

Cloning and sequencing of SOB3, a human gene coding for a sperm protein homologous to an antimicrobial protein and potentially involved in zona pellucida binding

Sonia Hammami-Hamza¹, Mireille Doussau¹, Jérôme Bernard², Edith Rogier², Clotilde Duquenne¹, Yolande Richard², Annick Lefèvre³ and Catherine Finaz^{1,4}

¹INSERM U 355, Maturation Gamétique et Fécondation, Institut Paris-sud sur les Cytokines, 32 rue des Carnets, 92140 Clamart, France, ²INSERM U 131, Cytokines et immunorégulation, Institut Paris-sud sur les Cytokines, 32 rue des Carnets, 92104 Clamart, France, ³INSERM U 418, Communications Cellulaires et Différenciation, Hôpital Debrousse, 29 rue Soeur Bouvier, 69322 Lyon, France

⁴To whom correspondence should be addressed at: INSERM U 355, 32 rue des Carnets, 92140 Clamart.
E-mail: catherine.finaz@inserm.ipsc.u-psud.fr

We have previously characterized an 18–19 kDa cationic protein, SOB3, that was detected in the epididymis and localized within the acrosome and on the neck region of human spermatozoa. We suggested that it is involved in secondary sperm binding to the zona pellucida. The present study describes its purification to homogeneity by preparative electrophoresis and non-equilibrium pH gradient electrophoresis. Degenerate primers deduced from microsequencing were used to amplify a specific fragment from human epididymal RNA by reverse transcription-polymerase chain reaction (RT-PCR). This 164 bp fragment was extended by 5' and 3'-RACE to obtain the 548 bp full length cDNA. The open reading frame encodes a 170 amino acid protein. SOB3 is a single copy gene. It is 98% identical to *prepro-FALL39* and 100% identical to *CAP18*, two human genes which were initially identified by screening a human bone marrow λ gt11 library, and which encode an antimicrobial protein. Northern blots of human tissues revealed a 1 kb transcript in corpus and cauda epididymis only, while RT-PCR showed presence of the mRNA in the three epididymal regions and also in round spermatids. The above results suggest that SOB3 has two roles in sperm protection and fertilization, depending on its dual origin and final sperm localization.

Key words: antimicrobial peptide/CAP18/fertilization/prepro-FALL-39/sperm protein

Introduction

Fertilization requires a cascade of cell-matrix and cell-cell interactions that bring spermatozoa into contact with the oocyte. Crossing of the zona pellucida is a particularly complex step in mammalian fertilization. The zona pellucida in most mammals, including humans, consists of three glycoproteins, the ZP3 and ZP2 heterodimers that are cross-linked by ZP1 molecules. The first event, primary binding, occurs when intact spermatozoa bind to ZP3 via specific receptors located on the anterior head surface. The spermatozoa then undergo the acrosome reaction; they loose the membrane surrounding the acrosome and release trypsin-like proteases that solubilize the zona pellucida frame. Spermatozoa must then pass through the zona pellucida while remaining anchored to it. This secondary binding is thought to involve receptors located on the newly exposed inner acrosomal membrane and ZP2. Several sperm proteins that may take part in this interaction have been identified, but mechanisms underlying this step are far from being understood in either animals or humans.

PH-20, a sperm glycoprotein initially identified in guinea-pigs, seems to be a good candidate for secondary binding to zona pellucida. The PH-20 protein is originally located on the posterior head plasma membrane, but, when spermatozoa undergo the acrosome reaction, it migrates to the inner acrosome membrane, where it joins pre-existing PH-20 molecules (Myles and Primakoff, 1984; Cowan *et al.*, 1991). Studies with specific function-blocking antibodies indicate that plasma membrane PH-20 has the hyaluronidase activity required for penetration of the cumulus layer, while the PH-20 molecules on the inner acrosomal membrane are involved in secondary binding to zona pellucida (Primakoff *et al.*, 1987; Lin *et al.*, 1994; Hunnicutt *et al.*, 1996). Recent studies (Yudin *et al.*, 1999) on Rhesus monkeys have confirmed this last observation. Nevertheless, the variation in PH-20 distribution over spermatozoa varies from one species to another, suggesting that it is not responsible by itself for binding to zona pellucida but that it has a role in the regulation of the specificity of sperm-zona binding (Benoff, 1997).

The proacrosin-acrosin pair has long been suggested to be involved in secondary binding. Proacrosin was first identified in the acrosomes of rabbit and boar spermatozoa. It is converted into a mature form by proteolysis and probably also has a dual function. In addition to its enzymatic activity, which is thought to lyse zona pellucida, it binds to zona pellucida via its polysulphate binding domain which is different from the enzymatic site (Richardson *et al.*, 1996; Jansen *et al.*, 1998). Other data suggest that the binding domain functions in both secondary binding and in enzyme activation (Moreno *et al.*, 1998). However, the observation that spermatozoa from a mouse homozygous for a targeted mutation in the acrosin gene penetrate the zona pellucida and effect fertilization has raised questions about its role (Baba *et al.*, 1994; Adham *et al.*, 1997).

Additionally, the role of SP-10, a sperm protein that was believed to be important in binding to the oocyte, has recently been clarified. Data obtained in humans showed that a monoclonal anti-SP-10 antibody (mAb) inhibits the binding of spermatozoa to the oolemma in the zona-free hamster egg penetration test, while it has no effect in the hemizona assay, as initially suggested (Hamatani *et al.*, 2000).

