Molecular Cloning and Characterization of *N*-Syndecan, a Novel Transmembrane Heparan Sulfate Proteoglycan

David J. Carey, Diane M. Evans, Richard C. Stahl, Vinod K. Asundi, Kimberly J. Conner, Peter Garbes, and Gunay Cizmeci-Smith

Sigfried and Janet Weis Center for Research, Geisinger Clinic 26-13, Danville, Pennsylvania 17822

Abstract. A cDNA clone coding for a membrane proteoglycan core protein was isolated from a neonatal rat Schwann cell cDNA library by screening with an oligonucleotide based on a conserved sequence in cDNAs coding for previously described proteoglycan core proteins. Primer extension and polymerase chain reaction amplification were used to obtain additional 5′ protein coding sequences. The deduced amino acid sequence predicted a 353 amino acid polypeptide with a single membrane spanning segment and a 34 amino acid hydrophilic COOH-terminal cytoplasmic domain. The putative extracellular domain contains three potential glycosaminoglycan attachment sites, as well as a domain rich in Thr and Pro residues. Analysis of the cDNA and deduced amino acid sequences revealed a high degree of identity with the transmembrane and cytoplasmic domains of previously described proteoglycans but a unique extracellular domain sequence. On Northern blots the cDNA hybridized to a single 5.6-kb mRNA that was present in Schwann cells, neonatal rat brain, rat heart, and rat smooth muscle cells. A 16-kD protein fragment encoded by the cDNA was expressed in bacteria and used to immunize rabbits. The resulting antibodies reacted on immunoblots with the core protein of a detergent extracted heparan sulfate proteoglycan. The core protein had an apparent mass of 120 kD. When the anti-core protein antibodies were used to stain tissue sections immunoreactivity was present in peripheral nerve, newborn rat brain, heart, aorta, and other neonatal tissues. A ribonuclease protection assay was used to quantitate levels of the core protein mRNA. High levels were found in neonatal rat brain, heart, and Schwann cells. The mRNA was barely detectable in neonatal or adult liver, or adult brain.

HEPARAN sulfate proteoglycans (HSPG)[1] are ubiquitous components of the plasma membranes of mammalian cells (Fransson, 1987; Kjellen and Lindahl, 1991). Biochemical studies have identified both peripherally associated and integral membrane proteoglycans. The latter group includes proteoglycans with transmembrane core proteins (Saunders et al., 1989; Marynen et al., 1989) as well as proteoglycans that are attached to membranes by glycosylphosphatidylinositol (GPI) membrane anchors (Carey and Stahl, 1990; David et al., 1990). The specific functions of membrane proteoglycans are not known. Some of them have been shown to bind in vitro to a variety of extracellular matrix (ECM) adhesive proteins (Saunders and Benfield, 1988; Carey et al., 1990; Sun et al., 1989; Elenius et al., 1990; Salmivirta et al., 1991) and to be capable of mediating cell adhesion to immobilized matrix proteins. There is also evidence that cell surface HSPGs serve as low-affinity binding sites for members of the heparin binding growth factor family, the best studied of which is basic fibroblast growth factor (bFGF) (Kiefer et al., 1990; Vlodavsky et al., 1991). Recent experiments indicate that binding of bFGF to these low affinity sites may be essential for binding to high-affinity signal transducing receptors (Yayon et al., 1991; Rapraeger et al., 1991).

cDNAs coding for two transmembrane cell surface HSPGs have been cloned and sequenced (Saunders et al., 1989; Marynen et al., 1989; Mali et al., 1990; Kiefer et al., 1990). The proteoglycan structures predicted from the cDNA sequences exhibit a similar overall structural organization, with large *N*-terminal extracellular domains containing attachment sites for four to five glycosaminoglycans, a single hydrophobic membrane-spanning segment, and a short COOH-terminal cytoplasmic domain. A comparison of the amino acid sequences of these core proteins reveals little sequence identity in the extracellular domains of these proteoglycans, but extensive identity in the transmembrane and cytoplasmic domains. This had led to the suggestion that these proteins represent a gene family of membrane proteoglycan core proteins (Mali et al., 1990).

The distribution of one of these proteoglycans, syndecan, which was originally identified in murine epithelial cells, has been investigated in some detail. Based on immunocytochemistry and Northern blot analysis it was suggested initially that syndecan expression is restricted to epithelial cells (Hayashi et al., 1989) and certain subclasses of lymphocytes

1. *Abbreviations used in this paper*: bFGF, basic fibroblast growth factor; ECM, extracellular matrix; GPI, glycosylphosphatidylinositol; HSPG, heparan sulfate proteoglycan; PCR, polymerase chain reaction.

(Sanderson et al., 1989). Recent experiments from several laboratories, however, indicate a more widespread distribution of syndecan (Vainio et al., 1989; Elenius et al., 1991; Smith, G., V. Asundi, R. Stahl, L. Teichman, D. Carey, manuscript submitted for publication). Data on the tissue distribution of the other transmembrane core protein, which was originally identified in and cloned from human lung fibroblasts, has not been reported.

A major unresolved question is whether there are additional members of this gene family that account for integral membrane proteoglycans present in other cells and tissues. We have been studying membrane HSPGs synthesized by rat Schwann cells, and have obtained biochemical evidence for hydrophobic HSPGs that are attached to membranes by GPI anchors as well as HSPGs that are attached by other mechanisms and are, presumably, transmembrane proteins (Carey and Evans, 1989). In this paper we report on the cloning and characterization of a transmembrane HSPG that is expressed by Schwann cells and in the central nervous system, and is a novel member of the syndecan family of transmembrane proteoglycan core proteins. Because of its isolation initially from neural cells and its high level of expression in the nervous system, we propose the name *N*-syndecan for this HSPG.

Materials and Methods

Library Construction and Screening

A λgt11 Schwann cell cDNA library was prepared and used to isolate the initial proteoglycan cDNA clone. Schwann cells were cultured from the sciatic nerves of 2–4-d old rats as described previously (Carey and Stahl, 1990). Total RNA was isolated by extraction with RNAzol (Cinna/Biotecx, Friendswood, TX) and polyA$^+$ RNA was isolated by adsorption to oligo dT cellulose (Invitrogen, San Diego, CA). 5 μg of polyA$^+$ RNA was used for first strand synthesis. A cDNA Synthesis System Plus kit (Amersham Life Science Products, Arlington Heights, IL) was used for cDNA synthesis reactions. Two separate reactions, primed by oligo dT and random hexanucleotides, were carried out. After completion of the second strand synthesis the reaction products were combined, ligated to EcoRI linkers, inserted into the EcoRI site of λgt11, and packaged as bacteriophage using the cDNA Cloning System kit (Amersham). The library was amplified one time before its use for screening.

For library screening *Escherichia coli* cells, strain Y1090, were infected with the bacteriophage and plated on LB agar plates with ampicillin. Filter lifts were prepared from 450,000 recombinant plaques onto MagnaGraph nylon filters (Micron Separations, Inc., purchased from Fisher Scientific, Pittsburgh, PA) and hybridized to a ^{32}P-labeled oligonucleotide probe (5′-GTAGCTGCCTTCGTCCTTCTTCTTCATCCGGTA) complementary to a conserved 33-base sequence in the cytoplasmic domains of the murine epithelial and human lung fibroblast transmembrane core protein sequences. The filters were hybridized overnight at 42°C in 5× SSC, 5× Denhardt's solution, 10 μg/ml yeast tRNA, 2.2 mg/ml tetrasodium pyrophosphate, and 0.1% SDS (1× SSC = 0.15 M NaCl, 0.015 M sodium citrate, pH 7.2). The filters were washed at 42°C with 2× SSC, 0.1% SDS, and exposed to X-Omat AR film (Eastman Kodak Co.) at −70°C with intensifying screens. A single positive clone was obtained that contained a cDNA insert of approximately 2 kb. The cDNA insert was excised from the bacteriophage DNA by digestion with EcoRI and subcloned into the EcoRI site of the plasmid pGEMz7-f (Promega, Madison, WI), to produce the plasmid pl04. Both strands of the cDNA insert were sequenced by the dideoxy chain termination procedure using T7 DNA polymerase (from Promega or U.S. Biochemicals, Cleveland, OH).

Primer-extended cDNA Isolation

Analysis of the cDNA sequence of plasmid pl04 revealed that the isolated clone did not contain the entire protein coding region of the mRNA. To obtain additional sequence from the 5′ end of the mRNA we carried out cDNA synthesis reactions using specific primers and amplified the sequences by PCR. Typically, 5 μg of polyA$^+$ Schwann cell RNA was primed with an oligonucleotide primer and used for first strand synthesis with reverse transcriptase (GIBCO-BRL, Gaithersburg, MD). A poly G tail was added to the 3′ end of the cDNA using terminal deoxynucleotidyl transferase and dGTP, and the resulting products were used as templates for PCR amplification with polyC and a sequence specific oligonucleotide as primers. The PCR products were blunt end cloned into the SmaI site of pGEM-7z and their sequences determined by dideoxy chain termination. Two overlapping clones of approximately 250 and 750 nucleotides were obtained in this way.

Northern Blot Analysis

Aliquots of polyA$^+$ or total RNA were fractionated on formaldehyde-agarose gels, transferred to nylon membranes, and hybridized to radiolabeled cDNA probes, as described previously (Asundi et al., 1990). The probes were prepared by random prime labeling of fragments of pl04 cDNA sequences. Final washes of the filters were with 0.1× SSC, 0.1% SDS at 50°C.

