Expression of a *Xenopus* counterpart of mammalian syndecan 2 during embryogenesis

Norman D. ROSENBLUM,*†§ Barbara B. BOTELHO†|| and Merton BERNFIELD‡

*Division of Nephrology, The Hospital for Sick Children, University of Toronto, Toronto, Ont. M5G 1X8, Canada, †Division of Nephrology, Children's Hospital, Harvard Medical School, Boston, MA 02115, U.S.A. and ‡Division of Newborn Medicine, Children's Hospital, Harvard Medical School, Boston, MA 02115, U.S.A.

We have identified a *Xenopus* cDNA, XS-2, by screening a *Xenopus* embryonic stage-22–24 cDNA library with a DNA probe encoding the transmembrane and cytoplasmic domains of mouse syndecan 1. The 1.4 kb cDNA consists of an open reading frame of 642 nucleotides encoding a protein of 191 amino acids. The predicted protein of 20869 Da contains a 25-amino acid putative transmembrane domain and a 32-amino acid putative cytoplasmic domain, both of which are highly similar to the corresponding regions of rat syndecan 2 (92 % identity) and to a lesser degree those of rat syndecans 1, 3 and 4 (62, 64 and 78 % respectively). The putative N-terminal ectodomain contains a possible attachment site for heparan sulphate, identical with the comparable glycosaminoglycan-attachment sequence of rat

syndecan 2. Polyclonal antisera raised against recombinant ectodomain of XS-2, expressed as a fusion protein, recognized a heparan sulphate proteoglycan in XTC cell-culture medium. This proteoglycan bound to DEAE-Sephacel and was eluted with 1 M NaCl; digestion with heparitinase but not chondroitinase ABC resulted in the identification of a 46 kDa protein by these antisera. Northern-blot analysis indicated that XS-2 identifies two *Xenopus* mRNA species approx. 4 and 2 kb in size in embryos ranging in maturation from the 64-cell stage to stage 54. These results demonstrate that a heparan sulphate proteoglycan, similar to syndecan 2, is expressed during *Xenopus* embryogenesis.

INTRODUCTION

The syndecans are a family of integral membrane cell-surface heparan sulphate proteoglycans. In mammals, four types of syndecan have been defined. Each is encoded by a unique gene (reviewed in ref. [1]). A single syndecan, that of *Drosophila*, has been identified in non-vertebrate species [2]. The core proteins of all known syndecan types share similar domain structures: a cytoplasmic domain ranging from 28 to 35 amino acids depending on the particular syndecan, a transmembrane domain consisting of 25 amino acids and an extracellular domain (ectodomain) the size and composition of which differs in each of the syndecan types. The cytoplasmic and transmembrane domains of each syndecan type are very similar but the extracellular domains diverge in size and amino acid composition. Each of the extracellular domains has putative glycosaminoglycan (GAG)-attachment sites as well as a putative proteolytic-cleavage site near the junction of the ectodomain and transmembrane domains. Each syndecan is modified by one or more heparan sulphate side chains; syndecan 1 also contains a chondroitin sulphate side chain [1].

Syndecans are capable of binding a variety of biological substrates involved in cell proliferation and differentiation. These include proteinase inhibitors [3], extracellular matrix components [4–6] and heparin-binding growth factors such as basic fibroblast growth factor (bFGF) [7–10]. The significance of syndecan binding to bFGF has been investigated in detail. The ability of bFGF to induce cell proliferation and differentiation depends on

specific binding to cell-surface heparan sulphate [8]. Binding of bFGF to heparan sulphate increases the affinity of bFGF for its cognate tyrosine kinase receptors [11]. Both syndecans 1 and 3 bind bFGF via their heparan sulphate side chains [7,10].

The spatial and cell-specific expression of syndecans during embryogenesis also suggests that these proteoglycans play a role in development. For example, syndecan 1 is expressed in an asymmetric pattern in cells and tissues during early embryogenesis. In the mouse embryo, syndecan 1 is expressed as early as the four-cell stage and is present on the surface of the blastomeres principally at the sites of cell–cell contact. Later in development it is expressed on the basolateral surface of the primitive ectodermal cells and their derivatives, the endoderm and undifferentiated mesodermal cells emerging from the primitive streak. Syndecan-1 mesodermal expression then becomes asymmetric as it is lost from the anterior mesoderm but maintained in the posterior and lateral mesoderm [12].

Syndecans are also expressed at critical periods during embryonic mesenchymal–epithelial interactions [13]. For example, in the mouse embryo at the onset of these interactions syndecan 1 is expressed by epithelial cells but not by mesenchymal cells. As development ensues, the epithelium undergoes a morphogenetic change and transiently loses its cell surface syndecan 1. Meanwhile, the mesenchymal cells proliferate, condense around the epithelium and express syndecan 1 *de novo*. The temporal and cell-specific expression of syndecans during these interactions suggests that they may be involved in interactions with growth factors, extracellular matrix molecules and adjacent cells. All of

Abbreviations used: GAG, glycosaminoglycan; bFGF, basic fibroblastic growth factor; TGF β , transforming growth factor β ; 1 $\times$ SSC, 0.15 M NaCl plus 0.015 M sodium citrate; poly(A)⁺, polyadenylated; GST, glutathione S-transferase.

§ To whom correspondence should be addressed at Division of Nephrology, The Hospital for Sick Children, 555 University Avenue, Toronto, Ont. M5G 1X8, Canada.

|| Present address: Division of Pediatric Nephrology, University of California, Sacramento, CA, U.S.A.

The sequence reported in this paper has been deposited in the EMBL/GenBank/DBJ Nucleotide. Sequence Databases under the accession number U24433.

these interactions may determine, in part, some or all stages of tissue morphogenesis.