Our group has generated monoclonal antibodies (mAb) directed against sperm surface proteins in order to identify proteins involved in the interaction of human gametes (Boué *et al.*, 1992). Using the LB5 mAb we have identified a cationic protein, SOB3 (sperm oocyte binding 3), that seems to be involved in the secondary binding to the zona pellucida (Martin Ruiz *et al.*, 1998). Support for this role in gamete recognition comes from the observation that the fragment antigene binding (Fab) of the LB5 mAb inhibited the binding of spermatozoa to the zona pellucida by 35.7% and that this inhibition was increased to 59.9% when spermatozoa had been induced to acrosome-react. This role was reinforced by the strong correlation observed between the percentages of acrosome-labelled and acrosome-reacted spermatozoa that indicated the presence of SOB3 within the acrosome. Western blotting of spermatozoa and cauda epididymal extracts revealed two bands of 18 and 19 kDa. A similar 18 kDa band was detected in extracts of erythrocytes and a 19 kDa band was found with B-lymphocytes.

We have now cloned and sequenced the epididymal cDNA corresponding to the human SOB3 sperm protein. The *SOB3* gene was characterized by Southern blotting, its expression was analysed along the male reproductive tract by Northern blotting and reverse transcription-polymerase chain reaction (RT-PCR). The sequence of protein SOB3 is very similar to those of two homologous antimicrobial proteins, prepro-FALL-39 (preproprotein FALL-39) and CAP18 (18 kDa cationic antimicrobial protein), that belong to the cathelicidin family and are present in the immune system. Our results show that the *SOB3* gene is transcribed in both the testis and the epididymis and suggest that the protein may have a double function, being involved in the secondary binding to zona pellucida and being homologous with antimicrobial proteins.

Materials and methods

Monoclonal antibody LB5 origin

The LB5 hybridoma was produced by immunising a BALB/c mouse with ejaculated human spermatozoa (Boué *et al.*, 1992; Martin Ruiz *et al.*, 1998). The resulting monoclonal antibody (mAb) is an IgG1.

Purification of human SOB3 protein

SOB3 was purified in three steps: extraction, preparative sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) and non-equilibrium pH gradient electrophoresis (NEPHGE). A total of 10×10^{10} human spermatozoa were washed with phosphate-buffered saline (PBS), suspended in 15 ml extraction buffer [0.01 mol/l Tris pH 7.4, 0.15 mol/l NaCl, 1% Nonidet NP-40, 10% glycerol and 2 mmol/l Pefabloc (Pentapharm, Interchim, Montluçon, France)], and placed on a rotary agitator at 4°C for 1 h. The suspension was ultracentrifuged (100 000 g) at 4°C for 1 h and the protein concentration in the supernatant was measured (Bio-Rad Protein Assay; Bio-Rad, Hercules, CA, USA).

Preparative SDS-PAGE and protein elution were performed as previously described (Martin-Ruiz *et al.*, 1998). Fractions in the range of 17–20 kDa were tested by Western blotting using the LB5 mAb and the positive fractions were pooled, concentrated and stored at –80°C. Proteins were precipitated with 80% acetone, 10% trichloroacetic acid and 0.12% dithiothreitol to remove the SDS (Görg *et al.*, 1997). The protein pellet was dissolved (2–3 µg/µl) in solubilization solution (9 mol/l urea, 4% (w/v) Chaps, 1% mercaptoethanol and a mix of ampholytes pH range 3–10.5 [1% Pharmalytes 3–10, 3.25% Pharmalytes 5–8 and 8–10.5 (Pharmacia Amersham Biotech, Uppsala, Sweden)], and heated for 1 h at 37°C.

NEPHGE was performed with a Mini Protean II Dual Slab Cell (Bio-Rad, Hercules, CA, USA) according to a published method (Robertson *et al.*, 1987). Gels were cast with 8% Readysol Plus (Pharmacia Biotech), 9 mol/l urea, 10% glycerol, 2% (w/v) Chaps, ampholyte mix in the same concentration as in the solubilization solution, 0.16% Temed and 0.04% (w/v) ammonium persulphate. The upper chamber was filled with 10 mmol/l H₃PO₄ as anode solution and the lower chamber with 20 mmol/l NaOH as cathode solution. Proteins (60 µg in 30 µl solubilization solution) were applied in each well and electrophoresis was run at 120 V for 20 min, 180 V overnight and 400 V for 30 min, at 10–12°C (3000–3500 volthours). The resulting gel was stained with Coomassie blue.

Western blotting

Proteins separated by NEPHGE (10 µg per lane) were electroblotted onto a PVDF membrane (Amersham Pharmacia Biotech) in 0.7% acetic acid according to a published method (Reinhart *et al.*, 1982). The membrane was immunostained to detect the SOB3 protein. It was first saturated with 10% non-fat milk in PBS for 1 h at 20°C, washed twice in PBS-0.05% Tween and incubated overnight at 4°C with either LB5 or an irrelevant G12 mAb supernatant (Lefèvre *et al.*, 1997) or culture medium as controls. The membrane was washed twice in PBS-Tween and incubated in a 1:3000 dilution of horseradish peroxidase-linked sheep anti-mouse IgG (Amersham Pharmacia Biotech) in PBS. Protein bands were visualized using the ECL chemiluminescence protocol according to the manufacturer's instructions (Amersham Pharmacia Biotech).

Amino acid sequencing

Microsequencing was performed by J.d'Alayer (Institut Pasteur, Paris). Patterns of the NEPHGE gel stained with naphthol blue-black and of the Western blotting were compared. The two bands corresponding to SOB3 were cut into small pieces and partially dehydrated in a Speed-

Vac. The pieces of gel were rehydrated in 200 µl 0.1 mol/l Tris-HCl, pH 8.6/0.01% Tween 20 and digested with 0.2 µg trypsin at 30°C for 18 h. The supernatant was recovered and injected onto a DEAE HPLC column linked to a C18 RP-HPLC and eluted with a 2–70% acetonitrile, 0.1% TFA gradient. Peptides were collected manually and frozen. They were sequenced on Applied Biosystems 473A and 494 sequencers.