Bacterial Expression and Purification of Proteoglycan Core Protein

A fragment of the pl04 cDNA was subcloned into the bacterial expression vector pEX2 (Clonetech Laboratories, Palo Alto, CA) to generate a β-galactosidase fusion protein for antibody production. In this vector the expression of the fusion protein is regulated by the λP_R promoter. The expression plasmid was grown in an *E. coli* strain that contains a temperature-sensitive allele of the λ repressor protein, cI857.

To generate the expression construct, the plasmid pl04 was digested with EcoRI and PstI to generate a 407-bp fragment of the cDNA insert coding for ~16 kD of proteoglycan core protein (corresponding to amino acids 229-353 in Fig. 1). This DNA fragment was gel purified and ligated to BamHI-EcoRI adapters (Boehringer-Mannheim Biochemicals, Indianapolis, IN) as described previously (Haymerle et al., 1986). After removal of excess linkers the insert DNA was phosphorylated using T4 polynucleotide kinase (Promega). The expression vector was digested with BamHI, ligated to BamHI-EcoRI adapters, digested with PstI and ligated to the phosphorylated and adapted insert DNA. Ampicillin resistant transformants of *E. coli* strain NS428 (American Type Culture Collection, Rockville, MD) were selected for growth at 30°C on LB-ampicillin plates. Transformants containing plasmids with inserts were identified by colony hybridization. One of several colonies that expressed a 131-kD β-galactosidase fusion protein was selected and used for large scale production and purification of the fusion protein.

For large scale expression of the fusion protein 5 × 200 ml cultures of LB broth with ampicillin were inoculated with cells containing the fusion plasmid and grown at 30°C to mid-log phase. Fusion protein synthesis was induced by growing the cells at 42°C for 2 h. The cells were harvested and lysed as described previously (Milman, 1987). The cell pellet was extracted with 4 M urea, 50 mM Tris-HCl, pH 7.4, 1 mM PMSF, followed by 8 M urea, 50 mM Tris-HCl, pH 7.4. The latter extract contained the fusion protein as determined by SDS gel electrophoresis. The fusion protein was purified by chromatography on Sepharose CL-6B and Q-Sepharose Fast-flow (Pharmacia-LKB Biotechnology, Piscataway, NJ).

Antibody Preparation

The purified fusion protein was emulsified with RAS adjuvant (RIBI ImmunoChem Research, Inc., Hamilton, MT) and injected into rabbits (300 μg/immunization) at approximately 2-wk intervals. The titer and specificity of the antisera were tested by their ability to stain the fusion protein on immunoblots.

The anti-core protein antibodies were affinity purified by chromatography on a column of immobilized fusion protein made by coupling purified fusion protein to CNBr-activated Sepharose (Pharmacia-LKB). The unbound proteins and antibodies were removed by washing with 0.15 M NaCl, 0.05 M sodium phosphate, pH 7.5. The specifically bound antibodies were eluted with 0.1 M glycine, pH 2.5.

Immunoblotting

Aliquots of detergent extracts from cells or tissues were subjected to SDS-gel electrophoresis in 7.5% polyacrylamide gels, and the proteins were

transferred to Immobilon-P membranes (Millipore Corp., Bedford, MA) by electrophoretic transfer in 12.5 mM Tris, 96 mM glycine for 2 h at 70 V. The membranes were blocked in 5% instant nonfat milk, 0.1 M NaCl, 0.02 M Tris-HCl, pH 7.5, and then incubated overnight with affinity-purified anti-core protein antibodies diluted in blocking buffer. After washing the membranes, the bound antibodies were visualized by incubating the filters in secondary goat anti-rabbit IgG-alkaline phosphatase conjugate (Promega) and production of alkaline-phosphatase reaction products using nitroblue tetrazolium (0.033%) and 5-bromo-4-chloro-3-indolylphosphate (0.017%) as substrates. The proteoglycan was subjected to deglycosylation by digestion with nitrous acid (Shively and Conrad, 1976), heparitinase (Carey and Stahl, 1990), or chondroitinase ABC (Hamati et al., 1989).

Immunofluorescence Microscopy

Immunofluorescent staining was carried out essentially as described previously (Carey and Stahl, 1990). Freshly dissected tissues from neonatal rats were frozen and sectioned with a cryostat at a thickness of 5-10 μm. The sections were fixed in 3% paraformaldehyde and then stained with the anti-core protein antibodies. The antibodies were diluted 1:50 to 1:100 in blocking buffer (see above) and incubated with the tissue sections at room temperature for 1 h. After rinsing with blocking buffer, the bound antibodies were visualized by incubating with affinity purified fluorescein-conjugated goat anti-rabbit IgG (Sigma Chemical Co., St. Louis, MO). The sections were examined on an inverted light microscope (Axiovert 35; Carl Zeiss, Oberkochen, Germany) equipped for epifluorescence.

Ribonuclease Protection Assays

For quantitative measurements of core protein mRNA levels a ribonuclease protection assay was used. Samples of total RNA (1-10 μg) isolated from cultured cells or tissues were hybridized in solution to a high sp act 400 nucleotide long anti-sense RNA probe complementary to a region in the 3'-untranslated portion of *N*-syndecan mRNA. The probe was synthesized by in vitro transcription using T7 RNA polymerase (Promega) in the presence of ^{32}P-UTP (800 Ci/mmole; NEN-DuPont, Boston, MA), with NsiI-digested pl04 as the template. Hybridization was carried out overnight at 42°C in buffer containing 80% formamide. After the hybridization step, single-stranded RNAs were digested with a mixture of RNAse and RNAse T1 (Ambion, Austin, TX). The samples were subjected to electrophoresis in 5% polyacrylamide gels containing 8 M urea. Radiolabeled bands representing the protected probe RNA molecules were visualized by autoradiography and quantitated with a radioanalytic scanner (Ambis Systems, San Diego, CA).

Results

Isolation and Sequence Analysis of N-syndecan cDNA

Inspection of the published cDNA and amino acid sequences of syndecan and the human lung fibroblast transmembrane proteoglycan core protein revealed a high degree of sequence homology in regions corresponding to the membrane and cytoplasmic domains (Saunders et al., 1989; Marynen et al., 1989). To determine whether these or similar proteins were expressed in rat Schwann cells, we synthesized a 33-base antisense oligonucleotide corresponding to the region of highest homology, and hybridized it to Schwann cell RNA that had been fractionated on a formaldehyde-agarose gel and transferred to a nylon membrane. Several bands that appeared to result from specific hybridization were visible after high stringency washes (not shown), indicating the presence of mRNAs in Schwann cells that contained this conserved transmembrane proteoglycan sequence.

To identify these proteins we used the oligonucleotide to screen a λgt11 Schwann cell cDNA library. A positive clone was obtained that when sequenced was found to contain a cDNA insert of 1,934 bp. Computer-assisted translation of the cDNA sequence revealed an open reading frame of 125 amino acids at one end, followed by ~1.55 kb of 3'-untranslated sequence, but lacking an apparent translational start site, polyadenylation signals, and poly A tail.

To obtain additional 5' protein coding sequences Schwann cell RNA was subjected to primer extension by reverse transcriptase and PCR amplification, as described in Materials and Methods. A set of two overlapping clones was obtained that provided an additional 670 bases of 5'-cDNA sequence. The cDNA and deduced amino acid sequences are shown in Fig. 1.

Predicted Structure of the Core Protein

The cDNA sequence obtained was 2,615-bp long, and contained an open reading frame of 1,059 bp plus 1,556 bp of 3'-untranslated sequence lacking a polyadenylation signal, and a poly A tail. The open reading frame sequence predicted a polypeptide of 353 amino acids. A proposed structural model of the protein is shown in Fig. 2. At the COOH-terminal end is an uninterrupted span of 25 hydrophobic amino acids followed by a hydrophilic domain of 33 amino acids, that includes a stretch of basic amino acids adjacent to the hydrophobic segment. These features suggested the protein was a type I transmembrane polypeptide with a short COOH-terminal cytoplasmic domain. In the presumed NH_2-terminal extracellular domain are several notable features, including a dibasic sequence just outside the hydrophobic membrane spanning segment, and 3 Ser-Gly sequences, that are potential sites of glycosaminoglycan attachment (Kjellen and Lindahl, 1991). These are clustered in the COOH-terminal one-third of the extracellular domain. No potential sites of Asn-linked oligosaccharide attachment are present in the sequence. The extracellular domain exhibited an unusual amino acid composition, with a high percentage of Thr and Pro and no Cys residues. A large block of over 200 amino acids in the center of the extracellular domain consists of 40% Thr and Pro. Because of the helix-breaking properties of these amino acids (Gibrat et al., 1987), this part of the polypeptide would be expected to exhibit an unusual three-dimensional structure.

There is an in-frame Met residue near the 5' end of the cDNA sequence followed by a basic residue and a short hydrophobic stretch that could comprise a membrane translocation signal peptide. There is a G residue in the DNA sequence at position -3 relative to this codon, thus making it a potential "strong" initiation codon (Kozak, 1989). The cDNA sequence is incomplete, however, based on the size of the mRNA detected by Northern blot analysis (see below). It appears that at least some of the missing cDNA sequence may be NH_2-terminal protein coding information. This conclusion is based on a comparison of the predicted mass of the deduced polypeptide (~35,000) and the apparent molecular mass of the immunoreactive HSPG core protein as determined by SDS gel electrophoresis (see below).