The *Xenopus* embryo is a model for early embryonic induction and cellular differentiation. In *Xenopus*, bFGF is a known inducer of mesodermal tissue. Addition of bFGF to animal cap explants results in the formation of both dorsal and ventral mesoderm [14]. Expression of truncated bFGF receptor in the embryo results in major deficiencies of the lateral and posterior mesoderm [15]. bFGF may also modulate the inducing activity of activin, a member of the transforming growth factor β superfamily and a potent inducer of anterior and dorsal mesoderm [16]. The activity of both of these growth factors appears to be dose-dependent, suggesting that mechanisms that titrate the amount of these growth factors presented to cells in a temporally and spatially specific manner must exist within the embryo. In the *Xenopus* embryo, the expression of cell-surface heparan sulphate proteoglycans may serve as a mechanism by which local concentrations of growth factors such as activin and bFGF are modulated. Although proteoglycans such as the syndecans have not previously been described in *Xenopus*, the conservation of expression of many developmentally important genes across species [17] strongly suggests that syndecan-type proteoglycans are also expressed in *Xenopus*.

This report describes the molecular cloning of a full-length cDNA, XS-2, encoding a putative *Xenopus* integral membrane proteoglycan. The domain structure of the putative translational product is strikingly similar to that of the syndecan family of proteoglycans with closest identity with mammalian syndecan 2. Polyclonal antisera raised against a recombinant ectodomain of XS-2, expressed as a fusion protein, recognized a fraction of XTC cell culture medium which binds to the DEAE-Sephacel and is eluted with 1 M NaCl. Digestion of this eluted fraction with heparitinase but not chondroitinase ABC results in the identification of a 46 kDa protein by these antisera. The XS-2 mRNA is expressed as early as the 64-cell stage and throughout embryogenesis. These studies demonstrate that a heparan sulphate proteoglycan similar in domain structure to the syndecan family of cell-surface proteoglycans is expressed during *Xenopus* embryogenesis.

MATERIALS AND METHODS

Isolation of cDNA clone, XS-2

A *Xenopus* embryonic stage 22–24 cDNA library (kindly provided by Dr. L. Zon) was screened with a DNA probe encoding the transmembrane and cytoplasmic domains of mouse syndecan 1 [18]. This probe was generated by PCR using a 5' sense primer, 5'-TAGTCCGCATCTTGAGGAGGACGAAC-3', a 3' antisense primer, 5'-CTCAGAGCCTTTTGACAGGAAGG-AAG-3' and a cDNA encoding the coding region of mouse syndecan 1 as substrate. Thirty cycles of amplification were performed with a Perkin-Elmer thermal cycler under the following conditions: 94 °C for 1 min, 55 °C for 2 min and 72 °C followed by an additional 7 min incubation at the end of the 30 cycles. The amplified 207 bp DNA fragment was labelled with [³²P]dCTP and used as a probe to screen 650 000 plaques of the *Xenopus* cDNA library. The conditions for library screening were as follows: prehybridization in 5 × SCC (1 × SCC is 0.15 M NaCl plus 0.015 M sodium citrate, pH 7.0) with 1 % sarkosyl and 20 µg/ml salmon sperm DNA for 30 min at 56 °C, hybridization for 16 h at 56 °C in the same mix followed by two washes with 3 × SSC/0.5 % sarkosyl and two washes with 3 × SSC at 56 °C. A single recombinant phage containing a 1.4 kb insert was identified and purified. The cDNA insert was isolated by PCR using flanking oligonucleotide primers encoding sequences

of λ gt10 (forward primer, 5'-AGCAAGTTCAGCCTGGTT-AAG-3'; reverse primer, 5'-CTTATGAGTATTTCTTCCA-GGGTA-3') and subcloned into a plasmid vector, pCRII (Invitrogen). Complete DNA sequences were obtained for both strands of the cDNA by the dideoxynucleotide chain-termination method using a modified T7 DNA polymerase (Sequenase 2.0; United States Biochemical). Oligonucleotide primers encoding regions within the phage T3 and T7 promoters as well as nested customized oligonucleotide primers were used to obtain a full-length sequence of both strands. Analyses of DNA sequences were performed using sequence analysis software (GCG Program) and by comparing sequences with published sequences in Genbank using the BLAST program.

Isolation of *Xenopus* RNA

RNA was isolated from whole embryos at different stages of development. Ovulation and *in vitro* fertilization were carried out as described previously [19]. Developmental stages were determined as described by Nieuwkoop and Faber [20]. Embryos at defined stages were snap-frozen in liquid N₂ and stored at -80 °C. Total RNA was subsequently isolated by the lithium chloride method [21]. The *Xenopus* embryonic cell line, XTC, was maintained in tissue culture at 22 °C in medium consisting of 70 % Liebovitz L-15 medium (Gibco-BRL), 10 % fetal bovine serum (Hyclone), 1 % HEPES, 1 % L-glutamine, penicillin (100 units/ml) and streptomycin (100 µg/ml) [22]. Confluent cells were treated with trypsin, washed in PBS, snap-frozen in liquid N₂ and stored at -80 °C. Poly(A)⁺ mRNA was subsequently isolated by the proteinase K method followed by binding to oligo(dT)-cellulose (Micro Fast Track; Invitrogen).

Northern-blot analysis

RNA isolated from *Xenopus* embryos and embryonic cells was fractionated by electrophoresis in agarose gels in the presence of formaldehyde/formamide after which the RNA was transferred to nylon membranes (Duralon UV; Stratagene). The membranes were hybridized with ³²P-labelled XS-2 under the conditions described above.

Construction of expression plasmids and isolation of the hybrid proteins

A plasmid was constructed that expressed the ectodomain of XS-2 as a fusion protein with glutathione S-transferase (GST) [23]. A DNA fragment encoding 115 amino acids of the XS-2 extracellular domain (residues 18–132) was generated by PCR using a 5' sense primer, 5'-GCGGATCCGGGGAAACGTGG-GCCCAAGCT-3', including a *Bam*HI site at the 5' end, and a 3' antisense primer, 5'-GCGAATTCTTATCTGTGGAACAGG-TTCTC-3', including the stop codon TAA and an *Eco*RI site at the 5' end. XS-2 was used as the substrate in the PCR under the conditions described above. The 364 bp amplification product was digested with *Bam*HI and *Eco*RI and ligated to the expression plasmid pGEX-2T (Pharmacia), digested with the corresponding enzymes. Transformed DH5 α cells were grown overnight at 37 °C on ampicillin plates. Colonies containing recombinant plasmids were detected by screening isolated plasmid DNA with ³²P-labelled XS-2 and by identifying GST fusion proteins in SDS/polyacrylamide gels. DNA sequences were obtained at the *Bam*HI site of vector-insert recombination to confirm that the coding sequence of the insert was in frame.