Cell recovery and purification

Ejaculated spermatozoa from proven fertile donors were washed three times in PBS containing 0.9 mmol/l CaCl₂ and 0.5 mmol/l MgCl₂. Round spermatids were recovered from testicular biopsies of patients undergoing infertility treatment by intracytoplasmic sperm injection (ICSI). They were purified as previously described using a FACScalibur flow cytometer (Becton Dickinson, San Jose, CA, USA) equipped with a cell sorting system (Ziyyat *et al.*, 1999).

Mononuclear cells obtained from the buffy coats of blood from healthy donors were isolated by density gradient centrifugation through Ficoll (Lymphoprep; Life Technologies, Cergy Pontoise, France). B-cells were obtained by positive selection using CD19 Dynabeads (DynaL AS, Oslo, Norway) and were ≥96% CD22⁺, ≤1% CD14⁺ and ≤1% CD3⁺ (*n* = 5). T-cells were recovered from the E⁺ fraction, and depleted of residual B-cells and monocytes using CD19 and CD14 Dynabeads M-450 (DynaL). Blood T-cells were ≥96% CD3⁺, ≤1% CD14⁺ and CD20⁺ cells (*n* = 5). Monocytes were obtained from mononuclear cells by depletion of B- and T-lymphocytes with CD19 and CD3 Dynabeads M-450 (DynaL) and were ≥95% CD14⁺ cells, ≤1% CD3⁺ and CD19⁺ cells (*n* = 5).

Mononuclear cells were also obtained from palatine tonsils removed from children with chronic tonsillitis by gentle dissociation with forceps. Total B-cells, obtained after one cycle of rosette formation and depletion of residual T-cells with CD2 Dynabeads M-450 (DynaL), were 93 ± 4% CD19⁺, 59 ± 6% IgD⁺, 81 ± 16% CD44⁺, 21 ± 6% CD38⁺ and 1% CD14⁺, CD3⁺ and DRC1⁺ cells (*n* = 5). In some experiments, total B-cells were separated into naïve B-cells (IgD⁺), memory B-cells (IgD⁻ CD44⁺) and germinal centre (GC) B-cell (IgD⁻ CD44⁻) populations as previously described (Krzysiek *et al.*, 1999).

The Raji Burkitt lymphoma and MOLT-4 cell lines were obtained from the American Type Culture Collection (Rockville Pike, MD, USA). They were propagated in Roswell Park Memorial Institute (RPMI) 1640 medium (Gibco BRL, Paisley, UK) containing 10 mmol/l HEPES, 2 mmol/l L-glutamine, 100 U/ml penicillin, 100 µg/ml streptomycin, 1 mmol/l sodium pyruvate and 10% heat-inactivated FCS, and subcultured three times a week. In some experiments, T-cells (10⁶/ml) were activated for 2 days with 5 µg/ml phytohaemagglutinin (PHA) (Murex, Dartford, UK).

Isolation of total RNA and mRNA

Organ samples removed from adult fertile men, purified bone marrow and blood cells were frozen immediately in liquid N₂ and stored. Total RNA was extracted by the guanidine thiocyanate/phenol/chloroform method, and 40 µg RNA samples were treated with 10 U RNase-free DNase for 15 min at 37°C (Boehringer, Mannheim, Germany). For 5' RACE, poly (A⁺) mRNA were isolated from total RNA with the Oligotex mRNA Mini Kit (Quiagen, Hilden, Germany).

For RT-PCR of round spermatids and spermatozoa, RNA was extracted by four successive rounds of freezing and thawing from 500 purified cells resuspended in 2 µl H₂O containing 25 IU RNase inhibitor (Boehringer). Samples were treated with 10 IU RNase-free DNase.

Cloning and sequencing of the human SOB3 cDNA, production of a SOB3 cDNA probe

A partial SOB3 cDNA was amplified by RT-PCR using epididymal mRNA as a template and the following degenerated primers deduced

from the microsequencing results: sense 5'-TGYATGGGIAC-NGTNAC-3' and antisense 5'-CCKCARRAARTCYTTDAT-3'. The PCR cycling parameters were 30 s at 94°C, 1 min at 40°C, 45 s at 72°C for 40 cycles, and 10 min at 72°C. A 164 bp band was found in the caput, corpus and cauda epididymal samples. It was purified by gel electrophoresis, cloned into the PCR 2-1 TOPO vector with the TOPO TA cloning kit (Invitrogen, Groningen, The Netherlands) and sequenced on both DNA strands.

The 5' RACE end of SOB3 transcript was obtained from 0.5 µg human epididymal poly (A⁺) mRNA using the SMART RACE cDNA amplification kit from Clontech (Palo Alto, CA, USA). The first strand cDNA was synthesized using the 5'CDS and the SMART II oligonucleotides provided in the kit as primers with MMLV reverse transcriptase (Life Technologies). This cDNA was then amplified in a primary PCR reaction with the upstream universal primer mix included in the kit and a SOB3-specific primer (designed from the 164 bp fragment: 5'-CTGATGTCAAAGGAGCCCCCTGGCC-3') using the Advantage 2 PCR Enzyme System from Clontech with the following thermal cycle profile: 5 s at 94°C and 3 min at 72°C for 5 cycles; 5 s at 94°C, 10 s at 70°C and 3 min at 72°C for 5 cycles, then 5 s at 94°C, 10 s at 68°C and 3 min at 72°C for 25 cycles. The resulting PCR product was amplified in a second round of PCR using the nested universal primer mix (supplied with the RACE kit) and a nested SOB3 primer (5'-CCTGGTTGAGGGTCACTGTCCCC-3') with the same thermal cycle parameters.