Sequence Homology to Other Transmembrane Proteoglycan Core Proteins

The predicted polypeptide exhibited an interesting pattern of sequence homology to two other transmembrane proteoglycans whose amino acid sequences have been reported. Fig. 3 shows the results of an alignment of the amino acid sequence of *N*-syndecan with those of murine syndecan and the human lung fibroblast proteoglycan core proteins. The ex-

```
CTT CGA GAG ACA GCC ATG CGG TTC ATT CCT GAC ATA GCC CTG GCT GCA CCC ACC GCA CCT GCC ATG CTA CCC ACG ACC GTT      81
L   R   E   T   A   M   R   F   I   P   D   I   A   L   A   A   P   T   A   P   A   M   L   P   T   T   V       27
ATC CAG CCC GTG GAT ACC CCG TTT GAG GAA CTC CTT TCT GAG CAT CCC GGC CCC GAA CCA GTC ACC AGT CCC CCA CTG GTG     162
I   Q   P   V   D   T   P   F   E   E   L   L   S   E   H   P   G   P   E   P   V   T   S   P   P   L   V       54
ACA GAG GTG ACA GAA GTT GTG GAA GAG CCC AGT CAG AGA GCC ACC ACT ATC TCC ACC ACC ACA TCT ACC ACT GCA GCC ACC     243
T   E   V   T   E   V   V   E   E   P   S   Q   R   A   T   T   I   S   T   T   T   S   T   T   A   A   T       81
ACC ACA GGG GCC CCA ACT ATG GCC ACA GCA CCT GCT ACA GCA GCC ACC ACT GCC CCT AGC ACT CCC GCG GCA CCC CCT GCC     324
T   T   G   A   P   T   M   A   T   A   P   A   T   A   A   T   T   A   P   S   T   P   A   A   P   P   A     108
ACG GCC ACC ACG GCT GAC ATA AGG ACC ACC GGC ATA CAA GGG CTG CTG CCT CTT CCC CTG ACC ACG GCT GCC ACA GCC AAG     405
T   A   T   T   A   D   I   R   T   T   G   I   Q   G   L   L   P   L   P   L   T   T   A   A   T   A   K     135
GCC ACT ACC CCA GCA GTA CCC TCA CCA CCC ACT ACT GTG ACT ACC TTG GAC ACA GAG GCC CCG ACA CCT AGG CTG GTC AAC     486
A   T   T   P   A   V   P   S   P   P   T   T   V   T   T   L   D   T   E   A   P   T   P   R   L   V   N     162
ACA GCT ACC TCT CGG CCA CGA GCC CTT CCT CGG CCA GTC ACC ACC CAG GAG CCT GAA GTT GCT GAG AGG AGT ACC CTG CCG     567
T   A   T   S   R   P   R   A   L   P   R   P   V   T   T   Q   E   P   E   V   A   E   R   S   T   L   P     189
TTG GGG ACC ACC GCT CCT GGA CCC ACA GAG GTG GCT CAG ACC CCA ACT CCA GAG TCC CTT CTG ACC ACG ACC CAG GAT GAG     648
L   G   T   T   A   P   G   P   T   E   V   A   Q   T   P   T   P   E   S   L   L   T   T   T   Q   D   E     216
CCA GAG GTG CCA GTA AGC GGG GGG CCC AGC GGG GAC TTT GAG CTG CAA GAA GAG ACC ACA CAG CCA GAC ACA GCC AAT GAA     729
P   E   V   P   V  (S)  G   G   P  (S)  G   D   F   E   L   Q   E   E   T   T   Q   P   D   T   A   N   E     243
GTG GTG GCG GTG GAA GGA GCC GCG GCC AAG CCA TCA CCT CCA CTG GGG ACA CTA CCC AAG GGT GCC CGC CCA GGC CTT GGC     810
V   V   A   V   E   G   A   A   A   K   P   S   P   P   L   G   T   L   P   K   G   A   R   P   G   L   G     270
CTC CAC GAC AAT GCC ATC GAC TCG GGC AGC TCT GCT GCC CAG CTC CCT CAG AAG AGC ATA CTA GAG CGG AAG GAG GTG CTC     891
L   H   D   N   A   I   D  (S)  G   S   S   A   A   Q   L   P   Q   K   S   I   L   E   R   K   E   V   L     297
GTA GCT GTG ATC GTA GGT GGC GTG GTG GGC GCC CTC TTC GCT GCC TTC CTG GTC ACG TTG CTC ATC TAC CGC ATG AAG AAG     972
V   A   V   I   V   G   G   V   V   G   A   L   F   A   A   F   L   V   T   L   L   I   Y   R   M   K   K     324
AAG GAC GAA GGC AGC TAC ACC TTG GAA GAG CCC AAG CAG GCA AGC GTC ACA TAC CAG AAG CCC GAC AAG CAG GAG GAG TTC    1053
K   D   E   G   S   Y   T   L   E   E   P   K   Q   A   S   V   T   Y   Q   K   P   D   K   Q   E   E   F     351
TAC GCT    TAGCAGAGAG AGCCACAGTG CCTCCTGCAG CCTCGACTCC GCCTTGCCCA GTCCCTGTCC CAACAACAGC CCAGXCCAAT CCTGGGCCTG    1149
Y   A                                                                                                         353
GGCCTGGGCC TGGGATGGAG CCTGGCCCTG CTTCCTTCTG CTCAGGCTGC TAGCTTAACA CAGACTGTCC TAAGGAGCAG AGGCGCCACC ATCTGCCCCA    1249
GACTGTGCCC GTATGACCCC TTCTCTAGGC CCATTCCCTC CAGCCTGGGG CTTCAGGATC TTGAGTCCCA TGGACAAGAG GAAGGAAGCC CTGGTTGCTG    1349
GTTGAAACAT GGGCGGGGCC AGGGTTAAGA TGGCCCACAG TGCTCTTCAC GCAGAGGCCA CCATGCTGGC TTCTAAAACC AACACATGGT ACATCTAGCC    1449
CTCCAATTCA ACATCATCCG GATGCTGAGG CCTCGAGCTG CTGTCCAAGG CCTCTCTCGA CAGGAGGGTG TCCTTTGTCA CCAGCCTGAG TCGTCGTGGA    1549
GTCACCGCCC CCTCTCCCCT GCCTAATGCA CATGTCCCCA GGCTGCACCT CTTTCTGCAC CTCCTCCCAG GGAAAGGGCT TCCTCAGTGC ACAGGGCAGT    1649
CAGTGGTAAG GGAGGCCACT CTAGCGTCCC TCCTGAGAGG GGATGTCACA TCGCCTTTGT CAACCACTGC TATAATCCTA TAATGCATTC ACACAGGAGA    1749
CAAAAACATA CCCAGTCCTG ACCACCCGGC CAGCGAGACA TCACACACAC GCACCACCTT CACGCAGTCT CAGACATCCT GACGCTTCTG CCACGTAGGC    1849
CACTGTCATC GTCACCAGTG GCAACCACGC CACCTTCCTC CCACTCTCTC TTTACACACA TTGGTCCTGC ATCTGGCTCC TCTAACCATC CGGGTCACTG    1949
AGGAAGGCAG GACCGAGTTG TAGGCATCAG CCCATATTGG GTCCCCCAGA GTCACCCTAT TCCACTTGGT CCCCACCATG ACACCTATAC CAGCCACACT    2049
GATGCTGATC CCAGCTGATG TCAGTCACAA CCCCACACAG ATGCAGGCTC GCCCTCCCCA CTGTGGCTCA TAGAGAAGTT TTGGGCTAGC CTCTGCCAGC    2149
CAAGTCTGAA GGAAGAAGGG AGCTAGGGTC TCCCAGGACC AAGGAGCTGT GTGGGTTGTC CCTGGTCTCT ATGTATTTCC TTTCTGCTCT GAGCCAGGAG    2249
CTGTCTCTGC CTCCCAGGAC ACGTGTCTCC AAAGTGCCTG CGAGGGCGGG CTCCGCCCAG GCCCTCTGTG TCCCCTGCCC TGCCCTGGCC TGGTGAGACC    2349
GACCTCAGGC CTACCTACTC TGGGTCCTCT CTGTGGACAC AACTGACCAG GCAACTTGGC AGTTCTGGCT ACAACTGGAT GGACTCATAC TCCAGCAGGC    2449
TCATCCCAAT GCTGTAGGCC CCAAGAGTCA TTGATGGGGA CAAGGGTGTC ACAGGCCTCA AGTCTACCAC CCTGTCTCTG TTATTCTTAG AAGTGGGGCT    2549
ATTCCAGAGG TATTTTAAGT GTGGGATCAC CATTCTCTTT GTGAGAGGGG TGTGCCGAGG GAATTC                                          2615
```

Figure 1. cDNA and predicted amino acid sequence of *N*-syndecan. The DNA sequence determined from overlapping cDNA clones coding for *N*-syndecan is shown; the deduced amino acid sequence is shown in single letter code below the DNA sequence. Potential glycosaminoglycan attachment sites are indicated by parentheses; the hydrophobic membrane spanning segment is underlined. This sequence is available from the EMBL/GenBank under accession number 63143.

tracellular domains show little sequence homology, with only 3% amino acid identity among the three proteins. In pairwise comparisons the homology was somewhat greater, with *N*-syndecan showing 25 and 17% amino acid identity with syndecan and the human lung fibroblast core protein, respectively. In contrast, there is striking sequence homology in the transmembrane and cytoplasmic domains, with 51% amino acid identity among the three proteins in these regions. In pairwise comparisons of *N*-syndecan with syndecan and the human lung fibroblast core protein there is 72 and 62% amino acid identity in these regions. Specific features that are conserved among the three polypeptides are dibasic amino acid sequences adjacent to the membrane insertion sites, and four Tyr residues in the cytoplasmic domains.

