

# Molecular Cloning of Amphiglycan, a Novel Integral Membrane Heparan Sulfate Proteoglycan Expressed by Epithelial and Fibroblastic Cells

Guido David, Bernadette van der Schueren, Peter Marynen, Jean-Jacques Cassiman, and Herman van den Berghe

Center for Human Genetics, University of Leuven, Leuven, Belgium

**Abstract.** We have synthesized an antisense oligonucleotide primer that matches a supposedly conserved sequence in messages for heparan sulfate proteoglycans with transmembrane orientations. With the aid of this primer we have amplified partial and selected full-length copies of a message from human lung fibroblasts that codes for a novel integral membrane heparan sulfate proteoglycan. The encoded protein is 198 amino-acids long, with discrete cytoplasmic, transmembrane, and amino-terminal extracellular domains. Except for the sequences that represent putative heparan sulfate chain attachment sites, the extracellular domain of this protein has a unique structure. The transmembrane and cytoplasmic domains, in contrast, are highly similar to the corresponding domains of fibroglycan and syndecan, the two cell surface proteoglycans that figured as models for the design of the antisense primer. This similarity includes the conservation of four tyrosine

residues, one immediately in front of the stop transfer sequence and three in the cytoplasmic segment, and of the most proximal and most distal cytoplasmic sequences. The cDNA detects a single 2.6-kb message in cultured human lung fibroblasts and in a variety of human epithelial and fibroblastic cell lines. Polyclonal and monoclonal antibodies raised against the encoded peptide after expression as a beta-galactosidase fusion protein react with the 35-kD coreprotein of a cell surface heparan sulfate proteoglycan of human lung fibroblasts and decorate the surface of many cell types. We propose to name this proteoglycan "amphiglycan" (from the Greek words *amphi*, "around, on both sides of" and *amphoo*, "both") referring to its domain structure which extends on both sides of the plasmamembrane, and to its localization around cells of both epithelial and fibroblastic origin.

**H**EPARAN sulfate proteoglycans prove to play an important role in the regulation of the activities and bioavailabilities of a growing number of enzymes, enzyme inhibitors, adhesion molecules, and growth factors (Jackson et al., 1991). Heparan sulfate proteoglycans that form part of the extracellular matrix function as binding sites for chymase (Sayama et al., 1987), antithrombin III (Pejler et al., 1987; de Agostini et al., 1990), protease nexin (Farrell et al., 1988), and FGFs (Gonzalez et al., 1990) which may provide the cells with matrix-bound reservoirs of these components. For basic FGF the interaction with the heparan sulfate of the matrix seems to metabolically stabilize the cytokine (Saksela et al., 1988), and to limit its diffusion and availability (Flaumenhaft et al., 1990), but active heparan sulfate growth factor complexes can be released from the matrix through limited proteolysis (Saksela and Rifkin, 1990) or partial degradation of the heparan sulfate residues (Ishai-Michaeli et al., 1990). Moreover, the glycosaminoglycan in the complex seems required for the interaction of the growth factor with its receptor at the cell surface (Yáyon et al., 1991; Rapraeger et al., 1991). Adherent cells, however, also express a number of proteins at their cell surfaces that are substituted with heparan sulfate and that seem to en-

gage themselves in similar interactions as reported for the matrix proteoglycans, acting as "receptors" for the same proteinase inhibitors (Marcum et al., 1986), matrix components (Koda et al., 1985; Saunders and Bernfield, 1988; Elenius et al., 1990), and FGFs (Kiefer et al., 1990; Brunner et al., 1991). This suggests that the various heparan sulfate ligands may also occur in a cell surface-bound pool, and that their interaction with the glycosaminoglycans of the matrix and cell surface may result in a tight spatial control of their activities.

Recently, the characterization of a few of these cell surface proteins has demonstrated the heterogeneity of the cell surface proteoglycans. Fibroglycan (Marynen et al., 1989), syndecan (Saunders et al., 1989; Mali et al., 1990), and glypican (David et al., 1990) are structurally distinct proteins, with different patterns of expression, membrane associations, and metabolic fates. Glypican is widely expressed in many different cell types, where it is linked to the cell surface through phosphatidylinositol. Fibroglycan and syndecan, in contrast, are transmembrane proteins with more restricted distributions (Lories et al., 1992). Moreover, in fibroblasts, where glypican is constitutively shed and released to the pericellular space, fibroglycan and syndecan

appear to be rapidly cleared from the cell surface through endocytosis, suggesting that each of these proteoglycans may be committed to specific processes (David et al., 1990).

The structural analysis of these proteins has also revealed some shared features that may support functions or pathways that are common to the different heparan sulfate proteoglycans. All three share a sequence of the type SGXG preceded by a few acidic residues, first identified in serglycin as a site for substitution with glycosaminoglycan (Bourdon et al., 1987). Fibroglycan and syndecan in addition show a remarkable sequence similarity in their transmembrane and cytoplasmic parts. These short domains are 52 and 66% identical, respectively, and together contain four tyrosines that have been conserved in the two proteins. The function of these residues is unknown, but they are reminiscent of sequences that have been implicated in the targeting of integral membrane proteins to specific membrane domains and endocytotic compartments (Ktistakis et al., 1990). Controlled expression of the cell surface proteoglycans in specific membrane domains could be a further way of regulating the activities of their associated ligands.

The structural heterogeneity of the cell surface heparan sulfate proteoglycans is, however, likely to be more extensive than what is illustrated by the three proteins that have been identified so far. In addition to glypican, fibroglycan and syndecan cultured human lung fibroblasts express at least two other heparan sulfate proteoglycans with membrane associations (Lories et al., 1989). Their core proteins, which have apparent molecular masses of 35 and 125 kD, resist a treatment with phospholipase C (David et al., 1990), which suggests that they may belong to the category of proteoglycans with transmembrane orientations. We have speculated that one or both of these proteoglycans would also display the features that are shared by fibroglycan and syndecan and designed oligonucleotide primers that might allow the amplification of their messages. This approach has proven partially successful, leading to the isolation of cDNA clones coding for the 35-kD core protein. This novel cell surface heparan sulfate proteoglycan belongs to the same molecular family as fibroglycan and syndecan, but appears to be more widely expressed than these two proteins.

## Materials and Methods

### Isolation of the 35K-cDNA Clones

The antisense oligonucleotide 5'ATAGCTTCCTTCATCCTTCTTCTCAT, the complement of the 48K cDNA sequence which codes for MRKKDEGSY (a peptide sequence in the cytoplasmic domain of human fibroglycan that closely resembles the sequence MKKKDEGSY in the cytoplasmic domain of syndecan), and an oligonucleotide terminating with the sequence 5'GAGTCTGAGAACTTT, representing the reverse translation of a potential core protein sequence (see Results), were used as primers in a polymerase chain reaction. The source of cDNA and reaction conditions were as reported before (David et al., 1990), except for the annealing temperature which was set at 45°C. The product of this amplification was blunted, kinased, and ligated to a dephosphorylated SmaI restricted pGEM3Z vector. After cloning and sequencing, this polymerase chain reaction product was <sup>32</sup>P-labeled by random oligonucleotide priming and used to screen a human lung fibroblast cDNA library composed of 6 × 10<sup>5</sup> independent reverse transcripts cloned in the lambda ZAP II vector (David et al., 1990). Five positive clones were plaque purified and converted to the corresponding Bluescript SK<sup>-</sup> plasmids by superinfection with R408 helper phage.