Human epididymal total RNA (1 µg) was used as template for generating the 3' end of human SOB3 with the same SMART RACE cDNA amplification kit. The first strand cDNA was obtained using the 3'CDS primer for the MMLV reverse transcriptase, then amplified with the upstream universal primer mix and a SOB3 specific primer designed from the 164 bp fragment (5'-CGGTGTATGGGACAGTGACC-3') under the conditions used for 5' RACE. A second round of PCR was performed using the nested universal primer mix (supplied with the RACE kit), a nested SOB3 primer (5'-GCCCTGCTGGGTGATTTCTTCCGG-3') and the same thermal cycle conditions.

The PCR products of 5' and 3' RACE PCR were cloned into the PCR 2-1 TOPO vector from Invitrogen and sequenced on both DNA strands. A 548 bp cDNA probe was amplified using the upstream (5'-GCCATG-AAGACCCAAAGGAATGG-3') and downstream (5'GCCAGAA-GCCTGAGCCAGGG-3') specific SOB3 primers deduced from the resulting sequence, and Taq Gold (Perkin Elmer) with the following thermal cycle: 10 min at 94°C, then 30 s at 94°C, 1 min at 60°C, 45 s at 72°C for 40 cycles, and 10 min at 72°C. This probe was purified by gel electrophoresis and used for Northern and Southern blotting.

Sequence data analysis

Sequence data were compiled and DNA and protein aligned using the Blast 2 and Fasta programs.

Southern blotting

Genomic DNA was isolated with a Qiamp DNA minikit (Quiagen). 10 µg samples were digested with 60 U *EcoRI*, *BamHI* or *HindIII* (Life Technologies), separated on a 0.8% (w/v) agarose gel, and transferred by capillary action to a Hybond-N membrane (Pharmacia Biotech, Uppsala, Sweden). The membrane was prehybridized at 42°C for 4 h in 50% (w/v) formamide, 5×standard saline citrate (SSC), 5×Denhardt's reagent, 0.5% SDS and 5% dextran sulphate containing 100 µg/ml salmon spermatozoa, and hybridized overnight in the same solution containing the purified 548 bp probe labelled with [α -³²P]dCTP by random-priming using the Rediprime II labelling system (Amersham, Freiburg, Germany). The membrane was then washed to a final stringency of 2×SSC and 0.5% SDS at 20°C, and exposed to Hyperfilm MP film (Amersham) at -80°C for 4 days.

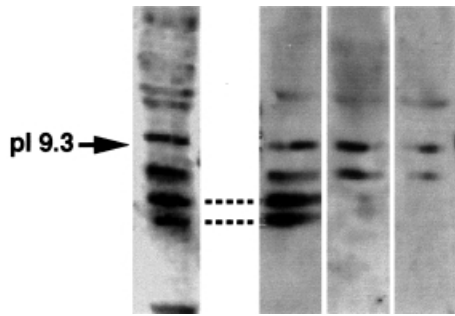


Figure 1. Non-equilibrium pH gradient electrophoresis pattern of the 17–20 kDa fraction from preparative sodium dodecyl sulphate-polyacrylamide gel electrophoresis of spermatozoa and immunoreactivity in Western blotting. From left, first lane: Coomassie blue staining; second lane: LB5 IgG1 blotting; third lane: control IgG1 blotting; fourth lane: control without first antibody. The amount of protein applied was 60 μ g for Coomassie blue staining and 10 μ g per lane for Western blots. The broken lines indicate the two SOB3 bands. Trypsinogen was used as a standard: its pI is indicated on the left.

Northern blotting

Total RNA from each tissue sample was separated on a 1.5% (w/v) agarose-formaldehyde gel and transferred to a Hybond-N membrane by upward capillary transfer. The membrane was air-dried and UV cross-linked, then prehybridized at 65°C for 4 h in 50% (w/v) formamide, 0.75 mol/l NaCl, 0.05 mol/l NaH₂PO₄, 0.5 mmol/l EDTA, 5 \times Denhart's reagent, 0.1 mg/ml salmon spermatozoa, and 0.5% SDS, and hybridized at 65°C overnight in the same solution containing 0.1% SDS only and the 548 bp cDNA insert labelled with [α^{32} P]dCTP by random-priming. The membrane was washed to a final stringency of 1 \times SSC at 65°C for 15 min and exposed to Hyperfilm MP film at –80°C for 4 days.

Reverse transcription-polymerase chain reaction (RT-PCR) analysis

Total RNA from epididymis or testis (10 μ g), blood and bone marrow cells (2 μ g) or 500 germ cells was treated with 10 U RNase-free DNase and cDNA was synthesized with MMLV reverse transcriptase (Life Technologies).

PCR amplifications were carried out with 100 ng somatic cell cDNA. PCR for round spermatids and spermatozoa was performed with cDNA corresponding to 125 cells. The SOB3 primers and the thermoprofile were the same as those used to amplify the 548 bp probe. The semi-nested PCR of spermatid and sperm cDNA was performed using the upstream first-round primer (5'-GGGGACAGTGACCCTCAACC-AGG-3') and the downstream internal primer (5'-GCCCAGAAGCC-TGAGCCAGGG-3') with the same thermoprofile in order to synthesize a 218 bp amplicon. Negative controls without cDNA matrix were included in all amplifications (data not shown).

Results

SOB3 purification and amino acid sequencing

The purity of the protein was assessed by Western blotting after NEPHGE (Figure 1). Incubation with LB5 mAb revealed two bands, A and B, with an estimated pI of 9.3–9.5.