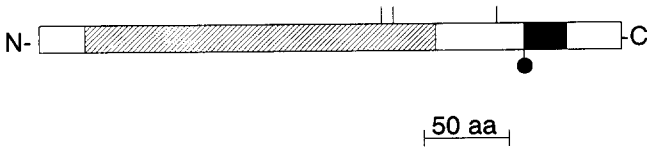

Figure 2. A model of *N*-syndecan structure. A schematic model of the proteoglycan, based on the deduced amino acid sequence, is shown. (■) Transmembrane segment. Vertical lines represent potential glycosaminoglycan attachment sites. (●) Dibasic amino acid motif conserved among the transmembrane proteoglycan sequences. (▨) Thr-Pro-rich portion of the extracellular domain.

There was little recognizable sequence homology among the putative glycosaminoglycan attachment sites except for the Ser-Gly sequences.

Characterization of the Proteoglycan

To provide evidence that the cloned cDNA encoded a proteoglycan core protein, and to investigate the properties of the proteoglycan, we produced antibodies in rabbits against a recombinant fragment of the protein expressed in bacteria. Fig. 4 shows the results that were obtained when these antibodies were affinity purified against the recombinant protein and used to stain immunoblots of proteins extracted by nonionic detergent from cultures of neonatal rat Schwann cells. The antibodies specifically stained a smear migrating near the top of the separating gel at >250,000 M_r (Fig. 4, lane 1). This is typical of the pattern that is obtained on immunoblots with proteoglycans. Pre-immune antibodies did not stain any material on the blots (Fig. 4, lane *2*). To verify that the immunostained smear was a proteoglycan and to identify the attached glycosaminoglycan chains the extracts were subjected to enzymatic and chemical degradation followed by immunoblot analysis. Treatment of the extracts with heparitinase resulted in a shift in the migration of the immunostained material to a less broad but still somewhat diffuse band migrating at an ~120,000 M_r (Fig. 4, lane *3*). An identical result was obtained when the extracts were subjected to nitrous acid digestion (Fig. 4, lane *4*). In contrast,

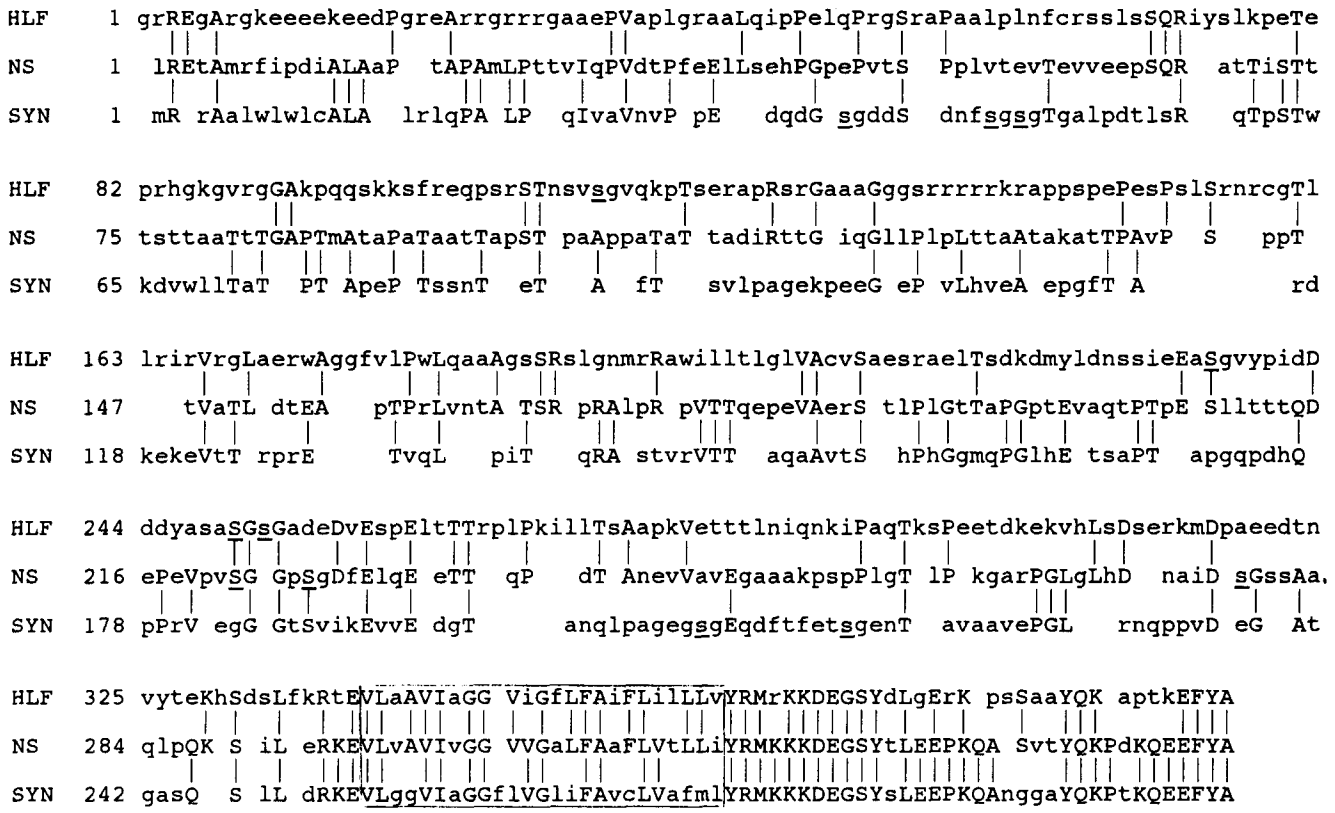

Figure 3. Amino acid sequence homology of transmembrane proteoglycan core proteins. The deduced amino acid sequences of the human lung fibroblast (*HLF*) HSPG core protein, *N*-syndecan (*NS*), and mouse syndecan (*SYN*) were compared using the multiple alignment program of IntelliGenetics Suite. Vertical lines and upper case letters indicate amino acid identities. The transmembrane segments are boxed. Gaps have been introduced to maximize alignment. The sequences were taken from references cited in the text.

digestion with chondroitinase ABC did not affect the migration in the gel of the immunostained material (Fig. 4, lane *5*). Similarly, digestion with *N*-glycanase or *O*-glycanase, either alone or in combination with heparitinase, did not affect the migration of the immunostained protein (not shown). Taken together, these results demonstrate that the antibodies recognized a heparan sulfate proteoglycan with a core protein of 120,000 M_r, and indicate that the cloned cDNA encodes a HSPG core protein.

Northern Blot Analysis

Fig. 5 shows the results of Northern blot analysis when the cloned cDNA was used as a probe. When polyA$^+$ RNA isolated from cultured neonatal rat Schwann cells was separated on a denaturing agarose gel and hybridized to a probe derived from the 3′-untranslated region of the cDNA, a single radiolabeled band was detected, migrating at an apparent size of ~5.6 kb (lane *1*). Identical results were obtained with probes derived from the combined transmembrane and cytoplasmic domains or the extracellular domain (not shown).

The cDNA probes also hybridized to mRNA species of the same size in samples obtained from newborn rat brain (Fig. 5, lane *2*), adult rat heart (lane *3*), or cultured vascular smooth muscle cells obtained from thoracic aortas of adult rats (lane *4*).

Immunocytochemistry

We used the affinity purified anti-recombinant core protein antibodies to investigate the distribution of the HSPG in several neural and nonneural tissues. When sections of neonatal rat sciatic nerve were stained with the antibodies specific staining was observed within the endoneurium (Fig. 6, *A* and *C*). The staining was not uniformly distributed and in some cases clearly delineated small ringlets. The pattern was distinctly different from that obtained when similar sections were stained with anti-laminin antibodies (Fig. 6 *D*) or antibodies against a GPI-anchored Schwann cell HSPG data (not shown; see Carey and Stahl, 1990). Laminin is present within the basement membrane that surrounds individual axon-Schwann cell units. These results suggest that *N*-syndecan is present in membranes of Schwann cells and/or axons at sites of axon–Schwann cell apposition, and not at sites of Schwann cell–basement membrane apposition, as has been observed for the GPI-anchored Schwann cell HSPG.

Northern blot analysis indicated that the mRNA coding for *N*-syndecan was present in newborn rat brain (see above). Immunofluorescent staining of sections of newborn rat brains confirmed this observation. Anti-*N*-syndecan staining was present throughout the brain cortex (Fig. 7 *A*) and appeared to outline the surface membranes of individual cells (Fig. 7 *B*). In other brain regions there was pronounced staining of specific cell layers, e.g., in the area of the choroid plexus (Fig. 7 *C*) and in the midbrain (Fig. 7 *D*). Immunoblot analysis of detergent extracts of neonatal rat brain produced results essentially identical to those obtained with Schwann cell extracts (not shown; see Fig. 4) confirming that the immunoreactive brain protein was *N*-syndecan.