### Analysis of the cDNA Clones

The resulting phagemids were compared by restriction mapping and by se-

quence analysis of both ends of the inserts. Clone 35K17 was selected, digested with suitable restriction enzymes, and subcloned in Bluescript SK<sup>-</sup> or pGEM3Z plasmids. The ends of the inserts of the 35K17S clones were sequenced by the dideoxy chain termination method, using supercoiled plasmids, modified T7 DNA polymerase, the appropriate primers, and both dGTP and 7-deaza dGTP (Pharmacia LKB Biotechnology, Uppsala, Sweden), covering the whole length of both strands of the parental clone 35K17. A search of the Genbank and EMBL databases and the sequence alignment were done with the Intelligenetics program release 5.4.

### Northern Blot Analysis

Total RNA was isolated from cultured cells by the acid guanidinium thiocyanate-phenol-chloroform extraction method (Chomczynski and Sacchi, 1987), fractionated by electrophoresis in denaturing agarose gels and blotted on Nylon, essentially as described before (Marynen et al., 1989). The blots were hybridized to the <sup>32</sup>P-oligolabeled insert from the 35K17 clone and washed at high stringencies as reported before (David et al., 1990).

### Construction and Analysis of the Expression Plasmid

A cDNA fragment spanning from the SmaI site in the 35K17 insert to the SalI site in the multiple cloning site of the Bluescript vector was ligated in a SmaI and SalI restricted pEX3 vector (Genofit, Geneva, Switzerland) and used to transfect POP 2136 *E. Coli* cells. The 35K-pEX3 transfectants and control cells transfected with pEX3 vector without insert were selected at 28°C on ampicillin plates. Exponentially growing cultures were initiated from single colonies, and induced by incubation at 42°C for 2 h. The cell pellets were sonicated, digested with lysozyme, extracted with 6 M urea and finally solubilized in hot SDS for analysis by electrophoresis and Western blotting using a monoclonal anti-beta galactosidase antibody (Promega Corporation, Madison, WI).

### Preparation of Poly- and Monoclonal Anti-core Protein Antibodies

Temperature-induced transfectant *E. Coli* cells were extracted with lysozyme, 6 M urea, and 4 M guanidinium chloride as described before (Marynen et al., 1989). The insoluble residue, consisting mostly of recombinant beta-galactosidase-35K fusion protein, was sonicated in saline, mixed with Freund's adjuvant, and injected subcutaneously into rabbits. Blood samples were collected by ear bleeding. Pooled sera from immune animals were affinity purified on fusion protein that had been coupled to CNBr-activated Sepharose, and cleared from contaminant specificities by incubation with CNBr-activated Sepharose substituted with extract from pEX3-transfected *E. Coli* cells, essentially as described before (Marynen et al., 1989).

Alternatively, the insoluble residue was rinsed with water, solubilized in hot SDS, fractionated by SDS-PAGE, and blotted onto pure nitrocellulose membranes. The blot was stained with Coomassie brilliant blue and the segment that contained the beta-galactosidase-35K fusion protein was excised from the membrane. After drying, the nitrocellulose support was solubilized in acetone and the insoluble protein was pelleted by centrifugation. After several rinses with acetone, the protein was sonicated in a (1:1) mixture of PBS and Freund's adjuvant and used to immunize Balb/c mice by repeated intraperitoneal injection. Blood samples were collected by eye bleeding, and the spleens from immune animals were collected for the establishment of hybridomas as described before (Deboeck et al., 1987).

### Isolation and Analysis of the Cell Surface Proteoglycans

Cell surface proteoglycans were isolated from confluent cultured human fetal lung fibroblasts and analyzed, essentially as described before (Lories et al., 1989). Briefly, the cells were extracted in ice-cold 0.5% Triton X-100 buffer in the presence of proteinase inhibitors. The hydrophobic proteoglycans were purified from this extract by ion-exchange chromatography on Mono Q, incorporation into liposomes, and gel filtration over Sepharose CL4B in 4 M guanidinium chloride buffer without detergent. The cell surface proteoglycan samples were concentrated, digested with heparitinase, fractionated by electrophoresis in SDS-polyacrylamide gradient gels, and blotted on Z-probe membranes. The blots were incubated with the rabbit or murine antibodies and developed with the appropriate HRP- or alkaline phosphatase-conjugated second antibodies, by staining, using 3,3' DAB, or by chemiluminescence and autoradiography, using AMPPD (Tropix, Bedford, Mass). The procedures for the metabolic labeling or radioiodination

of the proteoglycans, for the coupling of mAb 8G3 to CNBr-activated Sepharose, and for the immunopurification of the 8G3-reactive proteoglycan on the immobilized antibody were as described before (Lories et al., 1989).

### Immunohistochemistry

Cultured human lung fibroblasts and cryosections of human skin tissue were prepared and treated as reported before (David et al., 1989), and incubated for 1 h with the conditioned culture medium of hybridoma F94-8G3 or with the antibodies F58-2E9 (anti-syndecan), F58-10H4 (anti-fibroglycan), and S1 (anti-glypican) described before (De Boeck et al., 1987; Lories et al., 1989; David et al., 1990). After rinsing, the sections were further incubated for 1 h with an HRP conjugated rabbit anti-mouse Ig (Dakopatts, Glostrup, Denmark) and stained using 3,3' DAB and hydrogen peroxide.

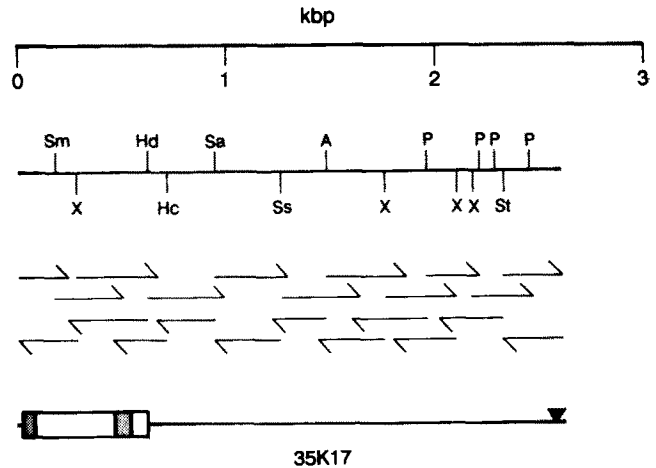
## Results

### Cloning of a Novel Cell Surface Proteoglycan

Fibroblasts express multiple heparan sulfate proteoglycans at their cell surfaces (Lories et al., 1989). At least two of these proteins, fibroglycan (Marynen et al., 1989) and syndecan (Lories et al., 1992), show significant sequence similarities in their membrane-spanning and cytoplasmic domains. Their ectodomains, in contrast, are dissimilar, except for the sequences that have been proposed as sites for substitution with glycosaminoglycan chains in these and other proteins (Bourdon et al., 1987; Mann et al., 1990). Speculating that conserved cytoplasmic structures may form the hallmark of a specific subset of the cell surface proteoglycans, we designed a "universal" antisense oligonucleotide primer that, in conjunction with appropriate sense primers, might allow the specific amplification of novel fibroglycan- and syndecan-related messages from the appropriate libraries. We used it successfully in combination with a primer that included the sequence GAGTCTGAGA ACTTT coding for the peptide sequence ESENF which mimics a cluster of amino acids that forms part of a glycanation site of the "serglycin type" (Bourdon et al., 1987) in syndecan.

The polymerase chain reaction product which was obtained with this primer combination using a fibroblast library as a source of cDNA, was cloned and sequenced and potentially represented a cDNA for a novel cell surface proteoglycan. This ~350-bp cDNA was used to screen the library from which it had been amplified to isolate corresponding full-length cDNA clones. Based on the restriction patterns and terminal sequences of the isolated clones, clone 35K17 was selected for further subcloning (Fig. 1) and complete sequencing (Fig. 2).

