Characterization and *N*-terminal sequence of a heparan sulphate proteoglycan synthesized by endothelial cells in culture

Carlos J. CASTILLO,* Patti COLBURN and Vincenzo BUONASSISI†

W. Alton Jones Cell Science Center, 10 Old Barn Road, Lake Placid, NY 12946, U.S.A.

We have isolated from the conditioned medium of an established endothelial cell line a heparan sulphate proteoglycan whose involvement in the inhibition of the extrinsic coagulation pathway was reported in previous studies [Colburn & Buonassisi (1982) *Biochem. Biophys. Res. Commun.* **104**, 220–227]. The proteoglycan was purified by gel filtration and ion-exchange chromatography, and appears to be free of contaminating proteins as determined by polyacrylamide-gel electrophoresis of the radioiodinated protein core before and after removal of the glycosaminoglycan chains by treatment with heparitinase. By this procedure the M_r of the protein core was estimated to be 22000. The *N*-terminal end was sequenced up to amino acid 25. The 21st residue is likely to be glycosylated. Analysis of the purified proteoglycan by gel-filtration chromatography yielded K_a values of 0.2 for the whole molecule and 0.35 for the glycosaminoglycan chains. The structure that emerges from these data is that of a heparan sulphate proteoglycan characterized by a relatively small protein core and few glycosaminoglycan chains.

INTRODUCTION

In a previous paper we reported on the presence of heparan sulphate proteoglycans at the cell surface and in the supernatant growth medium of primary endothelial-cell cultures and described the kinetics of synthesis of these molecules [1]. With other experiments we were able to show that, in the presence of degradative enzymes purified from *Flavobacterium heparinum*, breakdown products typical for heparan sulphate chains are liberated from the endothelial-cell surface [2]. More recently, we observed that, in the presence of an antiserum directed against these endothelial-cell products, endothelial-cell cultures lose the ability to inhibit the extrinsic coagulation pathway [3,4]. This functional impairment results in the deposition of platelet clumps and fibrin strands on the endothelial surface and the formation of a clot in the growth medium when plasma is added to endothelial-cell cultures [3,4]. Prompted by these findings, we have undertaken the isolation and the preliminary study of the structure of an endothelial-cell proteoglycan (band 1 [5]) as an essential step in our progress towards the elucidation of its site and mechanism of action.

EXPERIMENTAL

Materials

The established endothelial-cell line derived from rabbit aorta, and the conditions for its maintenance and growth, have been described previously [6]. F-12 tissue culture medium and fetal bovine serum were obtained from Gibco. $H_2^{35}SO_4$ (carrier-free), [3H]glucosamine (30 Ci/mmol) and $Na^{125}I$ (17.4 Ci/mg) were from New England Nuclear. The gels used for column chromatography were from Pharmacia. Acrylamide, bisacrylamide and 2-mercaptoethanol were from Bio-Rad

Laboratories, and SDS was from BDH Chemicals. Proteinase K was obtained from Beckman, heparitinase was from Miles Laboratories and all other materials were from Sigma Chemical Co.

Preparation of conditioned medium

To obtain the proteoglycan in amounts sufficient for a preliminary chemical analysis of the protein moiety, the cells, 3 weeks post-confluency, were incubated for 48 h at 37 °C in a CO_2 atmosphere with F-12 tissue-culture medium supplemented with 5% (v/v) fetal bovine serum pretreated, as previously described [4], to remove the high- M_r material. A tenth of the plates were labelled with $H_2^{35}SO_4$ (2 $\mu Ci/ml$). At the end of the incubation, the conditioned medium was collected, made 1 mM with respect to phenylmethanesulphonyl fluoride, centrifuged to remove floating cells and cellular debris and either stored at –70 °C or processed immediately. In a typical preparation, 200 100 mm $\times$ 20 mm tissue-culture dishes provided 1 litre of conditioned medium. For the analysis of the glycosaminoglycan chains, post-confluent cultures were incubated for 2 days in F-12 medium supplemented with 5% fetal bovine serum, containing either 45 μCi of [3H]glucosamine or 180 μCi of $H_2^{35}SO_4$ per ml. The medium was collected and processed as indicated above, and the heparan sulphate proteoglycan was purified according to the procedure outlined below.