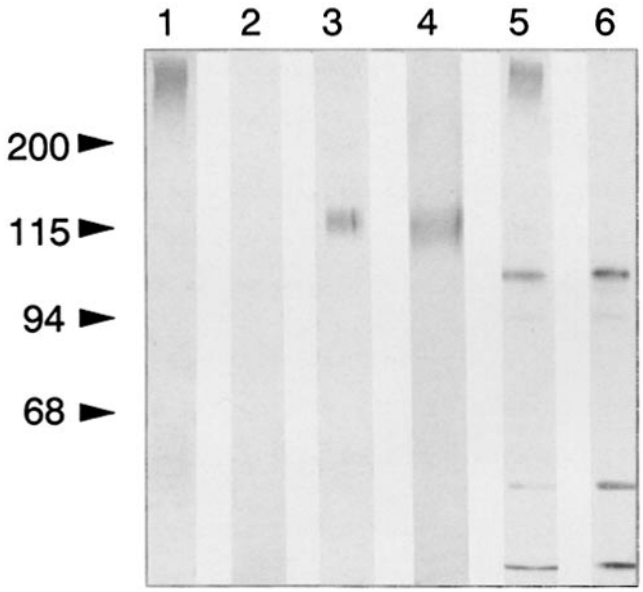

Figure 4. Immunoblot analysis of *N*-syndecan. Aliquots of detergent extracted proteins from cultures of neonatal rat Schwann cells were separated on 7.5% polyacrylamide gels, transferred to Immobilon-P membranes and stained with affinity-purified anti-*N*-syndecan antibodies (lanes *1*, and *3–5*) or preimmune serum (lane *2*) as described in Materials and Methods. The anti-*N*-syndecan antibodies were made by immunizing rabbits with purified recombinant *N*-syndecan fusion protein. Before electrophoresis some of the protein samples were subjected to either heparitinase (lane *3*), nitrous acid (lane *4*) or chondroitinase ABC (lane *5*) digestion. Lane *6* was the chondroitinase ABC enzyme preparation stained with the antibodies. The bands in lanes *5* and *6* at molecular weights of ~50,000 and 100,000 are bacterial proteins that appear to cross-react with antibodies against the β-galactosidase portion of the fusion protein that was used as immunogen.

In contrast to these results, when sections of embryonic (e.g., eighteenth day of gestation) or adult rat brains were examined, only faint staining with the antibodies was observed (not shown).

Consistent with the fairly wide distribution of core protein mRNA indicated by Northern blot analysis (see above), we observed anti-*N*-syndecan immunoreactivity in non-neural tissues as well. When sections of neonatal rat thoracic aorta were examined, staining was visible in the medial layer as circumferential bands that corresponded to the bands of smooth muscle cells present in the vessel wall (Fig. 8 *A*). Staining was also present in sections of neonatal rat heart throughout the myocardium (Fig. 8 *B*), as well as in sections of neonatal rat esophagus (Fig. 8 *C*) and tongue (Fig. 8 *D*). In the latter two tissues the most intense staining was in the stratified epithelial cell layers.

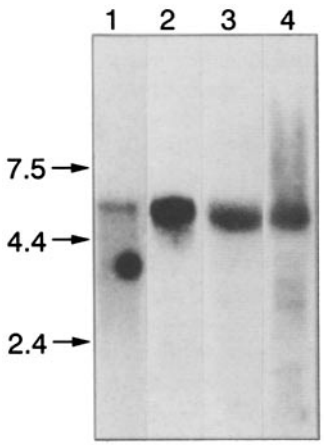

Figure 5. Northern blot analysis of *N*-syndecan mRNA. RNA samples were fractionated on formaldehyde-agarose gels, blotted onto nylon filters and hybridized to ^{32}P-labeled *N*-syndecan cDNA probes. (Lane *1*) Poly A$^+$ RNA from cultured rat Schwann cells; (lane *2*) total RNA from 2-d-old rat brain; (lane *3*) total RNA from adult rat heart; (lane *4*) total RNA from rat vascular smooth muscle cells. The probes were derived from the 3′-untranslated (lane *1*) or extracellular domain coding (lanes *2–4*) regions of *N*-syndecan cDNA.

Regulated Expression of Core Protein mRNA

The immunocytochemical studies suggested that, at least in the brain, expression of the core protein was significantly higher in newborn rats than in embryonic or adult animals. To investigate this further, we used a ribonuclease protection assay to quantitatively measure core protein mRNA levels in brain, liver, and heart tissue taken from embryonic, neonatal, and adult animals. We also compared these to mRNA levels in cultured neonatal rat Schwann cells and adult rat vascular smooth muscle cells. A radiolabeled antisense RNA probe synthesized by in vitro transcription of plasmid p104 was used for the assay. As shown in Fig. 9 *A* (lanes *1–4*) when increasing amounts of Schwann cell total RNA were used in the assay increasing amounts of protected radiolabeled probe were detected. Quantitation of the radioactivity in the gels revealed a linear relationship between input RNA and the amount of protected probe (Fig. 9 *B*), indicating that within this range the assay was accurate and linear. Results obtained with smooth muscle cell RNA produced similar results but with a slope that was only about one-tenth of the slope obtained with Schwann cell RNA (Fig. 9 *B*). Thus, the assay detected core protein mRNA in smooth muscle cells, but at a significantly reduced level relative to Schwann cells. Incubation of the probe with yeast total RNA did not result in the appearance of a protected probe band (Fig. 9 *A*, lane *14*).

Ribonuclease protection assay data for the tissue samples are shown in Fig. 9 *A* (lanes *5–13*) and Fig. 9 *C*. Consistent with the immunocytochemical data, the assay detected significant amounts of *N*-syndecan mRNA in neonatal rat brain but considerably reduced amounts in embryonic and adult brains. The *N*-syndecan mRNA level in neonatal rat brain exceeded that measured in a comparable amount of Schwann cell RNA. Results for heart RNA exhibited the same overall pattern, but the *N*-syndecan mRNA levels were lower in the heart than in the brain RNA samples from the same age animals. High levels of core protein mRNA were not detected in any RNA samples taken from liver. These data indicate there is both developmental and tissue specific regulation of core protein mRNA levels, with the highest levels occurring in neonatal rat brain, among the tissues examined.

Discussion

In this paper we have described the cloning of cDNA from neonatal rat Schwann cells that codes for a portion of a novel transmembrane HSPG. Based on its similarity to the previously described transmembrane proteoglycan syndecan (Saunders et al., 1989), its identification initially in neural cells, and the high level of expression of this HSPG in cells

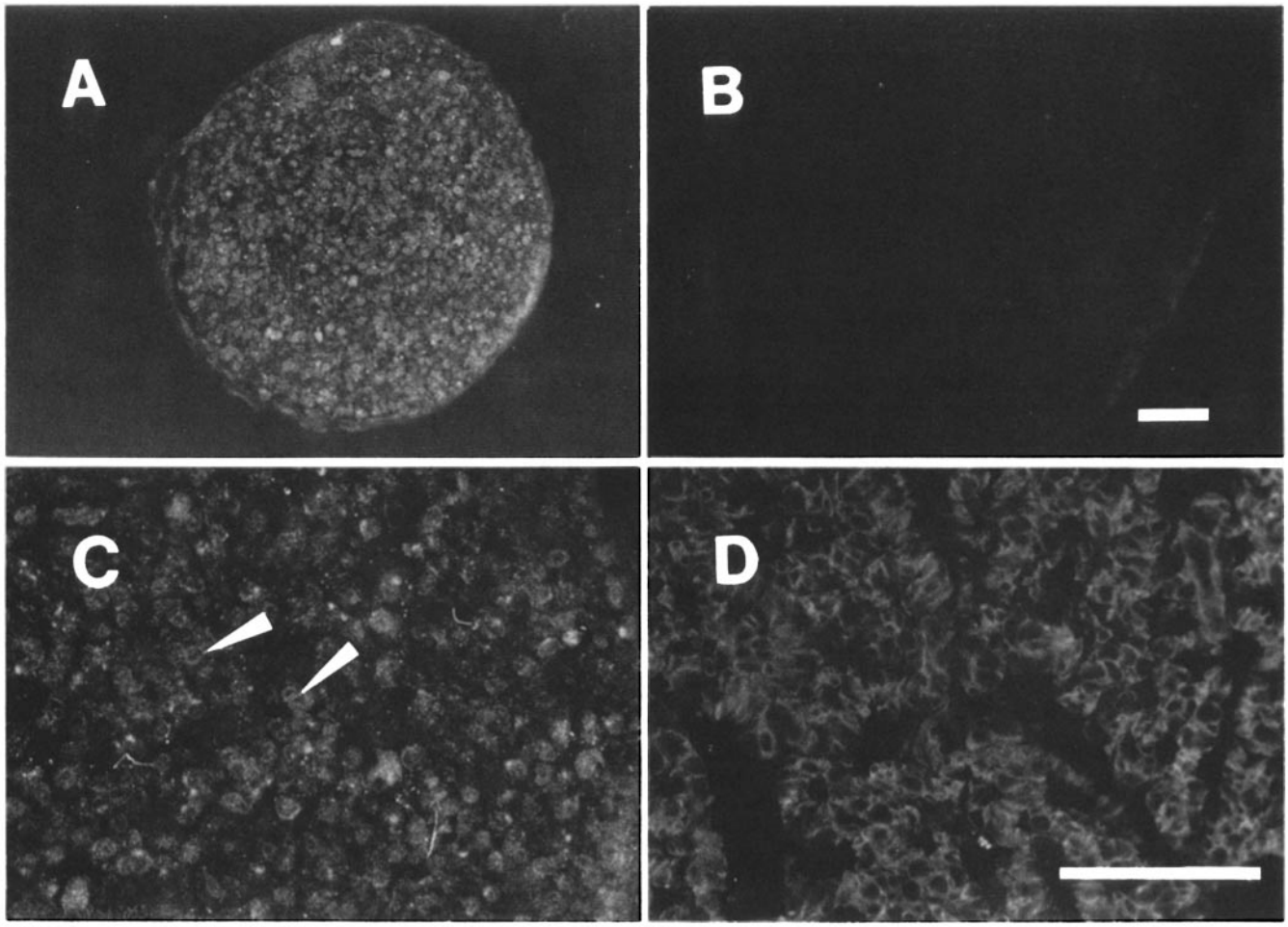

Figure 6. Immunofluorescent staining of *N*-syndecan in peripheral nerve. Cryostat sections of 2-d-old rat sciatic nerve were stained with affinity-purified anti-*N*-syndecan antibodies (*A* and *C*), pre-immune antiserum (*B*), or anti-laminin antibodies (*D*). *C* and *D* are high power views of the endoneurium; arrowheads in *C* point to small ringlets of anti-*N*-syndecan staining. Bars, 50 μm.

of both the peripheral and central nervous system we have given it the name *N*-syndecan. The cDNA was identified as one coding for a HSPG by its sequence similarity to previously identified transmembrane proteoglycan core proteins, and by the biochemical characterization of a macromolecule present in Schwann cell and brain extracts that reacted with affinity-purified antibodies made against a recombinant form of the protein expressed in bacteria. The immunoreactive macromolecule was sensitive to digestion by heparitinase and nitrous acid, but not chondroitinase ABC, and therefore appeared to possess only heparan sulfate glycosaminoglycan side chains.

