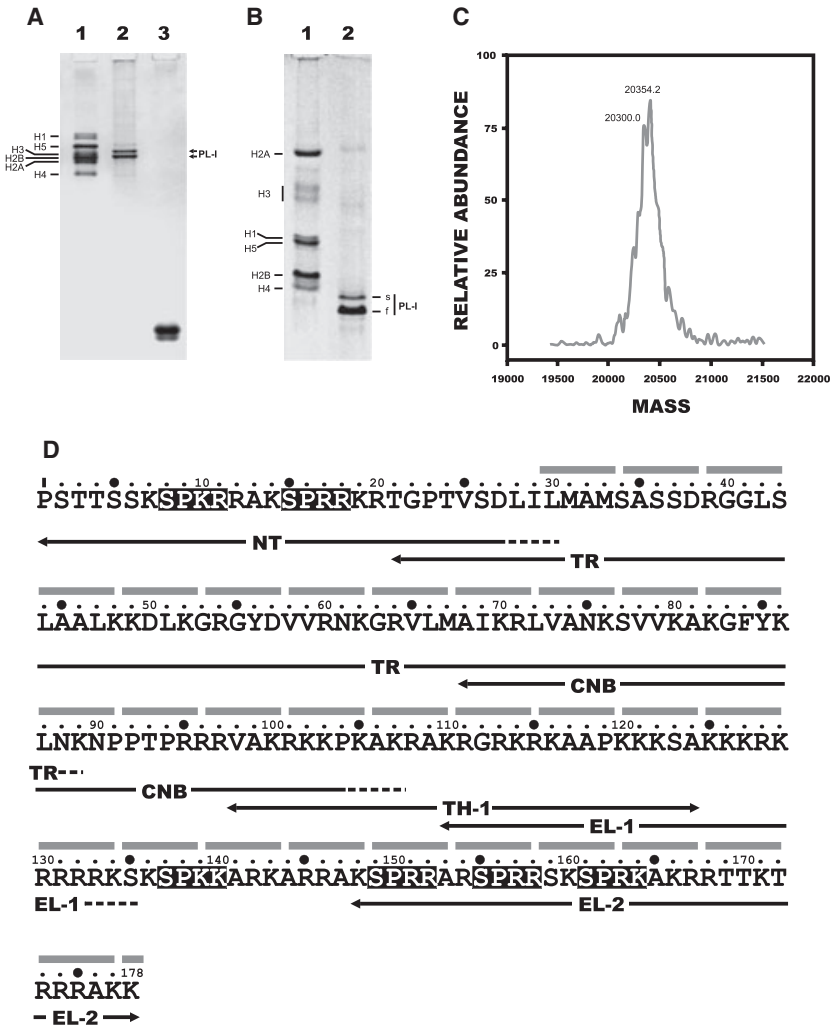

Fig. 1. Characterization of the *M. surmuletus* sperm nuclear basic proteins (SNBPs). (A) Acetic acid–urea (AU)/PAGE analysis of *M. surmuletus* SNBPs. Lane 1: Chicken erythrocyte histones used as a histone marker. Lane 2: *M. surmuletus* SNBPs. Lane 3: Protamine salmine from *Oncorhynchus keta* (chum salmon) used as a protamine marker. (B) Acetic acid–urea–Triton (AUT)/PAGE analysis of *M. surmuletus* SNBPs. Lane 1: Chicken erythrocyte histones used as a histone marker. Lane 2: *M. surmuletus* SNBPs. s and f denote the slow-moving and fast-moving PL-I protein bands, respectively. (C) MS analysis of *M. surmuletus* SNBPs carried out by MALDI-TOF on a Voyager Linear DE using a sinapinic acid matrix. (D) Primary structure of the main PL-I protein [fast electrophoretic 'f' component in (A) and (B), lane 2] determined by Edman degradation from overlapping peptides obtained by digestion with different proteolytic enzymes, chemical cleavage and 3′RACE PCR (accession number P84802). NT, information obtained by direct sequencing from the N-terminus; TR, trypsin-resistant peptide; CNBr, cyanogen bromide cleavage peptide; TH-1, thermolysin peptide; EL, elastase peptides. The gray dashed line indicates the amino acid sequence obtained from 3′RACE analysis. The SPBB motifs where B = K or R are highlighted in gray.

encountered in the past when attempting to use the Edman degradation approach to sequence other PL-I proteins. Therefore, 3′RACE PCR was performed, and the translated sequence was in perfect agreement with the sequence already obtained by Edman degradation and completed the missing section of the sequence (Fig. 1D). Four additional SPBB motifs are found in the C-terminal region.

The SNBPs of *M. surmuletus* are histone H1-related proteins of the PL-I type

In addition to the amino acid compositional similarity of the *Mullus* SNBPs and PL-I proteins (Table 1), further evidence of their true PL-I nature (and hence their relation to histone H1) was revealed by the presence of a trypsin-resistant core (Fig. 2A). The presence of a globular trypsin-resistant core is one of the characteristic features that distinguishes histone H1 from core histones.

The availability of the sequence of this core domain allowed us to take the crystallographic data available for the trypsin-resistant globular part of chicken erythrocyte histone H5 [33] and use it as a template to model the tertiary structure corresponding to this sequence (Fig. 2B). As can be seen in Fig. 2B, this domain still maintains the characteristic winged-helix domain [33].

Alignment of the sequences corresponding to the trypsin-resistant core of the consensus sequence for this domain in vertebrate replication-dependent (RD) and replication-independent (RI) histone H1s (H1/H5) (Fig. 2C) showed that the sequence of the globular domain of *M. surmuletus* PL-I protein is indeed more related to the vertebrate RD line than to the RI line. This was somewhat surprising, as in invertebrate PL-I

Table 1. Amino acid composition (mol%) of PL-I proteins from different organisms in comparison to calf thymus histone H1. C.i., *Ciona intestinalis* (tunicate); C.t., calf thymus; M.s., *M. surmuletus* (red mullet); M.t., *Mytilus trossulus* (mussel); P.a., *Pseudopleuronectes americanus* (winter flounder); S.m., *Styela montereyensis* (tunicate); S.s., *Spisula solidissima* (surf clam); tr., trace amounts.

Amino acid	C.t. [75] H1	S.s. [21] PL-I	M.t. [76] PL-I (PLII + PLIV)[a]	S.m. [22] PL-I	C.i. [22] PL-I	P.a. [52] PL-I	M.s. (slow band)	M.s. (fast band)
Lys	26.8	23.8	28.7	17.2	29.4	15.1	21.7	22.8
His	–	0.4	0.5	1.1	1.6	0.8	–	–
Arg	1.8	22.7	7.4	30.1	18.7	16.6	20.3	21.1
Asx	2.5	0.4	3.5	4.3	3.2	3.4	3.7	4.1
Thr	5.6	4.0	4.0	5.4	2.7	5.7	4.0	4.1
Ser	5.6	22.7	15.3	5.4	9.1	18.1	9.8	8.8
Glx	3.7	0.4	0.5	1.6	1.6	3.4	tr.	tr.
Pro	9.2	2.2	7.9	3.2	2.1	8.3	7.5	7.8
Gly	7.2	2.4	5.4	10.8	11.2	2.6	5.8	5.6
Ala	24.3	14.1	14.9	6.5	6.4	7.5	11.6	11.4
Cys	–	0.2	–	1.1	1.1	0.4	–	–
Val	5.4	2.4	3.5	3.2	3.7	4.9	4.6	4.2
Met	–	1.3	1.5	1.6	1.1	4.5	1.7	1.2
Ile	1.5	0.4	2.5	2.7	3.2	1.5	1.2	1.2
Leu	4.5	1.5	3.0	3.2	2.7	5.7	5.0	5.3
Tyr	0.9	0.2	0.5	2.2	1.6	0.4	1.3	1.3
Phe	0.9	0.2	1.0	0.5	0.5	1.1	1.3	tr.
Trp	–	0.2	–	–	–	–	–	–

[a] The *Mytilus trossulus* PL-I gene expresses two proteins, PL-II and PL-IV, as a result of post-translational cleavage [76].

proteins this domain has been shown to be much closer in sequence to the RI line [5,34].

