Domain structure of proteoglycan sulphate from confluent cultures of human embryonic skin fibroblasts

Lars-Åke FRANSSON, Lars CÖSTER, Ingemar CARLSTEDT and Anders MALMSTRÖM

Department of Physiological Chemistry 2, University of Lund, P.O. Box 94, S-221 00 Lund, Sweden

Radiolabelled proteoglycan sulphates were isolated from confluent monolayers of fibroblasts and from their spent media. The cell-surface-associated proteoglycan (M_r 350 000) has a core protein of M_r 180 000 that is cleaved by reduction of disulphide bonds into polypeptides of M_r 90 000, both of which can bind transferrin [Fransson, Carlstedt, Cöster & Malmström (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 5657–5661]. Thrombin digestion of the proteoglycan yielded two major fragments. The larger one contained the heparan sulphate chains and glycoprotein-type oligosaccharides, whereas the smaller one contained interchain disulphide bond(s) and had affinity for transferrin as well as for octyl-Sepharose. The larger thrombin fragment was cleaved by trypsin into fragments containing the heparan sulphate chains and the oligosaccharides respectively. The smaller proteoglycan sulphate derived from the culture medium (M_r 150 000) had a core protein of M_r 30 000, which contained heparan sulphate-attachment and oligosaccharide-attachment regions, but no domains for binding of transferrin or for hydrophobic interactions.

INTRODUCTION

Most adherent eukaryotic cells synthesize heparan sulphate-containing proteoglycans that are associated with the cell surface (for references see Carlstedt *et al.*, 1983). We have previously isolated and characterized radiolabelled proteoglycan sulphates from human skin fibroblasts in confluent cultures (Carlstedt *et al.*, 1983). The proteoglycan obtained from the monolayer has M_r 350 000, whereas that of the medium is significantly smaller (M_r 150 000). Cleavage of disulphide bonds in the cell-associated proteoglycan gave rise to a product with M_r 140 000 (referred to below as the 'proteoglycan subunit'), whereas the size of the medium-derived form was not changed. Heparan sulphate lyase treatment of the cell-associated proteoglycan yielded a core protein of M_r 180 000 (Cöster *et al.*, 1983), which, on reduction, afforded polypeptides of M_r 90 000. The intact proteoglycan as well as its core protein have affinity for transferrin (Fransson *et al.*, 1984). By using digestion with thrombin two major types of fragments have been isolated and characterized. A model for a domain structure of the cell-surface-associated proteoglycan sulphate is proposed.

EXPERIMENTAL

Materials

Human embryonic skin fibroblasts in confluent cultures were incubated with [^{3} H]glucosamine and $^{35}\text{SO}_4^{2-}$ or with [^{3} H]leucine and $^{35}\text{SO}_4^{2-}$ (Carlstedt *et al.*, 1981) for 72 h, and proteoglycan were isolated from the medium and from an extract of the cell layer (4 M-guanidinium chloride plus proteinase inhibitors). Proteoglycan sulphate was purified by using isopycnic density-gradient centrifugation in $\text{CsCl}/4$ M-guanidinium chloride followed by gel-permeation and ion-exchange chromatography after degradation of contaminating proteoglycan sulphates with chondroitin ABC lyase (Carlstedt *et al.*, 1983).

Heparan sulphate lyase (EC 4.2.2.8) from *Flavobacterium heparinum* was a product of Seikagaku Kogyo Co., Tokyo, Japan. Trypsin (EC 3.4.21.4) was purchased from Sigma Chemical Co., and bovine α -thrombin (EC 3.4.21.5) was a gift from Dr. Johan Stenflo, Malmö General Hospital, Malmö, Sweden. Sepharose gels, octyl-Sepharose, protein A-Sepharose and proteins for the calibration of gel columns were purchased from Pharmacia. Other chemicals or reagents, including transferrin and anti-transferrin, were obtained from sources listed previously (Carlstedt *et al.*, 1981, 1983; Fransson *et al.*, 1984).

Methods

Digestion with heparan sulphate lyase (6 munits/ml) was performed in 3 mM-calcium acetate buffer, pH 7.0, containing 0.5 mg of ovomucoid (egg trypsin inhibitor; Sigma Chemical Co.)/ml at 37 °C for 4 h in tubes that had been pretreated with a solution of ovalbumin (0.5 mg/ml). (One unit of enzyme releases 1 μmol of unsaturated hexuronate/min.)

