Characterization of Sp17: a ubiquitous three domain protein that binds heparin

Ying WEN[1], Richard T. RICHARDSON, Esther E. WIDGREN and Michael G. O'RAND[2]

Department of Cell Biology and Anatomy, CB#7090, 210 Taylor Hall, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, U.S.A.

Sp17 is a protein that was originally thought to be expressed exclusively in the testis and whose primary function was binding to the extracellular matrix of the oocyte. Several recent reports have implicated Sp17 as having a role in cell–cell adhesion and/or cell migration in transformed, lymphocytic and haematopoietic cells, possibly through its interaction with extracellular heparan sulphate. In the present study, we report that Sp17's central domain (amino acids 61–117), spanning exon 3, is critical for heparin binding. Sp17 has two additional functional domains, an N-terminal domain similar to the dimer-interaction site in the cAMP-dependent protein kinase IIα regulatory subunit and a C-terminal calmodulin-binding domain. The mouse gene for Sp17 is 6.5 kb and contains four exons. Although Sp17 expression is highest in the testis, it is present in all of the mouse somatic tissues examined and is highly conserved throughout all mammalian species. Sp17's central domain, which is necessary for heparin binding, exhibits the greatest sequence divergence of all three domains. The *Sp17* gene is induced in metastatic cells and during mucosal immune responses, and the protein appears to play an important role in cell migration and/or adhesion in somatic cells, as well as in male germ cells.

Key words: calmodulin, calmodulin-binding protein, heparan-binding protein, sulphated-carbohydrate-binding protein, syndecan.

INTRODUCTION

The extracellular matrix influences the behaviour of cells through its interaction with the cell's repertoire of surface receptors. Perhaps nowhere else is this more dramatic than in the sudden release of the apical granule (acrosome) of the sperm cell as it encounters the oocyte–cumulus complex, undergoing the acrosome reaction in preparation for penetrating the zona pellucida. To adhere to the extracellular matrix of the oocyte, spermatozoa employ one of the most adhesive, high-affinity (K_D approx. 5.6×10^{-13} M) binding molecules present after the acrosome reaction, the zona pellucida-binding protein, Sp17 [1,2]. This molecule, present in every mammalian species examined, was first described as a member of the rabbit sperm antigen family of proteins, with lectin-like properties that bound sulphated complex carbohydrates, such as fucoidin, dextran sulphate and heparin [2]. Originally thought to be expressed exclusively in the testis and to be present on the surface of acrosome-reacted spermatozoa, Sp17's primary function was binding to the extracellular matrix of the oocyte [1]. Moreover, the observation that Sp17-transfected COS cells expressed Sp17 on their surfaces and bound native zona pellucida [3] reinforced this hypothesis, in spite of the fact that Sp17 lacks a signal sequence or an obvious transmembrane domain.

Recently, new data about Sp17 have been reported from transformed cells and normal somatic cells. Using mRNA differential display, Dong et al. [4] found an increased expression of Sp17 in three metastatic cell lines derived from an *in vitro* transformed keratinocyte line (Pam 212). Lacy and Sanderson [5] reported Sp17 to be expressed by human haematopoietic cell lines, myeloma cells and normal peripheral blood mononuclear cells. Present on the surface of these cells, Sp17 evidently promotes cell–cell adhesion through its interaction with syndecan-1 heparan sulphate chains. Similarly, during a mucosal immune response to bovine rotavirus infection, *Sp17* gene expression is induced, leading to the hypothesis that Sp17 is involved in immune cell migration [6]. Consequently, Sp17 expression is neither testis-specific, as originally reported from Northern-blot data [1], nor is it exclusively for binding to the zona pellucida of the oocyte. Indeed, the three initial protein domains proposed for Sp17, the N-terminal dimer interaction domain, the sulphated carbohydrate-binding domain and the C-terminal calmodulin-binding domain [7], would all seem to be of potential importance in somatic cells; the dimer interaction site for interaction with A-kinase anchoring proteins (AKAPs) [8–10], and the calmodulin-binding site for the proteolytic release of calmodulin at the appropriate time and place [11]. The sulphated carbohydrate-binding domain appears to be of direct relevance to Sp17's binding to heparan sulphate chains of syndecan and its ability to aggregate certain transformed cells.

Because of the importance of Sp17's ability to bind and aggregate transformed cells, in the present paper we have examined the binding of Sp17 to the sulphated carbohydrate, heparin. We also report that Sp17 is present in all somatic tissues examined, although at a lower level than in the testis, and that the major sulphated carbohydrate-binding domain appears to span exon 3. From a comparison of Sp17 protein sequences from nine different species, it appears that the sulphated carbohydrate-binding domain is the most variable part of the protein.

MATERIALS AND METHODS

Materials

All chemicals and reagents were of molecular biology grade. Restriction enzymes and *Taq* DNA polymerase were purchased from Boehringer Mannheim. DNA was sequenced in both forward and reverse directions at the University of North

Abbreviations used: AKAP, A-kinase anchoring protein; ORF, open reading frame.

[1] Present address: Department of Anatomy and Cell Biology, College of Physicians and Surgeons, 630 West 168th Street, Columbia University, New York, NY 10032, U.S.A.

[2] To whom correspondence should be addressed (morand@unc.edu).

Carolina at Chapel Hill Automated DNA Sequencing Facility on a Model 373A DNA Sequencer (Applied Biosystems; Foster City, CA, U.S.A.) using the Taq DyeDeoxy™ Terminator Cycle Sequencing kit (Applied Biosystems).