Purification of the heparan sulphate proteoglycan

Gel-filtration chromatography. Conditioned medium to be used for gel-filtration chromatography was concentrated with an Amicon pressure cell with a PM-30 membrane (M_r cut-off 30000). The concentrated sample was made 0.5 M with respect to NaCl and 1 mM with respect to dithiothreitol and applied to a Sephacryl S-300 column (5 cm $\times$ 100 cm). The sample volume was 3% or

* On leave from the University of Puerto Rico, Mayaguez Campus, PR 00753, U.S.A.

† To whom correspondence should be addressed.

less of the bed volume. The sample was eluted with Tris buffer (25 mM-Tris/HCl buffer, pH 7.5, containing 0.5 M-NaCl and 1 mM-dithiothreitol) at a flow rate of 60 ml/h and 15 ml fractions were collected. Eluates were monitored by measuring, on-line, absorbancy at 280 nm and by determining the amount of radioactivity present in the column fractions. The amount of radioactivity applied to the column was sufficient to obtain peaks of at least 2000 c.p.m.

Ion-exchange chromatography. Fractions pooled from the Sephacryl S-300 column were diluted to give a concentration of NaCl of 50 mM, and made 7 M with respect to urea in 25 mM-Tris/HCl buffer, pH 7.5. The sample was then applied to a DEAE-Sephacel column (3 cm $\times$ 12 cm), and the column was washed with 100 ml of equilibration buffer. Bound material was eluted at a flow rate of 60 ml/h with a concave NaCl gradient (50 mM to 1 M) in 25 mM-Tris/HCl buffer, pH 7.5, containing 7 M-urea; 10 ml fractions were collected. The concave NaCl gradient was produced by using a cylindrical beaker for the less concentrated solution (900 ml) and a conical flask for the higher salt solution (500 ml). The NaCl molarities were determined from conductivity measurements of the fractions and of prepared standards. The fractions selected for further processing were pooled and freed from salt on an Isco electrophoretic concentrator. The buffer used, 10 mM-ammonium bicarbonate, pH 7.0, was exchanged several times to decrease urea concentration in the sample chamber by least 1000-fold. Generally, it took more than 24 h of electrophoresis at 3 W at 4 °C to concentrate and desalt the sample completely. More than 90 % of the radioactive sample was recovered.

Analytical gel-filtration chromatography

Analysis of the purified proteoglycan by gel-filtration chromatography was performed on a Sepharose CL-6B column (0.9 cm $\times$ 60 cm). Blue Dextran and Phenol Red were used to mark the void volume ($K_d = 0$) and the included volume ($K_d = 1$). The samples used for analysis contained at least 20000 c.p.m. The samples were made 4 M with respect to guanidinium chloride in 50 mM-sodium acetate buffer, pH 7.0, and eluted with the same buffer at a flow rate of 10 ml/h; 0.1 ml samples were applied and 1 ml fractions were collected.

Polyacrylamide-gel electrophoresis

Vertical slab-gel electrophoresis was done on separating gels made of a gradient of urea (2.2–6.6 M), acrylamide (3.3–11.4 %) and bisacrylamide (acrylamide/bisacrylamide molar ratio was 16:5) [7]. Stacking gel, SDS concentration, electrode buffers, sample preparation and running conditions were as described by Laemmli [8]. Diluted protein samples were concentrated by precipitation with trichloroacetic acid in the presence of deoxycholate [9]. The precipitate was resuspended in the sample buffer, and residual trichloroacetic acid was neutralized with NaOH. Gels were stained with Coomassie Blue [8] or with silver [10]. Autoradiographs of gels were obtained as described [4].