One tryptic peptide from band A and three peptides from band B were sequenced. Peptide IKDFLR from band A was similar to one of the peptides obtained from band B. All four peptides were identical to the deduced amino acid sequences of the cDNA of the precursor of an antimicrobial protein, prepro-FALL-39 or

CAP18. These cDNA sequences were simultaneously identified by two independent groups by screening of a human bone marrow λ gt11 cDNA library (Agerberth *et al.*, 1995; Larrick *et al.*, 1995).

Cloning and sequencing of epididymal SOB3 cDNA

The sequence of the SOB3 RNA expressed in the male genital tract was determined using RT-PCR and degenerate primers designed from the sequences of peptides CMGTVT and IKDFLR. Testis RNA revealed no band and a 164 bp product was obtained using epididymal RNA as a template. It contained the sequence encoding peptide FALLGDFFR. A 5' and 3' RACE approach was adopted using epididymal RNA. The full-length cDNA sequence (619 bp) and the corresponding deduced amino acid sequence are shown in Figure 2. The cDNA has a very short untranslated 5' region of 16 bp upstream of the ATG of the open reading frame; it contains neither a TATA box nor a Kozak consensus. The 3' end has an untranslated region of 89 bp, including a polyadenylation consensus signal, after the stop codon at position 527. The open reading frame encodes a protein of 170 amino acids with a 30 amino acid signal peptide. The deduced molecular weight of the processed protein is 19.04 kDa with a calculated pI of 9.59. The authenticity of the SOB3 cDNA is indicated by its homology with the microsequenced peptides, and the calculated molecular mass and pI of the deduced amino acid sequence.

The SOB3 sequence was found to be 98 and 100% homologous to the cDNA sequences of prepro-FALL-39 (GenBank accession no. Z38026) and CAP18 (GenBank accession no. U19970) respectively. The open reading frames of SOB3 and prepro-FALL-39 differ by two nucleotides at positions 16 and 117 of the SOB3 sequence. The SOB3 cDNA and the CAP18 genes both contain a similar short untranslated region of 16 nucleotides (Larrick *et al.*, 1996). In the prepro-FALL-39 cDNA, there is an untranslated region made up of only 11 nucleotides, thus differing from that of SOB3 (Agerberth *et al.*, 1995).

Southern blotting showed that human SOB3 is a single copy gene (Figure 3).

Tissue and cell distribution of SOB3 RNA

Northern blots of human tissues yielded a single 1 kb transcript in the corpus and cauda epididymis only (Figure 4). Northern blotting detected no transcripts in the other male genital tract samples, particularly in the testis.

RT-PCR amplification generated a strong 548 bp band with caput, corpus and cauda epididymis RNAs as templates (Figure 5). Two unexplained lighter bands were observed for caput epididymis. There was a faint band in the testis sample. No signal was seen with either 125 round spermatids or 125 spermatozoa under the same conditions, and semi-nested PCR was required to obtain the expected 218 bp band. The pattern obtained with RNA extracted from round spermatids was definitely stronger than that obtained with the same number of spermatozoa.

The SOB3 transcript was abundant in erythroblasts, granulocytes and B-cells (blood B- and Raji cells) (Figure 6A). It was absent from megakaryocytes and T-cells (blood T-

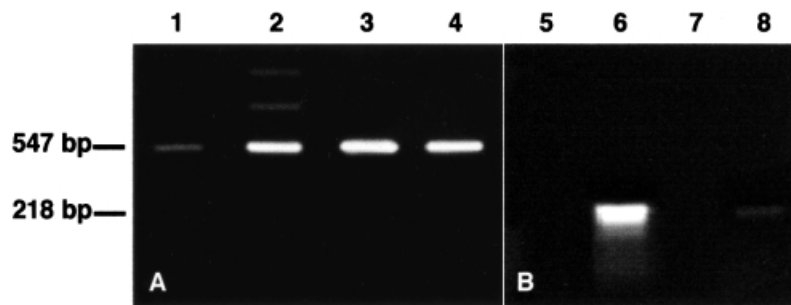


Figure 5. Reverse transcription-polymerase chain reaction (RT-PCR) analysis of SOB3 mRNA in the testis, epididymis, spermatids and spermatozoa. RT-PCR was performed with either 100 ng testis (A1) or epididymal caput (A2), corpus (A3) and cauda (A4) RNA, or RNA extracted from 125 spermatids (B5, B6) and spermatozoa (B7, B8). The whole SOB3 cDNA was amplified in the testis and epididymis samples and there was no visible product in spermatids and spermatozoa using first-round PCR (B5, B7). Semi-nested PCR amplified a 218 bp product in round spermatids (B6) and spermatozoa (B8). Molecular masses are indicated on the left.

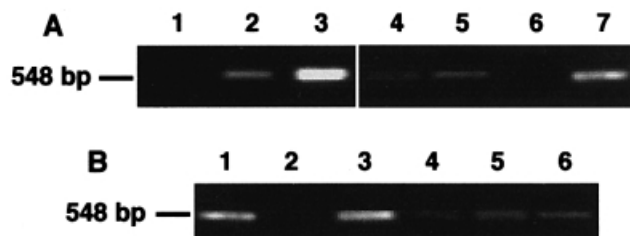


Figure 6. Reverse transcription-polymerase chain reaction (RT-PCR) analysis of SOB3 mRNA in hematopoietic cells. RT-PCR was performed with 100 ng total mRNA of bone marrow cells, blood cells and cell lines (A), and of tonsil cells (B). (A1) Megakaryocytes, (A2) erythroblasts, (A3) granulocytes purified from bone marrow, (A4) blood T- and (A5) B-cells, (A6) MOLT-4 and (A7) Raji cell lines. (B1) Total, (B2) naïve, (B3) GC and (B4) memory tonsil B-cells. (B5) Phytohaemagglutinin-stimulated and (B6) resting tonsil T-lymphocytes. The whole SOB3 cDNA was amplified.