Fusion proteins for immunization and for Western-blot analysis were isolated from exponentially growing cultures of

transformed cells induced with 0.1 mM isopropyl β -D-thiogalactopyranoside. Fusion proteins were isolated by addition of glutathione–Sephacel 4B beads (Pharmacia) to the cell lysate, precipitation of beads by centrifugation and subsequent elution of the fusion proteins by competition with reduced glutathione [23]. The size of the eluted proteins was compared with that of the parental GST in Coomassie stained SDS/polyacrylamide gels.

Generation of polyclonal anti-XS-2 sera

Purified pGEX-2T-XS-2-encoded fusion protein (200 μ g) was mixed with an equal volume of complete Freund's adjuvant and injected intradermally into rabbits. Additional injections of 100, 50 and 20 μ g of fusion proteins in incomplete Freund's adjuvant were performed subcutaneously at 28, 42 and 56 days respectively. Immune sera were obtained from ear bleedings and were tested on dot-blots of fusion proteins. The fusion-protein bands were detected in immunoblots at a dilution of 1:500–1:7500 with minimal background. Preimmune sera showed no reactivity.

Isolation and analysis of *Xenopus* proteoglycans

Proteoglycans were isolated from conditioned medium of *Xenopus* XTC cells (kindly provided by Dr. L. Zon, Children's Hospital, Boston, MA, U.S.A.) using a modification of previously published procedures [4]. Cell-culture medium was harvested 96 h after addition to XTC cultures. It was then prepared for proteoglycan extraction by adding urea and Triton X-100 to achieve final concentrations of 2 M and 0.1 % respectively and acidifying to pH 4.5 with acetic acid. DEAE-Sephacel (Pharmacia), equilibrated with column wash buffer (0.2 M NaCl, 2 M urea, 50 mM sodium acetate, pH 4.5, and 0.1 % Triton X-100), was added to the conditioned medium in a ratio of 1:75 (v/v) and incubated with gentle rocking for 16 h at 4 °C. The DEAE-Sephacel/conditioned medium mix was then loaded on to a column, washed with 10 bed vol. of column wash buffer and then 10 bed vol. of column wash buffer without urea. Material bound to the column was eluted in 3 bed vol. of a buffer containing 1 M NaCl, 0.1 % Triton X-100 and 50 mM sodium acetate, pH 4.5. Eluted material was precipitated in the presence of 3 vol. of 100 % ethanol containing 1.3 % potassium acetate at –20 °C for 16 h. The precipitate was isolated by centrifugation at 13000 *g* and then resuspended in enzyme-digestion buffer containing 100 mM Tris/HCl, pH 7.2, 1 mM PMSF, 10 mM *N*-ethylmaleimide and 5 mM EDTA. The final resuspension volume was 0.7 % of the starting volume of culture medium.

Enzyme digests

Enzyme digestions with heparitinase and chondroitinase ABC were performed in 100 μ l aliquots of resuspended proteoglycans. Proteoglycan was digested with either 150 m-units of heparitinase III (Sigma) or 25 m-units of chondroitinase ABC (Sigma) or both enzymes together for 1 h at 37 °C. Samples were then lyophilized and resuspended in a SDS loading buffer consisting of 2 % SDS, 10 % glycerol and 0.001 % Bromophenol Blue.

Immunoblot analysis of proteoglycan core proteins

Proteoglycans and their core proteins were separated by SDS/PAGE. Molecular size was estimated by comparison with globular molecular-mass standards (Amersham). After electrophoretic separation, proteins were transferred to a positively charged poly(vinylidene difluoride) membrane (Immobilon-N; Millipore) using a semi-dry electroblotting apparatus (Owl Scientific) in a buffer consisting of 20 mM Tris/HCl, 150 mM

glycine and 20 % methanol for 2 h at 400 mA. Membranes derived from dot-blotting of XTC proteoglycans or electro-transfer of electrophoresed proteins were blocked with 5 % non-fat dry milk in PBS containing 0.02 % NaN₃ overnight and then incubated with primary antibody for 1 h followed by a secondary antibody conjugated to alkaline phosphatase. Protein bands reactive with antibody were detected by a colour reaction using substrates for alkaline phosphatase supplied in a commercial kit (Protoblot Immunoscreening System; Promega Biotec).

RESULTS

XS-2 encodes a putative integral membrane cell-surface proteoglycan

The members of the syndecan family of cell-surface proteoglycans are strikingly similar in their transmembrane and cytoplasmic domains [1]. These similarities are maintained across species including human, rat, mouse and chicken. In contrast, the extracellular domains are dissimilar except for the sequences that encode putative sites for substitution with GAG chains. Given that the transmembrane and cytoplasmic sequences are highly specific for the syndecans, we used a DNA probe containing these sequences to screen for expression of syndecan family members during *Xenopus* development. Screening of a *Xenopus* embryonic stage-22–24 cDNA library with the mouse DNA probe encoding the transmembrane and cytoplasmic domains of syndecan 1 resulted in the identification of a single recombinant which contained a 1389 bp cDNA insert (Figure 1). The first ATG within this open reading frame occurs at position 459. This ATG is preceded by guanosine residues at positions –9 to –4 and +4 and +5, providing a favourable context for the initiation of translation [24]. The ATG at position 459 is the start of a 573 bp coding sequence encoding a 191-amino acid (20843 Da) protein product. Two canonical polyadenylation signals (AATAA) are contained within the 3' region of the cDNA (positions 1237 and 1279). Notably, neither of these is followed by a poly(dA) stretch, suggesting that there is more 3' sequence not contained within the cDNA.