Protein radioiodination

Radioiodination of the purified heparan sulphate proteoglycan was performed essentially according to the method of Greenwood *et al.* [11]. A 200 μ l portion of the proteoheparan sample containing approx. 100 ng of

protein was exposed to 500 μ Ci of Na¹²⁵I in the presence of chloramine-T. The reaction was stopped by the addition of sodium metabisulphite and unlabelled NaI. Unincorporated ¹²⁵I was removed from the preparation by applying the sample (250 μ l) to a Sephadex G-50 column (1.5 cm $\times$ 8.5 cm). The elution rate was 60 ml/h, and 1.0 ml fractions were collected with as eluent 0.1 M-ammonium acetate/0.1 M-sodium acetate buffer, pH 7.0.

β -Elimination reaction

To obtain free polysaccharide chains for gel-filtration analysis, the alkaline β -elimination reaction was carried out in the presence of NaBH₄. The proteoglycan sample was adjusted to 1.0 M-NaBH₄/0.05 M-NaOH and incubated at 45 °C for 24 h [12]. After neutralization with 1 M-acetic acid, salts were removed from the preparations (0.2 ml) by gel-filtration chromatography on a Sephadex G-75 column (1.5 cm $\times$ 20 cm). Samples were eluted with 0.2 M-ammonium acetate buffer, pH 7.0, supplemented with 10 % (v/v) ethanol at a flow rate of 30 ml/h; 1 ml fractions were collected.

Enzyme treatment

Proteinase K. A 50 μ l portion of the labelled proteoglycan sample in 50 mM-ammonium bicarbonate buffer, pH 7.0, was mixed with 5 μ l of proteinase K solution containing 20 μ g of enzyme/ μ l in 25 mM-Tris/HCl buffer, pH 7.5, containing 50 mM-NaCl and incubated for 12 h at 37 °C. After incubation, a portion of the preparation containing about 1000 c.p.m. of [³⁵S]sulphate label was mixed with 50 μ g of bovine serum albumin, made 10 % (w/v) with respect to trichloroacetic acid, incubated for 30 min at 4 °C and sedimented at 8000 *g* for 5 min. Proteinase K digestion was considered complete when no detectable radioactivity was found in the precipitate.

Digestion with trypsin was carried out under the same conditions except for the enzyme concentration, which was 10 μ g/ml.

Heparitinase. A 70 μ l portion of 0.2 M-sodium acetate buffer, pH 7.0, containing 2 mM-CaCl₂ and 10 μ l of the enzyme preparation containing 1 conventional unit in 0.1 M-sodium acetate buffer, pH 7.0, were added to 20 μ l of the proteoglycan preparation. The reaction mixture was incubated for 1 h at 37 °C. Under these conditions the enzyme produced, in 60 min on 10 mg of its substrate, 0.9 unit of 235 nm-absorbing material (90 % of the expected activity).

Protein concentration

Protein concentration was determined either by the Lowry method [14] or, for dilute solutions, by the Esen procedure [15].

Amino acid analysis

For the amino acid analysis, a pre-column derivatization method was used and the resulting phenylthiocarbamoyl derivatives were separated by reverse-phase chromatography [16]. The number of cysteine residues was determined in a separate analysis by oxidizing the sample with performic acid before hydrolysis with 6 M-HCl. Approx. 10 pmol of the heparan sulphate proteoglycan was used for each analysis. The analyses were performed at the Protein Sequencing Facility of Dr. G. Tarr, University of Michigan, Ann Arbor, MI, U.S.A. and by the Protein Chemistry

Laboratory of Dr. J. Crabb, W. Alton Jones Cell Science Center, Lake Placid, NY, U.S.A.

Amino acid sequence

The *N*-terminal amino acid sequence of the heparan sulphate proteoglycan was determined with automated Edman degradation on an Applied Biosystems 470A gas-phase sequencer. The repetitive yields of the automated procedure was 94 %. Two runs were performed; one, on 300 pmol of material, went for 35 cycles, and the second, on 250 pmol, confirmed the identity of the first 25 amino acid residues. Amino acid phenylthiohydantoin derivatives were separated on a Waters h.p.l.c. system. The column used was an Altex Ultrasphere ODS (5 μ m particle size) and the buffer system, acetate-buffered aqueous acetonitrile, was as specified by Black & Coon [17]. The identity of the first and second amino acid residues was confirmed by the partitioning method of manual sequencing, and the resulting amino acid phenylthiohydantoin derivatives were identified by h.p.l.c. [18]. This procedure was performed at the Protein Sequencing Facility in Ann Arbor, MI, U.S.A. Before the analysis, the sample (500 pmol) was freed of salts by precipitation from 88 % formic acid with 3 vol. of ethyl acetate.

