Identification of an extended *N*-acetylated sequence adjacent to the protein-linkage region of fibroblast heparan sulphate

Malcolm LYON,*† William P. STEWARD,* Ian N. HAMPSON† and John T. GALLAGHER*

Cancer Research Campaign *Department of Medical Oncology and †Department of Experimental Haematology, Christie Hospital and Paterson Laboratories, Wilmslow Road, Manchester M20 9BX, U.K.

The distribution of *N*-sulphate groups within fibroblast heparan sulphate chains was investigated. The detergent-extractable heparan sulphate proteoglycan from adult human skin fibroblasts, radiolabelled with [³H]glucosamine and [³⁵S]sulphate, was coupled to CNBr-activated Sepharose 4B. After partial depolymerization of the heparan sulphate with nitrous acid, the remaining Sepharose-bound fragments were removed by treatment with alkali. These fragments, of various sizes, but all containing an intact reducing xylose residue, were fractionated on Sephacryl S-300 and the distribution of the ³H and ³⁵S radiolabels was analysed. A decreased degree of sulphation was observed towards the reducing termini of the chains. After complete nitrous acid hydrolysis of the Sepharose-bound proteoglycan, analysis of the proximity of *N*-sulphation to the reducing end revealed the existence of an extended *N*-acetylated sequence directly adjacent to the protein-linkage sequence. The size of this *N*-acetylated domain was estimated by gel filtration to be approximately eight disaccharide units. This domain appears to be highly conserved, being present in virtually all the chains derived from this proteoglycan, implying the existence of a mechanism capable of generating such a non-random sequence during the post-polymeric modification of heparan sulphate. Comparison with the corresponding situation in heparin suggests that different mechanisms regulate polymer *N*-sulphation in the vicinity of the protein-linkage region of these chemically related glycosaminoglycans.

INTRODUCTION

Heparan sulphate is a complex polysaccharide synthesized by most mammalian cell types in culture. It is coupled to protein via a linkage sequence comprising GlcAβ1→3Galβ1→3Galβ1→4Xyl-Ser(peptide) (Knecht *et al.*, 1967; Fransson *et al.*, 1985) forming both cell-surface- and extracellular-matrix-associated proteoglycans (for a review see Gallagher *et al.*, 1986). The basic polymeric structure of heparan sulphate, as displayed by the initial biosynthetic product, is an alternating repeat sequence of →4GlcAβ1→4GlcNAcα1→. However, this simple repeat structure undergoes extensive enzymic modification within the Golgi apparatus involving sequential amino group sulphation (GlcNAc→GlcNH₃⁺→GlcNSO₃⁻) and uronate epimerization (GlcA→IdoA), on which is superimposed a complex pattern of *O*-sulphations, predominantly on the iduronate [i.e. IdoA(2-OSO₃)] and *N*-sulphoglucosamine [i.e. GlcNSO₃-(6-OSO₃)] residues (for a review see Lindahl & Kjellén, 1987). Incomplete utilization of the available modification sites at each step leads to a considerable degree of sequence variability within the polymer chain.

Nitrous acid-induced deamination cleaves the glycosidic linkages of *N*-sulphoglucosamine units of heparan sulphate, liberating inorganic sulphate. Analyses of the degradation products have shown that heparan sulphates from a variety of different sources all contain approximately equal numbers of *N*-sulphated and *N*-acetylated glucosamine residues, though the degree of polymer *O*-sulphation varies considerably in a cell-lineage-related

manner (Gallagher & Walker, 1985). *N*-Sulphated and *N*-acetylated disaccharides occur mainly, though not exclusively, in separate domains rather than in a predominantly alternating arrangement, and the *O*-sulphate groups are always in close proximity to the *N*-sulphate groups. It is not known if there is a defined sequence of sugar residues throughout the polymer chain, nor if any particular groups of *N*-acetylated disaccharides are consistently excluded from modification during heparan sulphate biosynthesis.

In the present study we have utilized an immobilized heparan sulphate proteoglycan, derived from adult human skin fibroblasts, to investigate further the distribution of sulphate groups along the chain, and in particular the occurrence of *N*-sulphation within the vicinity of the protein-linkage region of the proteoglycan.