Trypsin or thrombin digestions of proteoglycan sulphates were carried out in 1 ml of 137 mM-NaCl/3 mM-sodium phosphate buffer, pH 7.4, containing approx. 0.1 mg of enzyme at 37 °C for 24 h. The thrombin digests of the proteoglycan were first chromatographed on columns (1 cm $\times$ 3 cm) of DE-52 DEAE-cellulose (Whatman), which were eluted with 6 ml of 6 M-urea/0.2 M-sodium acetate buffer, pH 5.8 ('neutral' material), followed by 4 M-guanidinium chloride in the same buffer supplemented with 5 μg of ovalbumin/ml (anionic material).

Reduction of disulphide bonds was conducted in 4 M-guanidinium chloride/10 mM-dithiothreitol/50 mM-sodium acetate buffer, pH 5.8, at 37 °C for 5 h. Alkylation was subsequently achieved by the addition of a 2.5-fold molar excess of iodoacetamide over dithiothreitol and by keeping the solution in the dark overnight. Polyacrylamide-gel electrophoresis was performed in sodium

Fig. 1. Gel chromatography on Sepharose CL-4B of cell-associated proteoglycan sulphate (a), anionic material from a thrombin digest of the proteoglycan (b), neutral material from the same thrombin digest (c), the pooled anionic material in (b) after trypsin digestion (d), the pooled anionic material in (b) after reduction and alkylation (e) and the pooled material in (c) after reduction and alkylation (f).

The profiles show the superimposed gel chromatograms from two different experiments, one obtained with [^3H]leucine- ($\circ$) and $^{35}\text{SO}_4^{2-}$ -labelled proteoglycan ($\triangle$), the other with [^3H]glucosamine- ($\bullet$) and $^{35}\text{SO}_4^{2-}$ -labelled proteoglycan. The column (0.6 cm $\times$ 150 cm) was eluted with 4 M-guanidinium chloride/1% (v/v) Triton X-100/50 mM-sodium acetate buffer, pH 5.8, containing ovalbumin (5 $\mu\text{g}/\text{ml}$). The thrombin digests were separated into 'neutral' and anionic material as described in the Experimental section. The two fractions were concentrated by pressure dialysis (PM-10 filter) and chromatographed on Sepharose CL-4B (b and c respectively). The pooled fractions in (b) (see bar) were concentrated, dialysed against the appropriate buffer and digested with trypsin (d) or reduced and alkylated (e). The pooled fractions in (c) (see bar) were concentrated and subjected to reduction of disulphide bonds (f). Recovery was 70% or better in all cases.

dodecyl sulphate on 6% or 11% gels ($T = 6\%$ or 11% ; $C = 2.5\%$) as described previously (Carlstedt *et al.*, 1981).

To test for transferrin binding, samples were treated with human transferrin followed by rabbit immunoglobulins against human transferrin as described previously (Fransson *et al.*, 1984). The suspensions were applied to columns of protein A-Sepharose, and bound material was displaced with 4 M-guanidinium chloride/1% (v/v) Triton X-100.

Gel-permeation, ion-exchange and hydrophobic chromatography were performed as described elsewhere (Carlstedt *et al.*, 1981, 1983) or in the legends to the appropriate Figures. Radioactivity was measured in a Packard 2650 scintillation counter with automatic quench correction.

RESULTS

After incorporation of radioactive precursors cells were extracted with 4 M-guanidinium chloride at pH 5.8 and in the presence of *N*-ethylmaleimide to minimize the potential for disulphide-thiol exchange between proteo-

glycans and extraneous proteins. The proteoglycan was quantitatively bound to DEAE-cellulose in the presence of 6 M-urea/10 mM-dithiothreitol/0.2 M-sodium acetate buffer, pH 5.8 (measured as ^3H as well as ^{35}S radioactivity; results not shown), indicating that no extraneous radiolabelled proteins had been covalently associated with the proteoglycans. Furthermore, to test whether all radioactivity was part of the covalent structure the undegraded proteoglycan was chromatographed on Sepharose CL-4B both in 4 M-guanidinium chloride/1% (v/v) Triton X-100 (Fig. 1a) and in 1% (w/v) sodium dodecyl sulphate. No ^3H radioactivity was found in the more-retarded fractions, suggesting the absence of non-proteoglycan proteins (results not shown).