Analysis of the hydropathy characteristics of the derived XS-2 translation product suggests that the protein consists of distinct hydrophilic and hydrophobic domains (Figure 2). The Kyte–Doolittle hydropathy analysis [25] indicates that the first 18 amino acids after the translation start form a hydrophobic domain. The composition of this peptide is compatible with that of a signal peptide. This hydrophobic domain is followed by a 116-amino acid hydrophilic domain extending from position 19 to residue 134. This domain contains numerous putative glycosylation sites (see below). The presence of these putative glycosylation and protease-cleavage sites in the context of the other putative domains suggest that this domain is an extracellular domain (ectodomain). Further toward the carboxyl end of the protein is a 25-amino acid hydrophobic region with the size and amino acid composition of known transmembrane domains. Finally, the coding sequence concludes with a 32-amino acid hydrophilic domain that contains three tyrosine residues. The position of these residues is conserved in the cytoplasmic domains of other cell-surface integral membrane proteoglycans (syndecans) [1]. Their function is unknown but they have been implicated in the targeting of integral membrane proteins to specific membrane domains and endocytic compartments [26]. The hydropathy characteristics and the relative positions of the latter two domains suggest that they encode a transmembrane and C-terminal cytoplasmic domain respectively.

The presumptive ectodomain contains three Ser-Gly sequences (positions 17, 36 and 48) each of which may function as a site for

1	c c a a g c t c a a t c g a c t c a c t a t a g g g c c c g g t a c c g a g c t c a c t a g t t t a a t a a a a g c t t	60
61	a t c g g c c g a g g t g a g a a g g g t t a a g g t a c c t c g a g c g g c c g c a a t c t i a g c a a g t t c t t a t g a g t a t t	120
121	t c t t c c a g g g t a a g g t a c c t c g a g c g g c c g c a a t c t i a g c a a g t t c a g c c t g t t t a a g t c	180
181	c a a g c t g a a t t c c c g c c g c c g g g a t c a g c t c a g t c g a g c c g c c g c a g g a g a a g c a g t	240
241	g c c c a g t c c g a g c g g c g g a a g t t g g t g g a g c g g a a g g a l t c g g g a g t c g g a g a a a a g g g	300
301	g a t t t t g t t t g a g g a g a g t g a g t g c g g a g g g a g c g g g g g a g g a g g a c t c a c a g c g t g a a t	360
361	c c a g t g g g a a t c a c t c c a g c a a c a a g t g a g c g c g g g g c a g t g t g g g a c c t g c a a g t t c a t	420
421	c g c a c c g c c c g g g a c c g a c t c t c t t a c t c g g g g g g a l a a t g a g g a a c g t g t g g c t t a t c g	480
	M R N V W L I V	8
481	t t c c c t t t g c c c t g c t c g c t g c c t t g t c c g g g g a a a c g t g g g c c c a a g c t g a c a g a g a c c	540
	P F A L L A A L S G E T W A Q A D R D L	28
541	t a t a t a t c g a c a g c a c a g a a t c a t c a g g a a a c t a c c c a g t t g a t g a t g a t g a t t a c t c t t	600
	Y I D S T E S S G N Y P V D D D D Y S S	48
601	c t g g a t c t g g c t c t g g c a t c c c a g c a c g t g g t g a c g a l g a a g a c g a a a a t g t a g t a t t a a	660
	G S G S G I P A R G D D E A G D E N V V L T	68
661	c a a c t g t c c a g a c a c t t a t a a g c a g t c c a t c a a g t g a a a t g c c a t a t g t t g a g a c c a c a a	720
	T V Q T L I S S P S S E M P Y V E T T T	88
721	c g c t g a a a a c a c t a a c a a a g a t g g c a c c t g a a a c c a a g g a a c c t g g t g a a g t a g a g t c a a	780
	L K T L T K M A P E T K E P G E V E S T	108
781	c t a a c a c a g t t c t t g t c t a c g g a a a g a a g g a t a t c g t c c a g a c t g c g a c a c a c a c a g a g a	840
	N T V L V Y G K K D I V Q T A T H T E N	128
841	a c c t g t t c c a c a g a a c a g a a g t a c t t g c g g c t g t c a t t g c t g g g g g c g g c a t t g g c t t c c	900
	L F H R T E V L A A V I A G G G G I G F L	148
901	t g t t t t g c c g t a t t c t t a a t c c t c c t g t t g g t t a c c g c a t g a g a a a g a a g g a t g a a g g c a	960
	F A V F L I L L L V Y R M R K K D E G S	168
961	g c t a c g a c c t t g g c g a a c g c a a a c c a t c c a g g g c c g t c t a c c a a a a g g c a c c a a c t a a a g	1020
	Y D L G E R K P S R A V Y Q K A P T K E	188
1021	a g t t t t a t g c a t a a a a c t g a c g t c t t t a g t g t c t c t t t t a c g a g a t c a c t g a a c t t t t t	1080
	F Y A *	191
1081	c t t t t a a g t a a a g c t t t t g a t a g a t a a t g a a g a t t t t t t t c a t t t a a a a c a a a a a a a a	1140
1141	c a a a a a c c t t t a c g g t a a c t t t a t g a t a a a a a t c c c a t a t a t t t a a g a a a a a a a a c a a	1200
1201	a c a a a a c t t t c a t t c a t t t a t t c c t t t a a a a a g a a a a a a a a a t g t a a a g t c a a t t c t t c a	1260
1261	c a t t t t c a t t g t t g a g t g a a t a c c g c a g c a g g t a t t a t t t t c t a a g a c c a g g a a t t c t t	1320
1321	t t g c t t t t a c c c t g g a a g a a a t a c t a t a a g a a g c t t c g g c c g c t c g a g g t a c c t t a a a	1380
1381	c g g t t c t c a	

Figure 1 Sequence and derived protein sequence of cDNA, XS-2

The sequence consists of 1389 nucleotides. It has an open reading frame of 573 nucleotides encoding a 191-amino acid protein. The putative hydrophobic domain is in bold type and underlined. Putative GAG-attachment sites are identified by circles (○). An arrow marks the putative site of ectodomain cleavage from the cell surface. Poly(A)⁺ signal sequences are underlined.