Bacterial expression and purification of recombinant proteins

Recombinant rabbit Sp17 (full length) and Sp17-N-terminal (Sp17-$N_{1–117}$) were made as described previously [11]. To make recombinant rabbit Sp17-C-terminal (Sp17-$C_{61–146}$) and short Sp17-N-terminal (Sp17-$N_{1–60}$), DNA was amplified by PCR using specific primers containing 5′ *Bam*HI and 3′ *Kpn*I restriction sites. The sense primer for Sp17-$C_{61–146}$ was 5′-CGCGGATCC-GGGGCTAAGGTTGATGACCGCTTC-3′ and the antisense primer was 5′-CGGGGTACCGCCAGTGCCCTCAATTGT-3′. The sense primer for Sp17-$N_{1–60}$ was 5′-CGCGGATCCA-TGTCGATTCCATTTTCC-3′ and the antisense primer was 5′-CGGGGTACCTCATTCAGCTGGATCAAAGTT-3′. The amplified PCR products were cloned into pQE-30 pre-digested with *Bam*HI and *Kpn*I, and the correctness of the inserts verified by DNA sequencing. The recombinant DNA was transformed into bacterial cells and recombinant proteins purified using the Ni^{2+} nitrilotriacetate column, as described previously [1].

Heparin-binding assays

All binding assays were performed at room temperature. To perform the [^{125}I]heparin–albumin binding assay, 5 pmol of recombinant protein were air-dried on Immulon 'Removawell' strips (Dynatech Laboratories, Chantilly, VA, U.S.A.) and blocked for 90 min in the heparin-binding buffer (50 mM Tris, pH 7.5, 150 mM NaCl, 1 mM $CaCl_2$ and 0.05 % Tween 20) with 5 % BSA. The plate was incubated in the same buffer containing increasing amounts of [^{125}I]heparin–albumin for 150 min, washed with the heparin-binding buffer three times, and the amount of heparin bound was detected by removing individual wells from the strip and counting in a Pharmacia γ-radiation counter. To show the specificity of the binding, the plate was incubated with a saturating amount (0.8 μg) of [^{125}I]heparin–albumin in the presence of increasing amounts of unlabelled heparin.

PCR of Sp17 from multiple tissues

The sense primer for amplifying the Sp17 open reading frame (ORF) from Mouse Multiple Tissue cDNA Panel I (ClonTech Laboratories) was 5′-ATGTCGATTCCTTTCTCCAACACC-CAC-3′ and the antisense primer was 5′-TCAATTGTCTGCC-TCTTCTTTCAGATTCTC-3′. PCR was performed with the following conditions: cycle 1, 94 °C for 2 min 30 s, 58 °C for 1 min 30 s and 72 °C for 2 min; cycles 2–30, 94 °C for 45 s, 58 °C for 1 min 30 s and 72 °C for 2 min; cycle 31, 94 °C for 45 s, 58 °C for 1 min 30 s and 72 °C for 7 min. The PCR products were analysed on 1 % agarose gels (Fisher Scientific).

mRNA dot blot

Mouse RNA Master Blot was purchased from ClonTech Laboratories and probed with mouse testis Sp17 ORF. The probe was made using the North2South™ Biotin Random Prime kit (Pierce, Rockford, IL, U.S.A.) and hybridized and developed using the North2South™ Chemiluminescent Nucleic Acid Hybridization and Detection kit (Pierce). The level of Sp17 expression was analysed with GelExpert software (Nucleotech Corp., San Carlos, CA, U.S.A.).

Cloning of mouse heart and kidney Sp17

Mouse heart and kidney polyadenylated RNA were purchased from ClonTech Laboratories. Mouse heart and kidney Sp17 ORFs were obtained by performing 3′-rapid amplification of cDNA ends (Gibco BRL), using gene-specific primers to mouse testis Sp17.

Western blot

The mouse multiple tissue blot was purchased from Geno Technology (St Louis, MO, U.S.A.). Tissue lysate (75 μg) was loaded into each lane. The blot was probed with affinity-purified rabbit anti-rabbit Sp17, as described previously [11]. Rabbit anti-rabbit Sp17 was made as described by Richardson et al. [1], and affinity-purified using the SulfoLink Coupling Gel (Pierce) coupled with recombinant rabbit Sp17. The blot shown in Figure 3 was probed with rabbit anti-recombinant human Sp17 [12] and mouse antiserum against an Sp17 peptide (A9DT; AEWGAK-VED), prepared as described previously [13]. Mouse myeloma cell lysate was prepared by homogenization of 2×10^8 cells/ml in PBS (pH 7.4), containing proteasome inhibitors and protease inhibitor cocktail I (Calbiochem). Following centrifugation at 12000 ***g***, the supernatant was mixed with an equal volume of sample buffer and a portion loaded on to the SDS gel as described in [1].

Immunofluorescence

Mouse 3T3 cells growing on microscope slides were washed in PBS, fixed in methanol at −20 °C for 10 min, washed, and stained with antiserum or control serum (1/100 dilution) for 1.5 h at room temperature. Cells were viewed with a Zeiss fluorescence microscope with a 100× oil objective.

Sp17 genomic structure

A total of 16 primers were used to amplify potential intron-spanning Sp17 genomic DNA fragments from the mouse 129/SvJ lambda genomic library (Stratagene). Primers were designed according to the mouse Sp17 sequence reported by Kong et al. [14]. PCR was performed as described above. When the PCR product size was larger than the expected cDNA size, the product was purified using the QIAquick PCR Purification kit (Qiagen) and cloned into pCR2.1®-TOPO vector using the TOPO TA cloning kit (Invitrogen) for sequencing. The plasmid was purified using QIAprep Spin Miniprep kit (Qiagen). The exon–intron junctions were determined by comparing the genomic sequence with the cDNA sequence.