The phylogenetic relationships inferred from the complete protein sequences of several RD/RI H1 histones and histone H1-related SNBPs (Fig. 3) show that, indeed, *Mullus* PL-I protein clusters with invertebrate and chordate proteins of the PL type, in close proximity to the vertebrate spermatogenic transition proteins [35], a group of proteins with which PL proteins may have had an ontogenic relationship [4].

The chromatin of the sperm of *M. surmuletus* is organized in bundles of parallel DNA molecules

Mullus surmuletus sperm chromatin was digested with micrococcal nuclease and treated afterwards as described in Experimental procedures to obtain fractions SI, SII and PII. Figure 4 shows the results of the micrococcal nuclease digestion pattern of this unusual chromatin. After 45 min of digestion, the values for the soluble DNA in the SI fraction are almost identical to those for soluble DNA determined by perchloric acid (PCA) solubilization [36] (results not shown). This indicates that most of the absorbance at 260 nm (DNA released) beyond this point is due to small oligonucleotides that do not appear in Fig. 4C (our results do not exclude the possibility that the increase in absorbance is not also contributed to by nuclear RNA). Continuous degradation of DNA in this way results in the displacement of PL-I proteins to adjacent chromatin regions, contributing to the overall insolubility of fraction SII, whose solubility does not increase beyond 5%. This value most likely reflects the overall amount of nucleosomally organized chromatin, which eventually could come from some contaminating spermatid material. Indeed, no PL-I proteins are released in fraction SI, despite the increasing amounts of DNA being solubilized as the time of digestion increases (Fig. 4B). As can be seen in Fig. 4B, the proteins associated with this fraction contain only small amounts of histones and consist of a nuclease-resistant fragment of approximately the size (146 bp) that would correspond to the DNA protected from digestion in a nucleosome core particle (Fig. 4C). In contrast, PL-I protein accumulates in the SII and P fractions, where it is the only SNBP present (Fig. 4B). The DNA composition appears as a smear whose size distribution decreases with the digestion time. A broad diffuse band centered at about 100 bp is also seen, which may correspond to fragments protected from digestion by the interaction of individual PL-I protein molecules with DNA. These results are very reminiscent of those obtained by Young and Sweeney using SDS/dithiothreitol-decondensed sperm chromatin from rabbits and fowl [37]. They are also very similar to the DNase I digestion patterns obtained with human sperm chromatin [38] and those observed in invertebrate sperm chromatin consisting of PL-I proteins [20,39]. Nevertheless, a discrete repetitive nuclease digestion

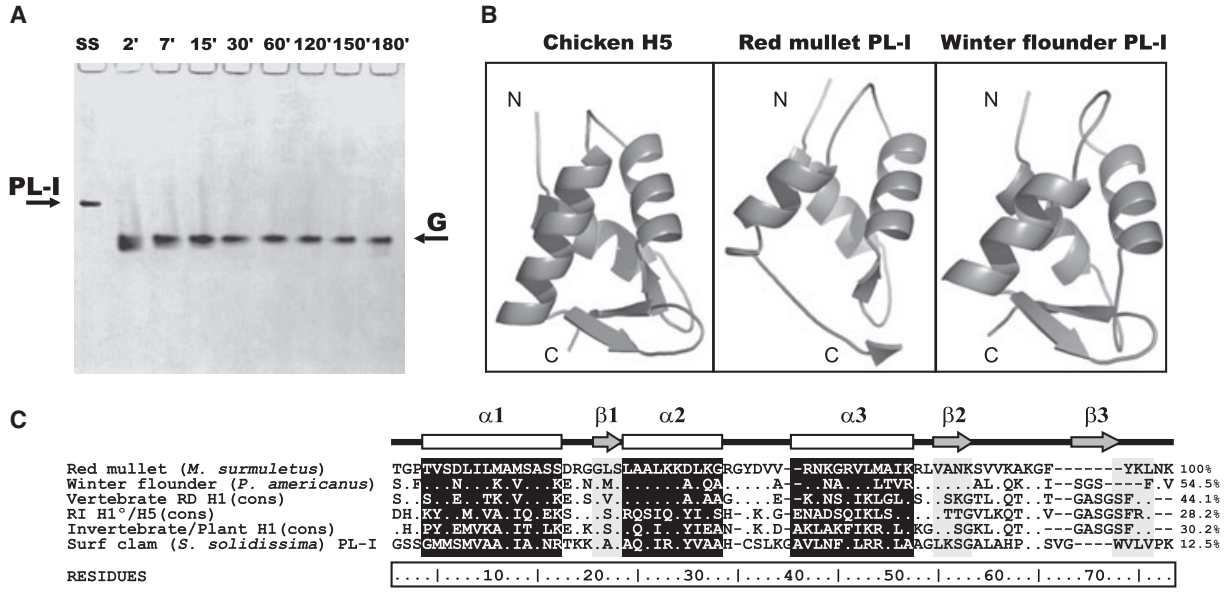

Fig. 2. The main sperm nuclear basic protein (SNBP) component of *M. surmuletus* contains a trypsin-resistant winged helix motif. (A) Acetic acid–urea (AU)/PAGE analysis of the time course of trypsin digestion of the major PL-I SNBP of *M. surmuletus* carried out in the presence of 2 M NaCl. The protein was digested at room temperature. SS denotes the starting sample before digestion; G indicates the resistant globular core of the protein. The digestion times (2, 7, 15, 30, 60, 120, 150 and 180 min) are indicated on top of the lanes. (B) Tertiary structure of the globular core of chicken erythrocyte histone H5 obtained from the coordinates determined in [33]. The structure was subsequently used as a template to model the three-dimensional structures of the globular part of the red mullet (*M. surmuletus*) and the winter flounder (*Pseudopleuronectes americanus*) PL-I proteins using the SWISS-MODEL server [69]. (C) Sequence alignment of the amino acid region corresponding to the globular domain of *M. surmuletus* PL-I protein in comparison to the corresponding domain in the *P. americanus* [52], vertebrate (replication-dependent) histone H1 consensus sequence [70,71], replication-independent histone H1/H5 consensus and invertebrate/plant consensus sequence [70,71] and *Spisula solidissima* (surf clam) PL-I [21]. The dots indicate identical amino acids and the dashes indicate deletions. The degree of homology (%) is indicated on the right. The GenBank accession numbers for the sequences are: *P. americanus*, AAC13878; *S. solidissima*, AY626224. In the schematic secondary structure assignment shown above, β-turns and strands are indicated by arrows, and α-helices are indicated by open boxes.

pattern such as that corresponding to nucleosome-organized chromatin in somatic tissues [40] or sperm with SNBPs of the H type [16] is not observed.