GAG attachment. The third of these sequences starting at position 48 is a likely acceptor site, as it consists of the tetrapeptide sequence Ser-Gly-Ser-Gly preceded by four aspartic acid residues. This sequence is in agreement with the recognition consensus amino acid sequence for the attachment of chondroitin/dermatan sulphate to core proteins (acidic-acidic-Ser-Gly-Xaa-Gly) proposed by Bourdon et al. [27].

Although there are five asparagine residues in the presumptive ectodomain (amino acid residues 3, 38, 64, 109 and 128), these are unlikely sites for N-linked glycosylation as they are not found in the context of the usual acceptor-site sequences, Asn-Xaa-Ser or Asn-Xaa-Thr. In contrast, there are numerous possible sites for O-linked glycosylation at serine or tyrosine residues. Clusters of hydroxylated amino acid residues (Ser-Thr, Ser-Ser, Thr-Thr) are present at residues 32, 47, 68, 75, 78, 86 and 107.

XS-2 is a member of the syndecan family of integral membrane cell-surface proteoglycans

Comparison of XS-2 amino acid sequences with those of other reported proteins in Genbank, the Brookhaven Protein Data Bank and the EMBL Data Bank revealed that the presumptive

transmembrane and cytoplasmic domains of XS-2 are highly related to those of each previously defined member of the syndecan family of integral membrane cell-surface proteoglycans (Figure 3). The amino acid composition of the transmembrane and cytoplasmic domains of rat syndecan 1 [28], syndecan 2 [29], syndecan 3 [10] and syndecan 4 [30] are compared with those contained within XS-2. A remarkable degree of similarity is observed among these proteins particularly in the transmembrane domain and the N- and C-terminal regions of the cytoplasmic domains. XS-2 shares 62, 92, 64 and 78% identity with rat syndecans 1, 2, 3 and 4 respectively in the transmembrane and cytoplasmic domains. XS-2 is the most similar to syndecan 2 with respect to amino acid composition, sequence and length of the transmembrane and cytoplasmic domains.

The XS-2 putative ectodomain is not similar to other proteins. Indeed, a comparison of these XS-2 amino acid sequences with those in computerized databases reveals significant similarities only between sequences in the putative GAG-binding domain (amino acids 42–53) (Figure 4). In this region XS-2 sequences were nearly identical with those in syndecan 2 [29], serglycin [31] and CD44 [32], each a proteoglycan. These similarities emphasize the probable identity of XS-2 as a proteoglycan.

Figure 2 Hydropathy plot of XS-2 translation product and schematic diagram of predicted protein domains

The method of Kyte and Doolittle [25] was used to generate hydropathy values. Hydropathic signals are shown above the line. The predicted protein domains are shown below the hydropathy plot and include the N-terminal hydrophobic signal peptide, the hydrophilic ectodomain (□) including the putative GAG-attachment sites (|), the hydrophobic transmembrane domain (■) and the C-terminal hydrophilic cytoplasmic domain (□). △, Putative cleavage site.

XS-2 codes for a core protein linked to heparan sulphate

To identify the translation product of the XS-2 gene, a DNA construct encoding the ectodomain of XS-2 was expressed as a fusion protein with GST. The ectodomain was chosen because the sequences contained within it are found in a limited number of proteins (Figure 4). High-titre rabbit polyclonal antisera were generated against this fusion protein and used to study the XS-2 translational product. The observation that XS-2 mRNA is expressed by the *Xenopus* embryonic cell line, XTC, provided the basis for studying the proteins/proteoglycans expressed by these cells (Figure 7, below). Initial analysis of proteoglycans isolated from XTC culture medium by dot-blot assay indicated that the antisera identified material which bound to DEAE-Sephacel and was eluted under conditions of high salt concentration (1 M NaCl (Figure 5). This suggested that the antisera recognizes a proteoglycan.

XS-2	42	D D D D Y S S G S G S G	53
Syndecan 2	57	D D D D Y S S A S G S G	68
Serglycin	61	D D Y G S G S G S G	70
CD44	1	D D D V S S G S S S	10

Figure 4 Putative GAG-attachment regions of XS-2

Amino acid sequences in the ectodomain of XS-2 were analysed by comparing them with known sequences in the Brookhaven Protein Data Bank, Genbank and the EMBL Data Bank using the BLAST program. Results indicate significant similarity ($P < 0.01$) between a 12-amino acid region in the XS-2 ectodomain and the putative GAG-attachment sites found within the heparan sulphate proteoglycans, rat syndecan 2, serglycin [31] and CD44 [32]. Amino acid residues that differ between these sequences are presented in plain type. The potential XS-2 GAG-attachment sequence between amino acid residues 42 and 53 shows the sequence Ser-Gly-Ser-Gly with acidic residues N-terminal to the serine. This sequence is in agreement with the consensus sequence described by Bourdon et al. [27].

Figure 5 Dot-blot analysis of secreted XTC products chromatographed against DEAE-Sephacel

Secreted XTC cell products chromatographed against DEAE-Sephacel and eluted in 1 M NaCl buffer were bound to positively charged nylon membranes (Biodyne; BRL-Gibco) using a dot-blot apparatus (BRL-Gibco) and incubated with a polyclonal antiserum raised against recombinant XS-2 ectodomain expressed as a bacterial fusion protein. A, 1 µg of fusion protein incubated with preimmune serum; B, 1 µg of fusion protein incubated with XS-2 polyclonal antiserum; C, 0.2 M NaCl/6 M urea wash of DEAE-Sephacel column exposed beforehand to XTC-secreted products; D, 0.2 M NaCl wash of DEAE-Sephacel column exposed beforehand to XTC-secreted products; E, 1 M NaCl eluate of XTC-secreted products previously bound to DEAE-Sephacel.