Multiple sequence alignment and phylogenetic analysis

Sp17 sequence alignment and phylogenetic analyses were performed using GCG Software packages (Genetic Computer Group, University of Wisconsin, Madison, WI, U.S.A.). Sequence alignment was generated by the PILEUP program. The PAUP program [15] was used to construct the phylogeny of Sp17. Sp17 has been identified and sequenced from rabbit (*Oryctolagus cuniculus*; [1]), mouse (*Mus musculus*; [14]), human (*Homo sapiens*; [12]), baboon (*Papio hamadryas*; [16]), macaque (*Macaca fascicularis*; [17]), rat (*Rattus norvegicus*, GenBank® accession number AJ131888), wallaby (*Macropus eugenii*, GenBank® accession number 3915005), *Monodelphis* (*Monodelphis domestica*, GenBank® accession number 3915006), and

marmoset (*Callithrix jacchus*, GenBank® accession number AF134585). In addition, *Macaca mulatta* Sp17 cDNA was sequenced and found to be 100% identical with *Macaca fascicularis* (Y. Wen, unpublished work). 'Zoo' Southern blots have shown that Sp17 is also present in dogs and cows, but not in chickens or yeast (*Saccharomyces cerevisiae*; [12]).

RESULTS AND DISCUSSION

Sulphated carbohydrate (heparin) binding

Because of Sp17's ability to aggregate transformed cells via syndecan heparan sulphate chains [5], and to bind spermatozoa to the zona pellucida [1], we have investigated the major sulphated carbohydrate-binding domain in the Sp17 sequence. Four recombinant proteins, Sp17, Sp17-$C_{61–146}$, Sp17-$N_{1–117}$ and Sp17-N_{1-60} (Figures 1 and Figure 2A), were examined for their heparin-binding activities. Full-length recombinant Sp17, Sp17-$N_{1–117}$ and Sp17-$C_{61–146}$ all bound [^{125}I]heparin–albumin in a saturable manner, but Sp17 bound somewhat more heparin than the shorter, mutant recombinant proteins. The binding activity of Sp17-$N_{1–60}$ was at background levels (Figure 2A). Scatchard-plot [18] analysis (Figure 2B) demonstrated that the Sp17, Sp17-$C_{61–146}$ and Sp17-$N_{1–117}$ binding affinities overlap and were not significantly different from one another (Figure 1). However, Sp17 had more heparin-binding sites than either Sp17-$C_{61–146}$ or Sp17-$N_{1–117}$ (Figure 2B). Therefore there appear to be at least two high-affinity heparin-binding sites in Sp17. The interaction between recombinant proteins and [^{125}I]heparin–albumin is specific since it can be competed away with increasing amounts of unlabelled heparin (Figure 2C). The potential heparin-binding motif, BBXB (where B is any basic amino acid) [19], occurs at amino acid 49–52 (KREK) and at amino acid 131–134 (KKIR), consequently heparin binding may be contingent upon at least one of these sites and the proper conformation of the region between amino acid residues 61–117, which is the overlapping region between Sp17, Sp17-$N_{1–117}$ and Sp17-$C_{61–146}$. This conclusion is supported by the observation that the short N-terminal fragment Sp17-$N_{1–60}$ does not have heparin-binding activity, although it does contain amino acid residues 49–52.

Expression of Sp17 in somatic tissues

Although Sp17 was originally described as a sperm- and testis-specific protein [1], it has now been reported in human metastatic cell lines derived from an *in vitro* transformed keratinocyte line (Pam 212) using mRNA differential display and in human haematopoietic cell lines, myeloma cells, normal peripheral blood mononuclear cells and sheep lymphocytes [4–6]. Sp17 is thought to promote cell–cell adhesion through its interaction with syndecan-1 heparan sulphate chains on the surface of these cells. To determine whether Sp17 is present in mouse somatic tissues,

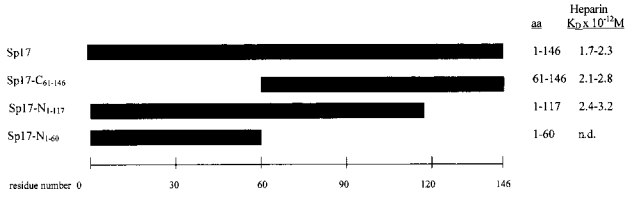

Figure 1 Heparin binding of Sp17 recombinant constructs

The K_D was calculated based on heparin–albumin having a molecular mass of 102–138 kDa. Abbreviations: aa, amino acid residues; n.d., not determined.

