Molecular Cloning of a Phosphatidylinositol-anchored Membrane Heparan Sulfate Proteoglycan from Human Lung Fibroblasts

Guido David,* Veerle Lories,* Benny Decock,‡ Peter Marynen,* Jean-Jacques Cassiman,* and Herman Van den Berghe*

*Center for Human Genetics and ‡Rega Institute for Medical Research, University of Leuven, B-3000 Leuven, Belgium

Abstract. Two mAbs raised against the 64-kD core protein of a membrane heparan sulfate proteoglycan from human lung fibroblasts also recognize a nonhydrophobic proteoglycan which accumulates in the culture medium of the cells. Pulse–chase studies suggest that the hydrophobic cell-associated forms act as precursors for the nonhydrophobic medium-released species. The core proteins of the medium-released proteoglycans are slightly smaller than those of the hydrophobic cell-associated species, but the NH_2-terminal amino acid sequences of both forms are identical. The characterization of human lung fibroblast cDNAs that encode the message for these core proteins and the effect of bacterial phosphatidylinositol-specific phospholipase C suggest that the hydrophobic proteoglycan is membrane-anchored through a phospholipid tail. These data identify a novel membrane proteoglycan in human lung fibroblasts and imply that the shedding of this proteoglycan may be related to the presence of the phospholipid anchor.

THE heparan sulfate proteoglycans from the cell surface can bind a variety of ligands, including cell adhesion molecules (Cole and Burg, 1989), matrix components (Koda et al., 1985; Saunders and Bernfield, 1988; Sun et al., 1989; LeBaron et al., 1989), growth factors (Segarini and Seyedin, 1988; Cheifetz et al., 1988; Saksela and Rifkin, 1990), enzymes (Cisar et al., 1989), and enzyme inhibitors (Marcum and Rosenberg, 1987). This suggests that many of the interactions of the cells with the extracellular environment may be influenced by the presence of proteoglycan, and that these interactions are modulated by factors that determine the accumulation and abundance of the proteoglycans at the cell surface (Ruoslahti, 1988; Gallagher, 1989).

Some of these proteoglycans are only peripherally associated with the cell surface, as they can be displaced by high salt concentrations or by polyanions, but others have tight membrane linkages (Höök et al., 1984). Because detergent is required to obtain their complete solubilization, and because they can bind to hydrophobic matrices and be incorporated into liposomes, these cell surface proteoglycans have been assumed to be integral membrane components. This assumption has been validated by the characterization of cDNA clones coding for syndecan, a cell surface proteoglycan from mouse and human mammary epithelial cells (Saunders et al., 1989; Mali et al., 1990), and of others coding for the core protein of a cell surface proteoglycan from human lung fibroblasts (Marynen et al., 1989). In both instances, the predicted core protein structures display ecto, transmembrane, and cytoplasmic domains. The cell surface of rat liver cells (Ishihara et al., 1987), Schwann cells (Carey and Evans, 1989), ovarian granulosa cells (Yanagashita and McQuillan, 1989), and embryonic heart cells (Chajek-Shaul et al., 1989), however, contain heparan sulfate proteoglycans which can be released by treatment with phosphatidylinositol-specific phospholipase C, suggesting that other proteoglycans may be surface anchored by covalent linkage to membrane phospholipid, or through strong interactions with other membrane components that have such anchors. In each of these instances only a fraction of the hydrophobic cell surface proteoglycans appears sensitive to the enzyme, implying that different modes of membrane insertion may coexist in a single cell type. The observation that hamster (Bretscher, 1985) and human fibroblasts (Lories et al., 1989; Schmidtchen et al., 1990), mouse mammary epithelial cells (David and Van den Berghe, 1989), and rat granulosa cells (Yanagashita and McQuillan, 1989) contain multiple cell surface proteoglycans with structurally distinct core proteins seems to leave room for this possibility.

Where investigated, the proteoglycans from the cell surface appear to be subject to rapid turnover (David and Bernfield, 1981; Iozzo, 1984; Yanagashita and Hascall, 1984; Jalkanen et al., 1987; Carey and Evans, 1989). Structural similarities, immunological cross-reactivities, and kinetic data imply that a fraction of these proteoglycans may disappear from the cell surface by release into the pericellular environment instead of undergoing endocytosis. The interpretation of these data is not always straightforward, however, since hydrophobic and nonhydrophobic proteoglycan forms coexist at the cell surface, and since structurally related proteoglycans may, for example, represent the translation of alternative transcripts of a single gene product. Shedding of membrane-embedded proteoglycan would require the activ-

© The Rockefeller University Press, 0021-9525/90/12/3165/12 $2.00

ity of phospholipases (Ishihara et al., 1987) or of proteinases (Jalkanen et al., 1987) which separate the proteoglycan or proteoglycan ectodomain from the membrane anchor.

In the present report we describe the molecular cloning of a heparan sulfate proteoglycan from human lung fibroblasts which is linked to the cell surface via a glycosyl-phosphatidylinositol anchor and which is rapidly, quantitatively, and selectively released to the culture media of the cells. These data provide direct evidence for the molecular heterogeneity of the cell surface proteoglycans and for the existence of alternative modes of association with the plasma membrane for these molecules.

Materials and Methods

Antibodies

The immunization protocols, hybridoma selection procedures, and the characterization of monoclonal antibody S1 have been reported before (De Boeck et al., 1987). A similar approach was followed to obtain and characterize mAb F81-1G11.

Cell Culture and Labeling

Normal human lung fibroblasts were maintained in DME containing 10% FCS (Gibco, Paisley, Scotland). For the pulse-labeling and pulse-chase experiments, the cells were grown to confluency in 10-cm^2 tissue culture dishes. Before labeling, the monolayers were preincubated for 30 min in culture medium containing one tenth of the normal sulfate concentration. After 10 or 60 min of labeling with 100 μCi of carrier-free [^{35}S]sulfuric acid (New England Nuclear, Boston, MA) per milliliter of low sulfate medium, the cultures were rinsed with standard culture medium and chased in label-free culture medium supplemented with 5 mM sodium sulfate. In other experiments, the cells were doubly labeled for 24 h with 50 μCi of [^{3}H]ethanolamine (Amersham International, Amersham, UK), and 50 μCi of [^{35}S]sulfuric acid or [^{35}S]methionine (Amersham International) per milliliter of medium.

Analysis of the Pulse-Chase Experiments

After varying times of chase, the culture media were withdrawn. The monolayers were rinsed once with 5 ml of PBS. Rinses and chase media were pooled. The cell monolayers were extracted for 30 min with ice cold detergent buffer (0.5% Triton X-100, 150 mM NaCl, 50 mM Tris-HCl, pH 7.4) supplemented with proteinase inhibitors (50 mM 6-amino hexanoic acid, 5 mM benzamidine, 5 mM *N*-ethyl maleimide, 10 mM EDTA, 3 μM pepstatin A, 50 μg/ml leupeptin, and 2 mM PMSF. The combined media and rinses, and the detergent extracts were applied on small columns of DEAE-Trisacryl M. Bound proteoglycan was eluted with 0.5% Triton X-100, 0.6 M NaCl, 50 mM Tris-HCl, pH 7.4. The eluates were diluted fourfold with 0.5% Triton X-100, 50 mM Tris-HCl, pH 7.4, supplemented with proteinase inhibitors, and incubated overnight with 5 μg of Sepharose 4B-coupled mAb S1. After incubation, the immunobeads were washed in 150 mM NaCl, 50 mM Tris-HCl, pH 7.4, and bound proteoglycan was eluted in 50 mM octylglucoside, 4 M guanidinium chloride, 50 mM Tris-HCl, pH 8.0. The immunopurified proteoglycan was tested for hydrophobicity by incorporation in liposomes as described below. The amount of ^{35}S-heparan sulfate in the extracts and in the immunopurified samples was assessed by cetylpyridinium chloride precipitation of proteinase K- and chondroitinase ABC-digested samples as described before (David and Van den Berghe, 1983).

Purification and Radioiodination of the Heparan Sulfate Proteoglycans

The cell surface-associated heparan sulfate proteoglycans were isolated as described before (Lories et al., 1987) with minor modifications of the procedure. The Triton X-100 extracts of the cells were concentrated and partially purified by adsorption on DEAE-Trisacryl M and Mono Q. After rinsing with 0.6 M NaCl in 0.5% Triton X-100, 6 M urea, 50 mM Tris-HCl, pH 8.0, to remove most of the contaminating protein, and with 50 mM octylglycoside, 6 M urea, 50 mM Tris-HCl, pH 8.0, to exchange detergents, the proteoglycans were eluted in 100 mM octylglucoside, 4 M guanidinium chloride, 50 mM Tris-HCl, pH 8.0, and mixed with 5 mg/ml of phosphatidylcholine. The mixture was dialyzed against 4 M guanidinium chloride buffer to allow the formation of liposomes, and applied on Sepharose CL4B in 4 M guanidinium chloride buffer to separate the liposome-intercalated proteoglycans, which elute in the void volume, from the non-intercalated forms, which elute in the included volumes (Lories et al., 1986). The lipophilic proteoglycans were supplemented with 1% Triton X-100, dialyzed against 0.5% Triton X-100, 6 M urea, 50 mM Tris-HCl, pH 8.0, and further purified by ion exchange chromatography on Mono Q in this buffer. The medium-released heparan sulfate proteoglycans were purified by ion-exchange chromatography, gel filtration on Sepharose CL4B, and CsCl density gradient centrifugation in the presence of 4 M guanidinium chloride as reported before (Lories et al., 1986; Heremans et al., 1988).

The purified cell surface and medium proteoglycans were radioiodinated using Na ^{125}I and chloramine T (Lories et al., 1987). The labeled proteoglycans were further immunopurified on mAb S1 for core protein sizing and peptide mapping, and assessed for their hydrophobicity before and after phospholipase C treatment.