To analyse further the material identified by the antisera, the material that eluted from DEAE-Sephacel in 1 M NaCl was separated by SDS/PAGE and transferred to a positively charged membrane. The undigested proteoglycan did not enter the 12% polyacrylamide gel used to identify low-molecular-mass proteins of the size predicted for the XS-2 core protein (21 kDa). However, digestion with the heparitinase before electrophoresis resulted in

XS-2	132	RTEVL AA VIAGG G IGFLFAVFLILLVYRMRKKDEGSY	DLGEG K- PSRAVYQKAPTK-- EFYA	191
Syndecan 2	152	RTEVL AA VIAGG V IGFLFAIFLILLVYRMRKKDEGSY	DLGER K- PSSAAYQKAPTK-- EFYA	211
Syndecan 4	147	RTEVL AA LIVGG V VGILFAVFLILLVYRMRKKDEGSY	DLG-K K P---IYKKAPT-- EFYA	202
Syndecan 3	293	RKEVL VA VIVGG V VGALFAAFVLTLLIYRMRKKDEGSY	TLEEP KQ A-SVTYQK-PDKQEEFYA	349
Syndecan 1	250	RKEVL GG VIAGG L VGLIFAVCLVAFMLYRMRKKDEGSY	SLEEP KQ ANGAYQK-PTKQEEFYA	306
<div style="display: flex; justify-content: space-around; align-items: center;"> <div style="border-top: 1px solid black; width: 40%;"></div> <div style="border-top: 1px solid black; width: 40%;"></div> </div> <div style="display: flex; justify-content: space-around; align-items: center;"> Transmembrane domain Cytoplasmic domain </div>				

Figure 3 Comparison of the amino acid sequence of the XS-2 putative transmembrane and cytoplasmic domains with corresponding sequences from syndecans 1, 2, 3 and 4

The amino acid sequences of the putative transmembrane and cytoplasmic domains of XS-2 are compared with sequences contained within the comparable domains of rat syndecan 2, rat syndecan 4, rat syndecan 3 and rat syndecan 1. Amino acid residues, presented in outline form, are conserved among all these proteins with respect to their position within their respective domains and their chemical characteristics (charge and philicity). Amino acid residues presented in plain type are not conserved by these criteria. Spaces have been inserted between amino acid sequences to generate the greatest degree of alignment for this analysis.

Figure 6 Western-blot analysis of proteins separated by SDS/PAGE

Proteins were electrophoresed in a 12% polyacrylamide gel under non-reducing conditions, transferred to a positively charged nylon membrane and incubated with polyclonal antiserum raised against a fusion protein encoding the XS-2 ectodomain. Migratory positions of globular molecular-mass markers are shown on the left. A, XTC-cell-secreted products eluted from DEAE-Sephacel in high salt and digested with heparitinase before electrophoresis; B, heparitinase alone; C, XTC-cell-secreted products eluted from DEAE-Sephacel in high salt and digested with chondroitinase ABC before electrophoresis; D, chondroitinase ABC alone.

Figure 7 Northern-blot analysis of poly(A)⁺ mRNA isolated from *Xenopus* embryonic cell line, XTC, and stage-7 embryos

Poly(A)⁺ mRNAs isolated from a *Xenopus* embryonic cell line and from stage-7 embryos were electrophoresed in a 0.8% formaldehyde denaturing gel, blotted on to a nylon membrane and hybridized with ³²P-labelled XS-2.

the identification of a protein band with an apparent molecular mass of 46 kDa that strongly reacted with the antisera and a protein band of 40 kDa that reacted weakly (Figure 6, lane A). Heparitinase alone did not react with the antisera (Figure 6, lane B). In contrast, no proteins were identified by the antisera after digestion with chondroitinase ABC alone (Figure 6, lane C). No change in the mobility of the core protein identified after heparitinase digestion was observed after simultaneous digestion with both heparitinase and chondroitinase ABC (results not shown). As demonstrated in Figure 6 (lane D), the anti-XS-2 antibodies recognized components present in the chondroitinase ABC preparation itself. These results suggest that the XS-2 translation product is modified by one or more heparan sulphate side chains and not by chondroitin/dermatan sulphate chains.

XS-2 mRNA is expressed during *Xenopus* development

Hybridization of ³²P-labelled XS-2 cDNA to *Xenopus* poly(A)⁺ mRNA isolated from whole embryos or embryonic cells (XTC) identified two mRNA species, about 4 and 2 kb in size (Figure 7). The intensity of hybridization with both of these species was similar. The finding that these XS-2 mRNAs are larger than the XS-2 cDNA is consistent with the observation that, although

Figure 8 Northern-blot analysis of total RNA isolated from embryos at different stages of development

Total RNA isolated from *Xenopus* embryos at different stages of development [20] was electrophoresed in a 0.8% formaldehyde denaturing gel, blotted on to nylon membrane and hybridized with ³²P-labelled XS-2. Comparison of the ethidium bromide-stained 28 S and 18 S ribosomal markers in each lane indicated a similar loading of total RNA in each lane. The developmental stage at which each sample of RNA was isolated is indicated at the top. The lane marked with an asterisk corresponds to stage-12 RNA which was partially degraded.

the 3' end of the cDNA contains two poly(A)⁺ signal sequences, it does not encode a poly(dA) sequence. Therefore it is likely that the mRNA extends further 3' than is encoded by the cDNA. This is in keeping with the general observation that oligo(dT) primers used for first-strand cDNA synthesis anneal to sequences other than poly(dA).

Northern-blot analysis of total RNA isolated from *Xenopus* embryos at a variety of developmental stages revealed that XS-2 is expressed in the blastula (64 cell and stage 7), gastrula (stage 10.5), neurula (stages 17 and 20/21), during the period of organogenesis (stages 26, 33/34 and 35/36) and in the tadpole (stage 54) (Figure 8). Total embryonic expression of XS-2 increases at stage 17, is then relatively unchanged at stages 20/21 and 26 and then increases at stage 33/34. This level of expression is similar in the stage-35/36 and -54 embryo.