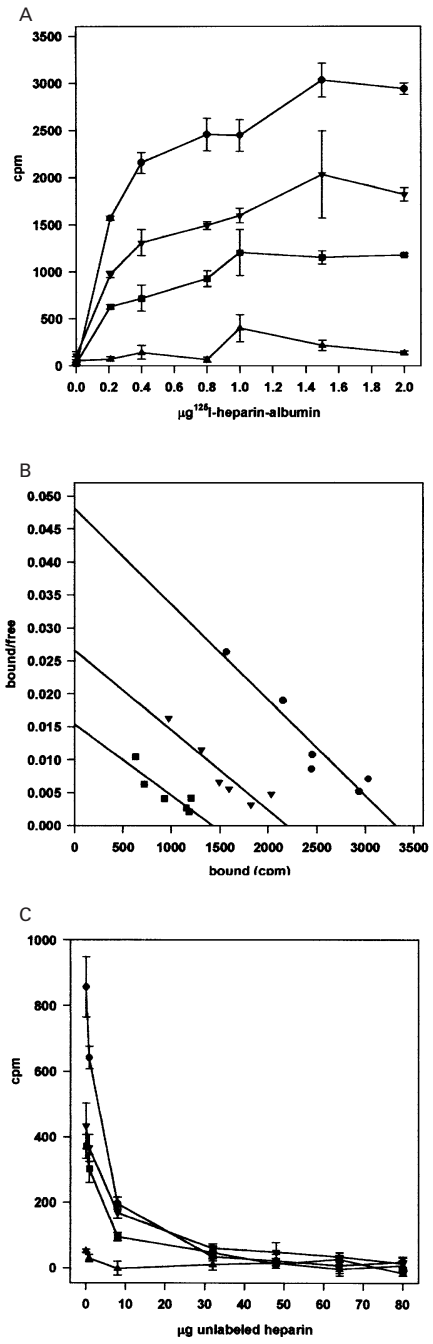

Figure 2 Heparin Sp17 binding

(**A**) Binding of [^{125}I]heparin–albumin to recombinant rabbit Sp17 constructs. Sp17 (●), Sp17-$N_{1–117}$ (■), Sp17-$N_{1–60}$ (▲) and Sp17-$C_{61–146}$ (▼). (**B**) Scatchard plot of [^{125}I]heparin–albumin binding to Sp17 (●), Sp17-$N_{1–117}$ (■) and Sp17-$C_{61–146}$ (▼). (**C**) Displacement of [^{125}I]heparin–albumin binding to Sp17 (●), Sp17-$N_{1–117}$ (■), Sp17-$N_{1–60}$ (▲) and Sp17-$C_{61–146}$ (▼) with increasing amounts of unlabelled heparin.

PCR was performed on a Mouse Multiple Tissue cDNA Panel I and on ovary and thymus cDNA, using primers to amplify the entire mouse Sp17 ORF. The amounts of cDNA were normalized to six different housekeeping genes in all tissues except ovary and thymus. The presence of Sp17, as shown by a 450 bp PCR product, is predominantly in testis, but is also in thymus, ovary, kidney, skeletal muscle, lung, brain and heart (Figure 3A).