Enzyme Treatment

The digestions of the proteoglycan fractions with heparitinase, chondroitinase ABC, and trypsin were as reported before (David and Van den Berghe, 1989; Lories et al., 1989). For the treatment with phosphatidyl inositol-phospholipase C (Boehringer Mannheim GmbH, Mannheim, FRG), the proteoglycan samples were dissolved in 0.1% Triton X-100, 10 mM EDTA, 50 mM Tris-HCl, pH 7.5, and incubated with 50 mU of enzyme (200 mU/ml) for 3 h at 37°C. After the digestion, the treated and control samples were absorbed on DEAE, rinsed to remove the Triton X-100, and eluted in 100 mM octylglucoside, 4 M guanidinium chloride, 50 mM Tris-HCl, pH 8.0.

Protein Sequencing

For the cyanogen (CNBr)[1] bromide fragmentation and NH_2-terminal amino acid sequence analyses, the medium and cell surface-associated heparan sulfate proteoglycans were purified by immunoaffinity chromatography on mAb S1. The proteoglycan from 10 liters of conditioned culture medium was concentrated by batch adsorption on DEAE-Trisacryl M (10 ml/liter), and incubated with 10 mg of Sepharose 4B-coupled antibody S1 in 0.5% Triton X-100, 0.3 M NaCl, 50 mM imidazol-HCl, pH 6.8 (buffer A). After overnight incubation the immunobeads were rinsed with four column volumes of buffer A supplemented with 100 μg of heparin/ml, with 4 vol of the same buffer supplemented with 600 mM NaCl, and finally again with buffer A. Bound proteoglycan was eluted with 0.5% Triton X-100, 4 M guanidinium hydrochloride, 50 mM imidazol-HCl, pH 6.8, and further purified by ion exchange chromatography on Mono Q in 0.5% Triton X-100, 6 M urea, 50 mM imidazol-HCl, pH 6.8. The elution of S1-reactive proteoglycan and of any contaminating mAb S1 was monitored by immunodot blotting (De Boeck et al., 1987). Cell surface-associated proteoglycan was purified by the same procedure, starting from the liposome-intercalated proteoglycan fraction from ~500 culture flasks.

The S1-immunopurified proteoglycan was reduced with 10 mM DTT for 3 h at room temperature and alkylated overnight with 40 mM iodoacetamide. The reduced and alkylated proteoglycan was precipitated with 3 vol of ethanol-1% sodium acetate and resuspended in 100 μl of 70% formic acid containing 50 mg CNBr/ml. After overnight incubation the sample was evaporated to dryness and resolubilized in 50 mM ammonium bicarbonate. After several cycles of evaporation/resolubilization, the sample was resuspended in 2% SDS/20 mM Tris-HCl, pH 6.8/5% glycerol, boiled, and submitted to SDS-PAGE in the Jovin buffer system (Moos et al., 1988). For the analysis of the NH_2 termini of the intact core proteins, the immunopurified medium and cell surface proteoglycans were treated with heparitinase before electrophoresis. After electrophoresis the core proteins and the CNBr fragments of these core proteins were electrotransferred to Immobilon membranes (Millipore/Continental Water Systems, Bedford, MA), and stained with Coomassie brilliant blue or with mAb 1G11. The Coomassie-stained bands were excised and transferred to the cartridge of a sequenator (model 477A; Applied Biosystems Inc., Foster City, CA). The phenylhydantoin derivatives of the amino acids were identified on line using an analyzer (model Applied 120A; Applied Biosystems Inc.).

1. *Abbreviation used in this paper*: CNBr, cyanogen bromide.

Table I. Protein Sequencing Data and Derived Oligonucleotide Sequences

	Amino acid sequences	Oligonucleotide primers
NCP M	DPASKSRSCGEVRQIYGAKGF- SLSDVPQAE	5′ GAC CCT GCC TCC AAG TCC CGG TCC TGT GGG GAG GTG CGG CAG ATC TAT GGG GCC AAG GGC TTC
NCP C	DPAS	
CNB 20	LATQLRSFDDHFQ- HLLXDS	5′ CTG GCC ACC CAG CTG CGG TCC TTT GAT GAC CAC TTC CAG + antisense complement
CNB 14	VLITDKFWGTSGVE- XVIGSV	5′ GTG CTG ATC ACA GAC AAG TTC TGG GGC ACC TCT GGG GTG GAG + antisense complement

The NH_2 termini of the native core proteins (*NCP*) isolated from the culture media (*M*) and from the cells (*C*), and of the 20- and 14-kD CNBr fragments (*CNB*) obtained from the medium proteoglycan were analyzed in a gas phase–automated protein sequencer. Gaps in the sequence are indicated by an X. The parts of the sequence that were reverse translated into oligonucleotide are underlined. The sequence of these oligonucleotide primers is given in the right column.

Amplification of Core Protein cDNA

Three sense (Table I) and two antisense oligonucleotide primers were synthesized. Their sequences were based on the reverse translation of the NH_2-terminal amino acid sequences of the intact core protein and of the CNBr core protein fragments (Table I). The primers (150 pmol) were mixed with 1 ng of cDNA obtained from the first strand synthesis reaction during the construction of the human lung fibroblast cDNA library (see below). The reaction mixture further contained 1.5 mM of each dNTP, 1× Taq polymerase buffer (Kogan et al., 1987), 10% (vol/vol) DMSO, and 2 U of Taq polymerase (Perkin-Elmer Corp. Norwalk, CT). After 30 thermal cycles (1 min denaturation at 94°C, 1 min annealing at 55°C, and 1 min extension at 68°C), the amplification products were analyzed in 1.1% agarose gels and detected by ethidium bromide staining. The ~900-, ~600-, and ~300-bp amplification products were electroeluted, ethanol precipitated, and ligated in the Sma I restriction site of pGEM-3Z (Promega Biotec, Madison, WI). After authentification by DNA sequencing, the 900-bp amplification product was labeled using [^{35}P]dCTP, random oligonucleotide primers, and Klenow enzyme (Feinberg and Vogelstein, 1983), and used to screen a human lung fibroblast cDNA library.

Isolation of Core Protein cDNA Clones

To construct the cDNA library, poly(A)$^+$ RNA was extracted from cultured human lung fibroblasts using 5 M guanidine isothiocyanate, and isolated by CsCl density gradient centrifugation and oligo(dT) cellulose chromatography as described before (Marynen et al., 1989). The cDNA prepared from this poly(A)$^+$ RNA using a cDNA synthesis kit (Gibco BRL, Gent, Belgium) was blunted, ligated to Eco RI adaptors (Pharmacia-LKB Biotechnology, Uppsala, Sweden), kinased, and size fractionated by gel permeation chromatography on a Bio-gel A-50m column or by electrophoresis in agarose gels. The cDNA larger than 2,000 bp was ethanol precipitated and ligated into Eco RI–prepared dephosphorylated λ ZAP II arms (Stratagene Cloning Systems, La Jolla, CA). The ligated DNA was packaged (Gigapak Gold, Stratagene), plated onto BB4 host cells, and screened without amplification.

Initially, 4 × 10^5 independent plaques were screened using the ^{32}P-oligolabeled 900-bp polymerase chain reaction amplification product. One positive clone (64K1) was plaque purified and converted to the Bluescript SK$^-$ plasmid by superinfection with R408 helper phage. The 64K1 insert was sequenced and used to further screen 6 × 10^5 plaques, yielding five additional cross-hybridizing clones, including clones 64K3 and 64K4. All clones were plaque purified, converted into Bluescript plasmids, and compared by restriction mapping.

Analysis of the cDNA Clones

cDNA fragments released by restriction enzyme digestion of the original Bluescript SK$^-$ plasmids were resolved by electrophoresis in agarose gels, electroeluted, and subcloned in Bluescript SK$^-$ or pGEM-3Z plasmids. The cDNA inserts were sequenced by the dideoxy chain termination method (Sanger et al., 1977), using supercoiled plasmid and a modified T7 DNA polymerase (Pharmacia-LKB Biotechnology). The full sequences of both strands were obtained by using T7 and T3 primers for the Bluescript SK$^-$ plasmids, and T7 and SP6 primers for pGEM-3Z plasmids. The complete sequence of larger fragments was obtained by using custom oligonucleotide primers synthesized in accordance to upstream sequences. All sequences were determined by using both dGTP and 7-deaza dGTP (Pharmacia-LKB Biotechnology). Some zones with severe compression were sequenced with 7-deaza dITP; some zones complicated by "full stops" were sequenced using Taq polymerase (Promega Biotec).

Construction of Expression Plasmids

A cDNA fragment spanning from the Bam HI site of the 64K3 insert to the Sal I site in the multiple cloning site of the Bluescript vector was ligated into the plasmid expression vector pEX$_2$ (Genofit, Geneva, Switzerland) digested with the corresponding restriction enzymes. In this construct the 64K coding sequences are linked, in frame, to the β galactosidase-coding sequences through the intermediate of the 50 bp normally untranslated, but open sequence preceding the ATG start codon (see Fig. 6). Transformed POP 2136 cells were grown overnight at 28°C on ampicillin plates. Exponentially growing cultures were initiated from single colonies, and induced by shifting the incubation temperature from 28 to 42°C for 2 h. After treatment with lysosyme and extraction with 6 M urea, as described before (Marynen et al., 1989), the cells were solubilized in SDS and the fusion proteins were analyzed by Western blotting using mAb 1G11 and antigalactosidase antibody (Promega Biotec).