DISCUSSION

The studies described in this report indicate that a cell-surface integral membrane heparan sulphate proteoglycan is expressed in the *Xenopus* embryo. The derived domain structure of the cDNA, XS-2, is highly suggestive of a core protein which consists of an extracellular domain containing GAG-attachment sites, a transmembrane domain and a cytoplasmic domain. The sequence similarity between the XS-2 transmembrane and cytoplasmic domains to those of the four mammalian syndecans provides compelling evidence that XS-2 is a syndecan and, more specifically, is the *Xenopus* counterpart of syndecan 2. The XS-2 transmembrane and cytoplasmic domains are 92% identical with those of rat syndecan 2 (rat fibroglycan) (Figure 3). The relatedness of XS-2 to syndecan 2 is further emphasized by the similarity between their putative GAG-attachment sites (Figure 4). The lack of similarity of the remainder of the XS-2 ectodomain to that of mammalian syndecans 1–4 and *Drosophila* syndecan is characteristic of the syndecan family. Whereas the cytoplasmic and transmembrane domains of all four mammalian types are highly conserved, the ectodomains diverge between family

members except near their N-termini where each shows a cluster of GAG-attachment sites of the general form Ser-Gly-(Xaa)_n-Ser-Gly-Ser-Gly [1], as does XS-2.

Evidence that the mature XS-2 translation product is a proteoglycan was obtained using polyclonal antiserum generated against the recombinant XS-2 ectodomain. This antiserum identifies a fraction of secreted XTC cell products which bind to DEAE-Sephacel under low-salt (0.2 M NaCl) conditions and which are eluted in high-salt (1 M NaCl) buffer. This fraction could not be detected in high-percentage polyacrylamide gels unless it was treated with heparitinase; a specific protein band of 46 kDa was then identified with the antiserum. In contrast, no reactive protein could be detected after digestion with chondroitinase ABC. These results indicate that the mature XS-2 translation product is a heparan sulphate proteoglycan. These results are also compatible with the presence of a sequence motif in the putative XS-2 ectodomain that closely resembles the Ser-Gly repeat and flanking sequences of other known heparan sulphate proteoglycans (rat syndecan 2, serglycin and CD44).

The XS-2 cDNA sequence predicts that the XS-2 translation product is approx. 21 kDa in size. This is identical with the predicted syndecan-2 protein product as cloned from rat liver [29]. However, the anti-XS-2 serum detected an approx. 46 kDa protein after digestion of the DEAE-Sephacel-bound fraction with heparitinase. A similar discrepancy in the size on SDS/PAGE of the syndecan-2 core protein isolated from rat liver plasma membranes has been reported [29]. Several possible explanations may account for this discrepancy. It has been generally observed that syndecan core proteins migrate as anomalously large proteins on PAGE [1]. This may be secondary to non-covalent dimerization of the core proteins, stripped of their chains, when separated under electrophoretic conditions. Alternatively, the glycosylation of the core proteins may lead to slower than expected migration. The finding of many possible O-linked glycosylation sites in XS-2 is in keeping with this possibility. The core proteins may have highly extended structures causing their spurious migration. Finally, it is possible that the antibody directed against the XS-2 ectodomain cross-reacts with another unrelated protein secreted by XTC cells. Although possible, this seems improbable given the uniqueness of the XS-2 ectodomain sequence, the dependence on prior heparitinase treatment and the similar behaviour of the rat syndecan-2 core protein when analysed by PAGE.

XS-2 appears to be expressed throughout embryogenesis. XS-2 mRNA is expressed as early as the 64-cell stage and as late as the tadpole stage of development. These results suggest that the expression of XS-2 may be different from that of mouse syndecan 1 and 2 during embryogenesis. XS-2 mRNA expression at the 64-cell stage of development indicates that it is expressed as a maternal transcript because zygotic transcription does not occur in *Xenopus* until just after the 12th cleavage [20]. Although murine syndecan 1 is similarly expressed in the early blastula, it is expressed as a zygotic not a maternal transcript [12]. More fundamental differences may exist between the expression of XS-2 and mouse syndecan 2. In contrast with XS-2 mRNA, the mouse syndecan-2 core protein is only detectable relatively late in embryogenesis, during organogenesis [33]. This suggests that whereas both XS-2 and mouse syndecan 2 may have a role in organogenesis, XS-2 may also be involved in earlier developmental processes, occurring in the blastula and gastrula. Further analysis of these potential differences between XS-2 and mouse syndecan-2 expression will depend on the study of mouse syndecan-2 mRNA expression and XS-2 core protein expression during embryogenesis.

From the Northern-blot analysis of XS-2 expression, XS-2

mRNA levels appear to increase during the neurula stage of development (stage 17) and again during organogenesis (stage 33/34). Interpretation of the significance of these changes in expression levels awaits studies of the spatial expression of XS-2 at these embryonic stages. Because syndecan-1 expression is both transcriptionally and post-transcriptionally regulated during murine metanephric differentiation [34], complementary analyses of XS-2 protein expression during *Xenopus* embryogenesis are also needed.

The domain structure and temporal expression of XS-2 provide a basis for speculation about the possible roles of this molecule during *Xenopus* development. The similarity between the XS-2 domain structure to that of the mammalian syndecans suggests that XS-2 expressed on the surface of embryonic cells may promote cell-cell interactions as well as interactions between cells and the surrounding extracellular matrix. Modification of the XS-2 core protein with heparan sulphate chains may provide a structural basis for XS-2 acting as a co-receptor for bFGF and other heparan sulphate-binding growth factors. Such a role for XS-2 is supported by several types of evidence. First, heparan sulphate is a necessary cofactor for bFGF signalling [8]. Secondly, both XS-2 and bFGF are expressed in the *Xenopus* blastula and gastrula [14,15]. Finally, heparitinase treatment to degrade heparan sulphate proteoglycans in the early blastula blocks normal mesodermal development and treatment of animal caps with heparitinase blocks normal responsiveness to induction to bFGF [35]. Several cell-surface heparan sulphate proteoglycans may be expressed during early *Xenopus* development; our studies suggest that XS-2 is one of them. Future studies aimed at determining the spatial distribution of XS-2 as well as that of growth factors such as bFGF will help to elucidate further the role of XS-2 during *Xenopus* development.