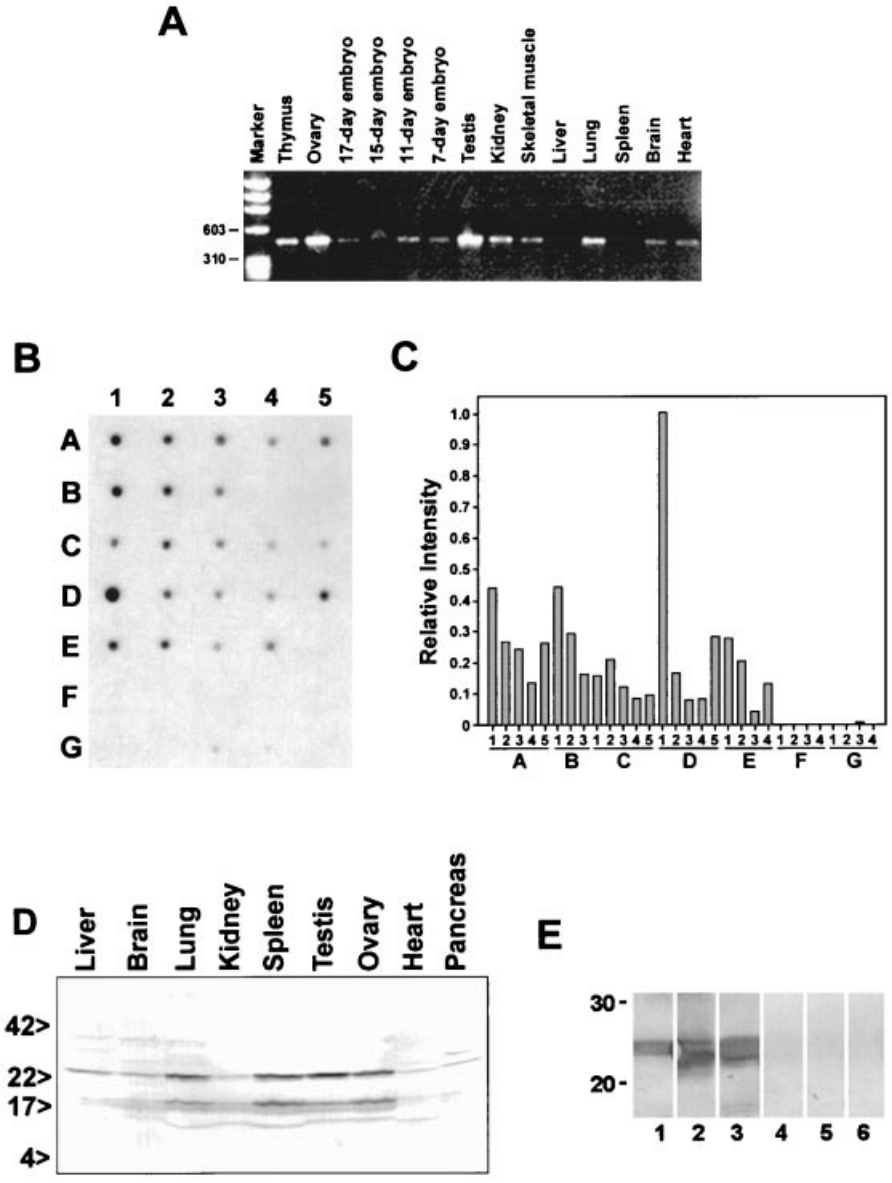

Figure 3 Somatic Sp17

(**A**) PCR of Sp17 from different tissues and different stages of embryos. The amounts of cDNA were normalized except in ovary and thymus. Tissues are identified above each lane. The markers are φX174 RF DNA/*Hae*III fragments (Gibco BRL). (**B**) Mouse RNA Master blot probed with Sp17 ORF DNA. The amounts of cDNA were normalized. The tissue identities are: A1, brain; A2, eye; A3, liver; A4, lung; A5, kidney; B1, heart; B2, skeletal muscle; B3, smooth muscle; C1, pancreas; C2, thyroid; C3, thymus; C4, submaxillary gland; C5, spleen; D1, testis; D2, ovary; D3, prostate; D4, epididymis; D5, uterus; E1, 7-day embryo; E2, 11-day embryo; E3, 15-day embryo; E4, 17-day embryo; F1, yeast total RNA; F2, yeast tRNA; F3, *E. coli* rRNA; F4, *E. coli* DNA; G1, polyadenylated RNA; G2, C_0t 1 DNA; G3 and G4, mouse DNA. B4, B5, E5, F5, and G5 are blank. (**C**) Relative intensities of *Sp17* gene expression. The intensity of testis Sp17 expression was arbitrarily assigned as 1.0 and the intensity of Sp17 expression from other tissues was compared relative with the testis. Tissue identities are the same as in (**B**). (**D**) Mouse multiple tissue Western blot. Tissue lysate (75 μg) was loaded into each lane. The blot was probed with affinity-purified rabbit anti-rabbit Sp17 (1:500). (**E**) Western blot following SDS/PAGE on a 4–15% gradient gel probed with mouse anti-Sp17 peptide (A9DT), lanes 1–3; and pre-immune control antibody, lanes 4–6. Lane 1, 20 μl of mouse myeloma cell lysate; lane 2, 5 μg of mouse recombinant Sp17; lane 3, 9.6 μg of human recombinant Sp17. The doublet band seen is at 24 kDa.

Additionally, Sp17 is present in 11-, 15- and 17-day embryos (Figure 3A). Sequencing of the Sp17 PCR products from mouse ovary and kidney revealed that somatic Sp17s have identical sequences with mouse testis Sp17 (GenBank® accession number 2842646; [14]).

To confirm the presence of Sp17 in somatic tissues, the expression of Sp17 mRNA was examined by probing a mouse RNA Master blot with the Sp17 ORF cDNA. Sp17 mRNA is expressed in all tissues, including liver and spleen, as well as in 11-, 15- and 17-day embryos (Figure 3B). It is not expressed in yeast or *Escherichia coli*. When the level of Sp17 mRNA expression was analysed, the highest expression is found in the testis as expected (Figure 3C). In comparison with testis, brain and heart expressed $> 40\,\%$ of the testis level, whereas eye, liver, kidney, skeletal muscle, thyroid and uterus expressed $> 20\,\%$ of the testis level. Low expression ($< 20\,\%$) of Sp17 mRNA is found in lung, smooth muscle, pancreas, thymus, submaxillary gland, spleen, ovary, prostate and epididymis. In embryos, Sp17

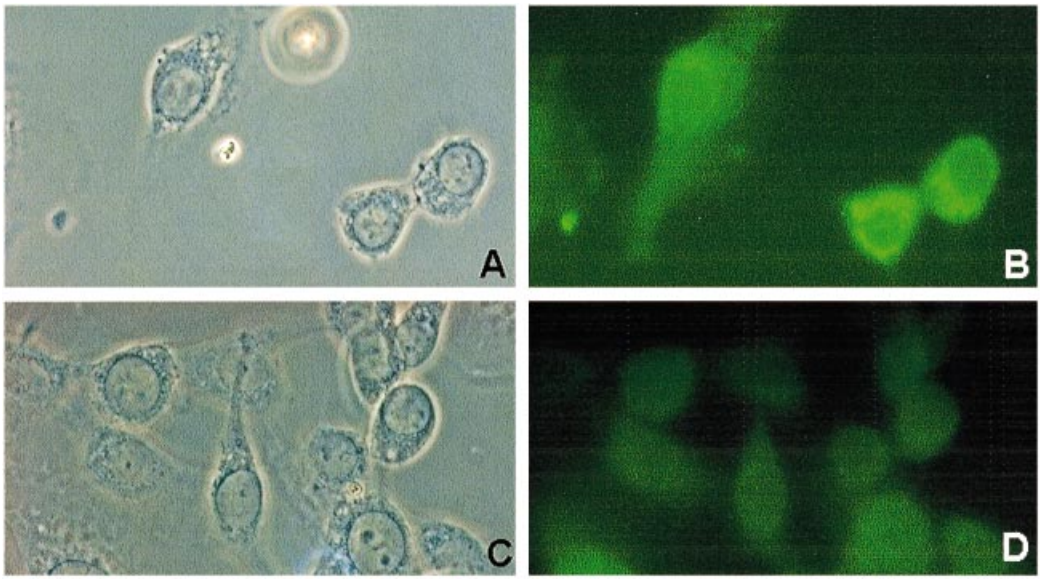

Figure 4 Immunofluorescent localization of Sp17 in mouse 3T3 cells with mouse anti-A9DT antiserum (B) and pre-immune control antibody (D)

Phase and contrast views are shown in (**A**) and (**C**) respectively. Magnification $\times 1000$.

mRNA expression is the highest in 7-day embryos and lowest in 15-day embryos. To further confirm Sp17 in somatic tissues, heart and kidney Sp17 mRNAs were obtained from mouse heart and kidney polyadenylated RNA and sequenced. The somatic sequences are identical with mouse testis Sp17 sequence [14]. Therefore we conclude that Sp17 mRNA is expressed in somatic mouse tissues and that it is identical with the testis form.

To confirm that Sp17 mRNA is translated in somatic tissues, a mouse multiple-tissue Western blot was probed with affinity-purified rabbit anti-rabbit Sp17 antibodies [11]. The antibody recognized bands at 24 kDa and 18 kDa (Figure 3D). The strongest antibody staining is seen in the testis, followed by lung, spleen and ovary. Weaker staining was seen in liver, brain, kidney, heart and pancreas. In the testis, the difference between the 24 kDa and 18 kDa bands is accounted for by the loss of the C-terminal calmodulin-binding domain [11].