RESULTS AND DISCUSSION

Purification of the endothelial-cell heparan sulphate proteoglycan

The first step in the purification procedure was the fractionation of conditioned medium on a Sephacryl-300 column (Fig. 1a). One-third of the initial radioactivity but only 0.2 % of the initial amount of protein are contained in the voiding fractions. These fractions were shown by gel electrophoresis to contain the high- M_r proteoglycan as well as a 220 000- M_r species. In fact, most of the 280 nm-absorbing material consisted of the second species. Sephacryl S-300 chromatography produced more than 100-fold purification, as shown in Table 1.

The material eluted with the void volume was subjected to DEAE-Sephacel chromatography (Fig. 1b). The [35 S]sulphate-labelled material retained in the column is eluted as a sharp peak at 400 mM-NaCl. Most of the protein present in the input sample by-passes the column or is eluted before 300 mM-NaCl.

Fractions containing the labelled material were concentrated by electrophoresis on an Isco concentrator and analysed by polyacrylamide-gel electrophoresis. Autoradiographs of the gel slabs show only the highly polydisperse species with the highest apparent M_r (Fig. 2 lane B). Also, the pattern produced by the radioiodinated sample (Fig. 2, lane A) is indistinguishable from that of the sample labelled with [35 S]sulphate. In both cases the molecule is highly polydisperse. This polydispersity is clearly to be attributed to the size heterogeneity of the carbohydrate chains (Fig. 2, lane C).

The purity of the preparation was ascertained by subjecting the radioiodinated molecule to electrophoretic analysis before and after treatment with heparitinase (Fig. 2, lanes B and D). This procedure facilitates the detection of contaminating proteins because after deglycosylation the electrophoretic mobility increases and the protein appears as a sharp band (M_r 24 000). The

sharpness of the band suggests that the glycosaminoglycan fragments that remain attached to the core protein have a uniform, perhaps minimal, size. The purification procedure and yields are summarized in Table 1. It can be seen that the total yield of the heparan sulphate proteoglycan from 1 litre of conditioned medium is about 20 μ g, which indicates that endothelial cells are releasing into the medium an amount in the order of 10^5 molecules per day per cell.

Gel-filtration analysis of the proteoglycan and its polysaccharide chain

When chromatographed on a Sepharose CL-6B column in 4 M-guanidinium chloride, the purified proteoglycan shows a broad elution profile peaking at a K_a of 0.2 (Fig. 3). On the basis of the values obtained by gel-filtration analysis for other proteoglycans whose M_r values had been determined also by sedimentation analysis [19–22] this K_a would correspond to an M_r of 100 000.

To obtain the elution profile of the glycosaminoglycan chains, the core protein was subjected to digestion with proteolytic enzymes. Treatment with trypsin, an enzyme that should find, on the average, a sensitive peptide bond every ten amino acid residues and therefore leave single glycosaminoglycan chains attached to an oligopeptide or more than one chain linked by one oligopeptide, did not affect noticeably either the chromatographic or the electrophoretic behaviour of the proteoglycan. It should be noted that, since the molecule is 80 % carbohydrate by weight, the loss of a protein fragment that is not part of the peptide linking the polysaccharide chains may not appreciably affect the chromatographic or electrophoretic behaviour of the [35 S]sulphate-labelled proteoglycan. The same enzyme thoroughly degrades the deglycosylated proteoglycan. These results suggest that there are no trypsin-sensitive sites between chains or that the chains are clustered or that the potentially sensitive peptide bonds located between them are protected, perhaps by the chains themselves, in the intact molecule. The absence of trypsin-sensitive sites in the peptide linking the polysaccharide chains has been noticed in other proteoglycans [23,24]. The heparan sulphate proteoglycan was also subjected to digestion with proteinase K. This enzyme should find sensitive peptide bonds, on the average, two or three amino acid residues apart [25], and therefore release chains capped with a very small oligopeptide. After digestion with proteinase K the [35 S]sulphate-labelled material is eluted with a K_a of 0.35 from a Sepharose CL-6B column under dissociating conditions (Fig. 3a). This K_a corresponds to an M_r of 55 000. Chromatographic analysis of the polysaccharide chains obtained by β -elimination shows a broad peak with a K_a of 0.35 (result not shown) identical with that obtained after proteinase K treatment.