Northern Blot Analysis

Poly(A)$^+$ RNA for Northern analysis was isolated from human lung fibroblasts as described above. 2.4-μg aliquots of denatured poly(A)$^+$ RNA per lane were submitted to electrophoresis in 1.2% agarose gels containing 0.2 M 3-(*N*-morpholino) propanesulfonic acid, 0.05 M sodium acetate, pH 7.0, 10 mM EDTA, and 6% formaldehyde (Maniatis et al., 1982). The RNA was transferred to Nytran membranes (Shleicher & Schuell, Inc., Keene, NH) and immobilized by UV cross-linking (Church and Gilbert, 1984). The blots were prehybridized for 4 h at 42°C in 50% formamide, 0.1% SDS, 0.1 mg/ml of heparin, 0.1 mg/ml denatured salmon sperm DNA, 5× SSPE, and 5× Denhardt's solution. Hybridization lasted overnight at 42°C in the same buffer containing 1× Denhardt's solution and 10^6 cpm ^{32}P-labeled cDNA/ml. The filters were washed twice for 15 min at room temperature and once for 30 min at 42°C in 0.1% SDS, 2× SSPE; and once for 30 min at 55°C, and once more for 30 min at 65°C in 0.1% SDS, 0.5% SSPE. The size of the RNA was estimated from the position of the 28S and 18S ribosomal RNA and that of RNA markers (Gibco BRL).

Results

Cell Surface Proteoglycan Shedding

Fetal human lung fibroblasts produce at least three distinct membrane-associated heparan sulfate proteoglycans. These

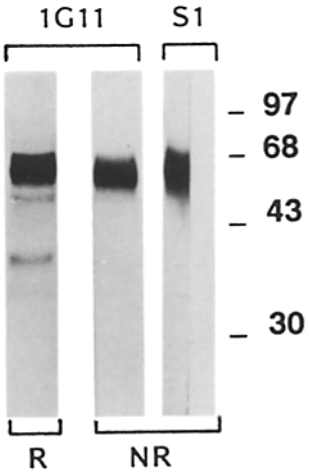

Figure 1. Specificity of the mAbs. The membrane proteoglycans from human lung fibroblasts were digested with heparitinase and submitted to SDS-PAGE under reducing (*R*) or nonreducing (*NR*) conditions. The fractionated digests were electroblotted onto a cationized nylon membrane and stained, using mAb 1G11 and horseradish peroxidase–conjugated rabbit anti-mouse Ig. One half of the third lane was stained with mAb S1, the other half with the second antibody only. Molecular weights are indicated on the right ($\times 10^{-3}$).

three proteoglycans differ in the sizes of their core proteins after heparitinase digestion, in their peptide maps, and in their immunological cross-reactivities (Lories et al., 1989). In preliminary screening experiments we observed that two of the mAbs that were raised against these membrane-associated proteoglycans also reacted with proteoglycan from conditioned culture media (not shown). Both antibodies reacted with the 64-kD core protein of the membrane-associated proteoglycan; one (mAb S1) exclusively with the unreduced form of this proteoglycan (De Boeck et al., 1987), the other (mAb 1G11) with both the reduced and unreduced forms (Fig. 1).

Metabolic labeling of the cells with $^{35}SO_4$ for brief periods, during which the hydrophobic heparan sulfate proteoglycans become preferentially labeled, followed by chase experiments in precursor-free media had suggested that at least some of the membrane proteoglycans act as precursors to the medium-released proteoglycans (Lories et al., 1986). We further investigated possible precursor-product relationships by using mAb S1 to immunoprecipitate $^{35}SO_4$-labeled proteoglycan recovered from cell and culture media fractions during pulse–chase experiments (Fig. 2). During a chase of briefly (10- or 60-min) labeled cells, the total ^{35}S-heparan sulfate was initially rapidly lost from the cells, but this rate of loss decreased after 2–3 h of chase. Concomitantly, ^{35}S-heparan sulfate was rapidly accumulating in the medium in the beginning of the chase (Fig. 2 *A*). After 24 h of chase ~40% of the label originally present in the cells had been released to the media, whereas only 20% remained with the cells indicating that heparan sulfate was lost from the cells both by release to the medium and by degradation. During chases of steady state–labeled cells, the amount of S1-immunoreactive ^{35}S-heparan sulfate proteoglycan in the cells also decreased rapidly during the early phases of the chase (Fig. 2 *B*). This lost label was virtually quantitatively recovered in the culture media (Fig. 2 *B*), indicating that this proteoglycan disappeared from the cell surface exclusively through "shedding."

Shedding Alters the Hydrophobic Membrane Anchor

The cellular proteoglycans consist of a heterogeneous mixture of hydrophobic and nonhydrophobic species, whereas the medium heparan sulfate proteoglycans lack hydrophobic properties (Lories et al., 1986). To evaluate the hydrophobic properties of the S1-reactive proteoglycans, we immunopurified these proteoglycans from $^{35}SO_4$-labeled cells at the beginning of the chase and from the chase media at the end of the chase, and tested their ability to interact with liposomes (Fig. 3). When the liposome–cellular proteoglycan mixtures were chromatographed over Sepharose CL4B in 4 M guanidinium chloride, in the absence of detergent, approximately one third of the label was excluded from the column where it coeluted with the liposomes (Fig. 3 *A*). When the mixture was supplemented with Triton X-100 and chromatographed

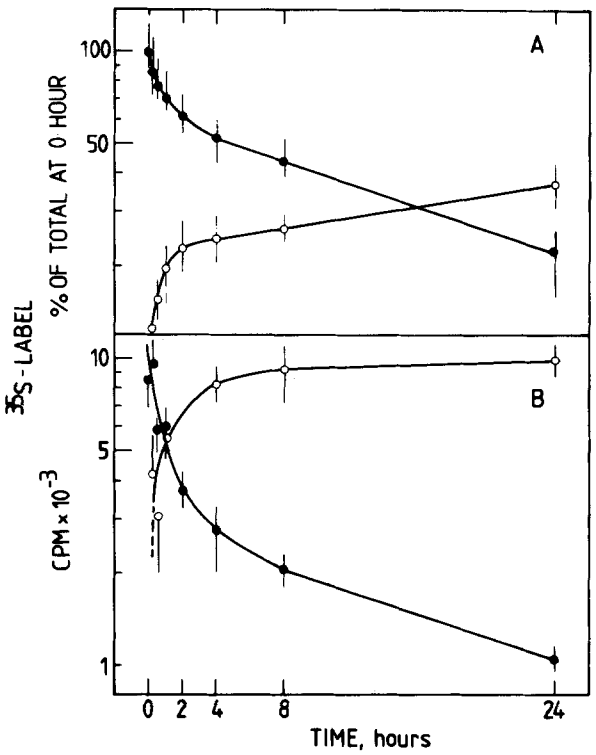

Figure 2. Loss of heparan sulfate proteoglycan from pulse-labeled cells. Dishes with confluent human lung fibroblasts were pulse labeled with $^{35}SO_4$ for 60 min (*A*), during which the membrane proteoglycans become preferentially labeled, or for 24 h (*B*), to reach near steady-state levels of label incorporation. The cultures were briefly rinsed and chased for varying lengths of time in precursor-free media. The amount of total ^{35}S-heparan sulfate (*A*) and the amount of mAb S1-bound ^{35}S-heparan sulfate (*B*) remaining with the cells (●) and accumulating in the culture media (○) were assessed by cetylpyridinium chloride precipitation of proteinase and chondroitinase ABC–digested samples. Each time point is the average value of three separate cultures. One out of three separate experiments is shown.

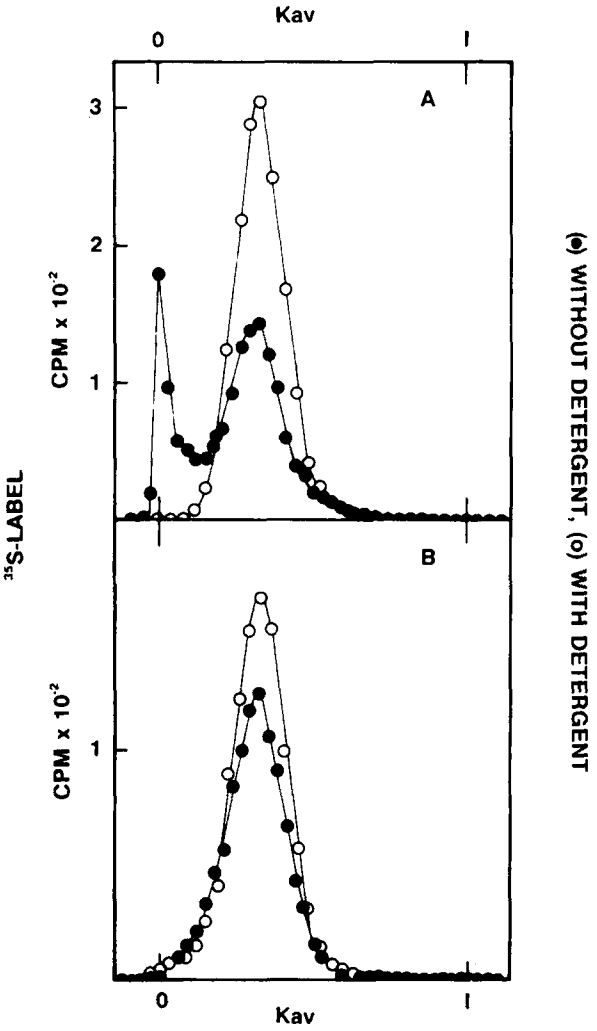

Figure 3. Assessment of the hydrophobicity of the cellular and medium proteoglycans. Proteoglycans extracted from unchased cells that had been pulse labeled with $^{35}SO_4$ for 24 h (*A*) and $^{35}SO_4$-proteoglycans accumulated in the chase media of pulse-labeled cells at the end of a 24-h chase (*B*) were immunopurified on mAb S1 and mixed with phosphatidylcholine. After dialysis, to remove the detergent, the liposome–proteoglycan mixtures were chromatographed over Sepharose CL4B in 4 M guanidinium chloride in the absence of detergent (●), or in the presence of 0.5% Triton X-100 to disrupt the liposomes (○). In the absence of detergent, ^{3}H-phosphatidylcholine–marked liposomes eluted in the void volume of the column.

in 4 M guanidinium chloride in the presence of this detergent, all label eluted as a single peak with a K_{av} of ~0.33 (Fig. 3 *A*). The S1-reactive proteoglycans from the medium, in contrast, were all nonhydrophobic, eluting as a single peak with K_{av} of ~0.33, whether analyzed in the presence or absence of detergent (Fig. 3 *B*). Since liposome-intercalatable and nonintercalatable forms were lost from the cells at similar rates (not shown), and since the lost proteoglycan was quantitatively recovered in the media (see Fig. 2 *B*), these data implied that the hydrophobic cellular proteoglycans also acted as precursors to the nonhydrophobic medium proteoglycans. This suggests that the membrane anchor of these proteoglycans is altered during the shedding process.