Mouse myeloma cell lysates were probed on a Western blot for Sp17 and exhibit the same doublet of antibody staining bands at 26 kDa (Figure 3E, lane 1) as mouse and human recombinant Sp17 (Figure 3E, lanes 2 and 3). The presence of Sp17 in mouse 3T3 cells was also shown by Western blots (results not shown) and confirmed by immunofluorescence (Figure 4B). Sp17 is present throughout the cytoplasm and is particularly concentrated around the nucleus. Whether Sp17 is present on the surface of 3T3 cells was not determined; however, Sp17 is present on the surface of B-lymphocytes and granulocytes. It is weakly expressed on the surface of T lymphocytes and monocytes, and present on the surface of most malignant B and T lymphocyte cell lines (H. M. Lacy and R. D. Sanderson, personal communication).

Mouse Sp17 genomic structure

Sp17 is a highly conserved protein whose sequence is identical in both somatic and testicular cells. In comparing mouse, rabbit and human sequences [12], a striking degree of identity at both the DNA and protein level was observed. To obtain a complete picture of Sp17, and the high-affinity heparin-binding site in Sp17 located between amino acids 61–117, we mapped the mouse *Sp17* genomic sequence. Figure 5 is a schematic diagram of the mouse *Sp17* genomic structure in the coding region. The gene is approx. 6.5 kb and contains four exons. The sizes of the exons and introns and exon–intron junctions are summarized in Table 1. The three exons in the ORF encode the highly conserved N-terminal domain (amino acids 1–75), a variable middle domain (amino acids 76–101) and a conserved calmodulin-binding C-terminal domain (amino acids 102–149) [11]. The mouse genomic structure approximates the structure–function protein motif, each exon encoding a functional domain of the protein. Unlike human Sp17, in which the N-terminal 75 amino acids are contained in two exons [12], the N-terminal 75 amino acids of mouse Sp17 contain only one exon (Figure 5). Therefore although mouse and human Sp17 N-terminal DNA (93% identity) and protein (94% identity) sequences are highly conserved, their genomic structures are not conserved.

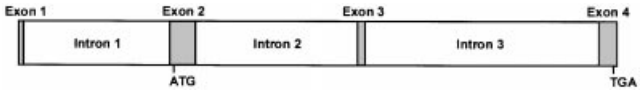

Figure 5 Schematic diagram of the mouse Sp17 genomic structure

Positions of the start (ATG) and the stop (TGA) codons are indicated.

Comparison of mammalian Sp17s

Using the mouse genomic structure as a model for viewing the alignment of Sp17 protein sequences, the structure–function motif can be clearly seen. Figure 6 shows the alignment of Sp17 protein sequences from eutherian and marsupial mammals, including rat, mouse, macaque, baboon, human, marmoset, rabbit, *Monodelphis* and wallaby. Exon 2 contains the highly conserved dimer interaction site, exon 3 contains the hypervariable domain and exon 4 the calmodulin-binding site. The

Table 1 Exon–intron junctions of the mouse *Sp17* gene

mRNA positions are as described in [14]. The intron sequences are shown in lower case letters and the exon sequences are shown in upper case.

Exon	Amino acids	mRNA position	Exon size (bp)	5′-Intron–exon junction	3′-Exon–intron junction	Intron size (bp)
1		−94 to −30	65		GAAGAG–gtgagg	1575
2	1–75	−29 to +225	254	ctttag-GTACCA	TTCAAG–gtatag	1760
3	76–101	+226 to +303	78	ttctag-GAACAA	CCCTTC–gtacgt	2549
4	102–149	+304 to +489	186	atgaag-GAGGAG		

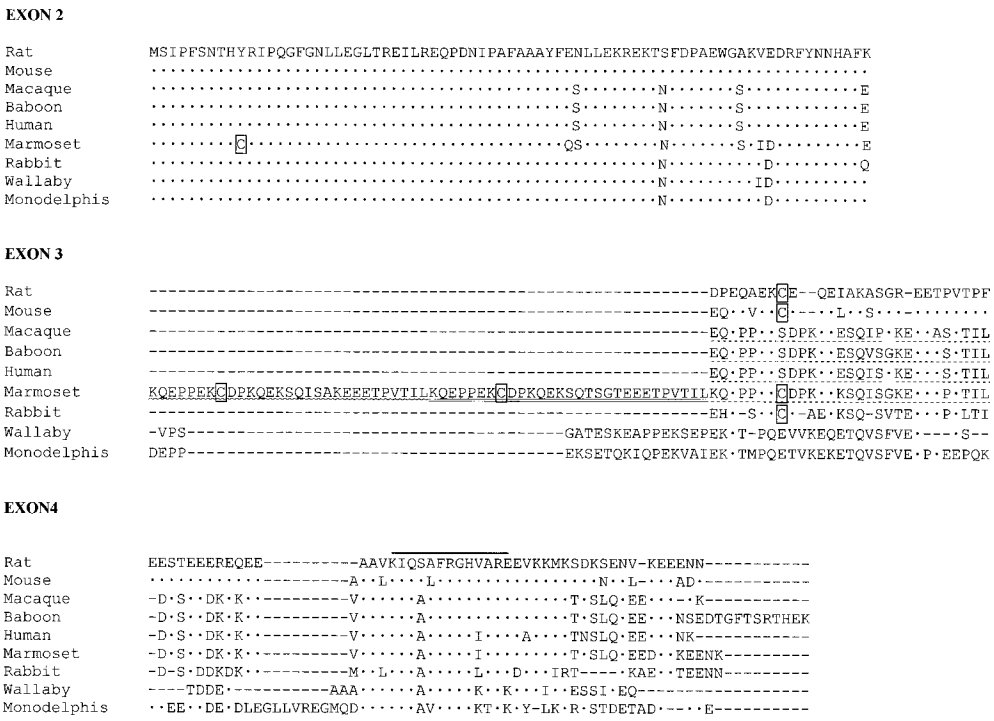

Figure 6 Alignment of Sp17 protein sequences

The species are identified by their common names. Residues identical to the top sequence are indicated by dots (·) and gaps inserted for improved sequence alignment are indicated by dashes (–). Positions of cysteines are boxed. The first sequence repeat in marmoset is wavy underlined, the second repeat is double underlined and the third repeat shared with other primates is dotted underlined. The IQ motif is overlined. Exons indicated are from the mouse sequence.