As shown above (Fig. 2, lane D), after heparitinase digestion the radioiodinated proteoglycan migrates on analytical polyacrylamide-gel electrophoresis as a sharp band with an M_r of 24 000. This finding, together with those obtained by gel-filtration analysis, suggest that the proteoglycan consists of a 22 000- M_r protein core (24 000 less 2000 for the residual glycosaminoglycan fragments) and two heparan sulphate chains with M_r in the order of 50 000. The molecule is highly polydisperse because of the size heterogeneity of the glycosaminoglycan moiety,

Fig. 1. Purification of the endothelial-cell heparan sulphate proteoglycan by gel-filtration and ion-exchange chromatography

(a) Gel-filtration chromatography. Concentrated conditioned medium was fractionated on a Sephacryl S-300 column as indicated in the Experimental section. Fractions eluted at the void volume (indicated by arrows) were pooled and further processed by ion-exchange chromatography. (b) Ion-exchange chromatography. The salt concentration of the pooled Sephacryl S-300 fractions was adjusted as indicated in the Experimental section and the sample was applied to a DEAE-Sephacel column. The material that bound to the column was eluted with a concave gradient of NaCl (■) (see the Experimental section). The arrows indicate the fractions selected for further processing. Key to symbols in (a) and (b): ○, A_{280} ; ●, radioactivity recovered.

Table 1. Purification of the endothelial-cell heparan sulphate proteoglycan

Experimental details are given in the text. Abbreviation: N.D., not determined.

Fraction	Volume (ml)	$10^{-3} \times$ Total radio-activity (c.p.m.)	Total protein (mg)	Specific radio-activity (c.p.m./mg)	Yield	Puri-fication (fold)
Conditioned medium	880	1720	1936	0.89	1.00	1.00
Amicon Concentration	47	1710	1880	0.91	1.00	1.02
Sephacryl S-300	140	546	3.9	140	0.32	156
DEAE-Sephacel	30	195	N.D.	—	—	—
Electrodialysis and concentration	1.2	178	0.016	11 125	0.10*	12 500

* Since in the conditioned medium the proteoglycan under study constitutes about 40 % of the [35]sulphate label incorporated into macromolecular material and in the purified sample 100 %, the adjusted yield would be 0.25.

and thus it is difficult, from the M_r data obtained by gel-filtration analysis, to rule out that more than two chains of diverse size are attached to the same protein core. Notice that chromatographic analysis of the [3 H]-glucosamine-labelled products of proteinase K digestion does not reveal the presence of low- M_r material that would correspond to oligosaccharide chains (Fig. 3b).

Amino acid analysis

The salient features of the amino acid composition of the proteoglycan (Table 2) are the high concentrations of

glutamate, aspartate and serine; glycine, an amino acid that has been reported to be part of the glycosylation site [26], is also present in an amount higher than in the average protein [27].