Shedding Alters the Structure of the Core Protein

The detergent-treated hydrophobic forms, the nonhydrophobic cellular forms, and the medium forms of the immunopurified proteoglycans were all similar in gross size (Fig. 3). In search of evidence for more subtle differences in structure we immunopurified the S1-proteoglycans from ^{125}I-labeled proteoglycan preparations. During the purification procedure before labeling, the S1 epitope copurified with the bulk of the ^{35}S-heparan sulfate proteoglycan in the membrane (detergent-extractable) fraction and in the media fractions, as monitored by immunodot blotting (not shown). In agreement with previously reported findings (Lories et al., 1987), the hydrophobic cellular proteoglycans yielded multiple core proteins after heparitinase digestion, the most prominent migrating as proteins of ~125, ~64, ~48, and 35 kD in SDS-PAGE (Fig. 4 *A*). The medium proteoglycans, in contrast, yielded only one major core protein with an apparent size of ~61 kD after heparitinase treatment (Fig. 4 *A*).

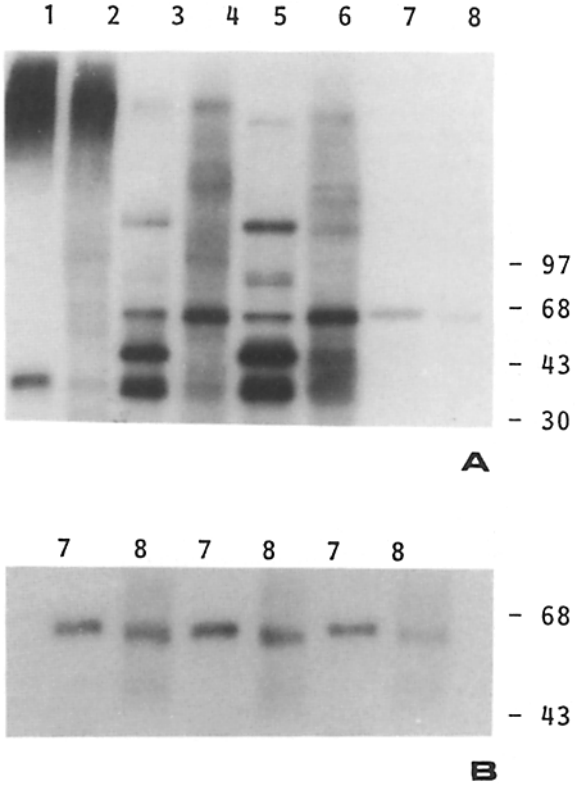

Figure 4. Comparison of the cell surface-associated and medium heparan sulfate proteoglycans after heparitinase digestion. The hydrophobic cell surface-associated heparan sulfate proteoglycans and the medium-released heparan sulfate proteoglycans from human lung fibroblasts were purified, ^{125}I-iodinated and analyzed by SDS-PAGE in (4–16%) gradient gels (*A*) or in (7%) homogeneous gels (*B*) under reducing conditions. The cellular (lanes *1*, *3*, *5*, and *7*) and medium (lanes *2*, *4*, *6*, and *8*) proteoglycans were analyzed before (lanes *1–6*) and after (lanes *7* and *8*) immunopurification on mAb S1; without enzyme digestion (lanes *1* and *2*), after heparitinase digestion (lanes *3*, *4*, *7*, and *8*), and after combined heparitinase and chondroitinase ABC digestion (lanes *5* and *6*). The bands were detected by autoradiography. Molecular weights are indicated on the right (× 10^{-3}).

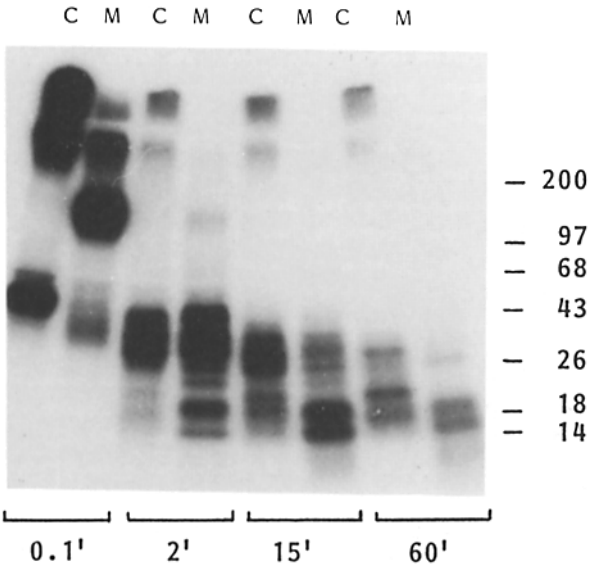

Figure 5. Comparison of the cell surface-associated and medium heparan sulfate proteoglycans after trypsin digestion. The immunopurified ^{125}I-iodinated hydrophobic cell surface (*C*) and medium (*M*) forms of the Sl-reactive proteoglycan were digested with trypsin for various lengths of time (0.1, 2, 15, and 60 min). The resulting fragments were fractionated by SDS-PAGE in (4–16%) gradients gels and analyzed by autoradiography. Molecular weights are indicated on the right ($\times 10^{-3}$).

The yield or apparent size of these major core proteins was not affected by further treatment with chondroitinase ABC (Fig. 4 *A*). A similar analysis of the Sl-reactive proteoglycan present in these iodinated preparations yielded a single ~64-kD core protein for the hydrophobic cellular forms, and a single, but slightly smaller core protein for the medium proteoglycans (Fig. 4 *B*). A similar slight core size difference was also found between hydrophobic and nonhydrophobic cellular Sl-purified proteoglycans (not shown). The immunopurified iodinated proteoglycans were also compared after treatment with trypsin for different lengths of time (Fig. 5). Some membrane- and medium-derived peptides ran at similar positions, but the partial digests of the medium proteoglycans contained a prominent doublet of ~15- and 11-kD bands, not seen in the digests of the cellular proteoglycans. The latter, instead, contained a doublet of ~18- and 14-kD bands. Since experiments with ^{35}S-proteoglycans had indicated that the glycosaminoglycan-substituted fragments did not run as sharp bands (not shown), these data suggested that the differences in the sizes of the cores and in the peptide patterns were due to differences in the protein and nonglycosaminoglycan constituents of these cores, rather than to differences in the number of glycosaminoglycan chains and glycosaminoglycan linkage regions (remaining after heparitinase treatment) that were attached to the core proteins of the cellular and medium forms of the Sl-reactive proteoglycans. Moreover, they suggested that shedding removed an ~3-kD domain from the core.

Shed and Membrane-bound Cores Have Similar NH_2 Termini

The shed form of the proteoglycan was purified from large volumes of conditioned culture media by affinity chromatography on mAb Sl and ion exchange chromatography on Mono Q. The proteoglycan was reduced and alkylated, heparitinase digested, and transferred to polyvinylidene difluoride membranes by Western blotting. The core protein band was detected by Coomassie staining, certified by staining with the reduction-insensitive antibody 1G11, excised, and sequenced. A sequence of 30 amino acids was obtained from ~37 pmol of sample (Table I). The membrane-associated form was analyzed in a similar fashion, starting from a liposome-incorporated cellular proteoglycan fraction. The yield of proteoglycan was much lower than for the shed form, and only a sequence of four residues could be obtained. These were identical to the first four residues of the shed form of the proteoglycan (Table I). These data suggested that the membrane proteoglycan was oriented with its NH_2 terminus outward, the shedding process involving its COOH-terminal end.