The sequence of the *N*-syndecan cDNA reported here represents only a portion of the *N*-syndecan mRNA, based on the size of the mRNA as determined by Northern blot analysis. Despite repeated attempts, we have been unable to obtain clones that contain the remaining cDNA sequence. Since our sequence contains neither a poly A tail nor a polyadenylation signal, it is not clear whether the missing sequence is from the 5′ or 3′ end of the mRNA, or both. Two other independently isolated *N*-syndecan cDNA clones obtained from different oligo dT primed cDNA libraries (from Schwann cells and vascular smooth muscle cells) terminated at the same 3′ site as our original clone p104 (unpublished observations). One possible explanation for these results is that *N*-syndecan mRNA lacks a poly A tail and that our clones contain the authentic 3′ end of the mRNA. Another possible explanation for these findings, and the one we favor, is that there is an unusually stable secondary structure in the 3′-untranslated portion of the *N*-syndecan mRNA that serves as a priming site for cDNA synthesis by reverse transcriptase. In this case, our clones would lack 3′ sequence between this priming site and the poly A tail. Several observations suggest that *N*-syndecan mRNA does possess a poly A tail. The cDNA is present in libraries constructed using poly A–selected RNA, and the mRNA is enriched by oligo dT selection as assayed by Northern blot analysis (data not shown). The discrepancy between the *N*-syndecan core protein molecular weights predicted from the cDNA sequence and determined by SDS-gel electrophoresis indicates that at least part of the missing sequence is from the 5′ protein coding portion of the mRNA. Other explanations for this discrepancy in size are possible, however, including an unrecognized posttranslational modification of the core protein, an unusual secondary structure that causes anomalous migration, or formation of stable SDS-resistant aggregates. We have observed that the recombinant β-galactosidase-*N*-syndecan fusion protein forms dimers and higher order aggre-

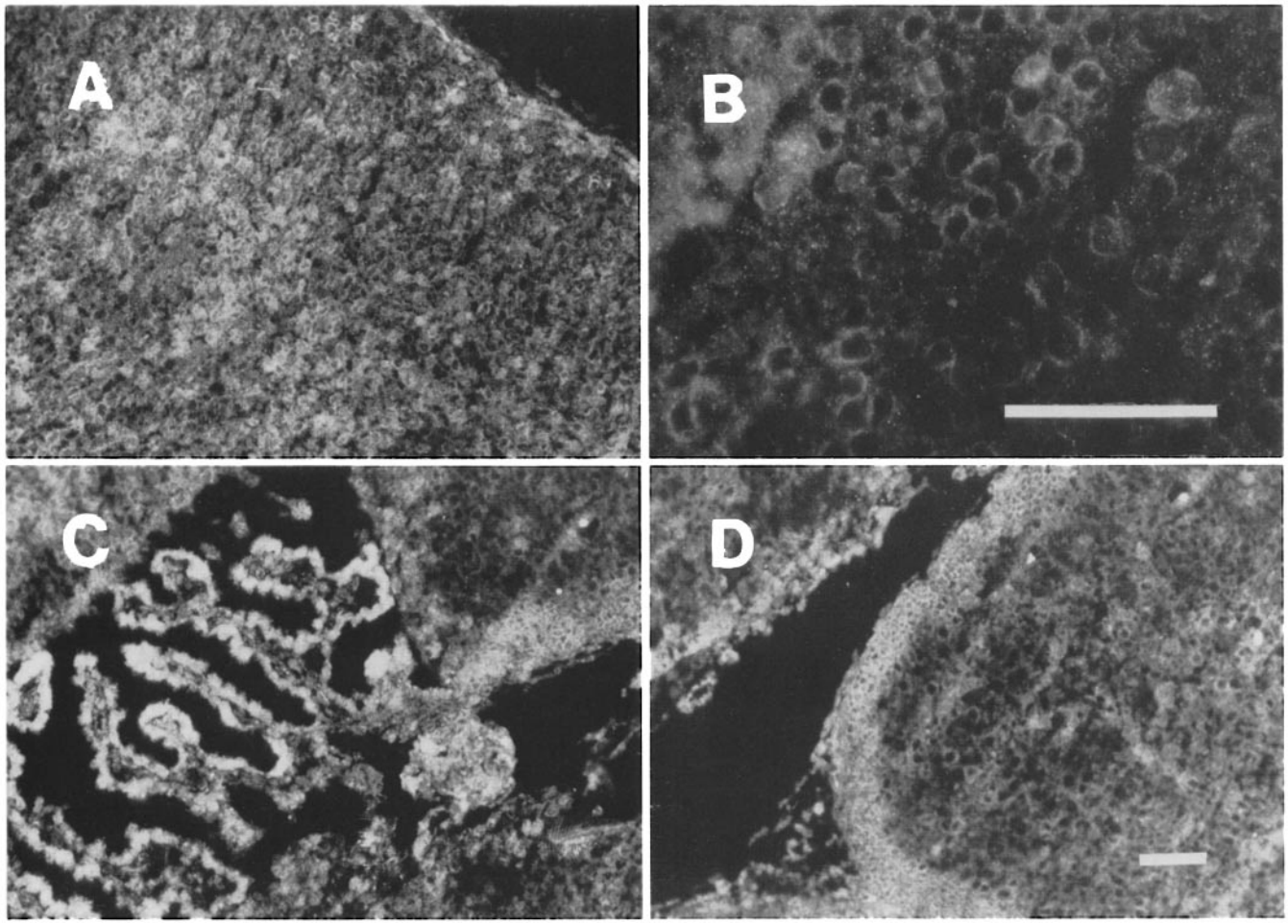

Figure 7. Immunofluorescent staining of *N*-syndecan in the central nervous system. Cryostat sections of brain tissue from a 2-d-old rat were stained with anti-*N*-syndecan antibodies. *A* and *B*, cortex; *C*, choroid plexus; *D*, mid-brain. *A*, *C*, and *D* are low power views; *B* is a high magnification view of the cortex showing staining on the surface membranes of individual cells. Bars, 50 μm.

gates that are stable in 2% SDS (not shown). Additional experiments will be required to adequately address these questions.

N-syndecan appears to represent a novel member of a gene family coding for transmembrane proteoglycan core proteins. Previously described members of this gene family are syndecan, first identified in murine epithelial cells (Saunders et al., 1989) and a proteoglycan first identified in human fibroblasts (Marynen et al., 1989). Interestingly, these core proteins show little structural homology in their extracellular domains, that would interact with extracellular molecules and other cells, but show a striking degree of sequence homology in their membrane and cytoplasmic domains. This comparison is based on the deduced amino acid sequences of the 3 gene family members from three different species (see Fig. 3); sequences for the rat homologues of the other members of the gene family have not been reported. This comparison would appear to be valid, however, based on the high degree of sequence homology of syndecans obtained from mouse (Saunders et al.), human (Mali et al., 1990), and hamster (Kiefer et al., 1990). This pattern in which the transmembrane and cytoplasmic domains exhibit a high degree of sequence homology strongly suggests these portions of the molecules carry out some important common function, perhaps related to cytoskeletal coupling or signal transduction. On the other hand, the fact that the extracellular domains are highly divergent in structure suggests these proteoglycans carry out distinct functions.

In keeping with this idea, there appear to be tissue specific patterns of expression of different members of this proteoglycan gene family. For example, we found that *N*-syndecan mRNA was present at highest levels in the developing nervous system, but was barely detectable in rat liver RNA. In contrast, it has been reported that the mRNA for syndecan, another member of the transmembrane proteoglycan gene family, is present at high levels in liver, but is undetectable in the nervous system (Saunders et al., 1989). In addition, both *N*-syndecan and syndecan appear to be present at higher levels in developing tissues than in adult tissues. The molecular mechanisms that account for regulation of expression of these molecules are not known.