The cell-associated proteoglycan sulphate was isolated after incorporation of ^3H from either [^3H]leucine or from [^3H]glucosamine, both in combination with $^{35}\text{SO}_4^{2-}$. In the former case the core protein and the heparan sulphate side chains are labelled, whereas in the second case heparan sulphate and glycoprotein-type oligosaccharides will contain the labelled precursors. The two preparations

Fig. 2. Hydrophobic chromatography on octyl-Sepharose of untreated (a), heparan sulphate lyase-treated (b) and reduced and alkylated cell-associated proteoheparan sulphate (c) as well as intact medium-derived proteoheparan sulphate (d)

[^3H]Leucine- and $^{35}\text{SO}_4^{2-}$ -labelled proteoglycans were isolated and treated as described in the Experimental section. The column (0.6 cm $\times$ 10 cm) was eluted at room temperature with 4 M-guanidinium chloride/50 mM-sodium acetate buffer, pH 5.8, followed (arrow) by the same solvent containing 1% (v/v) Triton X-100. Fractions were analysed for ^3H radioactivity (●) and ^{35}S radioactivity (○). Recovery was 80% or better.

were separately degraded with thrombin followed by separation into 'neutral' and anionic material by using ion-exchange chromatography. After thrombin digestion of [^3H]leucine- and $^{35}\text{SO}_4^{2-}$ -labelled proteoglycan 62% of the total ^3H was recovered with the 'neutral' material

whereas all ^{35}S was recovered with the anionic material. The latter material chromatographed on Sepharose CL-4B with a K_{av} of 0.3–0.4 (Fig. 1b). The 'neutral' material consisted of [^3H]glucosamine largely of K_{av} 0.5–0.7 (Fig. 1c). When [^3H]glucosamine was used as one of the precursors 92% of the ^3H and all ^{35}S was retained on the ion-exchanger after treatment with thrombin. This material also chromatographed with a K_{av} of approx. 0.3 (Fig. 1b, ● symbols). Subsequent trypsin digestion of the [^3H]glucosamine- and $^{35}\text{SO}_4^{2-}$ -labelled large thrombin fragment (see bar in Fig. 1b) yielded two components (Fig. 1d), a larger heparan sulphate-containing fragment and glycopeptides. The two major thrombin-generated fragments in Figs. 1(b) and 1(c) (see bars) were also treated with dithiothreitol, alkylated and re-chromatographed. The heparan sulphate- and oligosaccharide-containing fragment (Fig. 1b) was not affected (Fig. 1e), whereas the [^3H]leucine-labelled polypeptides (Fig. 1c) were much retarded (Fig. 1f). Thus the latter fragments might contain the disulphide bond(s) that join(s) the two 'proteoglycan subunits'.

The medium-derived form of proteoheparan sulphate had a K_{av} of 0.4 on Sepharose CL-4B (results not shown), corresponding to M_r approx. 150000. The core protein of this proteoglycan was obtained by treating the macromolecule with heparan sulphate lyase. Polyacrylamide-gel electrophoresis in sodium dodecyl sulphate (results not shown) gave a major component with M_r 30000 in reducing as well as non-reducing conditions. The medium-derived form of proteoheparan sulphate did not change its elution position on digestion with thrombin, and no low- M_r [^3H]leucine-labelled material was released (results not shown).

Chromatography on octyl-Sepharose was used to examine the presence of segments in the core proteins that are prone to hydrophobic interactions. Approx. 85% of the cell-associated proteoglycan population was bound to this ligand in the presence of 4 M-guanidinium chloride (Fig. 2a). The core protein obtained after treatment with heparan sulphate lyase was bound to the same extent (Fig. 2b). Since reduction and alkylation of the cell-associated proteoglycan did not affect binding (Fig. 2c), it is suggested that both polypeptides contain hydrophobic regions. In contrast, the medium-derived proteoglycan showed little or no binding to the affinity gel (Fig. 2d), nor did heparan sulphate chains obtained after proteolysis (results not shown).

As shown previously, the cell-associated proteoheparan sulphate is able to bind to transferrin (Fransson *et al.*, 1984). The two major thrombin-generated fragments were tested for their ability to bind to transferrin as well as to octyl-Sepharose. All of the [^3H]leucine-labelled material in Fig. 1(c) was retained on octyl-Sepharose, and 85% was bound to protein A in the presence of transferrin and antitransferrin, showing that these properties are unaffected by thrombin digestion. The heparan sulphate-containing thrombin fragment in Fig. 1(b) did not bind transferrin (results not shown).