EM analysis of chromatin spreads (Fig. 5A) showed that *Mullus* sperm chromatin is organized in bundles of somewhat variable diameter with an average diameter of 650 Å upon correction for the increase in thickness resulting from the platinum coating (Fig. 5A,C). This value is slightly higher than that for the bundles observed (250–500 Å) with invertebrate organisms consisting of related types of PL protein [27]. However, it is very similar to that of the fibrogranular 500 ± 100 Å structures previously observed during the spermiogenesis of *M. surmuletus* [29]. The fiber organization seen in Fig. 5A looks very similar to the toroidal structures obtained from complexes reconstituted with DNA and histone H1 and/or H1-related proteins [41–43]. Similar structures have been described in mammalian sperm, where chromatin consists exclusively of proteins of the P type [25,44].

The wide-angle X-ray diffraction pattern obtained with chromatin fibers pulled from *M. surmuletus* lysed sperm nuclei (Fig. 5B) shows a disoriented reflection at 3.3 Å and an equatorial reflection at 23.54 Å. The first ring corresponds to the distance between base pairs, indicating that the B-conformation (nominal base pair distance 3.4 Å) of DNA is not altered by its interaction with PL-I protein. The equatorial spacing at 23.54 Å corresponds to a distance of 27.2 Å between parallel DNA molecules organized in pseudohelicoidal bundles [39]. This dimension is also in very good agreement with the (23 ± 5 Å) cross-section of the fibers observed in mature spermatozoa of *M. surmuletus* upon coalescence of the 500 bundles [29]. None of the reflections characteristic of the low-angle X-ray diffraction patterns from fibers obtained from nucleosome-organized chromatin were observed [45,46]. The X-ray diffraction pattern of the *M. surmuletus* sperm chromatin is similar to that found in complexes of DNA with calf thymus histone H1 [47], although the latter show a slightly better orientation. It also bears a

A

```
Red mullet       ----------------------------------------------------------------------------------------------------
Winter flounder  MSNCFSIHLTESEKPSSVILRHSMSQPKSSSRSRATMRSKSPKRTVKTLKTRAKSTRRSKSPMRSRSPMTSKSRKRSRSLSRSKSPKRRV

Red mullet       ----------------------------------------------------PSTTSSKSPKRRAKS-------------------------
Winter flounder  KTLMTRAKSPGRSQSPMTSRSPRRSQSPKRRVKTPKMRAKSPGRSKLPMRSR.PRR....R..V.TPKMRAKSPKMRAKSPKTRANSPKR

Red mullet       ---PRRKRTGPTVSDLILMAMSASSDRGGLSLAALKKDLKGRGYDVVRNKGRVLMAIKRLVANKSVVKAKGF---YKLNKNPPTPRRRVA
Winter flounder  RAK.Q..KS.F...N...K.V...KERN.M.......A.QA.....A..NA...LTVR......AL.QK..ISGSF.V..KA.A.-----

Red mullet       KRKKPKAKRAKRGRKRKAAPKKKSAKKKRKRRRRKSKSPKKARKARRAKSPRRARSPRRSKSPRKAKRRTTKTRRRAKK
Winter flounder  ----------------------------------------------------------------------------------
```

B

Fig. 3. The *M. surmuletus* sperm nuclear basic protein (SNBP) is phylogenetically related to histone H1, and the globular domain shows a close sequence relationship with that of the winter flounder (*Pseudopleuronectes americanus*) histone H1-related SNBP. (A) Sequence alignment of the full amino acid sequences of *M. surmuletus* PL-I proteins in comparison to those of *P. americanus* [52]. The dots indicate identical amino acids and the dashes indicate deletions. The GenBank accession number for the *P. americanus* sequence is AAC13878. (B) Phylogenetic neighbor-joining tree [67] showing the evolutionary relationships between histone H1 and protamine-like proteins throughout the metazoans, reconstructed from the alignment of the corresponding 127 amino acid sequences (supplementary Table S1) using uncorrected *p*-distances with the complete-deletion option. The reliability of the groups defined by the topology was tested by the bootstrap (BS) and the interior branch test (IBT) methods, based on 1000 replications, and is only shown in the corresponding interior branches when the value is greater than 50% [72,73]. The tree was rooted with the H1-like protein from the protist *Entamoeba*, because it represents one of the most ancestral eukaryotes in which an H1 protein has been described [74]. The position of the replication-independent somatic H1 histones is indicated in the right margin of the tree and the group including the protamine-like proteins and the germinal H1 histones is highlighted in gray in the tree, where the protamine-like protein from *M. surmuletus* is indicated in bold.

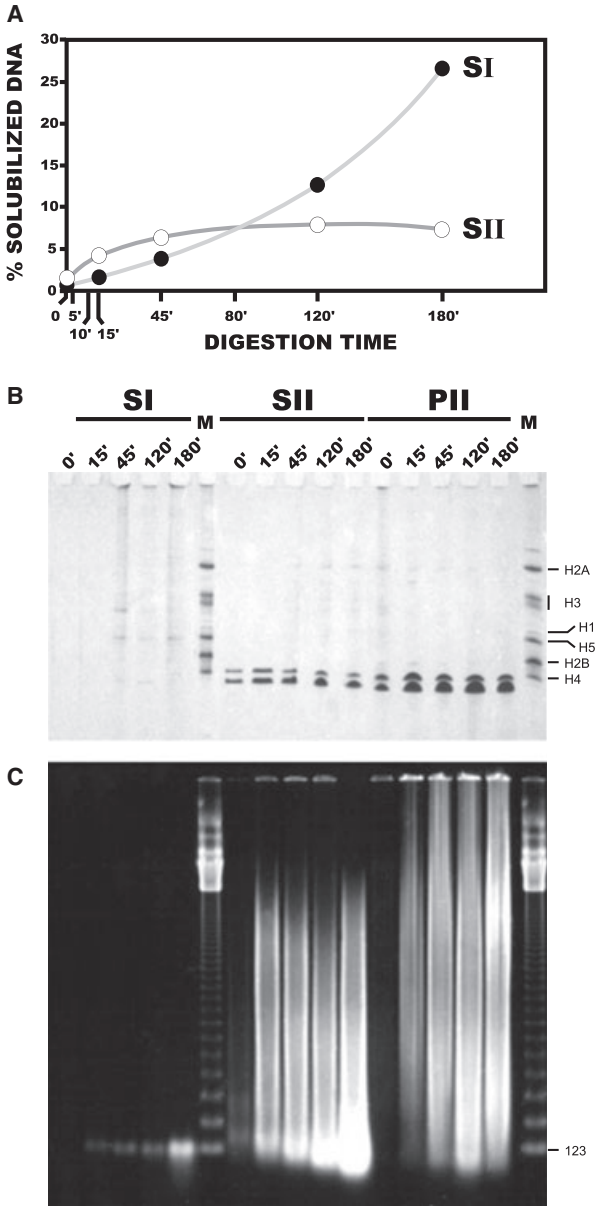

Fig. 4. Micrococcal nuclease digestion of *M. surmuletus* sperm chromatin. (A) Relative percentage of the solubilized DNA (either free or associated with proteins) present in the SI and SII fractions as a function of the micrococcal nuclease digestion time; 100% refers to the total initial amount of DNA. (B) Acetic acid–urea–Triton (AUT)/PAGE characterization of the sperm nuclear basic proteins (SNBPs) associated with the chromatin fragments released from *M. surmuletus* sperm chromatin upon digestion with micrococcal nuclease. (C) Agarose (1%) electrophoretic analysis of the DNA fragments associated with the chromatin fragments released from *M. surmuletus* sperm chromatin upon digestion with micrococcal nuclease. The supernatant SI/SII and pellet PII fractions were obtained as described in Experimental procedures. The times of digestion at 37 °C (0, 15, 45, 120 and 180 min) are indicated on top of the lanes. M, marker [chicken erythrocyte histones in (B) and 123 bp ladder in (C)].

strong resemblance to that obtained in a similar fashion in the blue mussel *Mytilus edulis* [39].