The insert of clone 35K17 was 2613 bp long. The first ATG in its sequence occurred at position 27. Surrounded by guanines at positions -3 and +4, this ATG occurred in a context favorable for the initiation of translation (Kozak, 1989) and formed the start of an open reading frame of 594 bases coding for a peptide sequence of 198 amino acids (Fig. 2). The coding sequence was terminated by a TGA stop codon and followed by 1,990 bases of 3' untranslated sequence terminating with an AATAAA polyadenylation signal and, 15 bases further, a poly A stretch. Northern blot analysis of a total RNA extract from human lung fibroblasts using the 35K17 insert as a probe revealed a single ~2.6-kb message, indicating that the 35K17 sequence represented a near complete transcript (Fig. 3). Comparison of the 35K17 sequence with the sequence of the amplified cDNA indicated that the amplification had resulted from the "legitimate" hy-



**Figure 1.** Restriction map and strategy used for the sequencing of clone 35K17. Only the restriction sites for enzymes used in the subcloning of clone 35K17 are indicated: *Sm*, SmaI; *X*, XbaI; *Hd*, HindIII; *Hc*, HincII; *Sa*, SacI; *Ss*, SstI; *A*, AsnI; *P*, PvuII; *St*, StuI. The arrows indicate the direction and length of the nucleotide sequences obtained from the analysis of the various 35K17s clones. The box represents the coding part of clone 35K17 and its shaded areas illustrate the sequences coding for hydrophobic peptide segments. The position of the single polyadenylation signal is indicated by the triangle.

bridization of the antisense primer to the target DNA sequence.

The peptide encoded by the open reading frame of the 35K17 insert had the features of a type I integral membrane protein, starting with an amino-terminal hydrophobic segment with the characteristics of a signal peptide (18 amino acids), followed by the presumptive extracellular domain of the mature protein (127 amino acids), a membrane-spanning segment (25 amino acids), and a short carboxy-terminal cytoplasmic domain (28 amino acids). The peptide had a predicted molecular weight of 21,607 in the precursor form, and of 19,766 after removal of the signal sequence. The presumptive ectodomain contained 4 SG sequences, in greater or lesser proximity of acidic residues, that could function as sites for the implantation of glycosaminoglycan chains (Bourdon et al., 1987; Mann et al., 1990). Two of these occurred as a repeat sequence in a context reminiscent of the structures of proposed glycanation sites in serglycin and in previously identified heparan sulfate proteoglycans (Fig. 4). The peptide lacked recognition sequences for substitution with N-linked oligosaccharides, but had one SST sequence that could serve for O-glycosylation. The ectodomain also contained two dibasic sequences that could form a substrate for trypsin-like enzymes. Most interestingly, at the presumptive start of the ectodomain the predicted sequence included a stretch of 25 amino acids (underlined in Fig. 2), which completely matched the sequence of the amino terminus of a heparan sulfate proteoglycan from rabbit aorta endothelial cells previously reported by others (Castillo et al., 1987). Moreover, the presumptive trans-membrane and cytoplasmic domains of the protein showed extensive sequence similarity to the corresponding domains of fibroglycan (73% identity) and syndecan (53% identity). This similarity included the conservation of four tyrosine residues (indicated by asterisks in Fig. 5). From these data we concluded that clone 35K17 indeed coded for a novel

35K17 SEQUENCE

```

GC GCG CCT TCT CCA GTC CGC GGT GGC ATG GCC CCC GCC CGT CTG TTC GCG CTG CTG CTG CTC TTC GTA GGC GGA GTC GCC GAG TCG ATC C 90
M A P A R L F A L L L L F V G G V A E S I R 22
GA GAG ACT GAG GTC ATC GAC CCC CAG GAC CTC CTA GAA GGC CGA TAC TTC TCC GGA GCC CTA CCA GAC GAT GAT GAT GTA GTG GGG CCC G 180
E T E V I D P Q D L L E G R Y F S G A L P D D E D V V G P G 52
GG CAG GAA TCT GAT GAC TTT GAG CTG TCT GGC TCT GGA GAT CTG GAT GAC TTG GAA DAC TCC ATG ATC GGC CCT GAA GTT GTC CAT CCC T 270
Q E S D D F E L S S G D L G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G G 82
TG GTG CCT CTA GAT AAC CAT ATC CCT GAG AGG GCA GGG TCT GGG AGC CAA GTC CCC ACC GAA CCC AAG AAA CTA GAG GAG AAT GAG GTT A 360
V P L D N H I P E R A G S G S Q V P T E P K K L E E N E V I 112
TC CCC AAG AGA ATC TCA CCC GTT GAA GAG AGT GAG GAT GTG TCC AAC AAG GTG TCA ATG TCC AGC ACT GTG CAG GGC AGC AAC ATC TTT G 450
P K R I S P V E E G S G D V S N K V S M S S S S S S S S S S S S S S S S S S S S S S S S S S S S S S S S S S S S S S S S S S 142
AG AGA ACG GAG GTC CTG GCA GCT CTG ATT GTG GGT GGC ATC GTG GGC ATC CTC TTT GCC GTC TTC CTG ATC CTA CTG CTC ATG TAC CGT A 540
R T E V L A A L I V G G I V G I L F A V F L I L L L M Y R M 172
TG AAG AAG AAG GAT GAA GGC AGC TAT GAC CTG GGC AAG AAA CCC ATC TAC AAG AAA GCC CCC ACC AAT GAG TTC TAC GGG TGA AGC TTG C 630
K K K D E G S Y D L L G K K P I Y K K A C P T N E F Y A * 198
TT GTG GGC ACT GGC TTG GAC TTT AGC GGG GAG GGA AGC CAG GGG ATT TTG AAG GGT GGA CAT TAG GGT AGG GTG AGG TCA ACC TAA TAC T 720
GA CTT GTC AGT ATC TCC AGC TCT GAT TAC CTT TGA AGT GTT CAG AAG AGA CAT TGT CTT CTA CTG TTC TGC CAG GTT CTT CTT GAG CTT T 810
GG GCC TCA GTT GCC CTG GCA GAA AAA TGG ATT CAA CTT GGC CTT TCT GAA GGC AAG ACT GGG ATT GGA TCA CTT CTT AAA CTT CCA GTT A 900
AG AAT CTA GGT CCG CCC TCA AGC CCA TAC TGA CCA TGC CTC ATC CAG AGC TCC TCT GAA GCC AGG GGG CTA AGC GAT GTT GTG TGG AGT C 990
CT GGC TGG AGG TCC TCC CCC AGT GGC CTT CCT CCC TTC CTT TCA CAG CCG TCT CTG CCA GGA AAT GGG GGA AGG AAC TAG AAC CAC C 1080
TG CAC CTT GAG ATG TTT CTG TAA ATG GGT ACT TGT GAT CAC ACT ACG GGA ATC TCT GTG GTA TAT ACC TGG GGC CAT TCT AGG CTC TTT C 1170
AA GTG ACT TTT GGA AAT CAA CCT TTT TTA TTT GGG GGG GAG GAT GGG GAA AAG AGC TGA GAT TTT ATG CTG AAA TGG ATT TAT AGA ATA T 1260
TT GTA AAT CTA TTT TTA GTG TTT GTT CGT TTT TTT AAC TGT TCA TTC CTT TGT GCA GAG TGT ATA TCT CTG CCT GGG CAA GAG TGT GGA G 1350
GT GCC GAG GTG TCT TCA TTC TCT CGC ACA TTT CCA CAG CAC CTG CTA AGT TTG TAT TTA ATG GTT TTT GTT TTT GTT TTT TTT TTT T 1440
TG AAA ATG AGA GAA GAG CCG GAG AGA TGA TTT TTA TTA ATT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT TTT 1530
TT CCC TTC TTT TTT CTT TGT TCT TAT AAT ATA TTA TTA TTC TTC CTT GGT TCT AGA AAA ATA GAT AAA TAT TTT TTT AGG AAA TAG TGT G 1620
GA TAT TTG TGT CTG TGT TCT TAT AAT ATA TTA TTA TTC TTC CTT GGT TCT AGA AAA ATA GAT AAA TAT TTT TTT AGG AAA TAG TGT G 1710
GT GTT TCC AGT TTG ATG TTG CTG GGT GGT TGA GTG AGT GAA TTT TCA TGT GGC TGG GTG GGT TTT TGC CTT TTT CTC TTG CCC TGT TCC T 1800
GG TGC CTT CTG ATG GGG CTG GAA TAG TTG AGG TGG ATG GTT CTA OCC TTT CTG CCT TCT GTT TGG GAC CCA GCT GGT GTT CTT TGG TTT G 1890
CT TTT TTC AGG CTC TAG GGC TGT GCT ATC CAA TAC AGT AAC CAC ATG CCG CTG TTT AAA GTT AAG CCA ATT AAA ATC ACA TAA GAT TAA A 1980
AA TTC CTT CCT CAG TTG CAC TAA CCA CGT TTT TAG AGG TGG ATG GTT CTA OCC TTT CTG CCT TCT GTT TGG GAC CCA GCT GGT GTT CTT TGG TTT G 2070
CT GTT GGA CAC CAC TAT TCT AGA GAA CTA AAC TGG CTT AAC GAG TCA CAG CCT CAG CTG TGG GAC CAC CCT TCT CCT GAG AGG GGA GGC CTG GTG GCT T 2160
GG GGG GGG AAT GGG GGA GGG CTG ATG AGG CCC CAG CTG GGG CCT GTT GTC TGG GAC CCT CCC TCT CCT GAG AGG GGA GGC CTG GTG GCT T 2250
AG CCT GCG CAG GTC GTG TCT CCT CCT CAG CCC AGT GGC TGC GGT GTC GGG AAC CAC CCT CCC TTG CTG CAG CAG TGG CCA TTA GCT CCC G 2340
TC ACC ACT GCA ACC CAG GGT CCC AGC TGG CTG GGT CTT CTG CCC CCA GTG CCC TTC CCC TTG GGC TGT GTT GGA GTG AGC ACC TCC T 2430
CT GTA GGC ACC TCT CAC ACT GTT GTC TGT TAC TGA TTT TTT TTG ATA AAA AGA TAA TAA AAC CTG GTA CTT TCT AAA AAA AAA AAA AAA A 2520
AA A 2610
    
```