blotting yielded two bands, A and B, with a pI ~9.4. Microsequencing of these two bands suggested homologies with two human antimicrobial proteins, prepro-FALL-39 and CAP18, whose sequences were deduced from cDNA obtained by screening a bone marrow gt11 library (Agerberth *et al.* 1995; Larrick *et al.*, 1995). The near identity between our SOB3 cDNA and the prepro-FALL-39 and CAP18 cDNA sequences suggests that a single gene encodes all three proteins and that there is no alternative splicing. The difference between prepro-FALL-39 and SOB3 results from several nucleotide discrepancies mainly localized in the promoter region. Finally, our cDNA sequence including the promoter is quite similar to the four exons of the CAP18 gene open reading frame (Larrick *et al.*, 1996). The human prepro-FALL-39 and CAP18 proteins are the human equivalent of two proteins purified from pig and rabbit leukocytes because of their antimicrobial characteristics (Ritonja *et al.*, 1989; Larrick *et al.*, 1993). They bind to lipopolysaccharides that are a major constituent of the outer membrane of gram-negative bacteria.

SOB3 was found only in the corpus and cauda epididymal regions using Northern blotting. We found no hybridization signal in any other tissue, even in the testis. RT-PCR showed that SOB3 mRNA is abundant in the epididymis, at a higher level in the corpus and cauda regions. A moderate expression was also detected in the testis. RT-PCR using the same

primers detected no mRNA in 125 spermatids and nested PCR was necessary to obtain a sound signal. This result is not surprising as the amount of RNA contained in 125 spermatids is only ~1 ng. Ejaculated spermatozoa yielded a faint band under the same conditions. These results therefore suggest that SOB3 is expressed in the germinal lineage at the round spermatid stage and in the epididymal epithelium. It is no more surprising that SOB3 was not detected in testis samples by Northern blotting as round spermatids are only a part of the testis cell content. Whether the low product in ejaculated sperm samples after semi-nested RT-PCR is the result of active gene expression or whether this technique detects residual SOB3 gene transcription only remains to be elucidated. It is also likely that spermatozoa in the epididymal lumen are probably not responsible for the abundant products observed by RT-PCR of the three epididymal regions.

Recently published data show that the CAP18 gene is strongly expressed in the epididymis and that there is CAP18 protein in the epididymal epithelium (Malm *et al.*, 2000), confirming our first results (Martin Ruiz *et al.*, 1998). The present study supports all these observations. Results on gene expression in the testis are conflicting, perhaps because of the sensitivity of the techniques used. No CAP18 gene expression in the testis has been detected by in-situ hybridization and immunochemistry (Malm *et al.*, 2000), and this is inconsistent with another study (Agerberth *et al.*, 1995) which detected a positive signal in human testis by Northern blotting. These authors improved the sensitivity of the Northern blots by using a membrane from Clontech blotted with Poly A⁺ RNA. Thus, these results and the present study indicate that the gene coding for prepro-FALL-39/CAP18/SOB3 is expressed in two regions of the male genital tract, in the testis by spermatids and in the epididymal epithelium.

Expression of the SOB3 gene was also detected by RT-PCR in erythroblasts, granulocytes and B-cells purified from human bone marrow and peripheral blood. The greatest expression in the B-cell pool purified from palatine tonsil was in GC cells, while naïve cells were negative. This last result suggests that SOB3 mRNA is restricted to recently activated B-cells *in vivo* (GC and memory B-cells) and, thus, this production may be induced in B-cells by activation. Within lymphoid organs, GC B-cells are in close contact with antigen trapped on follicular

dendritic cells. This may maintain SOB3 gene expression. Also, the Raji cell line, which is also positive for SOB3 in Western blot and RT-PCR experiments, is derived from GC since it originates from a Burkitt's tumour. These observations confirm and broadly extend the data on the distribution of the two homologous mRNA. The *prepro-FALL-39/CAP18* gene was shown to be expressed in purified granulocytes and bone marrow cells (Agerberth *et al.*, 1995; Larrick *et al.*, 1995, 1996; Malm *et al.*, 2000), while an electron microscopy study detected CAP18 in myelocytes and metamyelocytes and localized it in specific granules in neutrophils (Sorensen *et al.*, 1997).

Our RT-PCR experiments showed only one product in the testis and epididymis. The same 548 bp product was observed using cDNA from the testis, epididymis, erythroblast, B-lymphocyte, and tonsil cells. In a previous study, Western blots of sperm protein extracts yielded two immunoreactive bands of 18 and 19 kDa (Martin Ruiz *et al.*, 1998), and a single 18 kDa band was seen in erythrocyte extracts and a 19 kDa band in B-lymphocytes. The same doublet band was recently found in human seminal plasma and sperm extracts, while neutrophils displayed only the upper band (Malm *et al.*, 2000). As alternative splicing does not seem to account for the two molecular weights observed, these patterns may be due to post-translational modifications and cell-specific charged residues.

Our first studies using immunochemistry indicated two sperm locations for SOB3, in the acrosome or on the inner acrosomal membrane, and in the neck and tail regions (Martin Ruiz *et al.*, 1998). At that time, the protein seemed to be secreted by the epididymis and it was difficult to envisage how it successively bound to the spermatozoon, crossed the acrosomal membrane and settled inside the acrosome. The present RT-PCR analysis also detected *SOB3* gene activity in round spermatids, which sheds a new light on this problem and may explain the intra-acrosomal location of the protein.