The SNBPs of *My. edulis* consist of a mixture of PL proteins, including a PL-I protein that coexists with a reduced amount of other somatic-like histones. In both instances (*My. edulis* and *M. surmuletus*), the sperm DNA molecules appear to be organized in a parallel fashion in bundles that most likely correspond to those visualized by EM (Fig. 5A). The main difference between the two organisms stems from the variation in the distance between the parallel DNA molecules in these bundles: 27.2 Å in *M. surmuletus* compared to 29.3 Å in the mussel *My. edulis*. Part of this difference may be accounted for by the presence of additional SNBPs in the latter case [39]. Interestingly, the value of 23.54 Å for the spacing is very close to that of closely packed DNA molecules in the B-form, suggesting that the DNA molecules in the sperm chromatin bundles of *M. surmuletus* are closely packed. The X-ray diffraction pattern of Fig. 5B is very different from that of the semicrystalline DNA organization observed in the sperm of certain organisms with P-type SNBPs [48] and from that of the nucleosome DNA organization described in organisms with SNBPs of the H type [16]. No indication of nucleosome organization could be seen with this technique, in agreement with the nuclease digestion and the EM studies.

Discussion

Several sources of evidence detailed in the previous section indicate that the major SNBP components of *M. surmuletus* are related to histone H1 and phylogenetically cluster with members of the PL-I protein subfamily. Whereas the RI/RD identity of the members of the histone H1 family appears to be given by the specific sequence of their characteristic winged helix domain (Fig. 2), most of their tissue specificity resides in the intrinsically disordered domains of their N- and C-terminal tails. In the case of the PL-I proteins, there is a substantial increase of arginine within these regions (see Fig. 1D) when compared to their somatic histone H1 counterparts (Table 1). Histone H1 has been shown to bind to chromatin in a very dynamic way, with a residence time that largely depends on the structural features of the C-terminal domain [49]. Arginine can bind more effectively to DNA and most likely increases the residence time of PL-I proteins binding to DNA, making them suitable for sperm chromatin condensation.

As with histone H1, an important part of the PL-I protein molecule is intrinsically disordered. Indeed, in certain instances, such as in *Spisula solidissima* [21] and

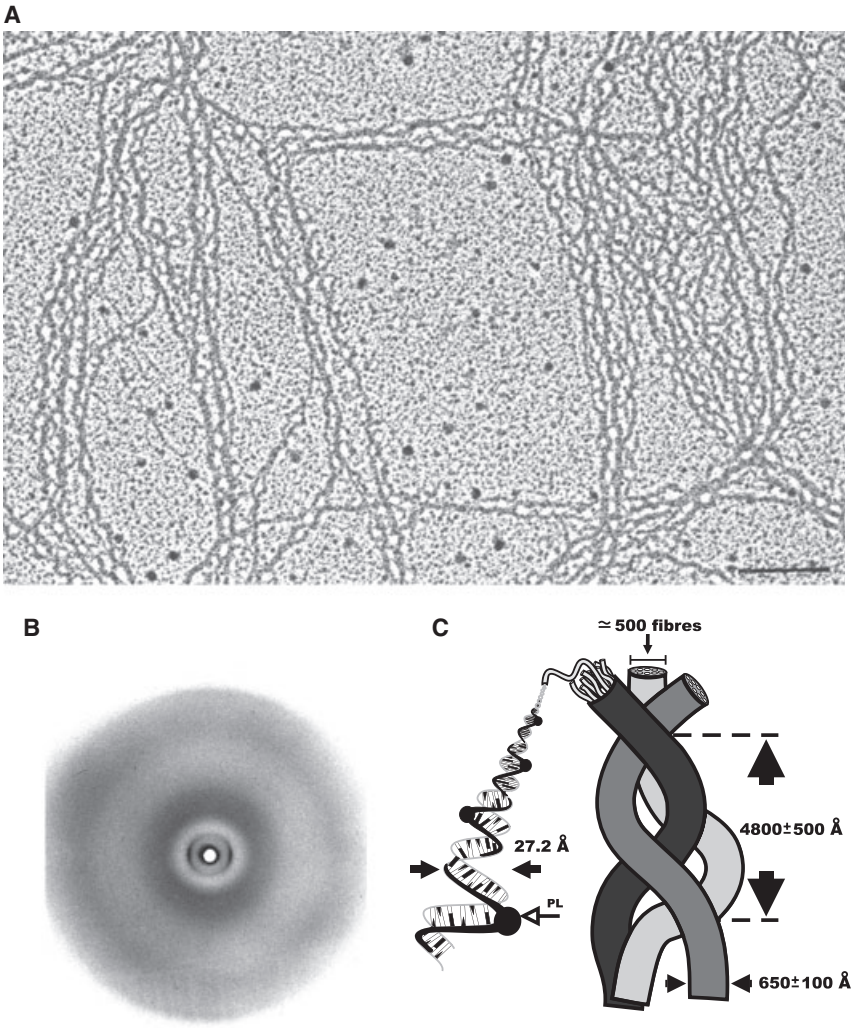

Fig. 5. The sperm chromatin of *M. surmuletus* consists of bundles of parallel DNA molecules. (A) Electron microscopy of spread sperm chromatin rotary shadowed with platinum. The bar is 2 μm. (B) Wide-angle X-ray diffraction pattern of fibers obtained from lysed *M. surmuletus* sperm nuclei. The relative humidity of the sample was 93%. A sharp equatorial reflection at 23.54 Å can be seen at the center. (C) Schematic representation of the sperm chromatin organization based on the information obtained from (A) and (B). As seen in (A), DNA bundles of approximately 650 ± 100 Å appear to be intertwined (giving the appearance of bubbles when spread flat in the carbon grid). From the packing distance of DNA fibers (27.2 Å) (B), it is possible to calculate the approximate number of DNA fibers per bundle as about 500.

in *Pseudopleuronectes americanus* [50], the intrinsically disordered domains of these molecules are much longer than those of the corresponding N-terminal and C-terminal domains of canonical H1 histones. As stated by Hansen *et al.* [51], there are many structural features, such as lower binding energy, greater flexibility and faster binding, that favor the selection of intrinsically disordered proteins (IDPs) over highly folded proteins, especially when the specificity of binding to DNA is low, such as in the case of the P-type and PL SNBPs. Thus, it is not surprising that these types of protein have been selected by evolution [19] over proteins with a lower entropy of folding (e.g. core histones) to tightly pack the DNA in the sperm chromatin. In this sense, it appears that histone H1, the least conserved of histones, has been repeatedly and independently used as the source of new SNBP models. The higher DNA-binding affinity of the SNBPs of the P and PL types would explain why they are often found in organisms at the tips of the phylogenetic tree [4].