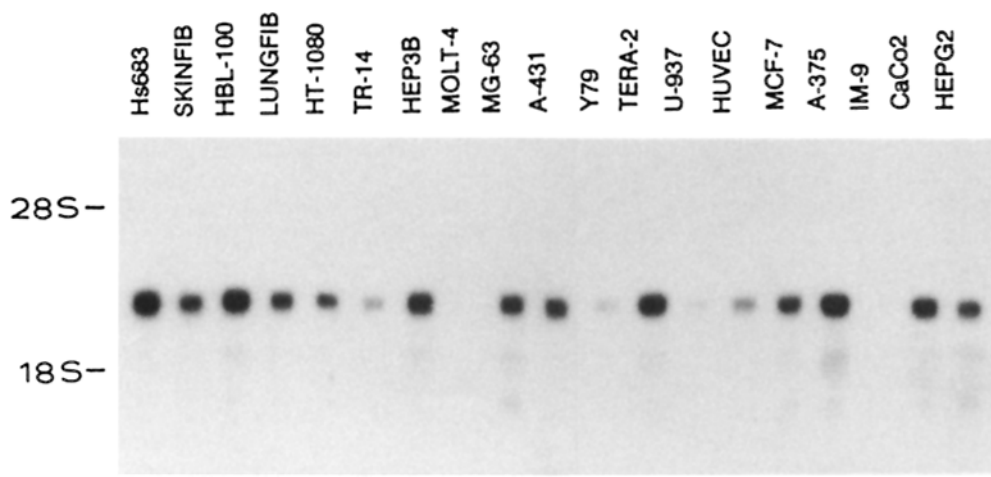
**Figure 2.** Sequence of clone 35K17 and predicted primary structure of amphiglycan. The part of the predicted protein sequence that matches the amino terminus of the endothelial proteoglycan has been indicated by thin underlining, the gap corresponding to a blank cycle obtained during the sequencing of the protein (Castillo et al., 1987). The double line marks the presumed transmembrane domain. The position of the four SG sequences is indicated by the triangles. The cluster of hydroxylated amino acids that may be O-glycosylated is indicated with a broken line. Residues in the ectodomain that form part of dibasic sequences have been encircled. The AAT-AAA polyadenylation signal is marked by the shaded box. These sequence data are available from EMBL/GenBank/DBJ under accession number X67016.

member of a distinctive family of cell surface heparan sulfate proteoglycans.

**Characterization of the Corresponding Fibroblast Proteoglycan**

To identify the cellular product which corresponded to the translation of the 35K sequence, the 35K17 insert was subcloned in the pEX3 prokaryotic expression vector and expressed as a beta-galactosidase fusion protein in *E. coli* host cells (Fig. 6). This fusion protein was used to immunize rab-

bits and mice, and to raise poly and monoclonal antisera directed against the presumptive core protein moiety of the recombinant peptide (Fig. 6). The polyclonal antiserum and one of the selected mAbs, designated F94-8G3, was used to analyze proteoglycan extracts from human lung fibroblasts in Western blotting experiments (Fig. 7). Both types of antibodies reacted with the native and heparitinase-digested proteoglycan fractions. The undigested immunoreactive proteoglycan migrated poorly and transferred poorly from the gels, but after a treatment with heparitinase (Fig. 7) or with both heparitinase and chondroitinase ABC (not shown) the



**Figure 3.** Northern blot analysis of the amphiglycan mRNA. 10-μg aliquots of total RNA from human HS 683 glioma (lane 1), skin fibroblasts (lane 2), HBL-100 breast (lane 3), lung fibroblast (lane 4), HT-1080 sarcoma (lane 5), TR-14 neuroblastoma (lane 6), HEP 3B hepatoma (lane 7), MOLT-4 acute lymphoblastic leukemia (lane 8), MG-63 osteosarcoma (lane 9), A-431 epidermoid carcinoma (lane 10), Y79 retinoblastoma (lane 11), Tera-2 embryonal carcinoma (lane 12), U-937 histiocytic lymphoma (lane 13), umbilical vein endothelial (lane 14),

MCF-7 breast adenocarcinoma (lane 15), A-375 malignant melanoma (lane 16), IM-9 myeloma (lane 17), CaCo-2 colon adenocarcinoma (lane 18), and HEP G2 hepatocarcinoma (lane 19) cell cultures were fractionated in agarose gels, blotted, and hybridized to the 35K17 probe.