most striking feature is that the N-terminal 75 amino acids in mouse exon 2 are highly conserved, 68 out of 75 (90 %) residues are identical. Completely identical N-terminal sequences are found in the order Rodentia and in the primate suborder Catarrhini (macaque, baboon and human). Marmoset has a cysteine in position 10 that is not present in any other species. In this region, all Sp17s showed similarities to the dimer interaction site of the cAMP-dependent protein kinase IIα regulatory subunit [8–10]. This domain also has similarities with the N-terminal of the rhophilin-binding protein, ropporin [20], an AKAP-binding protein. Based on the RIIα [10] and Sp17 sequences, these data suggest that Sp17 probably dimerizes and binds to AKAPs. AKAP-bound Sp17 could provide calmodulin and/or sulphated carbohydrates in the correct location for other cell functions.

The middle domain of Sp17, which is mouse exon 3 (amino acids 76–101; Table 1), contains the central region critical for the conformation of the high-affinity heparin-binding sites (Figure 1) and exhibits the greatest sequence divergence among all the sequences studied. As shown in Figure 6, this region in the marmoset is composed of three repeated sequences. The first two repeats (positions 76–104 and 105–133) are unique (underlined in Figure 6), whereas the third repeat (positions 134–162) is shared by other primates (dotted underline, Figure 6). This domain may be under more selective pressure than the other domains because of the different sulphated carbohydrates that each species encounters, which may, in turn, require variability and sequence repeats.

Both wallaby and *Monodelphis* have stretches of amino acids (positions 119–133) that are not found in other species. Additionally, *Monodelphis* Sp17 has a unique stretch of amino acids (positions 175–181) not found in any other species, including wallaby. It should also be noted that the cysteine content of Sp17 found at position 151 in Rodentia, marmoset and rabbit is not found in Catarrhini or marsupials.

The C-terminal domain of Sp17 sequences contains the calmodulin-binding site [11]. Characterized by the signature IQ motif, IQXXXRGXXXR, it is found in all species (Figure 6). In the rabbit this site is functional [11] and therefore likely to be

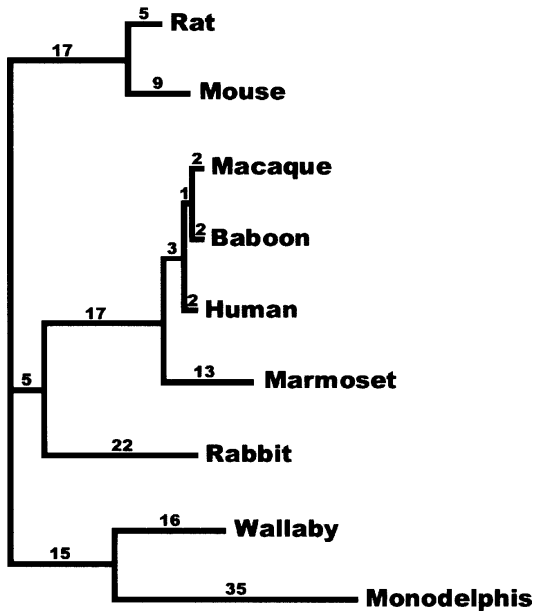

Figure 7 Phylogenetic tree constructed using Sp17 protein sequences

The species are identified by their common names. The tree has a consistency index of 0.9451. Numbers above tree branches indicate numbers of changes that occur along the branch.

functional in all the other species. The PKC phosphorylation site found in human growth-associated protein-43 is not present in the Sp17 IQ motif, consequently rabbit Sp17 and calmodulin interaction was not inhibited by PKC phosphorylation *in vitro* (results not shown). However, PKC phosphorylation of Sp17 was inhibited by calmodulin. Sp17 sequences showed considerable variability downstream of the IQ motif (Figure 6).

Phylogeny of Sp17

A phylogeny of Sp17 was constructed with Sp17 protein sequences by maximum-parsimony analysis using the exhaustive search method (Figure 7; [15]). Ten optimal trees with the same consistency index were obtained. The phylogeny shown in Figure 7 was chosen because the same branch pattern was generated by maximal-likelihood analysis using Sp17 DNA sequences. The grouping of rabbits with primates is consistent with a previous report [20] demonstrating that Lagomorpha is more closely related to primates and tree shrews than it is to rodents [21].

CONCLUSIONS

Sp17 is a ubiquitous molecule with four exons (Figure 5) that express three functional domains: an N-terminal domain similar to the dimer interaction site of the cAMP-dependent protein kinase IIα regulatory subunit [8–10], a C-terminal calmodulin-binding domain [11], and a central sulphated carbohydrate-binding domain (Figures 1 and 2). Although Sp17 expression is highest in the testis, it has been found in all of the tissues examined (Figures 3 and 4) and is conserved throughout all the mammalian species examined (Figures 6 and 7). The gene is induced in metastatic cells and during mucosal immune responses [4,6] and appears to play an important role in cell migration and/or adhesion as well as in fertilization [1,2,5]. This is further supported by the recent detection of Sp17 on the surface of normal B-lymphocytes and granulocytes and weakly on the surface of T lymphocytes and monocytes (H. M. Lacy and R. D. Sanderson, personal communication).