Isolation of Core Protein cDNA Clones

The purified shed proteoglycan was treated with cyanogen bromide, and the fragments were fractionated by electrophoresis. Two prominent core peptides of 20 and 14 kD were sequenced as above (Table I). The NH_2-terminal amino acid sequences of the intact core protein, and of the two CNBr fragments of this core formed the basis for the synthesis of oligonucleotide primers (Table I) for use in a polymerase chain reaction amplification procedure. The sense primer corresponding to the NH_2 terminus of the native core protein in combination with the antisense primers corresponding to the NH_2 termini of the 20- and 14-kD fragments, yielded amplification products of 266 (Pl) and 863 bp (P3), respectively (Fig. 6). Consistently, the combination of the "20 K" sense and "14 K" antisense primers yielded a 635-bp product (P2). The combination of the "14 K" and "20 K" sense primers, used as a negative control, was nonproductive (not shown). Sequencing of the amplification products indicated that the amplified nucleotide sequences were extending the open reading frames of the sense primers, coding for amino acids that matched, and extended the results obtained by protein sequencing. This confirmed that Pl, P2, and P3 were the amplification of core protein-related cDNA struc-

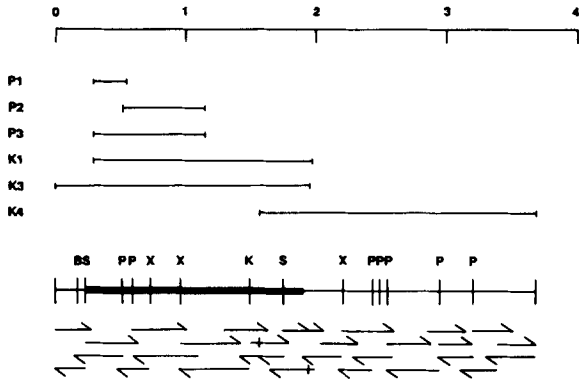

Figure 6. Alignment and restriction map of the polymerase chain reaction products and 64K cDNA clones. The strategy followed to sequence the 64K3 and 64K4 clones is illustrated by the arrows. The scale of the abscissa is in kbps. *B*, Bam HI; *S*, Sal I; *P*, Pst I; *X*, Xho I; *K*, Kpn I.

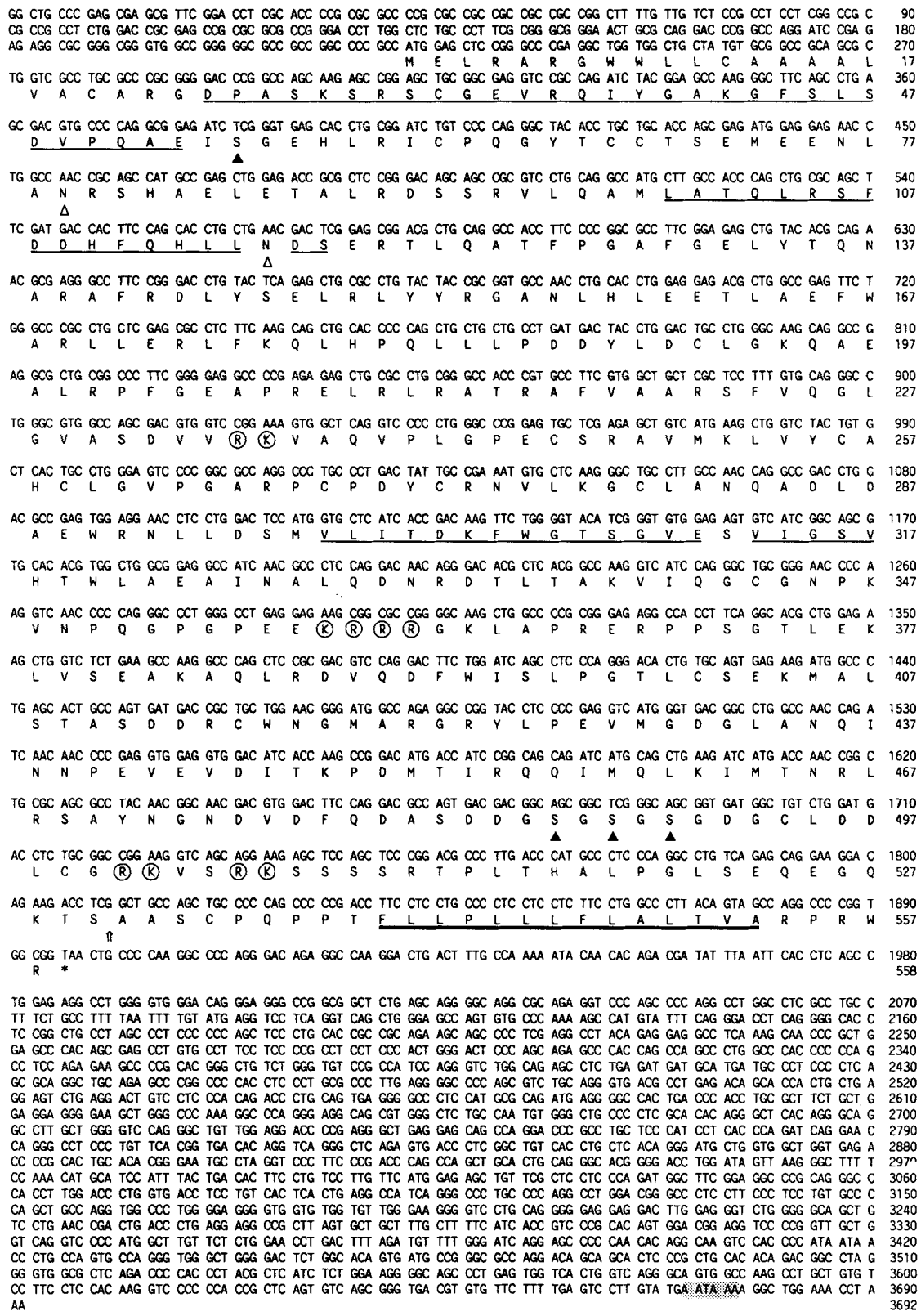

Figure 7. Merged sequences of the 64K3 and 64K4 cDNA clones and corresponding core protein sequence. The parts of the predicted protein sequence that match the sequence information obtained by direct analysis of the NH_2 termini of the mature protein, and of the two CNBr fragments are indicated by single underlining. The gaps in the lines correspond to blanc values obtained during protein sequencing. A stretch of hydrophobic amino acids near the COOH terminus of the protein is indicated by double underlining. A potential cleavage site in the protein serving for the attachment of a glycosyl-phosphatidylinositol tail is indicated by the arrow. Potential asparagine-linked glycosylation sites are indicated by open arrowheads; potential glycosaminoglycan attachment sites by full arrow heads. Residues that form dibasic or extended basic sequences have been encircled. The AATAAA polyadenylation signal is indicated by the shaded box.

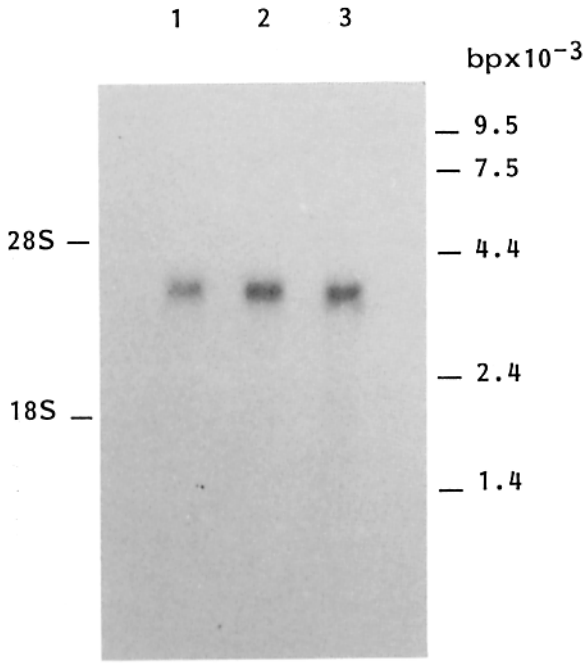

Figure 8. Northern blot analysis of fibroblast mRNA. Poly(A)$^+$ RNA from human lung fibroblasts was size fractionated in denaturing formaldehyde agarose gels, blotted to nylon membranes, and analyzed by hybridization to P3 (lane *1*), 64K1 (lane *2*), and a fragment of the 64K4 probe (obtained by digestion with PstI) extending from residue 2555 to residue 2938 of the 64K sequence given in Fig. 7 (lane *3*).

tures (not shown). P3 was then used to screen an unamplified human lung fibroblast cDNA library by DNA hybridization analysis. One positive clone, 64K1, was plaqued purified, subcloned, and completely sequenced. Clone 64K1 contained an insert of 1,687 bp, consisting of an open reading frame that encompassed and extended the open reading frames identified in the amplification products, but that lacked an initiation ATG codon (Fig. 6). Further screening of the library with clone 64K1 yielded five additional cross-hybridizing clones, which were characterized by restriction mapping and sequencing of their termini. All clones appeared to contain inserts corresponding to different parts of the same message. Clones 64K3 and 64K4 provided the largest amounts of additional 5′ and 3′ sequence information, respectively (Fig. 6). Both clones were completely sequenced. Finally, the sequence of clone 64K3 was confirmed to encode the 1G11 epitope by subcloning in pEX_2 and expression in POP cells as a β-galactosidase–core protein fusion product (not shown).

cDNA Sequence

The merged sequences of the sense strands of clones 64K3 and 64K4 are given in Fig. 7. The strategy used to sequence both strands of clones 64K3 and 64K4 is illustrated in Fig. 6. The sequence of clone 64K3 extended from residue 1 to residue 1947; that of clone 64K4 from residue 1565 to 3692. The sequence of clone 64K1 was in complete agreement, extending from residue 285 to 1972 of the merged 64K3–64K4 sequences. The first ATG codon (at position 222) in the sequence of 64K3 occurred in a context favorable for initiation of translation, featuring a G at positions −3 and +4 with respect to the translation start site (Kozak, 1989). It was followed by an open reading frame of 1,170 bases and a TAA stop codon at position 1896. Clone 64K4 ended with an AATAAA polyadenylation signal (at position 3671 of the merged sequence), followed 14 bases later by a sequence of 3A's but not by a typical poly(A) stretch. Northern blot analysis of the poly(A)$^+$ RNA from human lung fibroblasts using P3, 64K1, and a fragment of 64K4 (extending from the Pst I cleavage site at residue 2555 to the Pst I cleavage site at residue 2939 of the composite sequence) as hybridization probes all revealed a single message of ~3.8 kb (Fig. 8). The composite sequence of 3692 bases given in Fig. 7, therefore, seems to represent the structure of a near full-length cDNA.