DISCUSSION

The cell-associated proteoheparan sulphate (M_r 350000) appears to yield two subunits of M_r approx. 150000 on reduction of disulphide bonds (Carlstedt *et al.*, 1983). Removal of the heparan sulphate side chains from the cell-associated proteoglycan by enzymic degradation

Fig. 3. Proposed structure for proteoheparan sulphate

(a) The various domains or features are: I, 'cluster' of heparan sulphate chains; II, polypeptide with 'cluster' of oligosaccharides; III, interchain disulphide bond joining the proteoglycan 'subunits'; IV, transferrin-binding site; V, hydrophobic peptide region. (b) Proposed domain structure of proteoheparan sulphate from the cell layer.

affords a core protein of M_r 180 000 before and 90 000 after reduction (Cöster *et al.*, 1983). The 'proteoglycan subunits' are retained on DEAE-cellulose, heparan sulphate affinity gels (Fransson *et al.*, 1983) or octyl-Sepharose, and bind to transferrin (Fransson *et al.*, 1984). Thus both polypeptides are characterized by the presence of at least five distinct functional domains (Fig. 3a). Previous studies have suggested that the heparan sulphate side chains (four to six per polypeptide) as well as the glycoprotein-type oligosaccharides are clustered in separate domains (I and II in Fig. 3a) of the polypeptide (Carlstedt *et al.*, 1983). The order of arrangement of the heparan sulphate-attachment region (I), the oligosaccharides (II), the disulphide bond(s) (III), the transferrin-binding domain (IV) and the hydrophobic (and presumably [^{3} H]leucine-rich) peptide region (V) can only be assessed after various specific degradations of the core protein. Treatment of the proteoglycan with thrombin generated two major fragments, one of which (Fig. 1b) contained both heparan sulphate and oligosaccharides. The other (Fig. 1c) contained disulphide bond(s), a site for transferrin binding and hydrophobic region(s). The former fragment was further degraded by trypsin into separate heparan sulphate-carrying and oligosaccharide-carrying polypeptides.

From the above results we propose that proteoheparan sulphate from the cell surface of confluent fibroblasts has

the overall structure depicted in Fig. 3(b). It consists of two subunits of similar or identical size that are linked by disulphide bonds. Each polypeptide contains a cluster of heparan sulphate chains and oligosaccharide-substituted regions. Some of the latter may be located close to the heparan sulphate chains. In addition, there are domains that bind transferrin and that contain hydrophobic sequences, both in the vicinity of interchain disulphide bond(s). The precise order in which domains I-V are arranged cannot be determined from the present results. The glycosylated domains are presumably located to one end of the molecule and the hydrophobic portion to the other. Some, maybe all, of the oligosaccharides could alternatively be located between the end of the polypeptide and the heparan sulphate-attachment region. The proposed structure for the cell-surface-associated proteoheparan sulphate (Fig. 3b) implies that the molecule is cleaved by thrombin somewhere between the glycosylated region and the putative transferrin-binding site. Additional cleavage sites, giving rise to small fragments, are likely to be present, but the major cleavage products would be the two glycosylated fragments (estimated M_r approx. 150 000; see Fig. 1b) and the non-glycosylated polypeptide pool (estimated M_r 50 000–75 000; see Fig. 1c). Moreover, these fragments appear to account for the major part of the intact proteoglycan (M_r 350 000).

As suggested previously (Fransson *et al.*, 1984), the

core polypeptides of the proteoglycan and the transferrin receptor could be structurally related or even identical. Further, the amino acid sequence of the transferrin receptor has been deduced from the nucleotide sequence of the coding region of its mRNA (Schneider *et al.*, 1984). The sequence comprises a membrane-spanning segment 62 amino acid residues from the *N*-terminus, and eight potential sites for disulphide bridges. Those at positions 89 and 98 could be joining the two polypeptides. Three asparagine residues that could carry *N*-linked oligosaccharides are present at positions 251, 317 and 727. The latter is only 33 amino acid residues from the *C*-terminus. The positions of the transferrin-binding site, thrombin-cleavage sites and possible *O*-linked oligosaccharides are not known. Inspection of the sequence suggests that thrombin-sensitive sites could be present between a putative interchain disulphide bond at position 98 and the first *N*-linked oligosaccharide (at 251). As the extent of glycosylation and other steric factors will restrict the action of the enzyme, it is impossible to predict exactly where the cleavage sites would be. More interestingly, Ser-Gly or Gly-Ser combinations are present at six positions in the *C*-terminal half of the polypeptide. Two of those are present as Gly-Ser-Gly-Ser, a combination that appears to be common within the glycosaminoglycan-attachment regions of proteoglycans (Bourdon *et al.*, 1985). These six serine residues could be linkage sites for heparan sulphate side chains if the transferrin-receptor polypeptide is identical with the core polypeptide of proteoglycan sulphate from confluent fibroblasts. So far, there is no information on the structure of the transferrin receptor that rules out the possibility of a close relationship with proteoglycan sulphate.