At the chromatin level, this work represents the first time in which a histone H1-related PL-I protein replacing the somatic histone complement has been described. Comparison of the PL-I protein in the winter flounder (*P. americanus*) and the PL-I protein in red mullet (*M. surmuletus*), which appear to have closely

related sequences of their trypsin-resistant cores (Fig. 2C), provides some interesting insights into the functional aspects in relation to the evolution of chromatin in organisms with PL-I proteins. *Pseudopleuronectes americanus* and *M. surmuletus* PL-I proteins can indeed be phylogenetically traced to a common origin (Fig. 3). Not only do the proteins share a winged helix domain, but they both contain large amounts of the repetitive motif SPBB [52] in their IDP domains.

In *P. americanus*, PL-I proteins consist of a heterogeneous mixture of proteins of high M_r (average 110 000) [53]. In contrast to *M. surmuletus*, which lacks a histone complement, the PL-I protein of *P. americanus* coexists with a full complement of core and linker histones [54]. Appearance of these proteins during spermiogenesis results in an increase of the average nucleosomal repeat length from approximately 195 bp to 222 bp in the mature sperm [54]. This suggests that part of the interaction of these 'additional' PL-I proteins takes place (as with histone H1) in the linker chromatin regions connecting adjacent nucleosomes, which explains the need for a winged helix domain. However, the long IDP domains of the molecule extend to several adjacent nucleosomes, making the sperm chromatin more resistant to nuclease digestion [54]. *M. surmuletus* PL-I proteins interact with non-nucleosomally constrained DNA, but still retain the phylogenetic signature of the winged helix domain, a signature that is completely gone by the time that the transition in the evolution of SNBPs in other animals occurs from the PL to the P type [4,19]. The lesser need for a canonical winged helix domain in the transition is evidenced by the structural deficiencies observed in the case of *M. surmuletus* PL-I (Fig. 2B). Interestingly, the structurally changed region of the winged helix domain observed corresponds to a region that has been recently demonstrated to be important for the interaction of linker histones with the nucleosome in a native setting [55].

The results of micrococcal nuclease digestion (Fig. 4), as well as those from EM and X-ray diffraction of sperm chromatin fibers (Fig. 5), provide biochemical and biophysical evidence for the lack of a nucleosomally organized chromatin in the sperm of *M. surmuletus*. The results of these experimental approaches all support the notion that the organization of *M. surmuletus* chromatin resulting from the association of PL-I proteins and DNA consists of irregular bundles of parallel DNA molecules with an average cross-sectional diameter of 650 Å. Figure 5C shows a model based on the EM and X-ray crystallography findings. The DNA bundle structures are highly reminiscent of the toroidal DNA bundles [25,26] that are present in the sperm chromatin of vertebrate organisms with SNBPs of the P type. However, the relationship to these structures and/or their involvement in the organization of chromosomal territories, as occurs in mammalian sperm [56], await further elucidation.

Experimental procedures

Organisms and biological material

Samples of the striped red mullet *M. surmuletus* and the red mullet *M. barbatus* were collected from different locations along the Mediterranean Sea during the spawning season. The extent of gonadal maturity was microscopically assessed, and mature sperm was obtained from the spontaneous flow generated by abdominal massage of the organisms.

Sperm nuclei preparation

Sperm or minced testicular tissue was suspended in 0.25 M sucrose, 5 mM $CaCl_2$, 10 mM Tris/HCl (pH 7.4), 10 mM benzamidine chloride and homogenized in a Dounce homogenizer (Wheaton, Millville, NJ). The homogenate was centrifuged at 2000 ***g*** for 10 min at 4 °C in a Sorvall RC-B5 (DuPont Instruments, Wilmington, DE), and the pellets were homogenized again in the same buffer containing additionally 0.5% Triton X-100. After incubation for 10 min on ice, the homogenate was centrifuged as before. The pellets were resuspended in the starting buffer without Triton X-100 and centrifuged again. The nuclear pellets thus obtained were used immediately for further analysis or resuspended in 50% glycerol in starting buffer and stored at −20 °C.

Extraction of SNBPs

SNBPs were obtained from the nuclear pellets by direct extraction with 0.4 M HCl using a Dounce homogenizer [15]. In some instances, the HCl extraction was preceded by a 35% acetic acid extraction in order to selectively extract the histone component. The proteins in the acid extracts were precipitated by addition of trichloroacetic acid to a final concentration of 20% (v/v) on ice for 10 min. The protein precipitate was collected by centrifugation at 12 000 ***g*** in a Sorvall RC-B5 (DuPont Instruments), and the pellet was rinsed once with acidified cold acetone (acetone/0.1 M HCl, 6 : 1, v/v), and once with cold acetone, and finally dried.

Chromatographic purification of SNBPs

Mullus SNBPs were dissolved in 0.9 M NaCl/50 mM sodium acetate (pH 6.7) and applied to a CM-Sephadex

C-25 (General Electric, Baie d'Urfé, Quebec, Canada) column previously equilibrated in the same buffer. Protein fractions were eluted with a 0.9–1.3 M NaCl gradient in the same buffer. Fractions corresponding to the major PL-I component were pooled and further purified by RP-HPLC using a Vydac C_{18}, 5 μm 4.6 × 250 mm column (Vydac, Hesperia, CA). Elution was with a 25–35% solution A–B gradient, where solution A was 0.1% trifluoroacetic acid and solution B was 100% acetonitrile.

Proteolytic digestions and chemical cleavage

Elastase (EC 3.4.21.36)

Mullus surmuletus PL-I protein was digested at room temperature with elastase (type IV) (Sigma-Aldrich, Oakville, ON) in 0.1 M ammonium bicarbonate (pH 8.0). The concentration of the protein was 2–5 $mg \cdot mL^{-1}$ at an enzyme/substrate (E/S) ratio of 1 : 100 (w/w).

Thermolysin (EC 3.4.24.4)

Digestion of PL-I protein with this enzyme was carried out using thermolysin (type X) (Sigma-Aldrich) at an E/S ratio of 1 : 500 (w/w) in 100 mM ammonium bicarbonate (pH 8.0) at a concentration of 5 $mg \cdot mL^{-1}$ as described elsewhere [57].

Trypsin (EC 3.4.21.4)

The major PL-I SNBP component from *M. surmuletus* was digested with trypsin (type III) (Sigma-Aldrich). Digestions were carried out in 2 M NaCl, 25 mM Tris/HCl (pH 7.5) [34] buffer at an E/S ratio of 1 : 375 (w/w) at room temperature. For analytical digestions, aliquots of the digestion were collected at different times, mixed with 2× gel electrophoresis sample buffer and immediately frozen and kept until used for AU/PAGE (see below). For preparative purposes, 500 μg of PL-I protein was digested for 60 min under the conditions described, and the reaction was stopped by addition of soybean trypsin inhibitor (type I-S) (Sigma-Aldrich) at a 1 : 4 (w/w) enzyme/inhibitor ratio. The mixture was then precipitated by addition of trichloroacetic acid to 20% (v/v), rinsed with acetone/HCl and acetone, and dried. RP-HPLC purification was carried out as described above but using a 25–45% A–B solution gradient.

Cyanogen bromide

Mullus surmuletus PL-I protein was dissolved in 0.1 M HCl at a concentration of approximately 7 $mg \cdot mL^{-1}$, and was hydrolyzed with CNBr using a CNBr/protein ratio of 1 : 3 (w/w). The reaction was allowed to proceed for 24 h at room temperature in the dark.