Sperm surface proteins that may be involved in mammalian gamete interactions could be processed during spermiogenesis or epididymal maturation. Both situations have been described. SP-10 is expressed at the spermatid stage in humans (Kurth *et al.*, 1993), while P34H is secreted by the epididymal epithelium and added to spermatozoa as they pass through the epididymis (Légaré *et al.*, 1999). Another protein, PH-20, is attached to the membrane via a GPI anchor and probably migrates from the post-acrosomal plasma membrane to the inner acrosomal membrane during sperm maturation by diffusion, as observed in guinea-pig and Rhesus monkey spermatozoa (Phelps *et al.*, 1990; Overstreet *et al.*, 1995). Lastly, P26h, the hamster homologue of P34H, originates both from the testicular fluid as a secretory product of the seminiferous tubules and from the epididymal epithelium (Gaudreault *et al.*, 1999).

Our observations show that the SOB3 protein is produced at two sites and logic suggests that it has two different roles. When synthesized during spermiogenesis, SOB3 seems to be localized in the acrosome and might be involved in the secondary binding to zona pellucida. Alternatively, when synthesized in the epididymal lumen, SOB3 might help protect

spermatozoa against bacterial injury, as proposed by studies on CAP18 and prepro-FALL-39. It is interesting to note that defensins belong to the family of antimicrobial peptides that includes SOB3/prepro-FALL-39/CAP18 and have been strongly conserved during evolution. In mammals, they are also involved in cell mobilization and contact as they are potent chemo-attractants for dendritic cells and T-lymphocytes (Yang *et al.*, 1999).

The antimicrobial function of CAP18 is likely to depend on an α -helical structure in the C-terminal domain, which binds to lipopolysaccharides (Larrick *et al.*, 1993). It has been suggested that, because of its attachment to sperm surface detected by flow cytometry, CAP18 also has a role in conception (Malm *et al.*, 2000). In the light of our previous results, SOB3 seems to be involved in binding to the zona pellucida. The great conservation of the NH₂ domain among species described by Agerberth *et al.* (1995) suggests that this region plays such an essential role. Further research is clearly required to answer this question.

Acknowledgements

We thank Dr Ahmed Ziyat (INSERM U257) and Fawzia Louache (INSERM U362) for providing RNA from spermatids and bone marrow cells and for fruitful discussions. We are grateful to the technicians of the Laboratoire d'Assistance Médicale à la Procréation of the American Hospital of Paris (92202 Neuilly, France) for providing sperm samples. Thanks are also due to Owen Parkes for English reading of the manuscript.

References

- Adham, M.I., Nayiermia, K., Engel, W. *et al.* (1997) Spermatozoa lacking acrosin protein show delayed fertilization. *Mol. Reprod. Dev.*, **46**, 370–376.
- Agerberth, B., Gunne, H., Odeberg, J. *et al.* (1995) FALL-39, a putative human peptide antibiotic, is cysteine-free and expressed in bone marrow and testis. *Proc. Natl. Acad. Sci. USA*, **92**, 195–199.
- Baba, T., Azuma, S., Kashiwabara, S. *et al.* (1994) Sperm from mice carrying a targeted mutation of the acrosin gene can penetrate the oocyte zona pellucida and effect fertilization. *J. Biol. Chem.*, **269**, 31845–31849.
- Benoff, S. (1997) Carbohydrates and fertilization: an overview. *Mol. Hum. Reprod.* **3**, 599–637.
- Boué, F., Lassalle, B., Duquenne, C. *et al.* (1992) Human sperm proteins from testicular and epididymal origin that participate in fertilization: modulation of sperm binding to zona-free hamster oocytes using monoclonal antibodies. *Mol. Reprod. Dev.*, **33**, 470–480.
- Cowan, A.E., Myles, D.G. and Koppel, D.E. (1991) Migration of the guinea pig sperm membrane protein PH-20 from one localized surface domain to another does not occur by a simple diffusion-trapping mechanism. *Dev. Biol.*, **144**, 189–198.
- Gaudreault, C., Légaré, C., Bérubé, B. and Sullivan, R. (1999) Hamster sperm protein, P26h: a member of the short-chain dehydrogenase/reductase superfamily. *Biol. Reprod.*, **61**, 264–273.
- Görg, A., Obermaier, C., Boguth, G. *et al.* (1997) Very alkaline immobilized pH gradients for two-dimensional electrophoresis of ribosomal and nuclear proteins. *Electrophoresis*, **18**, 328–337.
- Hamatani, T., Tanabe, K., Kamei, K. *et al.* (2000) A monoclonal antibody to human SP-10 inhibits *in vitro* the binding of human sperm to hamster oolemma but not to human zona pellucida. *Biol. Reprod.*, **62**, 1201–1208.
- Hunnicut, G.R., Primakoff, P. and Myles, D.G. (1996) The sperm surface protein PH-20 is bifunctional: one activity is a hyaluronidase and a second, distinct activity is required in secondary sperm-zona binding. *Biol. Reprod.*, **55**, 80–86.
- Jansen, S., Jones, R., Jenneckens, I. *et al.* (1998) Site-directed mutagenesis of boar proacrosin reveals residues involved in binding of zona pellucida glycoproteins. *Mol. Reprod. Dev.*, **51**, 184–92.