Core Protein Structure

Starting at the first ATG codon, the core protein cDNA coded for a protein of 558 amino acids with a molecular weight of 61,649. The first 23 amino acids of the predicted sequence formed a typical signal peptide, which respected the −3,−1 rule (von Heijne, 1986) by featuring a C at posi-

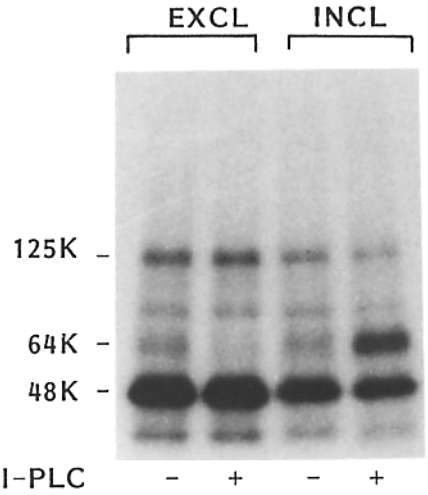

Figure 9. Effect of phosphatidylinositol-specific phospholipase C on the hydrophobicity of the membrane heparan sulfate proteoglycans. The ^{125}I-labeled hydrophobic heparan sulfate proteoglycans of human lung fibroblasts were treated with bacterial phosphatidylinositol-specific phospholipase C, or left untreated and mixed with phosphatidyl choline. The samples were dialyzed to remove the detergent and the liposome-proteoglycan mixtures were chromatographed over Sepharose CL4B in 4 M guanidinium chloride buffer without detergent. In the absence of phospholipase C treatment, the label eluted as a sharp slightly asymmetric peak, with 77% of the label present in the excluded fractions, and the remainder of the label trailing in the included fractions. In phospholipase C–treated samples, only 70% of the label eluted in the excluded fractions. The remainder of the label, eluting in the included fractions, formed a distinct shoulder in the tail of the first peak. For both types of samples the recovery of label was similar and greater than 85%. Equal amounts of label from excluded (hydrophobic) and included (trailing and nonhydrophobic) fractions, from phospholipase C–treated (+) and control (−) fractions were digested with heparitinase, submitted to SDS-PAGE, and analyzed by autoradiography. The numbers indicate the apparent molecular weights of the major core protein bands.

tion 20 and a G at position 23, to be followed by the stretch of amino acids identified as the NH_2 terminus of the mature core protein by protein sequencing. A second cluster of hydrophobic residues, followed by a brief polar sequence with three arginines, occurred at the COOH terminus of the protein. This hydrophobic sequence was interrupted by some polar residues, including a glutamine, rendering it unlikely that this part of the core would function as a membrane-spanning domain. Although rich in leucine (12.5% of the residues), the protein did not feature the type of leucine-rich repeats present in the proteoglycans biglycan (Fisher et al., 1989; Neame et al., 1989) and decorin (Krusius and Ruoslahti, 1986). The nascent core was predicted to contain 19 cysteines. Three of these are probably removed by further processing of the precursor; two occurring in the NH_2-terminal signal peptide, one in the signal peptidelike COOH-terminal hydrophobic sequence (see below). The cDNA coded for 6 ser-gly sequences, but only four of these occurred in a context which matched proposed consensus sequences serving for the attachment of glycosaminoglycan chains. The first of these was present near the NH_2 terminus of the protein, the three remainder occurring as a triplet repeat near the COOH terminus (Fig. 7). Finally, two asparagines occurred as NXS sequences and could therefore serve for the attachment of N-linked oligosaccharide. No extensive homologies to other known proteins included in the translated Genbank 61 and PIR 22 data bases were found.

Phospholipid Tailing of the Core Protein

The structure of the COOH terminus of the core protein, as predicted from the cDNA sequences, was atypical for proteins with a transmembrane orientation. We therefore investigated the possibility of a membrane insertion through a phosphatidylinositol anchor by treating the ^{125}I-iodinated membrane proteoglycan fraction with bacterial phosphatidylinositol-specific phospholipase C and comparing the ability of phospholipase-treated and untreated samples to intercalate into phospholipid vesicles. The enzyme treatment reduced the amount of lipid-intercalated label by only 10%, but SDS-PAGE analysis, after heparitinase treatment, indicated that the 64-kD core protein had been specifically depleted from the hydrophobic fractions, to be recovered, as a ~64-kD protein, in the trailing and nonhydrophobic fractions (Fig. 9). The distribution of the major 48-kD core protein in this preparation, for whom the cDNA sequence predicted a transmembrane orientation (Marynen et al., 1989), in contrast, was not affected by the treatment, consistent with the expectations and suggesting that nonspecific effects due to proteinases that might contaminate the phospholipase preparation could be excluded. Confirmatory experiments using affinity-purified materials indicated that up to 85% of the S1-reactive proteoglycans had lost their ability to intercalate into liposomes after the phospholipase C treatment. Finally, in double metabolic labeling experiments, both [^{3}H]ethanolamine and [^{35}S]sulfate, or [^{35}S]methionine could be incorporated in the hydrophobic S1-reactive proteoglycans (yielding ^{35}S/^{3}H ratios of 3.3 and 2.9 respectively). The S1-reactive proteoglycans from the media of these cultures, in contrast, had a 10–15-fold lower relative ^{3}H content than the hydrophobic proteoglycans (with ^{35}S/^{3}H ratios of 50.6 and 28, respectively). Taken together these data suggest that the hydrophobic proteoglycan is linked to the cell surface through a glycosyl-phosphatidylinositol anchor, and that this anchor and the bridging ethanolamine are missing from the shed forms.

Discussion

We report the isolation and sequencing of cDNA clones for a heparan sulfate proteoglycan from human lung fibroblasts. The predicted core protein structure and the effects of phospholipase C treatment suggest that this proteoglycan is linked to the cell surface via a glycosyl-phosphatidylinositol anchor. This proteoglycan is rapidly, quantitatively, and selectively shed from the cell surface by cultured fibroblasts.

A Novel Cell Surface Proteoglycan

The membrane-associated heparan sulfate proteoglycan from human lung fibroblasts that carries the S1 and 1G11 epitopes migrates as a protein of ~64 kD after heparitinase treatment, contains N-linked carbohydrate, and harbors at least one intrachain disulfide bond which is required for the expression of the S1 epitope (De Boeck et al., 1987), but not for the expression of the core protein epitope which is recognized by the antibody 1G11 (see Fig. 1). The selected cDNA clones can account for all these features. They code for a protein with a predicted molecular weight of 61,649 that contains several cysteines, that features potential acceptor sites for N-linked oligosaccharide and for glycosaminoglycan chain attachment, and that contains the 1G11 epitope when expressed as recombinant fusion protein. Moreover, the protein sequence predicted from the analysis of the cDNA clones includes the partial sequences obtained by direct analysis of the NH_2 termini of the cell-associated and shed forms of the native core proteins and of the CNBr fragments of these core proteins, providing direct and conclusive evidence for the authenticity of the cDNA clones. The primary structure of this core protein appears unique. It is distinct, more particularly, from the 48K heparan sulfate proteoglycan core protein previously identified in human lung fibroblasts (Marynen et al., 1989), providing final evidence for the structural heterogeneity and polymorphism of the cell surface proteoglycans of these cells (Lories et al., 1989). It is also distinct from syndecan, a cell surface proteoglycan from mouse (Saunders et al., 1989) and human (Mali et al., 1990) mammary epithelial cells, which may be indicative for additional heterogeneity amongst proteoglycans from different cell types. One striking difference with syndecan and the 48K core, is the absence of a characteristic transmembrane domain suggestive for an alternative mode of association with the cell surface.

Nature of the Membrane Anchor

The structure of the COOH terminus of the core protein and the abolishment of the hydrophobic properties of the core by treatment with phosphatidylinositol-specific phospholipase C suggest that this cell surface association is mediated by a glycosyl-phosphatidylinositol tail. Proteins that are membrane-bound by phosphatidylinositol appear to be synthesized as precursor proteins with short hydrophobic amino acid sequences at their COOH termini, and with insignificant or nonexistent "cytoplasmic" domains (Low, 1987; Fer-

guson and Williams, 1988). Both these features are present in this fibroblast core protein, where the last five residues (RPRWR) could function as a stop transfer sequence, but where the stretch of hydrophobic amino acids that precedes these charged residues appears too short to function as a transmembrane domain. Phospholipid tailing seems to involve the cleavage of the precursor protein at 10–12 residues from the start of the hydrophobic COOH-terminal sequence and the transfer of the NH_2-terminal moiety to a preassembled glycosyl-phosphatidylinositol anchor (Low, 1987). This process has been termed "glypiation" (Cross, 1987), and is presumably catalyzed by a transamidase. Since signal peptides functioning in membrane translocation or random hydrophobic sequences can substitute for these hydrophobic COOH-terminal ends (Caras and Weddell, 1989), the structural requirements for the anchor attachment presumably depend on the overall hydrophobicity of the domain rather than on the presence of specific sequences. Consistently, there seems to be no sequence conservation or clear consensus in these hydrophobic regions and at the processing sites, except that the two amino acids found immediately after the cleavage point include combinations of A, N, D, G, and S residues and that these are similar to the residues at the anchor attachment sites (Ferguson and Williams, 1988). At 10 residues before the start of a continuous run of hydrophobic amino acids the fibroblast core protein features the sequence SAA which could comply with this "rule." Cleavage at this site would remove the 28 most COOH-terminal amino acids from the precursor protein and, in concert with the removal of the NH_2-terminal leader sequence, reduce the molecular weight of the peptide core of the proteoglycan to 56,047. This value is in good agreement with previous estimates of ~56 kD, reported for samples treated with heparitinase and *N*-glycanase (De Boeck et al., 1987) or with trifluoromethanesulfonic acid (Lories et al., 1989). As a whole, these data are in agreement with evidence for the existence of phosphatidylinositol-linked heparan sulfate proteoglycan in rat liver (Ishihara et al., 1987), ovarian granulosa (Yanagashita and McQuillan, 1989), and Schwann cells (Carey and Evans, 1989). Further experiments should clarify the relationship of the 64K core protein to these proteoglycans.