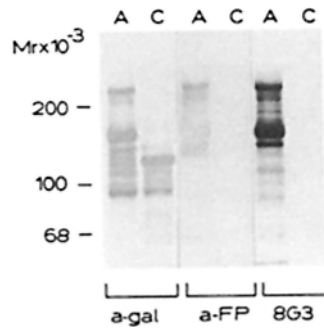
SERGLYCIN	L S D Y S G	F S	100
GLYPICAN	A S D G S G	S D	492
SYNDECAN	D S N F S G	S A A	51
FIBROGLYCIN	D D Y A S A	S S A	59
AMPHIGLYCAN	E S D F E L	S S D	65
	L L G R Y F	A L P	43
	I P R A G	S Q V	99
GLYPICAN	T D K F W G T	V E S	312
	P R R P P	T L E	376
SYNDECAN	D V P Q A I	H L	59
	P P D Q G	D S	41
	Q L P A A G	Q D	210
	Q D F T F T	N T	220
FIBROGLYCIN	N S S I E A	Y P	45
BETAGLYCAN	A P S P G S	W P D	543
	G Y D L V	N G	554
CD44E	P T D D V S	S S S	184
	S S R S S T	G Y I	194
	C N V N R S L	Q D	358
	D Q T F H P	G S H	367
	E G G A N T T	P I R	394

**Figure 4.** Comparison of putative glycanation sites in heparan sulfate proteoglycans. The putative glycanation sites in amphiglycan are compared to similar sequences in human serglycin (Stevens et al., 1988), syndecan (Mali et al., 1990), fibroglycan (Marynen et al., 1989), and glypican (David et al., 1990), which are known to be substituted with heparin or heparan sulfate chains and, in syndecan, also with chondroitin sulfate. They are also compared to similar sequences in rat betaglycan (Lopez-Casillas et al., 1991; Wang et al., 1991) and human CD44-E (Stamenkovic et al., 1991; Brown et al., 1991), which are primarily substituted with chondroitin sulfate but also some heparan sulfate chains. All are variations on a sequence motif which includes the clustering of acidic residues and SG. Two types of sequences, Acidic-XXXXSGXG and Acidic-XSG-Acidic, and hybrid forms of these sequences seem to occur. Seemingly conserved residues that appear characteristic for these glycanation "cassettes" have been boxed. Small gaps have been allowed to maximize the sequence similarity. All SG sequences in the proteins are shown, except for serglycin (where only the start of the repeat is shown) and betaglycan (only the SG sequences in proximity of acidic residues). Protein sequencing data suggest that in human lung fibroblasts the serine at position 308 in the glypican sequence is not used for glycanation (David et al., 1990).

antisera stained a discrete protein band which migrated with an apparent molecular mass of 35 kD. In contrast, staining the blots with mAb F69-3G10, an antibody that reacts with the unsaturated glucuronates generated by the cleavage of the heparan sulfate chains with heparitinase, yielded the characteristic series of bands in this preparation, including the core identified by the anti-35K antibodies, the cores of fibrogly-

Syn	245	GLLDRKEVLGGVIAAGLVGLIFAVCLVGFMLYRHKKDEGSYSLEEPKQANGGAYOK	PTKQEEFYA
F1b	138	SLFKRTEVLAAVIAGGVIGFLFAIFLILLVYRMRKKDEGSYDLGERKPSSA	AYQKAPTKEFYA
35K	139	NIFERTEVLAALIVGGIVGILFAVFLILLVYRMRKKDEGSYDLG	KKPI YKKAPTNEFYA
consensus		-1f-RtEVLaaVlaGG-vG-1FAVfL1111-YRmKkKDEGSYdLge-Kp-----aYqKaPTk---EFYA	

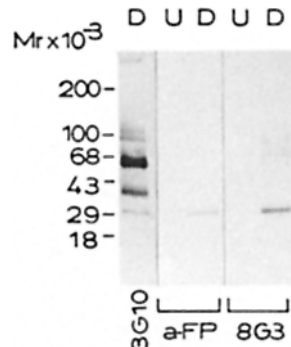
**Figure 5.** Alignment of the terminal ectodomain, transmembrane and cytoplasmic sequences of syndecan, fibroglycan and amphiglycan. Gaps have been included to allow for the maximum of sequence identity. The numbers indicate the position of the start of these sequences in the proteins (Mali et al., 1990; Marynen et al., 1989). The deduced consensus sequence shows a dash where the three sequences disagree, small lettering where only two out of the three show the same residue, and capital letters where the sequence of all three proteins is identical.



**Figure 6.** Expression of amphiglycan as a beta galactosidase fusion protein. Clone 35K17 was subcloned in the pEX3 prokaryotic expression vector. Lysates of heat-induced POP 2136 *E. Coli* cells that had been transfected with 35K17-pEX3 (A) or with an empty pEX3 vector (C) were fractionated by electrophoresis in SDS-polyacrylamide gels and blotted onto a nitrocellulose membrane. The membranes were inactivated and stained with a monoclonal anti-beta galactosidase antibody (lanes 1 and 2), with the polyclonal anti-fusionprotein antibody (lanes 3 and 4), and with the mAb F94-8G3 (lanes 5 and 6).

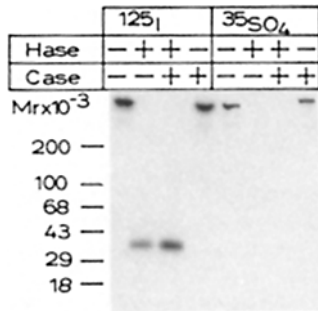
can (~48 kD), glypican (~64 kD), and syndecan (~85 kD), and finally the ~125-kD core of a still unidentified, but possibly syndecan-related heparan sulfate proteoglycan (Lories et al., 1992).

To confirm these results, mAb 8G3 was coupled to Sepharose CL4B and used to purify the immunoreactive proteoglycan from <sup>125</sup>I- and <sup>35</sup>SO<sub>4</sub>-labeled cell surface proteoglycan fractions from human lung fibroblasts. After the appropriate enzyme digestions, analysis of the iodinated proteoglycans by SDS-PAGE and autoradiography revealed multiple coreproteins in the unfractionated samples (not shown), but only a single coreprotein in the immunopurified proteoglycan (Fig. 8, lanes 1-4). Without enzyme treatment or after a digestion with only chondroitinase ABC, the 8G3-reactive proteoglycan barely entered the gel. The corresponding heparitinase-digested or doubly heparitinase and chondroitinase ABC-treated proteoglycan samples, in contrast, were entirely included in the gel and migrated as single discrete bands with apparent sizes of ~35 kD. The susceptibility of the immunopurified <sup>35</sup>SO<sub>4</sub>-labeled materials to heparitinase and their resistance to chondroitinase ABC also suggested that the proteoglycan carried mainly or only heparan sulfate chains (Fig. 8, lanes 5-8). Heparan sulfate chains



**Figure 7.** Reactivity of the cell surface proteoglycans from human lung fibroblasts with the anti-amphiglycan antibodies. The liposome-intercalatable heparan sulfate proteoglycans from human lung fibroblasts were left untreated (U) or digested with heparitinase (D), and fractionated by electrophoresis in SDS-polyacrylamide gradient gels. After blotting onto Nylon membranes the proteoglycan

fractions were stained with the polyclonal (a-FP) and monoclonal (8G3) anti-amphiglycan antibodies and with the anti-"heparan sulfate stub" antibody F69-3G10 (lane 1) to reveal all coreproteins in the samples. Undigested proteoglycans transfer poorly from the gel, but react with the antibodies as revealed by dot blot experiments (not shown). The arrowheads indicate the positions of glypican (G) and fibroglycan (F).



**Figure 8.** Enzyme-susceptibility of amphiglycan. Amphiglycan was immunopurified from  $^{125}\text{I}$ -labeled (lanes 1-4) or  $^{35}\text{SO}_4$ -labeled (lanes 5-8) cell surface proteoglycan fractions by adsorption on immobilized mAb 8G3 and analyzed by SDS-PAGE and autoradiography, either directly or after digestion with heparitinase (*Hase*), chondroitinase ABC (*Case*), or both enzymes.