Chromatin digestions

Mullus surmuletus sperm nuclei, prepared as described earlier, were washed twice [suspension followed by quick centrifugation at 2000 ***g*** at 4 °C using an Eppendorf S415 microfuge (Brinkmann, Westbury, NY)] in 0.15 M NaCl, 10 mM Tris/HCl (pH 8.0), 0.5 mM $CaCl_2$, 0.2 mM phenylmethylsulfonyl fluoride. The second pellet was gently suspended in the same buffer to an A_{260} of 20 (determined as described in [58]). Micrococcal nuclease (Sigma-Aldrich) digestions were carried out at 37 °C at 0.58 units·(mg DNA)$^{-1}$. Two digestion aliquots were withdrawn at selected times. One of them was immediately added to two volumes of 2 M PCA/2 M NaCl, incubated for 20 min on ice, and centrifuged at 16 000 ***g*** for 10 min at 4 °C using an Eppendorf S415 microfuge (Brinkmann). The A_{260} of the supernatant was used to determine the percentage of PCA-soluble DNA [36]. The second aliquot was added to tubes containing excess EDTA (so that the final EDTA concentration was 10 mM) and quickly vortexed and kept on ice. The samples thus obtained were subsequently centrifuged at 16 000 ***g*** for 20 min at 4 °C using an Eppendorf S415 microfuge (Brinkmann), giving a supernatant (SI) and a pellet (PI). The pellet was suspended and hypotonically lysed overnight in 0.25 mM EDTA (pH 7.5), 0.2 mM phenylmethanesulfonyl fluoride at 4 °C. On the next day, the lysate was centrifuged as before to produce a supernatant SII and a pellet PII. The A_{260} of the SI and SII fractions was measured, and the amount of DNA present in the samples was determined using an extinction coefficient of $A_{260} = 20\ cm^2 \cdot mg^{-1}$. The DNA and protein contents of the SI, SII and PII fractions were also electrophoretically analyzed (see below). To this end, each fraction at a given digestion time was divided into two aliquots. One aliquot was used for SNBP extraction with 0.4 M HCl as described above, and the second was used to extract the DNA. In this latter instance, the different time aliquots were brought to 1 M NaCl, 0.5% SDS, extracted with chloroform/isoamyl alcohol (24 : 1, v/v), and precipitated with ethanol.

Electrophoretic analyses

DNA samples were analyzed on (1%) agarose gels in Tris/borate/EDTA buffer according to the method of Sambrook *et al.* [59]. Protein samples were analyzed on either AU or AUT polyacrylamide gels. The former were prepared according to the method of Panyim & Chalkley [60] as modified by Hurley [61], and following the protocol described in Ausió [18]. The latter were prepared as described by Zweidler [62], with the following final composition: 15% polyacrylamide (acrylamide/bis-acrylamide, 150 : 1, w/w), 0.9 M acetic acid, 5.75 M urea, 6 mM Triton X-100.

MS

MS analyses of *M. surmuletus* SNBPs were carried out by MALDI-TOF on a Voyager Linear DE (PerSeptive Biosystems Inc., Foster City, CA) using a sinapinic acid matrix following the protocol described elsewhere [7].

Amino acid composition and protein sequencing

Amino acid analyses were carried out as described elsewhere [28]. Primary structure determination was carried out by automated Edman degradation on an ABI 473 A protein sequencer according to the method of Jutglar *et al.* [63].

cDNA sequence obtained from RACE PCR

Mature testes from *M. surmuletus* used for RNA extraction were stored in RNAlater RNA-stabilizing reagent (Qiagen Inc., Mississauga, ON). Total RNA was extracted using Trizol reagent (Gibco BRL, Burlington, ON). Subsequently, cDNA was transcribed for 3′RACE using the FirstChoice RLM-RACE Kit (Ambion, Austin, TX), following the manufacturer's directions. PCR was performed using a PCRsprint thermal cycler (Hybaid, Teddington, UK) with the cDNA as a template. The following nested degenerate primers were designed for RACE PCR based on the determined amino acid sequences: MsF1, 5′-GTGTCC GAYYTGATMCTGATG-3′; MsF2, 5′-GAYCTGATMC TGATGGCTATG-3′; and MsF3, 5′-TCYCTGGCAGCC YTGAAGAA-3′.

For DNA sequencing, agarose gel-purified PCR products were cloned into pCR 2.1-TOPO vectors (Invitrogen, Burlington, ON) and transformed into TOP10 competent cells (Invitrogen). Plasmids were isolated by QIAprep Miniprep purification (Qiagen Inc.), and the inserts were sequenced at the DNA Sequencing Facility, Centre for Biomedical Research, University of Victoria.

Phylogenetic analysis

Sequences retrieved from databases were subsequently corrected for errors in accession numbers and nomenclature, and aligned on the basis of their translated amino acid sequences using the CLUSTAL_X [64] and BIOEDIT programs with the default parameters as previously described by [65]. All molecular evolutionary analyses in the present work were carried out using the program MEGA ver. 3.1 [66]. The extent of amino acid divergence between sequences was estimated by means of the uncorrected differences (*p*-distance), as this distance is known to give better results than more complicated methods when the number of sequences is large and the number of positions used is relatively small, because of its smaller variance. Distances were estimated using the complete-deletion option, and standard errors were calculated by the bootstrap method with 1000 replicates. The neighbor-joining tree-building method [67] was used to reconstruct the phylogenetic tree. We decided to combine the bootstrap and the interior-branch test methods in order to test the reliability of the obtained topologies, producing the bootstrap probability (BP) and confidence probability (CP) values for each internal branch, assuming BP > 80% and CP ≥ 95% to be statistically significant.

EM

Nuclei from *M. surmuletus* were washed with 10 mM Tris/HCl, 60 mM EDTA (pH 7.5) and incubated overnight in the same buffer. Carbon-coated electron microscopy grids (300 mesh) were placed on drops of the suspension consisting of swollen nuclei, and adsorption of the nuclear material to the grids was allowed to proceed for 5 min. Grids were then removed, washed with double-distilled water, stained with uranyl acetate, and washed again by passing them sequentially through three ethanol drops. Finally, they were rotary shadowed with platinum Pt-C and visualized in an EM 301 Philips electron microscope (Philips, Eindhoven, the Netherlands).

X-ray analysis of sperm nuclear fibers

Sperm cell nuclei from *M. surmuletus* were sequentially washed in 0.15 M NaCl and rinsed with double-distilled water. The excess water was removed and the sample was allowed to dry while chromatin fibers were pulled out from the lysed nuclei. Small fragments of the fibers thus obtained were placed in a capillary glass tube at a controlled relative humidity of 93%. The diffraction patterns were obtained as described previously [39,48,68], using an X-ray tube with a Cu anode and an Ni filter.

Acknowledgements

This work was supported by Natural Science and Engineering Research Council (NSERC) of Canada Grant OGP 0046399-02 to JA, by grants BIO2002-00317 from the Spanish MCYT and 2001/SGR/00250 from the Generalitat de Catalunya to JAS, and by Grant BFU 2005-00123 from the Spanish MCYT PGC-FEDER and a NATO Collaborative Linkage Grant LST.CLG 976607 to MC. JME-L is a recipient of a Postdoctoral Marie Curie International Fellowship within the 6th European Community Framework Programme. LJB is a recipient of an NSERC postgraduate scholarship – doctoral.