N-Terminal sequence

The N-terminal amino acid sequence of the heparan sulphate proteoglycan is shown in Fig. 4(a). Amino acid 21 from the N-terminus appears as a 'blank' during automatic sequencing and is followed by a glycine,

Fig. 2. Analysis by polyacrylamide-gel electrophoresis of the purified heparan sulphate proteoglycan before and after treatment with different enzymes

The autoradiographs are shown. Lane A, purified heparan sulphate proteoglycan labelled with ^{125}I ; lane B, [^{35}S]-sulphate-labelled purified proteoglycan; lane C, [^{35}S]-sulphate-labelled purified proteoglycan after treatment with proteinase K; lane D, purified proteoglycan labelled with ^{125}I and treated with heparitinase; lane E, protein M_r markers (from top to bottom, fibronectin, phosphorylase *b*, bovine serum albumin, ovalbumin, carbonic anhydrase, soya-bean trypsin inhibitor and lactoglobulin A).

suggesting that the 21st residue may be glycosylated serine. The *N*-terminal sequence of the proteoglycan is unique, i.e. no significantly homologous amino acid sequences were found by Intelligenetics computer programs in their protein data bank. There is no homology between the *N*-terminal sequence of our heparan sulphate proteoglycan and other proteoglycan *N*-terminal sequences known [24,28].

Secondary structure of the *N*-terminal sequence

We have deduced a secondary structure for the *N*-terminal sequence of the proteoglycan (Fig. 5). It is based on the statistical tables of amino acid propensities to participate in α -helices, β -sheets or β -bends, and on the empirical rules for secondary-structure nucleation and termination published by Chou & Fasman [29]. Amino acids 10 (D), 11 (P), 12 (Q) and 13 (D) almost certainly participate in a β -bend ($66 \times 10^{-6} \gg 24 \times 10^{-6}$). This bend

Fig. 3. Gel-filtration analysis on a Sephadex CL-6B column of the purified heparan sulphate proteoglycan and its polysaccharide moiety

(a) Purified [^{35}S]sulphate-labelled proteoglycan (●) and purified proteoglycan after treatment with proteinase K (▲). (b) Purified proteoglycan labelled with [^3H]-glucosamine after digestion with proteinase K. For this particular analysis 22000 c.p.m. was applied to the Sephadex CL-6B column.

is preceded by an α -helix of nine amino acid residues, amino acids 1–9, and followed by a short helical segment, amino acids 14 (L), 15 (L) and 16 (E). The rest of the amino acid residues cannot be assigned into any secondary structure. This well-defined secondary structure agrees with that predicted by the Intelligenetics computer programs (Fig. 4b). In contrast, the *N*-terminal sequence of a dermatan sulphate proteoglycan is very rich in proline residues, which precludes formation of α -helices or β -sheets [28]. An attractive symmetry is revealed by this proposed secondary structure: two contiguous hydrophobic pairs, two glutamate residues and serine and threonine are equidistant from the β -bend. The two arginine residues are almost equidistant. In order to see whether this notable distribution would constrain the three-dimensional structural possibilities into few distinct conformations, we built a ball-and-stick model at 1 nm $\equiv$ 250 mm. Although no verification is possible at the moment, we consider that the two helical segments could pack in such a way that the four hydrophobic residues form a core surrounded by the polar amino acid residues. Acidic residues would predominate on one side of the structure, and the two arginine residues would be localized towards the opposite side. In partial agreement

22. Heinegård, D., Paulson, M., Inerot, S. & Carlström, C. (1981) *Biochem. J.* **197**, 355–366
 23. Carlstedt, I., Cöster, L., Malmström, A. & Fransson, L. (1983) *J. Biol. Chem.* **258**, 11629–11635
 24. Yanagishita, M. & Hascall, V. (1983) *J. Biol. Chem.* **258**, 12857–12864
 25. Ebeling, W., Hennrich, N., Klockow, M., Metz, H., Orth, H. D. & Lang, H. (1974) *Eur. J. Biochem.* **47**, 91–97
 26. Bella, A., Jr. & Danishefsky, I. (1968) *J. Biol. Chem.* **243**, 2660–2664
 27. Doolittle, R. F. (1981) *Science* **214**, 149–159
 28. Pearson, C. H., Winterbottom, N., Fackre, D. S., Scott, P. G. & Carpenter, M. R. (1983) *J. Biol. Chem.* **258**, 15101–15104
 29. Chou, P. Y. & Fasman, G. D. (1974) *Biochemistry* **13**, 222–245
-

Received 25 February 1987/7 May 1987; accepted 24 July 1987