The medium-derived proteoglycan sulphate has a core protein of M_r 30 000, which contains a cluster of four to six heparan sulphate chains as well as oligosaccharides (Carlstedt *et al.*, 1983). However, it does not bind transferrin (Fransson *et al.*, 1984), appears to lack interchain disulphide bonds and has no affinity for octyl-Sepharose (Fig. 2*d*). This is consistent with a structure that incorporates only features I and II in Fig. 3(*a*), i.e. similar to the large fragment generated by thrombin digestion of the cell-associated proteoglycan. The medium-derived proteoglycan sulphate may thus have been produced from the cell-associated one by limited proteolysis.

Proteoglycan sulphates have been isolated from mouse mammary epithelial cells (David & Bernfield, 1981). These cells produce both cell-associated and

secreted forms of the proteoglycan. The former, which is the larger one (K_{av} 0.3 on Sepharose CL-4B), associates quantitatively and vectorially with lipid vesicles, even in the presence of 4 M-guanidinium chloride, whereas the medium-derived form (K_{av} approx. 0.5) does not (Rapraeger & Bernfield, 1983). Although there is no further information as to the nature of the core proteins, the proteoglycans studied by Bernfield and co-workers may be very similar to those described in the present work.

The membrane-associated proteoglycan sulphate isolated from hepatocytes is conspicuously dissimilar to the fibroblast proteoglycan (see Oldberg *et al.*, 1979; Mutoh *et al.*, 1980; Kjellén *et al.*, 1981). The hepatic proteoglycan is smaller, and has a core protein of M_r 30 000 with a hydrophobic site for membrane intercalation. It is not known whether the hepatic proteoglycan can bind transferrin.

This work was supported by grants from the Swedish Medical Research Council (567), Kocks Stiftelser, Gustaf V:s 80-årsfond, Österlunds Stiftelse and the Medical Faculty, University of Lund. We thank Ms. Lena Åberg, Ms. Birgitta Havsmark and Ms. Helén Lindgren for technical assistance, Ms. Birgitta Jönsson for art work and Ms. Ruth Lovén and Ms. Inger Carlsson for secretarial aid.

REFERENCES

- Bourdon, M. A., Oldberg, A., Pierschbacher, M. & Ruoslahti, E. (1985) *Proc. Natl. Acad. Sci. U.S.A.* **82**, 1321–1325
- Carlstedt, I., Cöster, L. & Malmström, A. (1981) *Biochem. J.* **197**, 217–225
- Carlstedt, I., Cöster, L., Malmström, A. & Fransson, L.-Å. (1983) *J. Biol. Chem.* **258**, 11629–11635
- Cöster, L., Malmström, A., Carlstedt, I. & Fransson, L.-Å. (1983) *Biochem. J.* **215**, 417–419
- David, G. & Bernfield, M. (1981) *J. Cell Biol.* **91**, 281–286
- Fransson, L.-Å., Carlstedt, I., Cöster, L. & Malmström, A. (1983) *J. Biol. Chem.* **258**, 14342–14345
- Fransson, L.-Å., Carlstedt, I., Cöster, L. & Malmström, A. (1984) *Proc. Natl. Acad. Sci. U.S.A.* **81**, 5657–5661
- Kjellén, L., Pettersson, I. & Höök, M. (1981) *Proc. Natl. Acad. Sci. U.S.A.* **78**, 5371–5375
- Mutoh, S., Funakoshi, I., Ui, N. & Yamashina, I. (1980) *Arch. Biochem. Biophys.* **202**, 137–143
- Oldberg, Å., Kjellén, L. & Höök, M. (1979) *J. Biol. Chem.* **254**, 8505–8510
- Rapraeger, A. C. & Bernfield, M. (1983) *J. Biol. Chem.* **258**, 3632–3636
- Schneider, C., Owen, M. J., Banville, D. & Williams, J. G. (1984) *Nature (London)* **311**, 675–678

Received 15 April 1985/10 June 1985; accepted 8 July 1985