Glycosylation of the Core Protein

Besides the removal of the NH_2- and COOH-terminal "signal" sequences, the posttranslational processing of the core protein also includes substitution with N-linked oligosaccharide and with heparan sulfate chains. The shift in apparent molecular weight from ~64 to ~56 kD after treatment with N-glycanase (De Boeck et al., 1987) suggests that, in most of the proteoglycans, both potential acceptor asparagines in the sequence are substituted with oligosaccharide. Evidence in support for the use of the second acceptor site was provided by the obtention of a blanc value, instead of an N, in the 17th cycle during the sequencing of the NH_2 terminus of the 20K CNB fragment (Table I). Cyanogen bromide cleavage of the fibroblast proteoglycan (Lories et al., 1989) and trypsin treatment (Fig. 5), on the other hand, yield a series of heparan sulfate–free core protein fragments up to a size of 52 kD, indicating that at least in some of the proteoglycans the heparan sulfate chains are implanted near the ends of the core proteins. This would be consistent with the use of the sequences EISGE near the NH_2-terminal end and DDGSGSGSG near the COOH-terminal end of the mature core protein as glycosaminoglycan attachment sequences. Accepting that both sequences are modified in the same proteins (see also below), the cDNA predicts that incomplete cleavage at the intervening methionine residues would yield a series of glycosaminoglycan-free CNBr peptides, of which the largest possible would have a molecular weight of 43,699. This size would be further increased to ~50,000 by the presence of the two N-linked oligosaccharides in this segment of the core protein. Since the yield of ~50-kD peptide was not enhanced in the CNBr fragmentation of heparitinase-treated proteoglycan samples (Lories et al., 1989) it would seem that few, if any, of the proteoglycans carry heparan sulfate on the SG sequences that occur near the middle of the core protein. This was supported, at least for one of these sequences, by a normal yield of serine in the 11th cycle of the sequence of the 14K CNBr fragment (Table I). Consistently, the NH_2-terminal EISGE sequence matches the consensus structure Acidic-Xaa-SG-Acidic proposed to function for the attachment of heparan sulfate in syndecan (Saunders et al., 1989) and of chondroitin sulfate in versican (Zimmerman and Ruoslahti, 1989). The COOH-terminal DDGSGSGSG sequence, on the other hand, closely resembles the DDYSGSGSG and DYGSGSGSG sequences found in rat serglycin (Bourdon et al., 1986) and the closely related mouse mastocytoma core protein (Kjellén et al., 1989) respectively, the sequence DNFSGSGTG found in syndecan (Saunders et al., 1989), and the sequence DYASASGSG found in the core protein of the "48K" human fibroblast proteoglycan (Marynen et al., 1989). The two SG sequences that occur near the middle of the core, in contrast, do not resemble previously identified or proposed acceptor sequences. Since the molecular weight of the native SI-reactive proteoglycans is larger than 200,000 (Lories et al., 1989) and that of their isolated heparan sulfate chains close to 40,000 (Lories et al., 1986), the core proteins of these proteoglycans must be substituted with several glycosaminoglycan chains, indicating that, at the minimum, the COOH-terminal serglycin-like sequence, which itself may accommodate up to three chains, must be involved in heparan sulfate biosynthesis. The serglycin-like proteins are substituted with chondroitin sulfate in L2 rat yolk sac tumor cells (Oldberg et al., 1981), but carry heparin chains in mouse mastocytoma cells (Lindahl et al., 1986), and both types of chains in rat basophilic leukemia cells (Seldin et al., 1985). The related sequences in syndecan have been proposed to be substituted with chondroitin sulfate (Saunders et al., 1989), but for the human 64K and 48K core proteins there is only evidence for substitution with heparan sulfate (Fig. 4 *A*). All these data suggest that there is some flexibility in the primary structural contexts that allow the implantation of heparan sulfate chains. They also imply that signals which dictate the extension of the xylose-gal-gal linkage sequence by heparosan or chondroitin sequences may reside in elements of secondary and tertiary structure which may even vary for a single protein in different cell types or under different conditions.

Shedding of the Proteoglycan

In confluent cultured human lung fibroblasts the phospholipid-tailed proteoglycan is rapidly shed into the culture medium. It is the only cell surface proteoglycan which appears

to undergo this fate in these cells, and the shedding process only removes a small hydrophobic domain from the core. Cleavage of the phospholipid membrane anchor as a mechanism of proteoglycan shedding could account for the selectivity of the process if the 64K core would be the only lipid-tailed species, and would only create a small difference in apparent molecular weight between shed and membrane-bound forms. Yet, the small (≤3 kD) but noticeable difference in molecular weight between membrane and shed forms of the 64K core (Fig. 4 *B*) seems too large to only represent the removal of diacylglycerol or diacylglycerol phosphate (compare Fig. 4 with Fig. 9). Moreover, the experiments in which hydrophobic and shed proteoglycans from [^{3}H]ethanolamine-labeled cells are compared to each other are also in support of COOH-terminal proteolytic processing of the core. Syndecan, for which the cDNA sequence predicts the existence of a transmembrane segment, appears also to be shed to the culture medium of subconfluent cells (Jalkanen et al., 1987). Whereas the possibility remains that, in analogy with other membrane components, glycolipid-tailed syndecan variants may occur that would account for its shedding, this finding implies that a proteinase activity may be at the basis of the shedding of syndecan, and by extension, of other proteoglycans. The sequence of syndecan features a dibasic sequence close to the membrane insertion domain which may function as a substrate for trypsinlike enzymes (Saunders et al., 1989) and several dibasic sequences are also present in the 64K core protein, some close to the presumptive COOH-terminal end of the tailed core (Fig. 7). Yet, the evidence in favor of a proteolytic step does not rule out that a direct relationship may exist between the shedding process and the presence of the lipid anchor. For example, additional modification may have occurred in the COOH terminus after the removal of the lipid anchor as a primary step in the release process. Alternatively, the phospholipid anchor may direct the proteoglycan to a unique proteolytic compartment, from where it is released to the pericellular space. Recent observations suggest that the glycosylphospholipid-anchored folate receptor is recycling during folate internalization, and that it moves in and out of the cell using an uncoated pit pathway that does not merge with the clathrin-coated pit endocytic machinery (Rothberg et al., 1990). Interestingly, the folate receptor is also shed in the medium of cultured cells (Lacey et al., 1989). This raises the possibility that glypiated proteins are turned over by shedding after intracellular proteolysis in alternative endocytic compartments and recycling to the cell surface. A phospholipase D activity specific for the phosphatidyl anchor of cell surface proteins has, however, been identified in a variety of tissues, and is abundant in mammalian plasma (Low and Prasad, 1988) indicating that at least under certain conditions shedding through cleavage of the anchor may occur. It has been suggested that the phospholipid anchor of the rat hepatocyte cell surface proteoglycan, is cleaved by an insulin-activated phospholipase, and that the released proteoglycan is then reinternalized via a cell surface receptor that recognizes the exposed myoinositol phosphate moiety on the core protein (Ishihara et al., 1987).

In human lung fibroblasts the released proteoglycan appears to be metabolically stable, suggesting that, if cleavage of the anchor occurs as assumed, the mechanism of endocytosis is not operative in these cells. The observations on the rat liver heparan sulfate proteoglycan, and the effects of serum and insulin in the glycosyl-phosphatidylinositol-anchored proteins of myocytic and adipocytic cell lines (Lisanti et al., 1989), however indicate that the cell surface expression of these components may be acutely regulated by external stimuli. Further experiments should indicate whether in the case of the heparan sulfate proteoglycans this may, for example, be a way to reverse cell-adhesive processes, to set the level of pericellular proteolysis, or to cause the release of modulators of cell growth, as an alternative to or in concert with mechanisms that are based on proteolytic processing. In the meantime we propose to refer to this glypiated proteoglycan by the name "glypican."

We thank H. Ceulemans, Ch. Coomans, A. Rayé, and M. Willems for their expert technical assistance. We are very grateful to Dr. J. Van Damme for access to protein sequencing facilities.

Guido David is an Onderzoeksleider of the Nationaal Fonds voor Wetenschappelijk Onderzoek, Belgium. Peter Marynen is a Bevoegdverklaard Navorser of the Nationaal Fonds voor Wetenschappelijk Onderzoek, Belgium. This investigation was supported by grant 3.0066.87 of the Fonds voor Geneeskundig Wetenschappelijk Onderzoek, Belgium; by the Interuniversity Network for Fundamental Research sponsored by the Belgian government (1987–1991) and by a grant Geconcerteerde Acties from the Belgian government.

Received for publication 8 June 1990 and in revised form 22 August 1990.

References

Bourdon, M. A., M. Shiga, and E. Ruoslahti. 1986. Identification from cDNA of the precursor form of a chondroitin sulfate proteoglycan core protein. *J. Biol. Chem.* 261:12534–12537.

Bretscher, M. S. 1985. Heparan sulphate proteoglycans and their polypeptide chains from BHK cells. *EMBO (Eur. Mol. Biol. Organ.) J.* 4:1941–1944.

Caras, I. W., and G. N. Weddel. 1989. Signal peptide for protein secretion directing glycophospholipid membrane anchor attachment. *Science (Wash. DC).* 243:1196–1198.