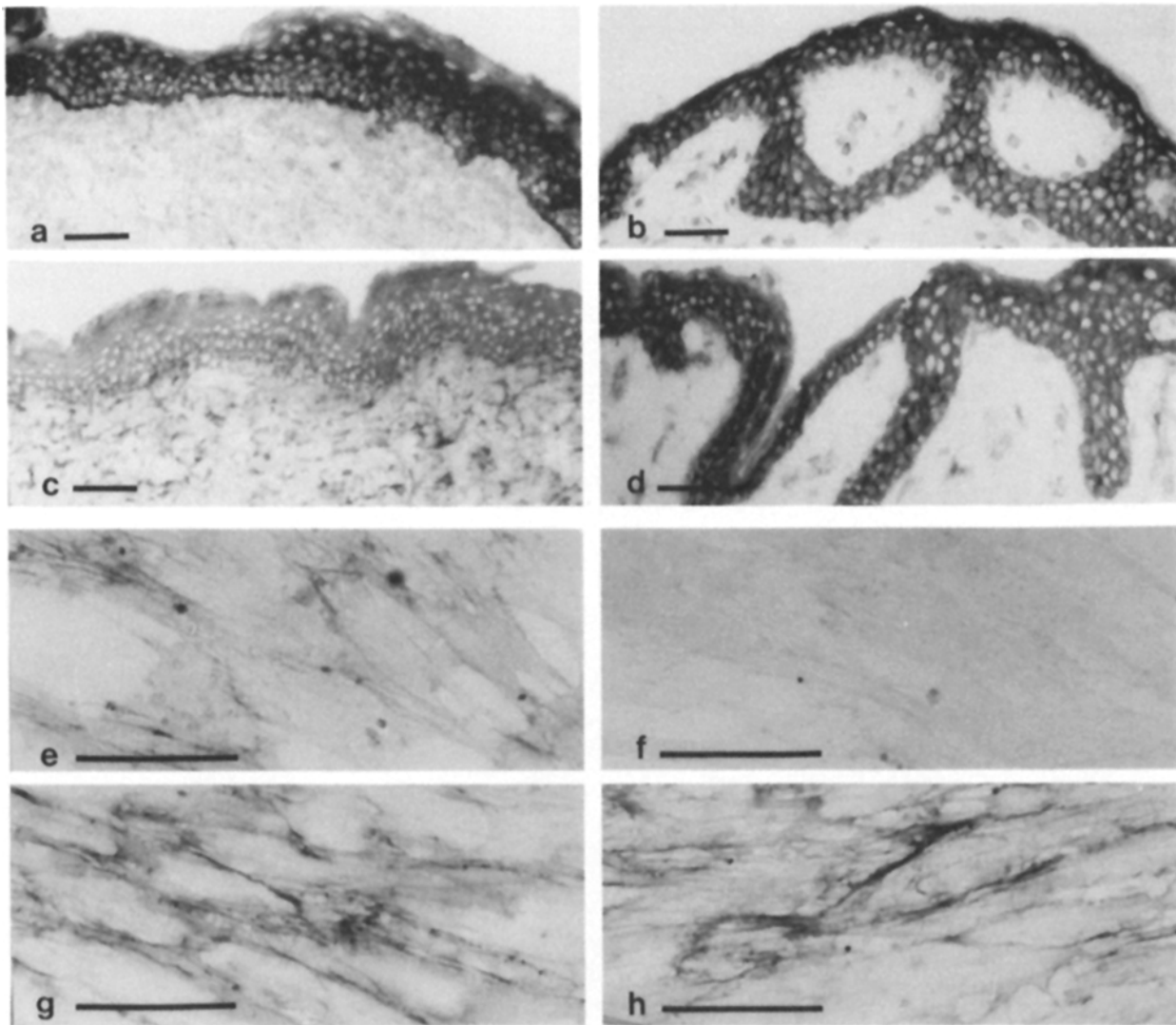
isolated from these materials by alkaline-borohydride treatment migrated with  $K_{av}$  value of 0.51-0.54 during gel filtration over Sepharose CL4B, which corresponded with an average  $M_r$  of 58-50 kD. The intact proteoglycan, in contrast, yielded a  $K_{av}$  value of 0.28 (apparent  $M_r$  > 200), which sug-

gested that the 35-kD protein core was substituted with several heparan sulfate chains.

### Expression in Other Cell Types

Northern blot analysis of total RNA extracts from several cultured cell lines using the 35K17 insert as probe revealed a single 2.6-kb mRNA in most of the samples (Fig. 3). Exceptions were some cells of hematogenous origin, where no message could be detected. Several of the cell lines with a strong 35K signal had previously been found to be relatively low expressors of mRNA for syndecan, fibroglycan, or glypican (Lories et al., 1992). Analysis of the detergent extracts of several of these cell lines by immuno-dot blotting confirmed the expression of the 8G3 epitope in these cells (not shown).

Immunohistochemically, the 8G3 epitope could be detected in both epithelial and connective tissue cells (Fig. 9). For example, in human skin the contours of the cells of the basal cell layer and of the stratum spinosum of the epidermis



**Figure 9.** Expression of the cell surface proteoglycans in human skin tissues. Sections of human skin (a-d) and whole mounts of cultured human skin fibroblasts (e-h) were stained using an indirect immunoperoxidase technique after incubation with either anti-amphiglycan (a and e), anti-syndecan (b and f), anti-fibroglycan (c and g), or anti-glypican (d and h) monoclonal first antibodies. Bars, 50  $\mu\text{m}$ .

were decorated by the antibody, but the most superficial layers of the stratum granulosum and the keratinized cells were unstained (Fig. 9 *a*). mAbs directed against syndecan (Fig. 9 *b*) or glypican (Fig. 9 *d*) also stained the epidermis. However, the anti-syndecan antibody (and to a lesser extent the anti-glypican antibody) stained mostly areas of cell-cell contact where the 8G3 antibody also stained the interface between the basal cells and the underlying matrix. A mAb directed against fibroglycan, in contrast, stained primarily cells in the dermis (Fig. 9 *c*). Unlike the anti-syndecan, but like the anti-fibroglycan and anti-glypican antibodies, antibody 8G3 reacted also with the surface of unpermeabilized cultured fibroblasts (Fig. 9, *e-h*), consistent with an occurrence at the surface of several different cell types. These data suggest that the 34K core protein may have a rather ubiquitous distribution, in contrast with syndecan and fibroglycan which, in the adult, appear to be preferentially expressed by epithelial and fibroblastic cells, respectively.

## Discussion

The heparan sulfate proteoglycans of the cell surface and extracellular matrix have the potential to modulate a variety of processes by affecting the compartmentalization, conformation, and reactivities of their ligands. They catalyze the inhibition of proteinases, stabilize and potentiate the activities of growth factors, and affect the localization and translocation of several structural proteins and enzymes. The heparan sulfate chains that mediate these effects are carried by a number of distinct cell surface and matrix proteins. The present report describes the molecular cloning of a novel cell surface proteoglycan which we have named "amphiglycan." Together with fibroglycan and syndecan it appears to belong to a distinctive family of integral membrane proteins that share conserved transmembrane and cytoplasmic domains and characteristic sequences in their ectodomains that may serve for the attachment of heparan sulfate chains. These three different proteins may assume similar functions but in different cellular and molecular contexts.