- Jury, J.A., Perry, A.C.F. and Hall, L. (1999) Identification, sequence analysis and expression of transcripts encoding a putative metalloproteinase, eMDC II, in human and macaque epididymis. *Mol. Hum. Reprod.*, **5**, 1127–1134.
- Krzysiek, R., Lefèvre, E.A., Zou, W. *et al.* (1999) Antigen receptor engagement selectively induces macrophage inflammatory protein-1 (MIP-1) and MIP-1 chemokine production in human B cells. *J. Immunol.*, **162**, 4455–4463.
- Kurth, B.E., Wright, R.M., Flickinger, C.J. and Herr, J.C. (1993) Stage-specific detection of mRNA for the sperm antigen SP-10 in human testes. *Anat. Rec.*, **236**, 619–625.
- Larrick, J.W., Hirata, M., Shimomura, Y. *et al.* (1993) Antimicrobial activity of rabbit CAP18-derived peptides. *Antimicrob. Agents Chemother.*, **37**, 2534–2540.
- Larrick, J.W., Hirata, M., Balint, R.F. *et al.* (1995) Human CAP18: a novel antimicrobial lipopolysaccharide-binding protein. *Infect. Immun.*, **63**, 1291–1297.
- Larrick, J.W., Lee, J., Ma, S. *et al.* (1996) Structural, functional analysis and localization of the human CAP18 gene. *FEBS Lett.*, **398**, 74–80.
- Lassalle, B., Ziyat, A., Testart, J. *et al.* (1999) Flow cytometric method to isolate round spermatids from mouse testis. *Hum. Reprod.*, **14**, 101–107.
- Lefèvre, A., Martin Ruiz, C., Chokomian, S. *et al.* (1997) Characterization and isolation of SOB2, a human sperm protein with a potential role in oocyte membrane binding. *Mol. Reprod. Dev.*, **3**, 507–516.
- Légaré, C., Gaudreault, C., Saint-Jacques, S. and Sullivan, R. (1999) P34H sperm protein is preferentially expressed by the human corpus epididymis. *Endocrinology*, **140**, 3318–3327.
- Lin, Y., Mahan, K., Lathrop, W.F. *et al.* (1994) A hyaluronidase activity of the sperm plasma membrane protein PH-20 enables to penetrate the cumulus cell layer surrounding the egg. *J. Cell Biol.*, **125**, 1157–1163.
- Malm, J., Sorensen, O., Persson, T. *et al.* (2000) The human cationic antimicrobial protein (hCAP-18) is expressed in the epithelium of human epididymis, is present in seminal plasma at high concentrations, and is attached to spermatozoa. *Infect. Immun.*, **68**, 4297–4302.
- Martin Ruiz, C., Duquenne, C., Treton, D. *et al.* (1998) SOB3, a human sperm protein involved in zona pellucida binding: physiological and biochemical analysis, purification. *Mol. Reprod. Dev.*, **49**, 286–297.
- Moreno, R.D., Sepulveda, M.S., de Ioannes, A. *et al.* (1998) The polysulfate binding domain of human proacrosin/acrosin is involved in both the enzyme activation and spermatozoa-zona pellucida interaction. *FEBS Lett.*, **6**, 75–83.
- Myles, D.G. and Primakoff, P. (1984) Localized surface antigens of guinea pig sperm migrate to new regions prior to fertilization. *J. Cell Biol.*, **99**, 1634–1641.
- Myles, D.G. and Primakoff, P. (1997) Why did the sperm cross the cumulus? To get to the oocyte functions of the sperm surface proteins PH-20 and fertilin in arriving at, and fusing with, the egg. *Biol. Reprod.*, **56**, 320–327.
- Overstreet, J.W., Lin, Y., Yudin, A.I. *et al.* (1995) Location of PH-20 protein on acrosome-intact and acrosome-reacted spermatozoa of cynomolgus macaques. *Biol. Reprod.*, **52**, 105–114.
- Phelps, B.M., Koppel, D.E., Primakoff, P. and Myles, D.G. (1990) Evidence that proteolysis of the surface is an initial step in the mechanism of formation of sperm cell surface domains. *J. Cell Biol.*, **111**, 1839–1847.
- Primakoff, P., Hyatt, H. and Myles, D.G. (1987) A role for the migration sperm surface antigen PH-20 in guinea pig sperm binding to the egg zona pellucida. *J. Cell Biol.*, **104**, 141–149.
- Reinhart, M.P. and Malamud, D. (1982) Protein transfer from isoelectric focusing gels: the native blot. *Anal. Biochem.*, **123**, 229–235.
- Richardson, R.T. and O’Rand, M.G. (1996) Site directed mutagenesis of rabbit proacrosin. Identification of residues involved in zona pellucida binding. *J. Biol. Chem.*, **271**, 24069–24074.
- Ritonja, A., Kopitar, M., Jerala, R. and Turk, V. (1989) Primary structure of a new cysteine proteinase inhibitor from pig leucocytes. *FEBS Lett.*, **255**, 211–214.
- Robertson, E.F., Dannelly, H.K., Malloy, P.J. *et al.* (1987) Rapid isoelectric focusing in a vertical polyacrylamide minigel system. *Anal. Biochem.*, **167**, 290–294.
- Sorensen, O., Arnljots, K., Cowland, J.B. *et al.* (1997) The human antibacterial cathelicidin, hCAP-18, is synthesized in myelocytes and localized to specific granules in neutrophils. *Blood*, **90**, 2796–2803.
- Westbrook, V.A., Diekman, A.B., Klotz, K.L. *et al.* (2000) Spermatid-specific expression of the novel X-linked gene product SPAN-X localized to the nucleus of human spermatozoa. *Biol. Reprod.*, **63**, 469–481.
- Yang, D., Chertov, O., Bykovskaia, S.N. *et al.* (1999) Beta-defensins: linking innate and adaptive immunity through dendritic and T cell CCR6. *Science*, **286**, 525–528.
- Yudin, A.I., Vandervoort, C.A., Li, M.W. *et al.* (1999) PH-20 but not acrosin is involved in sperm penetration of the macaque zona pellucida. *Mol. Reprod. Dev.*, **53**, 350–362.
- Ziyat, A., Lassalle, B., Testart, J. *et al.* (1999) Flow cytometry isolation and reverse transcriptase-polymerase chain reaction characterization of human round spermatids in infertile patients. *Hum. Reprod.*, **14**, 379–387.

Received on December 29, 2000; accepted on April 2, 2001