The precise functions of the transmembrane proteoglycans are not known. In general, they appear to provide cell surface binding sites for extracellular proteins, such as ECM adhesive proteins, growth factors, and cell–cell adhesion molecules. These functions appear to be mediated largely by the glycosaminoglycan chains, especially heparan sulfate, of the proteoglycans, that bind to heparin-binding domains of

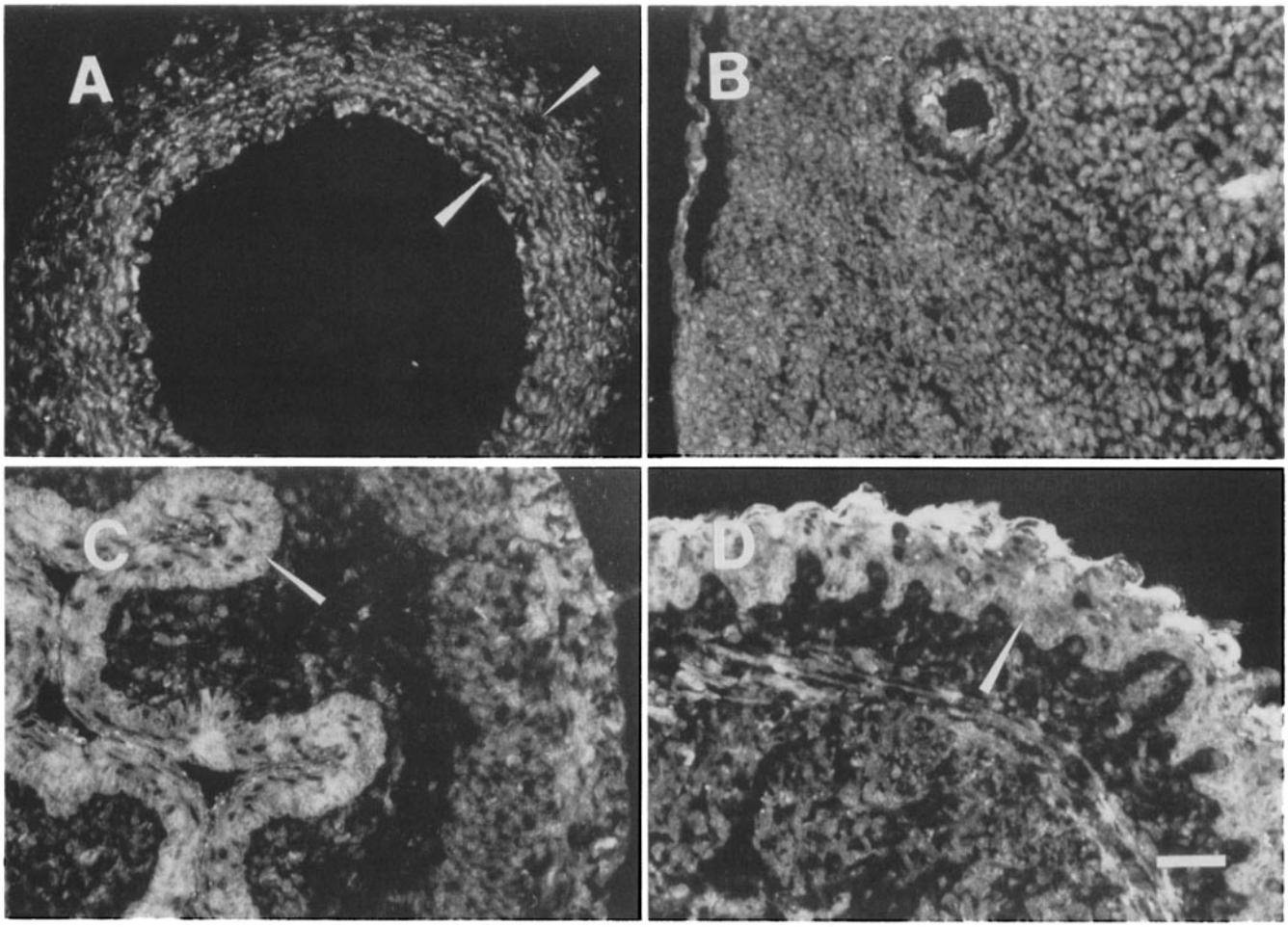

Figure 8. Immunofluorescent staining of *N*-syndecan in rat tissues. Cryostat sections of tissues from a 2-d-old rat were stained with anti-*N*-syndecan antibodies. *A*, Cross section of thoracic aorta. Staining is visible in bands of vascular smooth muscle cells in the medial layer of the vessel wall (the area bounded by the *arrowheads*), *B*, Heart; *C*, esophagus; *D*, tongue. In the latter two tissues the most intense staining is in the stratified epithelial layers (*arrowheads*).

the ligand proteins. Many ECM adhesive proteins, e.g., fibronectin (Barkalow and Schwarzbauer, 1991), laminin (Kouzi-Koliakos, 1989), and thrombospondin (Frazier, 1989), members of the heparin-binding growth factor family (e.g., bFGF) (Rifkin and Moscatelli, 1989) and cell adhesion molecules (e.g., N-CAM) (Reyes et al., 1990) possess such domains that bind heparin with moderate to high affinity in vitro and presumably bind to cell surface HSPGs. What is not clear at this point, however, is the functional significance of the structural diversity of the membrane proteoglycans, or the contributions of the core proteins to their specific functions. There are some indications of functional specificity. For example, syndecan binds fibronectin but not laminin in vitro (Saunders and Bernfield, 1988), whereas a GPI-anchored HSPG isolated from Schwann cells binds both laminin and fibronectin (Carey et al., 1990). A direct comparison of the binding activities of different members of the transmembrane proteoglycan gene family has not been carried out, however.

The localization of *N*-syndecan as determined by immunocytochemistry suggests its function is not to mediate cell–ECM interactions. In the developing neural cortex *N*-syndecan is detected on the surfaces of individual cells, and not at sites where there is significant accumulation of known ECM molecules such as laminin or fibronectin. In the sciatic nerves of newborn rats *N*-syndecan immunoreactivity is present at sites of Schwann cell–axon apposition, where there is no visible ECM (Bunge et al., 1980). In contrast, a GPI-anchored HSPG is present in sciatic nerves on the outer Schwann cell membrane that is in direct contact with the laminin rich basement membrane that surrounds the individual axon-Schwann cell units (Carey and Stahl, 1990). This pattern of localization of *N*-syndecan suggests its function may be to mediate cell–cell adhesion, although direct evidence for this is lacking. Several findings are consistent with this idea, however. N-CAM binds heparin (Reyes et al., 1990) and is present in sciatic nerves at sites similar to *N*-syndecan. Moreover, in in vitro binding assays the heparin binding domain appears to be important for N-CAM activity (Cole et al., 1985). An HSPG in developing chick brains that co-purifies with N-CAM has a core protein with a 120,000 M_r after heparitinase digestion, as determined by SDS-gel electrophoresis (Cole and Burg, 1989). This is identical to what we have observed for *N*-syndecan. The relationship of this HSPG to *N*-syndecan is not known. We have found, however, that *N*-syndecan and N-CAM distributions, deter-

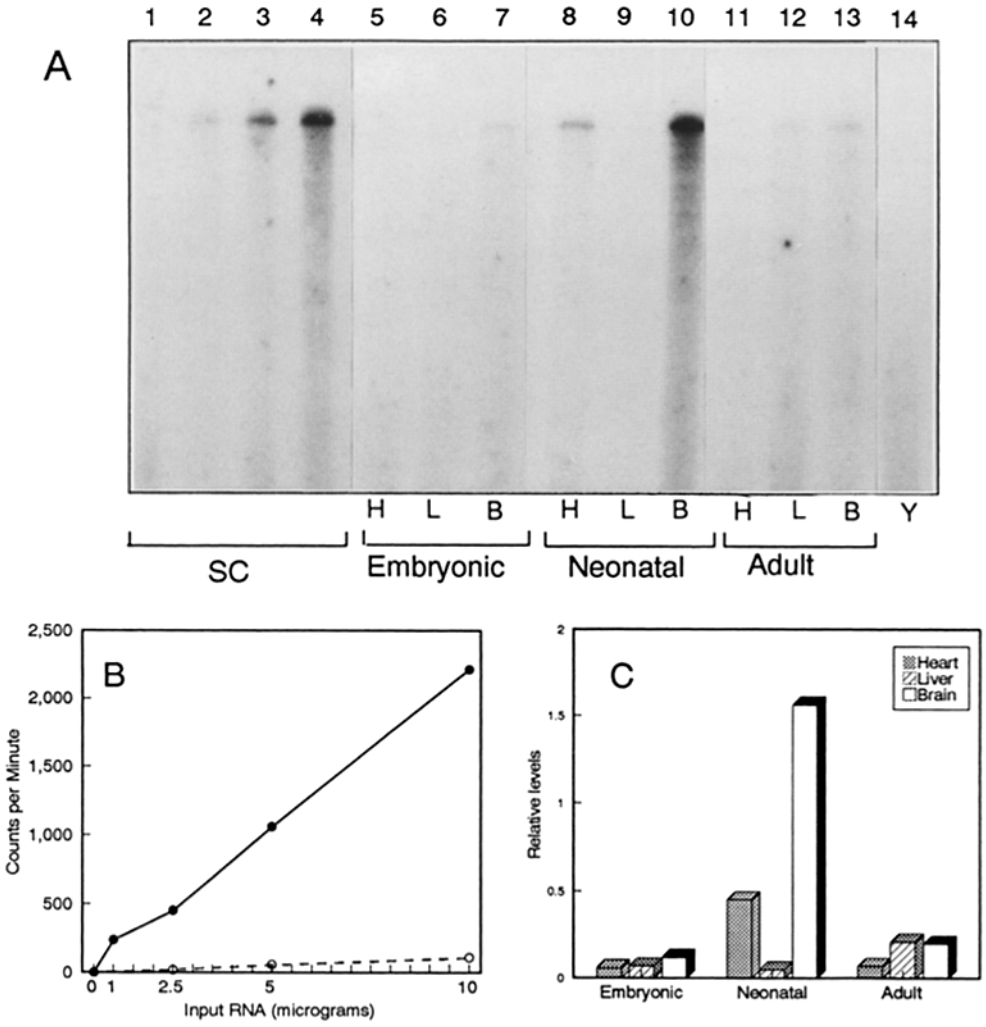

Figure 9. Ribonuclease protection assay for *N*-syndecan mRNA. Samples of total RNA from cultured cells or tissues were assayed for *N*-syndecan mRNA by an RNAse protection assay as described in Materials and Methods. (*A*) Autoradiographs showing the protected probe. (Lanes *1–4*) Schwann cell (*SC*) RNA at 1, 2.5, 5, and 10 μg per assay; (lanes *5–13*) heart (*H*), liver (*L*), and brain (*B*) RNA samples (10 μg per assay) isolated from embryonic (day 18) rats, neonatal (2-d old) rats, or adult rats; (lane *14*) 10 μg of yeast (*Y*) total RNA. (*B*) Data for Schwann cell RNA (●) and vascular smooth muscle RNA (○) are shown graphically as a function of the quantity of input RNA. (*C*) Data for the tissue samples are shown graphically and expressed relative to the signal obtained with 10 μg of Schwann cell RNA.