References

1 Strahl BD & Allis CD (2000) The language of covalent histone modifications. *Nature* **403**, 41–45.
2 Lewis JD, Abbott DW & Ausio J (2003) A haploid affair: core histone transitions during spermatogenesis. *Biochem Cell Biol* **81**, 131–140.
3 Churikov D, Zalenskaya IA & Zalensky AO (2004) Male germline-specific histones in mouse and man. *Cytogenet Genome Res* **105**, 203–214.
4 Ausió J (1999) Histone H1 and evolution of sperm nuclear basic proteins. *J Biol Chem* **274**, 31115–31118.
5 Ausió J (1995) Histone H1 and the evolution of the nuclear sperm specific proteins. In *Advances in Spermatozoal Phylogeny and Taxonomy*, **Vol. 166** (Jamieson BGM, Ausió J & Justine JL, eds), pp. 501–514. Memoires de Museum National d'Histoire Naturelle, Paris.
6 Lewis J, Song Y, de Jong M, Bagha S & Ausió J (2003) A walk though vertebrate and invertebrate protamines. *Chromosoma* **111**, 473–482.
7 Hunt JG, Kasinsky HE, Elsey RM, Wright CL, Rice P, Bell JE, Sharp DJ, Kiss AJ, Hunt DF, Arnott DP *et al.* (1996) Protamines of reptiles. *J Biol Chem* **271**, 23547–23557.
8 Oliva R & Dixon GH (1991) Vertebrate protamine genes and the histone-to-protamine replacement reaction. *Prog Nucleic Acid Res Mol Biol* **40**, 25–94.
9 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 *Comparative Spermatology 20 Years After* (Baccetti B, ed.), pp. 27–30. Raven Press (Serono Symposia Publications, New York.
10 Giménez-Bonafé P, Ribes E, Sautière P, Gonzalez A, Kasinsky H, Kouach M, Sautiere PE, Ausió J & Chiva M (2002) Chromatin condensation, cysteine-rich protamine, and establishment of disulphide interprotamine bonds during spermiogenesis of *Eledone cirrhosa* (Cephalopoda). *Eur J Cell Biol* **81**, 341–349.
11 van Holde KE (1988) *Chromatin*. Springer-Verlag, New York.
12 Poccia D (1995) Male germ line specific histones of sea urchins and sea stars. In *Advances in Spermatozoal Taxonomy and Phylogeny* (Jamieson BGM, Ausió J & Justine J-L, eds), pp. 475–490. Mémoires du Muséum National d'Histoire Naturelle, Paris.
13 Poccia DL, Simpson MV & Green GR (1987) Transitions in histone variants during sea urchin spermatogenesis. *Dev Biol* **121**, 445–453.
14 Munoz Guerra S, Colom J, Ausió J & Subirana JA (1982) Histones from spermatozoa of the horseshoe crab. *Biochim Biophys Acta* **697**, 305–312.
15 Wang X & Ausió J (2001) Histones are the major chromosomal protein components of the sperm of the nemerteans *Cerebratulus californiensis* and *Cerebratulus lacteus*. *J Exp Zool* **290**, 431–436.
16 Munoz-Guerra S, Azorin F, Casas MT, Marcet X, Maristany MA, Roca J & Subirana JA (1982) Structural organization of sperm chromatin from the fish *Carassius auratus*. *Exp Cell Res* **137**, 47–53.
17 Ausió J (1986) Structural variability and compositional homology of the protamine-like components of the sperm from bivalve molluscs. *Comp Biochem Physiol* **85B**, 439–449.
18 Ausió 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.
19 Eirin-Lopez JM, Frehlick LJ & Ausio J (2006) Protamines, in the footsteps of linker histone evolution. *J Biol Chem* **281**, 1–4.
20 Ausió J & van Holde KE (1987) A dual chromatin organization in the sperm of the bivalve mollusc *Spisula solidissima*. *Eur J Biochem* **165**, 363–371.
21 Lewis JD, McParland R & Ausio J (2004) PL-I of *Spisula solidissima*, a highly elongated sperm-specific histone H1. *Biochemistry* **43**, 7766–7775.
22 Lewis JD, Saperas N, Song Y, Zamora MJ, Chiva M & Ausio J (2004) Histone H1 and the origin of protamines. *Proc Natl Acad Sci USA* **101**, 4148–4152.
23 Saperas N, Ausió J, Lloris D & Chiva M (1994) On the evolution of protamines in bony fish: alternatives to the 'retroviral horizontal transmission' hypothesis. *J Mol Evol* **39**, 282–295.
24 Casas MT, Munoz-Guerra S & Subirana JA (1981) Preliminary report on the ultrastructure of chromatin in the histone containing spermatozoa of a teleost fish. *Biol Cell* **40**, 87–92.
25 Ward WS & Coffey DS (1991) DNA packaging and organization in mammalian spermatozoa: comparison with somatic cells. *Biol Reprod* **44**, 569–574.
26 Ward WS & Zalensky AO (1996) The unique, complex organization of the transcriptionally silent sperm chromatin. *Crit Rev Eukaryot Gene Expr* **6**, 139–147.
27 Casas MT, Ausió J & Subirana JA (1993) Chromatin fibers with different protamine and histone compositions. *Exp Cell Res* **204**, 192–197.
28 Saperas N, Lloris D & Chiva M (1993) Sporadic appearance of histones, histone-like proteins and protamines in sperm chromatin of bony fish. *J Exp Zool* **265**, 575–586.
29 Saperas N, Ribes E, Buesa C, Garcia-Hegart F & Chiva M (1993) Differences in chromatin condensation during spermiogenesis in two species of fish with distinct protamines. *J Exp Zool* **265**, 185–194.
30 Rall SC & Cole RD (1971) Amino acid sequence and sequence variability of the amino-terminal regions of lysine-rich histones. *J Biol Chem* **246**, 7175–7190.
31 Garcia BA, Busby SA, Barber CM, Shabanowitz J, Allis CD & Hunt DF (2004) Characterization of phos-