Direct evidence that the third member of this family is a heparan sulfate proteoglycan has been provided by raising antibodies against a recombinant version of this protein and demonstrating that these antibodies react with a specific core protein in proteoglycan extracts from human lung fibroblasts. Moreover, the predicted amino-terminal sequence of this protein is identical to that of a heparan sulfate proteoglycan isolated some years ago by other investigators (Castillo et al., 1987) from the culture medium of an established rabbit aorta endothelial cell line. The molecular mass of the endothelial proteoglycan core, however, was estimated at 22 kD, which is smaller than the  $M_r$  of the mature fibroblast core protein (35 kD) and seems to be in better agreement with the mass predicted from the cDNA sequence ( $\sim$ 20 kD). Whether these discrepancies reflect extensive posttranslational modifications of the protein core in fibroblasts and the existence of amphiglycan variant forms is not clear. All heparan sulfate proteoglycans of this family (syndecan, fibroglycan, as well as amphiglycan) in fibroblasts have core protein sizes that appear significantly larger than predicted from the cDNA sequences. If explainable by glycosylation and differences in glycosylation, then the added mass and apparent size differences for amphiglycan between fibroblasts and en-

dothelial cells should reside in O-linked residues since the protein contains no asparagine that seems amenable to glycosylation but one cluster of hydroxylated amino acids that may serve for that purpose. O-linked carbohydrate chains may range in size from 1 to more than 20 sugars, and the presence of even relatively short O-glycosylated sequences is known to have profound effects on the size and shape of glycoproteins (Jentoft, 1990). In this context it may also be noted that fibroglycan, which has the same predicted core size as amphiglycan but one site for N-glycosylation and two serine-threonine clusters, migrates as a 40–50-kD protein (Marynen et al., 1989). Another possibility that may account for the different size estimates in endothelial and fibroblastic cells is a cleavage of the endothelial core protein, e.g., by trypsin-like enzymes at the site of the dibasic sequences in the ectodomain, causing its release to the medium, whereas the size estimate for the fibroblast proteoglycan pertains to the membrane-associated form of the molecule.

The three members of this proteoglycan family and glypican, the phosphatidylinositol-linked heparan sulfate proteoglycan, share a sequence motif in the ectodomain that strongly resembles the serine-glycine repeat and flanking sequence of serglycin, the protein that forms the core of the heparin proteoglycan in mast cells (Kjellén et al., 1989), and therefore may serve for the attachment of heparin and heparan sulfate chains to these proteins. The consensus sequence of this "glycanation signal" could be read as Acidic-XXX-XSGXG (see Fig. 4). Interestingly, amino acid 21 of the amino terminus of the endothelial proteoglycan (corresponding to the serine of the first SG sequence in amphiglycan) appeared as a blank during protein sequencing (Castillo et al., 1987), which suggests that the consensus structure of the glycanation signal may perhaps be simplified to the sequence Acidic-XXXXXSG or even shorter versions (Acidic-XSG) of this sequence. Which sites are glycanated in amphiglycan in lung fibroblasts has not been established, but the apparent size of the intact proteoglycan in fibroblasts is such ( $>$ 200 kD) that it is likely to contain several heparan sulfate chains. On the other hand, neither amphiglycan nor fibroglycan, which carry only heparan sulfate chains where syndecan also carries chondroitin sulfate, have a sequence of the type Acidic-XSG-Acidic encountered in syndecan (Saunders et al., 1989). It should be noted, however, that in none of these cell surface proteoglycans have the glycosylation sites been directly identified.

The most striking structural similarities between syndecan, fibroglycan, and amphiglycan, which seem to group these molecules as members of the same family, occur in the membrane-spanning and cytoplasmic segments of these proteins and includes the conservation of four characteristic tyrosines. The functional significance of these structures is not known. However, in polarized epithelial monolayers the expression of syndecan becomes confined to the basolateral cell surfaces (Rapraeger et al., 1986), and in confluent human lung fibroblasts the different cell surface proteoglycans appear to be rapidly cleared from the cell surface and degraded, except glypican, which is shed (David et al., 1990). These observations imply that the conserved domains may contain signals for the targeting of these proteins (and possibly of their associated ligands) to specific membrane and endocytotic compartments, and that as for other membrane proteins, tyrosine residues may form essential parts of

these signals (Hopkins, 1991). The most strictly conserved parts of the sequences occur between the first and second and around the fourth tyrosine and might therefore support functions that are common to all proteoglycan forms. Transfection and in situ mutagenesis experiments may help to find experimental support for and clarification of these possibilities.

High levels of fibroglycan appear characteristic for fibroblastic cells (Lories et al., 1992), and syndecan, although transiently expressed in condensing mesenchymes (Bernfield and Sanderson, 1990; Vainio et al., 1991), seems to prevail in epithelia (Hayashi et al., 1987), but amphiglycan shows no clear predilection and is an abundant proteoglycan in both cell types. Whether this pattern of expression implies that amphiglycan may assume a housekeeping function where the other proteoglycans support differentiated functions is unknown. The endothelial proteoglycan from which the amino-terminal sequence data were obtained had been implicated in the inhibition of the extrinsic coagulation pathway (Castillo et al., 1987), suggesting that amphiglycan in these, and perhaps in other cells as well, may be involved in the activation of proteinase inhibitors. Results on the binding of the different cell surface proteoglycans of human umbilical vein endothelial cells to antithrombin III, however, suggest that this is a property which is neither unique nor typical for amphiglycan (Mertens et al., manuscript in preparation). Moreover, in fibroblasts we have failed to detect differences between any of the three members of the family in their abilities to bind to type I collagen or fibronectin (not shown), suggesting that specifications in their respective functions, if any, may reside on (yet unidentified) properties or associations of the core proteins or in the specific nature of the heparan sulfate-dependent processes that are at work in the cells where these proteoglycans are expressed. It will also be important to determine to what extent the expression of the different proteins may be modulated and by what signals, as differential regulatory mechanisms may endow the cells with greater flexibilities in their responses to various effector systems that tune the level of heparan sulfate at the cell surface and the processes that depend on it.

We thank Helga Ceulemans, Christien Coomans, Gisèle Degeest, Staf Doucet, Lut Meekers, An Rayé, and Marleen Willems for their expert technical assistance.

Guido David is a Research Director of the "Nationaal Fonds voor Wetenschappelijk Onderzoek" of Belgium.

These investigations have been supported by grant 3.0073.91 of the "Nationaal Fonds voor Wetenschappelijk Onderzoek" of Belgium, by a grant "Geconcerteerde Onderzoeks Acties" from the Belgian Government, and by the Interuniversity Network for Fundamental Research (1987-1991).

Received for publication 4 February 1992.

**Note Added in Proof.** Since the submission of this manuscript a cDNA sequence (available from EMBL/GenBank/DBJ under accession number M81785) which codes for a heparan sulfate proteoglycan from rat microvascular endothelial cells and which is probably the equivalent of the 35K17 sequence has been published (Kojima, T., N. W. Shworak, and R. D. Rosenberg. 1992. *J. Biol. Chem.* 267:4870-4877).

## References

Bernfield, M., and R. D. Sanderson. 1990. Syndecan, a morphogenetically regulated cell surface proteoglycan that binds extracellular matrix and growth factors. *Phil. Trans. R. Soc. Lond.* 327:171-186.