mined by immunocytochemistry, do not closely parallel one another, at least in the neonatal rat central nervous system (unpublished observations). Thus, the precise function of *N*-syndecan in nervous system development remains elusive.

We thank Kathy Knarr for help with the manuscript and Brian Shoop for photography.

This work was supported by National Institutes of Health grant NS21925 and a grant-in-aid from the American Heart Association, Pennsylvania Affiliate.

Received for publication 23 September 1991 and in revised form 20 December 1991.

References

Asundi, V., K. Cowan, D. Matzura, W. Wagner, and K. Dreher. 1990. Characterization of extracellular matrix proteoglycan transcripts expressed by vascular smooth muscle cells. *Eur. J. Cell Biol.* 52:98–104.

Barkalow, F. J. B., and J. E. Schwarzbauer. 1991. Localization of the major heparin-binding site in fibronectin. *J. Biol. Chem.* 266:7812–7818.

Bunge, M. B., A. K. Williams, P. M. Wood, J. Uitto, and J. J. Jeffrey. 1980. Comparison of nerve cell and nerve cell plus Schwann cell cultures, with particular emphasis on basal lamina and collagen formation. *J. Cell Biol.* 84:184–202.

Carey, D. J., and D. M. Evans. 1989. Membrane anchoring of heparan sulfate proteoglycans by phosphatidylinositol and kinetics of synthesis of peripheral and detergent-solubilized proteoglycans in Schwann cells. *J. Cell Biol.* 108:1891–1897.

Carey, D. J., and R. C. Stahl. 1990. Identification of a lipid-anchored heparan sulfate proteoglycan in Schwann cells. *J. Cell Biol.* 111:2053–2062.

Carey, D. J., D. M. Crumbling, R. C. Stahl, and D. M. Evans. 1990. Association of cell surface heparan sulfate proteoglycans of Schwann cells with extracellular matrix proteins. *J. Biol. Chem.* 265:20627–20633.

Cole, G. J., and M. Burg. 1989. Characterization of a heparan sulfate proteoglycan that copurifies with the neural cell adhesion molecule. *Exp. Cell Res.* 182:44–60.

Cole, G. J., D. Schubert, and L. Glaser. 1985. Cell-substratum adhesion in chick neural retina depends upon protein-heparan sulfate interactions. *J. Cell Biol.* 100:1192–1199.

David, G., V. Lories, B. Decock, P. Marynen, J.-J. Cassiman, and H. Van den Berghe. 1990. Molecular cloning of a phosphatidylinositol-anchored membrane heparan sulfate proteoglycan from human lung fibroblasts. *J. Cell. Biol.* 111:3165–3176.

Elenius, K., M. Salmivirta, P. Inki, M. Mali, and M. Jalkanen. 1990. Binding of human syndecan to extracellular matrix proteins. *J. Biol. Chem.* 265: 17837–17843.

Elenius, K., S. Vainio, M. Laato, M. Salmivirta, I. Thesleff, and M. Jalkanen. 1991. Induced expression of syndecan in healing wounds. *J. Cell Biol.* 114:585–595.

Fransson, L.-A. 1987. Structure and function of cell-associated proteoglycans. *Trends Biochem. Sci.* 12:406–411.

Frazier, W. A. 1987. Thrombospondin: a modular adhesive glycoprotein of platelets and nucleated cells. *J. Cell Biol.* 105:625–637.

Gibrat, J.-F., J. Garnier, and B. Robson. 1987. Further developments of protein secondary structure prediction using information theory. *J. Mol. Biol.* 198:425–433.

Hamati, H. F., E. L. Britton, and D. J. Carey. 1989. Inhibition of proteoglycan synthesis alters extracellular matrix deposition, proliferation, and cytoskeletal organization of rat aortic vascular smooth muscle cells in culture. *J. Cell Biol.* 108:2495–2505.

Hayashi, K., M. Hayashi, M. Jalkanen, J. H. Firestone, R. L. Trelstad, and M. Bernfield. 1987. Immunocytochemistry of cell surface heparan sulfate proteoglycan in mouse tissues. A light and electron microscopic study. *J.*

Histochem. Cytochem. 35:1079–1088.

Haymerle, H., J. Herz, G. M. Bressan, R. Frank, and K. K. Stanley. 1986. Efficient construction of cDNA libraries in plasmid expression vectors using adaptor strategy. *Nucleic Acids Res.* 14:8615–8624.

Kiefer, M. C., J. C. Stephans, K. Crawford, K. Okino, and P. J. Barr. 1990. Ligand-affinity cloning and structure of a cell surface proteoglycan that binds basic fibroblast growth factor. *Proc. Natl. Acad. Sci. USA.* 87:6985–6989.

Kjellen, L., and U. Lindahl. 1991. Proteoglycans: structures and interactions. *Annu. Rev. Biochem.* 60:443–475.

Kouzi-Koliakos, K., G. G. Koliakos, E. C. Tsilibary, L. T. Furcht, and A. S. Charonis. 1989. Mapping of three major heparin-binding sites on laminin and identification of a novel heparin-binding site on the B1 chain. *J. Biol. Chem.* 264:17971–17978.

Kozak, M. 1989. The scanning model for translation: an update. *J. Cell Biol.* 108:229–241.

Mali, M., P. Jaakkola, A.-M. Arvilommi, and M. Jalkanen. 1990. Sequence of human syndecan indicates a novel gene family of integral membrane proteoglycans. *J. Biol. Chem.* 265:6884–6889.

Marynen, P., J. Zhang, J.-J. Cassiman, H. Van den Berghe, and G. David. 1989. Partial primary structure of the 48- and 90-kilodalton core proteins of cell surface-associated heparan sulfate proteoglycans of lung fibroblasts. *J. Biol. Chem.* 264:7017–7024.

Milman, G. 1987. Expression plasmid containing the λP_L promoter and cI857 repressor. *Methods Enzymol.* 153:482–491.

Rapraeger, A. C., A. Krufka, and B. A. Olwin. 1991. Requirement of heparan sulfate for bFGF-mediated fibroblast growth and myoblast differentiation. *Science (Wash. DC).* 252:1705–1708.

Reyes, A. A., R. Akeson, L. Brezina, and G. J. Cole. 1990. Structural requirements for neural cell adhesion molecule-heparin interaction. *Cell Regulation.* 1:567–576.

Rifkin, D. B., and D. Moscatelli. 1989. Recent developments in the cell biology of basic fibroblast growth factor. *J. Cell Biol.* 108:1–6.

Salmivirta, M., K. Elenius, S. Vainio, U. Hofer, R. Chiquet-Ehrismann, I. Thesleff, and M. Jalkenen. 1991. Syndecan from embryonic tooth mesenchyme binds tenascin. *J. Biol. Chem.* 266:7733–7739.

Sanderson, R. D., P. Lalor, and M. Bernfield. 1989. B lymphocytes express and lose syndecan at specific stages of differentiation. *Cell Regulation.* 1:27–35.

Saunders, S., and M. Bernfield. 1988. Cell surface proteoglycan binds mouse mammary epithelial cells to fibronectin and behaves as a receptor for interstitial matrix. *J. Cell Biol.* 106:423–430.

Saunders, S., M. Jalkanen, S. O'Farrell, and M. Bernfield. 1989. Molecular cloning of syndecan, an integral membrane proteoglycan. *J. Cell Biol.* 108:1547–1556.

Shively, J. E., and H. E. Conrad. 1976. Formation of anhydrosugars in the chemical depolymerization of heparin. *Biochemistry.* 15:3932–3942.

Vainio, S., M. Jalkanen, and I. Thesleff. 1989. Syndecan and tenascin expression is mediated by epithelial-mesenchymal interactions in embryonic tooth mesenchyme. *J. Cell Biol.* 108:1945–1954.

Vlodavsky, I., R. Bar-Shavit, R. Ishai-Michaeli, P. Bashkin, and Z. Fuks. 1991. Extracellular sequestration and release of fibroblast growth factor: a regulatory mechanism? *Trends Biochem. Sci.* 16:268–271.

Yayon, A., M. Klagsbrun, J. D. Esko, P. Leder, and D. M. Ornitz. 1991. Cell surface, heparin-like molecular are required for binding of basic fibroblast growth factor to its high affinity receptor. *Cell.* 64:841–848.