phorylation sites on histone H1 isoforms by tandem mass spectrometry. *J Proteome Res* **3**, 1219–1227.
32 Suzuki M (1989) SPKK, a new nucleic acid-binding unit of protein found in histone. *EMBO J* **8**, 797–804.
33 Ramakrishnan V, Finch JT, Graziano V, Lee PL & Sweet RM (1993) Crystal structure of globular domain of histone H5 and its implications for nucleosome binding. *Nature* **362**, 219–223.
34 Ausió J, Toumadje A, McParland R, Becker RR, Johnson WC & van Holde KE (1987) Structural characterization of the trypsin-resistant core in the nuclear sperm-specific protein from *Spisula solidissima*. *Biochemistry* **26**, 975–982.
35 Meistrich ML, Brock WA, Grimes SR, Platz RD & Hnilica LS (1978) Nuclear protein transitions during spermatogenesis. *Fed Proc* **37**, 2522–2525.
36 Clark RJ & Felsenfeld G (1974) Chemical probes of chromatin structure. *Biochemistry* **13**, 3622–3628.
37 Young RJ & Sweeney K (1979) The structural organization of sperm chromatin. *Gamete Res* **2**, 265–282.
38 Sobhon P, Chutatape C, Chalermisarachai P, Vongpayabal P & Tanphaichitr N (1982) Transmission and scanning electron microscopic studies of the human sperm chromatin decondensed by micrococcal nuclease and salt. *J Exp Zool* **221**, 61–79.
39 Ausió J & Subirana JA (1982) Nuclear proteins and the organization of chromatin in spermatozoa of *Mytilus edulis*. *Exp Cell Res* **141**, 39–45.
40 Hewish DR & Burgoyne LA (1973) Chromatin substructure. The digestion of chromatin DNA at regularly spaced sites by a nuclear deoxyribonuclease. *Biochem Biophys Res Commun* **52**, 504–510.
41 Clark DJ & Thomas JO (1986) Salt-dependent co-operative interaction of histone H1 with linear DNA. *J Mol Biol* **187**, 569–580.
42 Garcia-Ramirez M & Subirana JA (1994) Condensation of DNA by basic proteins does not depend on protein composition. *Biopolymers* **34**, 285–292.
43 Rodriguez AT, Perez L, Moran F, Montero F & Suau P (1991) Cooperative interaction of the C-terminal domain of histone H1 with DNA. *Biophys Chem* **39**, 145–152.
44 Balhorn R, Cosman M, Thornton K, Krishnan VV, Corzett M, Bench G, Kramer C, Lee IVJ, Hud NV, Allen M *et al.* (1999) Protamine mediated condensation of DNA in mammalian sperm. In *The Male Gamete: from Basic Knowledge to Clinical Applications* (Gagnon C, ed.), pp. 55–70. Cache River Press, Vienna, IL.
45 Pardon JF, Wilkins MH & Richards BM (1967) Super-helical model for nucleohistone. *Nature* **215**, 508–509.
46 Widom J & Klug A (1985) Structure of the 300A chromatin filament: X-ray diffraction from oriented samples. *Cell* **43**, 207–213.
47 Azorin F, Vives J, Campos JL, Jordán A, Lloveras J, Puigjaner L, Subirana JA, Mayer R & Brack A (1985) Interaction of DNA with lysine-rich polypeptides and proteins: the influence of polypeptide composition and secondary structure. *J Mol Biol* **185**, 371–387.
48 Suau P & Subirana JA (1977) X-ray diffraction studies of nucleoprotamine structure. *J Mol Biol* **117**, 909–926.
49 Th'ng JP, Sung R, Ye M & Hendzel MJ (2005) H1 family histones in the nucleus. Control of binding and localization by the C-terminal domain. *J Biol Chem* **280**, 27809–27814.
50 Watson CE, Gauthier SY & Davies PL (1999) Structure and expression of the highly repetitive histone H1-related sperm chromatin proteins from winter flounder. *Eur J Biochem* **262**, 258–267.
51 Hansen JC, Lu X, Ross ED & Woody RW (2006) Intrinsic protein disorder, amino acid composition, and histone terminal domains. *J Biol Chem* **281**, 1853–1856.
52 Watson CE & Davies PL (1998) The high molecular weight chromatin proteins of winter flounder sperm are related to an extreme histone H1 variant. *J Biol Chem* **273**, 6157–6162.
53 Kennedy BP & Davies PL (1980) Acid-soluble nuclear proteins of the testis during spermatogenesis in the winter flounder. Loss of the high mobility group proteins. *J Biol Chem* **255**, 2533–2539.
54 Kennedy BP & Davies PL (1982) Chromatin reorganization during spermatogenesis in the winter flounder. *J Biol Chem* **257**, 11160–11165.
55 Brown DT, Izard T & Misteli T (2006) Mapping the interaction surface of linker histone H1(0) with the nucleosome of native chromatin in vivo. *Nat Struct Mol Biol* **13**, 250–255.
56 Mudrak O, Tomilin N & Zalensky A (2005) Chromosome architecture in the decondensing human sperm nucleus. *J Cell Sci* **118**, 4541–4550.
57 Rocchini C, Rice P & Ausió J (1995) Complete sequence and characterization of the major sperm nuclear basic protein from *Mytilus trossulus*. *FEBS Lett* **363**, 37–40.
58 Ausió J, Dong F & van Holde KE (1989) Use of selectively trypsinized nucleosome core particles to analyze the role of the histone 'tails' in the stabilization of the nucleosome. *J Mol Biol* **206**, 451–463.
59 Sambrook J, Fritsch EF & Maniatis T (1989) *Molecular Cloning, a Laboratory Manual*, 2nd edn. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.
60 Panyim S & Chalkley R (1969) High resolution acrylamide gel electrophoresis of histones. *Arch Biochem Biophys* **130**, 337–346.
61 Hurley CK (1977) Electrophoresis of histones: a modified Panyim and Chalkley system for slab gels. *Anal Biochem* **80**, 624–626.

62 Zweidler A (1978) Resolution of histones by polyacrylamide gel electrophoresis in presence of nonionic detergents. *Methods Cell Biol* **17**, 223–233.
63 Jutglar L, Borrell JI & Ausió J (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.
64 Thompson JD, Gibson TJ, Plewniak F, Jeanmougin F & Higgins DG (1997) The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Res* **25**, 4876–4882.
65 Eirín-López JM, Lewis JD, Howe L & Ausió J (2006) Common phylogenetic origin of protamine-like (PL) proteins and histone H1: evidence from bivalve PL genes. *Mol Biol Evol* **23**, 1304–1317.
66 Kumar S, Tamura K & Nei M (2004) MEGA3: integrated software for Molecular Evolutionary Genetics Analysis and sequence alignment. *Brief Bioinform* **5**, 150–163.
67 Saitou N & Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* **4**, 406–425.
68 Ausió J & Suau P (1983) Structural heterogeneity of reconstituted complexes of DNA with typical and intermediate protamines. *Biophys Chem* **18**, 257–267.
69 Schwede T, Kopp J, Guex N & Peitsch MC (2003) SWISS-MODEL: an automated protein homology-modeling server. *Nucleic Acids Res* **31**, 3381–3385.
70 Eirin-Lopez JM, Fernanda Ruiz M, Gonzalez-Tizon AM, Martinez A, Sanchez L & Mendez J (2004) Molecular evolutionary characterization of the mussel *Mytilus* histone multigene family: first record of a tandemly repeated unit of five histone genes containing an H1 subtype with 'orphon' features. *J Mol Evol* **58**, 131–144.
71 Schulze E & Schulze B (1995) The vertebrate linker histones H1 zero, H5, and H1M are descendants of invertebrate 'orphon' histone H1 genes. *J Mol Evol* **41**, 833–840.
72 Felsestein J (1985) Confidence limits on phylogenies: an approach using the bootstrap. *Evol Int J Org Evolution* **39**, 783–791.
73 Sitnikova T (1996) Bootstrap method of interior-branch test for phylogenetic trees. *Mol Biol Evol* **13**, 605–611.
74 Kasinsky HE, Lewis JD, Dacks JB & Ausió J (2001) Origin of H1 histones. *FASEB J* **15**, 34–42.
75 Mayes ELV & Johns EW (1982) Accumulated data. In *The HMG Chromosomal Proteins* (Johns EW, ed.), pp. 223–247. Academic Press, New York.
76 Carlos S, Hunt DF, Rocchini C, Arnott DP & Ausió J (1993) Post-translational cleavage of a histone H1-like protein in the sperm of *Mytilus*. *J Biol Chem* **268**, 195–199.

Supplementary material

The following supplementary material is available online:

Table S1. Proteins used in the phylogenetic analysis shown in Fig. 3. The accession number for the sequence and the organism's source are also indicated.

This material is available as part of the online article from http://www.blackwell-synergy.com