- Bourdon, M. A., T. Krusius, S. Campbell, N. B. Schwartz, and E. Ruoslahti. 1987. Identification and synthesis of a recognition signal for the attachment of glycosaminoglycans to proteins. *Proc. Natl. Acad. Sci. USA.* 84:3194-3198.
- Brown, T. A., T. Bouchard, T. St. John, E. Wayner, and W. G. Carter. 1991. Human keratinocytes express a new CD44 core protein (CD44E) as a heparan-sulfate intrinsic membrane proteoglycan with additional exons. *J. Cell Biol.* 113:207-221.
- Brunner, G., J. Gabrilove, D. B. Rifkin, E. L. Wilson. 1991. Phospholipase C release of basic fibroblast growth factor from human bone marrow cultures as a biologically active complex with a phosphatidylinositol-anchored heparan sulfate proteoglycan. *J. Cell Biol.* 114:1275-1283.
- Castillo, C. J., P. Colburn, and V. Buonassisi. 1987. Characterization and N-terminal sequence of a heparan sulphate proteoglycan synthesized by endothelial cells in culture. *Biochem. J.* 247:687-693.
- Chomczynski, P., and N. Sacchi. 1987. Single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. *Anal. Biochem.* 162:156-159.
- David, G., V. Lories, A. Heremans, B. Van Der Schueren, J.-J. Cassiman, and H. Van den Berghe. 1989. Membrane-associated chondroitin sulfate proteoglycans of human lung fibroblasts. *J. Cell Biol.* 108:1165-1175.
- 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.
- de Agostini, A. I., S. C. Watkins, H. S. Slayter, H. Youssoufian, and R. D. Rosenberg. 1990. Localization of anticoagulant active heparan sulfate proteoglycans in vascular endothelium: antithrombin binding on cultured endothelial cells and perfused rat aorta. *J. Cell Biol.* 111:1293-1304.
- De Boeck, H., V. Lories, G. David, J.-J. Cassiman, and H. Van den Berghe. 1987. Identification of a 64 kDa heparan sulphate proteoglycan core protein from human lung fibroblast plasma membranes with a monoclonal antibody. *Biochem. J.* 247:765-771.
- 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.
- Farrell, D. H., S. L. Wagner, R. H. Yuan, and D. D. Cunningham. 1988. Localization of protease nexin-1 on the fibroblast extracellular matrix. *J. Cell. Physiol.* 134:179-188.
- Flaumenhaft, R., D. Moscatelli, and D. B. Rifkin. 1990. Heparin and heparan sulfate increase the radius of diffusion and action of basic fibroblast growth factor. *J. Cell Biol.* 111:1651-1659.
- Gonzalez, A. M., M. Buscaglia, M. Ong, and A. Baird. 1990. Distribution of basic fibroblast growth factor in the 18-day rat fetus: localization in the basement membranes of diverse tissues. *J. Cell Biol.* 110:753-765.
- 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.
- Hopkins, C. R. 1991. Polarity signals. *Cell.* 66:827-829.
- Ishai-Michaeli, R., A. Eldor, and I. Vlodavsky. 1990. Heparanase activity expressed by platelets, neutrophils, and lymphoma cells releases active fibroblast growth factor from extracellular matrix. *Cell Reg.* 1:833-842.
- Jackson, R. L., S. J. Busch, and A. D. Cardin. 1991. Glycosaminoglycans: Molecular properties, protein interactions, and role in physiological processes. *Physiol. Rev.* 71:481-539.
- Jentoft, N. 1990. Why are proteins O-glycosylated? *TIBS (Trends Biochem. Sci.)* 15:291-294.
- Kiefer, M. C., J. C. Stephans, K. Crawford, K. Okino, and P. J. Barr. 1990. Ligand-affinity cloning and structure of a cell surface heparan sulfate proteoglycan that binds basic fibroblast growth factor. *Proc. Natl. Acad. Sci. USA.* 87:6985-6989.
- Kjellén, L., I. Pettersson, P. Lillhager, M.-L. Steen, U. Pettersson, P. Lehtonen, T. Karlsson, E. Ruoslahti, and L. Hellman. 1989. Primary structure of a mouse mastocytoma proteoglycan core proteins. *Biochem. J.* 263:105-133.
- Koda, J. E., A. Rapraeger, and M. Bernfield. 1985. Heparan sulfate proteoglycans form mouse mammary epithelial cells. *J. Biol. Chem.* 260:8157-8162.
- Kozak, M. 1989. The scanning model for translation: an update. *J. Cell Biol.* 108:229-241.
- Ktistakis, N. T., D.N. Thomas, and M. G. Roth. 1990. Characteristics of the tyrosine recognition signal for internalization of transmembrane surface glycoproteins. *J. Cell Biol.* 111:1393-1407.
- López-Casillas, F., S. Cheifetz, J. Doody, J. L. Andres, W. S. Lane, and J. Massagué. 1991. Structure and expression of the membrane proteoglycan betaglycan, a component of the TGF- $\beta$  receptor system. *Cell.* 67:785-795.
- Lories, V., J.-J. Cassiman, H. Van den Berghe, and G. David. 1989. Multiple distinct membrane heparan sulfate proteoglycans in human lung fibroblasts. *J. Biol. Chem.* 264:7009-7016.
- Lories, V., J.-J. Cassiman, H. Van den Berghe, and G. David. 1992. Differential expression of cell surface heparan sulfate proteoglycans in human mammary epithelial cells and lung fibroblasts. *J. Biol. Chem.* 267:1116-1122.
- 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.



- Mann, D. M., Y. Yamaguchi, M. A. Bourdon, and E. Ruoslahti. 1990. Analysis of glycosaminoglycan substitution in decorin by site-directed mutagenesis. *J. Biol. Chem.* 265:5317-5323.
- Marcum J. A., D. H. Atha, L. M. S. Fritze, P. Nawroth, D. Stern, and R. D. Rosenberg. 1986. Cloned bovine aortic endothelial cells synthesize anticoagulant active heparan sulfate proteoglycan. *J. Biol. Chem.* 261:7507-7517.
- 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.
- Pejler, G., G. Backstrom, U. Lindahl, M. Paulsson, M. Dziadek, S. Fujiwara, and R. Timpl. 1987. Structure and affinity for antithrombin of heparan sulfate chains derived from basement membrane proteoglycans. *J. Biol. Chem.* 262:5036-5043.
- Rapraeger, A., M. Jalkanen, and M. Bernfield. 1986. Cell surface proteoglycan associates with the cytoskeleton at the basolateral cell surface of mouse mammary epithelial cells. *J. Cell Biol.* 103:2683-2696.
- Rapraeger, A. C., A. Krufka, and B. B. Olwin. 1991. Requirement of heparan sulfate for bFGF-mediated fibroblast growth and myoblast differentiation. *Science (Wash. DC)*. 252:1705-1708.
- Saksela, O., and D. B. Rifkin. 1990. Release of basic fibroblast growth factor-heparan sulfate complexes from endothelial cells by plasminogen activator-mediated proteolytic activity. *J. Cell Biol.* 110:767-775.
- Saksela, O., D. Moscatelli, A. Sommer, and D. B. Rifkin. 1988. Endothelial cell-derived heparan sulfate binds basic fibroblast growth factor and protects it from proteolytic degradation. *J. Cell Biol.* 107:743-751.
- 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.
- Sayama, S., R. V. Iozzo, G. S. Lazarus, and N. M. Schechter. 1987. Human skin chymotrypsin-like proteinase chymase. Subcellular localization to mast cell granules and interaction with heparin and other glycosaminoglycans. *J. Biol. Chem.* 262:6808-6815.
- Stamenkovic, I., A. Aruffo, M. Amiot, and B. Seed. 1991. The hematopoietic and epithelial forms of CD44 are distinct polypeptides with different adhesion potentials for hyaluronate-bearing cells. *EMBO (Eur. Mol. Biol. Organ.) J.* 10:343-348.
- Stevens, R. L., S. Avraham, M. C. Gartner, G. A. P. Bruns, K. F. Austen, and J. H. Weis. 1988. Isolation and characterization of a cDNA that encodes the peptide core of the secretory granule proteoglycan of human promyelocytic leukemia HL-60 cells. *J. Biol. Chem.* 263:7287-7291.
- Vainio, S., M. Jalkanen, A. Vaahtokari, C. Sahlberg, M. Mali, M. Bernfield, and I. Thesleff. 1991. Expression of syndecan gene is induced early, is transient, and correlates with changes in mesenchymal cell proliferation during tooth organogenesis. *Dev. Biol.* 147:322-333.
- Wang, X.-F., H. Y. Lin, E. Ng-Eaton, J. Downward, H. F. Lodish, and R. A. Weinberg. 1991. Expression cloning and characterization of the TGF- $\beta$  type III receptor. *Cell.* 67:797-805.
- Yayon, A., M. Klagsbrun, J. D. Esko, P. Leder, and D. M. Ornitz. 1991. Cell surface, heparin-like molecules are required for binding of basic fibroblast growth factor to its high affinity receptor. *Cell.* 64:841-848.