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.

Chajek-Shaul, T., O. Halimi, M. Ben-Naim, O. Stein, and Y. Stein. 1989. Phosphatidylinositol-specific phospholipase C releases lipoprotein lipase from the heparin releasable pool in rat heart cell cultures. *Biochim. Biophys. Acta.* 1014:178–183.

Cheifetz, S., J. L. Andres, and J. Massagué. 1988. The transforming growth factor-β receptor type III is a membrane proteoglycan: domain structure of the receptor. *J. Biol. Chem.* 263:16984–16991.

Church, G. M., and W. Gilbert. 1984. Genomic sequencing. *Proc. Natl. Acad. Sci. USA.* 81:1991–1995.

Cisar, L. A., A. J. Hoogewerf, M. Cupp, C. A. Rapport, and A. Bensadoun. 1989. Secretion and degradation of lipoprotein lipase in cultured adipocytes: binding of lipoprotein lipase to membrane heparan sulfate proteoglycans is necessary for degradation. *J. Biol. Chem.* 264:1767–1774.

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.

Cross, G. A. M. 1987. Eukaryotic protein modification and membrane attachment via phosphatidylinositol. *Cell.* 48:179–181.

David, G., and M. Bernfield. 1981. Type I collagen reduces the degradation of basal lamina proteoglycan by mammary epithelial cells. *J. Cell Biol.* 91:281–286.

David, G., and H. Van den Berghe. 1983. Transformed mouse mammary epithelial cells synthesize undersulfated basement membrane proteoglycan. *J. Biol. Chem.* 258:7338–7344.

David, G., and H. Van den Berghe. 1989. Cell-surface heparan sulfate and heparan-sulfate/chondroitin-sulfate hybrid proteoglycans of mouse mammary epithelial cells. *Eur. J. Biochem.* 178:609–617.

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.

Feinberg, A. P., and B. Vogelstein. 1983. A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. *Anal. Biochem.* 132:6–13.

Ferguson, M. A. J., and A. F. Williams. 1988. Cell-surface anchoring of proteins via glycosyl-phosphatidylinositol structures. *Annu. Rev. Biochem.* 57:285–320.

Fisher, L. W., J. D. Termine, and M. F. Young. 1989. Deduced protein sequence of bone small proteoglycan I (biglycan) shows homology with proteoglycan II (decorin) and several nonconnective tissue proteins in a variety of species. *J. Biol. Chem.* 264:4571–4576.

Gallagher, J. T. 1989. The extended family of proteoglycans: social residents of the pericellular zone. *Curr. Opin. Cell Biol.* 1:1201–1218.

Heremans, A., J. J. Cassiman, H. Van den Berghe, and G. David. 1988. Heparan sulfate proteoglycan from the extracellular matrix of human lung fibroblasts. *J. Biol. Chem.* 263:4731–4739.

Höök, M., L. Kjellén, S. Johansson, and J. Robinson. 1984. Cell-surface glycosaminoglycans. *Annu. Rev. Biochem.* 53:847–869.

Iozzo, R. V. 1984. Biosynthesis of heparan sulfate proteoglycan by human colon carcinoma cells and its localization at the cell surface. *J. Cell Biol.* 99:403–417.

Ishihara, M., N. S. Fedarko, and H. E. Conrad. 1987. Involvement of phosphatidylinositol and insulin in the coordinate regulation of proteoheparan sulfate metabolism and hepatocyte growth. *J. Biol. Chem.* 262:4708–4716.

Jalkanen, M., A. Rapraeger, S. Saunders, and M. Bernfield. 1987. Cell surface proteoglycan of mouse mammary epithelial cells is shed by cleavage of its matrix-binding ectodomain from its membrane-associated domain. *J. Cell Biol.* 105:3087–3096.

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 protein. *Biochem. J.* 263:105–113.

Koda, J. E., A. Rapraeger, and M. Bernfield. 1985. Heparan sulfate proteoglycans from mouse mammary epithelial cells: cell surface proteoglycan as a receptor for interstitial collagens. *J. Biol. Chem.* 260:8157–8162.

Kogan, S. C., M. Doherty, and J. Gitschier. 1987. An improved method for prenatal diagnosis of genetic diseases by analysis of amplified DNA sequences. *N. Engl. J. Med.* 317:985–1025.

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

Krusius, T., and E. Ruoslahti. 1986. Primary structure of an extracellular matrix proteoglycan core protein deduced from cloned cDNA. *Proc. Natl. Acad. Sci. USA.* 83:7683–7687.

Lacey, S. W., J. M. Sanders, K. G. Rothberg, R. G. W. Anderson, and B. A. Kamen. 1989. Complementary DNA for the folate binding protein correctly predicts anchoring to the membrane by glycosyl-phosphatidylinositol. *J. Clin. Invest.* 84:715–720.

LeBaron, R. G., A. Höök, J. D. Esko, S. Gay, and M. Höök. 1989. Binding of heparan sulfate to type V collagen: a mechanism of cell-substrate adhesion. *J. Biol. Chem.* 264:7950–7956.

Lindahl, U., D. S. Feingold, and L. Rodén. 1986. Biosynthesis of heparin. *Trends Biochem. Sci.* 11:221–225.

Lisanti, M. P., J. C. Darnell, B. L. Chan, E. Rodriguez-Boulan, and A. R. Saltiel. 1989. The distribution of glycosyl-phosphatidylinositol anchored proteins is differentially regulated by serum and insulin. *Biochem. Biophys. Res. Commun.* 164:824–832.

Lories, V., G. David, J. J. Cassiman, and H. Van den Berghe. 1986. Heparan sulfate proteoglycans of human lung fibroblasts: occurrence of distinct membrane, matrix and secreted forms. *Eur. J. Biochem.* 158:351–360.

Lories, V., H. De Boeck, G. David, J. J. Cassiman, and H. Van den Berghe. 1987. Heparan sulfate proteoglycans of human lung fibroblasts: structural heterogeneity of the core proteins of the hydrophobic cell-associated forms. *J. Biol. Chem.* 262:854–859.

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.

Low, M. G. 1987. Biochemistry of the glycosyl-phosphatidylinositol membrane protein anchors. *Biochem. J.* 244:1–13.

Low, M. G., and A. R. S. Prasad. 1988. A phospholipase D specific for the phosphatidylinositol anchor of cell-surface proteins is abundant in plasma. *Proc. Natl. Acad. Sci. USA.* 85:980–984.

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.

Maniatis, T., E. F. Fritsch, and J. Sambrook. 1982. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY. 545 pp.

Marcum, J. A., and R. D. Rosenberg. 1987. Anticoagulantly active heparan sulfate proteoglycan and the vascular endothelium. *Sem. Thromb. Hemostosis.* 13:464–474.

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.

Moos, M., N. Y. Nguyen, and T.-Y. Liu. 1988. Reproducible high yield sequencing of proteins electrophoretically separated and transferred to an inert support. *J. Biol. Chem.* 263:6005–6008.

Neame, P. J., H. U. Choi, and L. C. Rosenberg. 1989. The primary structure of the core protein of the small, leucine-rich proteoglycan (PG I) from bovine articular cartilage. *J. Biol. Chem.* 264:8653–8661.

Oldberg, Å., E. G. Hayman, and E. Ruoslahti. 1981. Isolation of a chondroitin sulfate proteoglycan from a rat yolk sac tumor and immunochemical demonstration of its cell surface localization. *J. Biol. Chem.* 256:10847–10852.

Rothberg, K. G., Y. Ying, J. F. Kolhouse, B. A. Kamen, and R. G. W. Anderson. 1990. The glyco-phospholipid-linked folate receptor internalizes folate without entering the clathrin-coated pit endocytic pathway. *J. Cell Biol.* 110:637–649.

Ruoslahti, E. 1988. Structure and biology of proteoglycans. *Annu. Rev. Cell Biol.* 4:229–255.

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.

Sanger, F., S. Nicklem, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA.* 74:5463–5467.

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.* 18:1547–1556.

Schmidtchen, A., I. Carlstedt, A. Malmström, and L.-Å. Fransson. 1990. Inventory of human skin fibroblast proteoglycans: identification of multiple heparan and chondroitin/dermatan sulphate proteoglycans. *Biochem. J.* 265:289–300.

Segarini, P. R., and S. M. Seyedin. 1988. The high molecular weight receptor to transforming growth factor-β contains glycosaminoglycan chains. *J. Biol. Chem.* 263:8366–8370.

Seldin, D. C., K. F. Austen, and R. L. Stevens. 1985. Purification and characterization of protease-resistant secretory granule proteoglycans containing chondroitin sulfate di-B and heparin-like glycosaminoglycans from rat basophilic leukemia cells. *J. Biol. Chem.* 260:11131–11139.

Sun, X., D. F. Mosher, and A. Rapraeger. 1989. Heparan sulfate-mediated binding of epithelial cell surface proteoglycan to thrombospondin. *J. Biol. Chem.* 264:2885–2889.

von Heijne, G. 1986. A new method for predicting signal sequence cleavage sites. *Nucleic Acids Res.* 14:4683–4690.

Yanagashita, M., and V. C. Hascall. 1984. Metabolism of proteoglycans in rat ovarian granulosa cell culture. *J. Biol. Chem.* 259:10260–10269.

Yanagashita, M., and D. J. McQuillan. 1989. Two forms of plasma membrane-intercalated heparan sulfate proteoglycans in rat ovarian granulosa cells. *J. Biol. Chem.* 264:17551–17558.

Zimmerman, D. R., and E. Ruoslahti. 1989. Multiple domains of the large fibroblast proteoglycan, versican. *EMBO (Eur. Mol. Biol. Organ.) J.* 8:2975–2981.