Support for this project was provided by NIH grant U54HD29099. The authors thank Professor R. D. Sanderson, Arkansas Cancer Research Center, Little Rock, AR, U.S.A., for making the submitted manuscript (H. M. Lacy and R. D. Sanderson, personal communication) available to us. The λZAP tammar wallaby (*Macropus eugenii*) testis cDNA library was a gift from Dr Karen Mate and Dr Merrilee Harris (Macquarie University, New South Wales, Australia).

REFERENCES

1 Richardson, R. R., Yamasaki, N. and O'Rand, M. G. (1994) Sequence of a rabbit sperm zona pellucida binding protein and localization during the acrosome reaction. Dev. Biol. **165**, 688–701
2 O'Rand, M. G., Widgren, E. E. and Fisher, S. J. (1988) Characterization of the rabbit sperm membrane autoantigen, RSA, as a lectin-like zona binding protein. Dev. Biol. **129**, 231–240
3 Yamasaki, N., Richardson, R. T. and O'Rand, M. G. (1995) Expression of the rabbit sperm protein Sp17 in COS cells and interaction of recombinant Sp17 with the rabbit zona pellucida. Mol. Reprod. Dev. **40**, 48–55
4 Dong, G., Loukinova, E., Smith, C. W., Chen, Z. and Van Waes, C. (1997) Genes differentially expressed with malignant transformation and metastatic tumor progression of murine squamous cell carcinoma. J. Cell. Biochem. **Suppl 28/29**, 90–100
5 Lacy, M. and Sanderson, R. (1999) Sperm protein Sp17 promotes cell-cell adhesion through interaction with syndecan-1 heparan sulfate chains. Mol. Biol. Cell **10**, 459a
6 Tatlow, D., Brownlie, R., Babiuk, L. A. and Griebel, P. (2000) Differential display analysis of gene expression during the induction of mucosal immunity. Immunogenetics **52**, 73–80
7 O'Rand, M. G. and Widgren, E. E. (1994) Identification of sperm antigen targets for immunocontraception: B-cell epitope analysis of Sp17. Reprod. Fertil. Dev. **6**, 289–296
8 Taylor, S., Buechler, J. A. and Yonemoto, W. (1990) cAMP dependent protein kinase: Framework for a diverse family of regulatory enzymes. Ann. Rev. Biochem. **59**, 971–1005
9 Huang, L. J., Durick, K., Weiner, J. A., Chun, J. and Taylor, S. S. (1997) D-AKAP2, a novel protein kinase A anchoring protein with a putative RGS domain. Proc. Natl. Acad. Sci. USA **94**, 11184–11189
10 Newlon, M. G., Roy, M., Morikis, D., Hausken, Z. E., Coghlan, V., Scott, J. D. and Jennings, P. A. (1999) The molecular basis for protein kinase A anchoring revealed by solution NMR. Nature Structure Biol. **6**, 222–227
11 Wen, Y., Richardson, R. T. and O'Rand, M. G. (1999) Processing of the sperm protein Sp17 during the acrosome reaction and characterization as a calmodulin binding protein. Dev. Biol. **206**, 113–122
12 Lea, I. A., Richardson, R. T., Widgren, E. E. and O'Rand, M. G. (1996) Cloning and sequencing of cDNAs encoding the human sperm protein, Sp17. Biochim. Biophys. Acta **1307**, 263–266
13 Lea, I. A., van Lierop, M. J. C., Widgren, E. E., Grootenhuis, Y., Wen, Y., van Duin, M. and O'Rand, M. G. (1998) A chimeric sperm peptide induces antibodies and strain-specific reversible infertility. Biol. Reprod. **59**, 527–536
14 Kong, M., Richardson, R. T., Widgren, E. E. and O'Rand, M. G. (1995) Sequence and localization of the mouse sperm autoantigenic protein, Sp17. Biol. Reprod. **53**, 579–590
15 Swofford, D. L. (1993) PAUP: Phylogenetic analysis using parsimony, version 3.1.1, Smithsonian Institution, Washington, D. C.
16 Adoyo, P. A., Lea, I. A., Richardson, R. T., Widgren, E. E. and O'Rand, M. G. (1997) Sequence and characterization of the sperm protein Sp17 from the baboon. Mol. Reprod. Dev. **47**, 66–71
17 Lea, I. A., Kurth, B. and O'Rand, M. G. (1998) Immune response to immunization with sperm antigens in the macaque oviduct. Biol. Reprod. **58**, 794–800
18 Scatchard, G. (1949) The attraction of proteins for small molecules and ions. Ann. N. Y. Acad. Sci. **51**, 660–672
19 Hileman, R. E., Fromm, J. R., Weiler, J. M. and Linhardt, R. J. (1998) Glycosaminoglycan-protein interactions: definition of consensus sites in glycosaminoglycan binding proteins. BioEssays **20**, 156–167
20 Fujita, A., Nakamura, K., Kato, T., Watanabe, N., Ishizaki, T., Kimura, K., Mizoguchi, A. and Narumiya, S. (2000) Ropporin, a sperm-specific binding protein of rhophilin, that is localized in the fibrous sheath of sperm flagella. J. Cell Sci. **113**, 103–112
21 Graur, D., Duret, L. and Gouy, M. (1996) Phylogenetic position of the order Lagomorpha (rabbits, hares and allies). Nature **379**, 333–335

Received 14 February 2001/16 March 2001; accepted 2 April 2001