We thank Olga Goldberger for assistance in preparing syndecan-1 oligonucleotide primers, Celia Lee, M.Sc., for expert technical assistance and Dr. Masato Kato, Dr. Jurg Spring and Dr. Leonard Zon for helpful discussions during the course of these studies. Christie Chantler provided expert secretarial assistance in the preparation of this manuscript. This work was supported by a National Institutes of Health Individual Physician Scientist Award to N.D.R., NIH grants 5R37HD06763 and 5R01CA28735 to M.B., and institutional NIH Training Grant T32HD07268.

REFERENCES

- Bernfield, M., Kokenyesi, R., Kato, M. et al. (1992) *Annu. Rev. Cell Biol.* **8**, 1–40
- Spring, J., Paine-Saunders, S. E., Hynes, R. O. and Bernfield, M. (1994) *Proc. Natl. Acad. Sci. U.S.A.* **91**, 3334–3338
- Marcum, J. A., Atha, D. M., Fritze, L. M. S., Nawroth, P., Stern, D. and Rosenberg, R. D. (1986) *J. Biol. Chem.* **261**, 7507–7517
- Koda, J. E., Rapraeger, A. and Bernfield, M. (1985) *J. Biol. Chem.* **260**, 8157–8162
- Elenius, K., Salmivirta, M., Inki, P., Mali, M. and Jalkanen, M. (1990) *J. Biol. Chem.* **265**, 17837–17843
- Saunders, S. and Bernfield, M. (1988) *J. Cell Biol.* **106**, 423–430
- Kiefer, M. C., Stephans, J. C., Crawford, K., Okino, K. and Barr, P. J. (1990) *Proc. Natl. Acad. Sci. U.S.A.* **87**, 6985–6989
- Rapraeger, A. C., Krupka, A. and Olwin, B. B. (1991) *Science* **252**, 1705–1708
- Yayon, A. M., Klagsburn, M., Esko, J. D., Leder, P. and Ornitz, D. M. (1991) *Cell* **64**, 841–848
- Chernousov, M. A. and Carey, D. J. (1993) *J. Biol. Chem.* **268**, 16810–16814
- Roghani, M., Mansukhani, A., Dell'Era, P. et al. (1994) *J. Biol. Chem.* **269**, 3976–3984
- Sutherland, A. E., Sanderson, R. D., Mayers, M. et al. (1991) *Development* **113**, 339–351
- Trautman, M. S., Kimmelman, J. and Bernfield, M. (1991) *Development* **111**, 213–220
- Kimelman, D. and Maas, A. (1992) *Development* **114**, 261–268
- Amaya, E., Musci, T. J. and Kirschner, M. W. (1991) *Cell* **66**, 257–270
- Green, J. B. A., New, H. V. and Smith, J. C. (1992) *Cell* **71**, 731–739
- Rubin, G. M. (1988) *Science* **240**, 1453–1459
- Saunders, S., Jalkanen, M., O'Farrell, S. and Bernfield, M. (1989) *J. Cell Biol.* **108**, 1547–1556

-
- 19 Condie, G. C. and Harland, R. M. (1987) *Development* **101**, 93–105
- 20 Neuwkoop, P. and Faber, J. (1967) *Normal Table of Xenopus Laevis* (Daudin), North-Holland, Amsterdam
- 21 Auffray, C. and Rougeon, F. (1980) *Eur. J. Biochem.* **107**, 303–314
- 22 Smith, J. C. and Tara, J. R. (1991) in *Xenopus laevis: Practical Use in Cell and Molecular Biology* (Kay, B. K. and Peng, H. B., eds.), pp. 635–653, Academic Press, San Diego
- 23 Smith, D. B. and Johnson, K. S. (1988) *Gene* **67**, 31–40
- 24 Kozak, M. (1989) *J. Cell Biol.* **108**, 229–241
- 25 Kyte, J. and Doolittle, R. F. (1982) *J. Mol. Biol.* **157**, 105–132
- 26 Ktistakis, N. T., Thomas, D. N. and Roth, M. G. (1990) *J. Cell Biol.* **111**, 1393–1407
- 27 Bourdon, M. A., Krusius, T., Campbell, S., Schwartz, N. B. and Rouslahti, E. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 3194–3198
- 28 Cizmeci-Smith, G., Asundi, V., Stahl, R. C. et al. (1992) *J. Biol. Chem.* **267**, 15729–15736
- 29 Pierce, A., Lyon, M. and Hampson, I. N. (1992) *J. Biol. Chem.* **267**, 2894–3900
- 30 Kojima, T., Shworack, N. W. and Rosenberg, R. D. (1992) *J. Biol. Chem.* **267**, 4870–4877
- 31 Stevens, R. L. S., Avraham, S., Gartner, M. C., Bruns, G. A. P., Austen, K. F. and Weis, J. H. (1988) *J. Biol. Chem.* **263**, 7287–7291
- 32 Brown, T. A., Bouchard, T., St. John, T., Wayner, E. and Carter, W. G. (1991) *J. Cell Biol.* **113**, 207–221
- 33 David, G., Bai, X.-M., Van der Schueren, B., Marynen, P., Cassiman, J.-J. and Van den Berghe, H. (1993) *Development* **119**, 841–854
- 34 Vainio, S., Jalkanen, M., Bernfield, M. and Saxen, L. (1992) *Dev. Biol.* **152**, 221–232
- 35 Brickman, M. C. and Gerhart, J. C. (1994) *Dev. Biol.* **164**, 484–501
-

Received 23 December 1994/24 February 1995; accepted 27 February 1995