MATERIALS AND METHODS

Materials

D-[1-³H]Glucosamine hydrochloride (specific radioactivity 2.6 Ci/mmol) and Na₂³⁵SO₄ (specific radioactivity 25–40 Ci/mg) were obtained from Amersham International (Amersham, Bucks., U.K.). Chondroitin ABC lyase (EC 4.2.2.4) and heparan sulphate lyase (EC 4.2.2.8) were from Miles Chemicals (Slough, Berks., U.K.). Bovine testicular hyaluronidase (EC 3.2.1.35; type VI-S) and neuraminidase (EC 3.2.1.18; type X from *Clostridium perfringens*) were from Sigma Chemical Co.

Abbreviations used: GlcA, D-glucuronic acid; IdoA, L-iduronic acid; GlyA, glycuronic acid; GlcNAc, *N*-acetyl-D-glucosamine; GlcNSO₃⁻, *N*-sulpho-D-glucosamine; -OSO₃, sulphate ester; Gal, D-galactose; Xyl, D-xylose; aMan, 2,5-anhydro-D-mannose.

† To whom correspondence and requests for reprints should be addressed.

(Poole, Dorset, U.K.). Sephacryl S-300 and CNBr-activated Sepharose 4B were obtained from Pharmacia Fine Chemicals (Uppsala, Sweden). DEAE-cellulose (DE-52) was from Whatman Chemical Co. (Maidstone, Kent, U.K.) and Bio-Gel P-30 was from Bio-Rad Laboratories (Watford, Herts, U.K.). Eagle's Minimal Essential Medium, donor calf serum and all other reagents for cell culture were from Gibco Biocult (Paisley, Strathclyde, U.K.).

Cell culture

Cultures of adult human skin fibroblasts were kindly provided by Dr. S. Schor. These were maintained at 37 °C in a humidified atmosphere of 5% CO₂ in air. The culture medium comprised Eagle's Minimal Essential Medium supplemented with 15% (v/v) donor calf serum, 2 mM-glutamine, 1 mM-sodium pyruvate, non-essential amino acids, 100 units of penicillin/ml and 100 mg of streptomycin/ml. Confluent cultures were biosynthetically radiolabelled by incubation with 10 µCi of [³H]glucosamine hydrochloride/ml and 40 µCi of Na₂³⁵SO₄/ml for 72 h.

Extraction of cell-associated heparan sulphate proteoglycan

The medium was removed from radiolabelled confluent cultures, and the adherent cell layer was washed twice with 0.15 M-NaCl/20 mM-sodium phosphate buffer, pH 7.4. The washed cell layer was extracted with 1% (v/v) Triton X-100/1 M-NaCl/20 mM-sodium phosphate buffer, pH 7.4 (containing the following proteinase inhibitors: 0.1 M-6-aminohexanoic acid/20 mM-benzamidine hydrochloride/10 mM-EDTA/5 mM-*N*-ethylmaleimide/0.25 mM-phenylmethanesulphonyl fluoride) for 30 min at 4 °C. The extract was carefully removed and centrifuged at 1000 *g* for 5 min. The supernatant was concentrated, dialysed against 1% (v/v) Triton X-100/0.15 M-NaCl/20 mM-sodium phosphate buffer, pH 7.4 (containing the above proteinase inhibitors) and applied to a column (1 cm × 5 cm) of DEAE-cellulose (DE-52) in the same solution. The column was eluted stepwise with 0.25 M-NaCl and 1 M-NaCl, both in 0.1% (v/v) Triton X-100/20 mM-sodium phosphate buffer, pH 7.4. The 1 M-NaCl wash, which contained all of the ³⁵S-labelled proteoglycans, was then digested with a combination of 200 units of testicular hyaluronidase and 0.5 unit of neuraminidase in 0.1% (v/v) Triton X-100/0.15 M-NaCl/0.1 M-sodium acetate buffer, pH 5.5 (containing proteinase inhibitors at 10% of the standard concentration), at 37 °C for 16 h. The digest was dialysed against 0.1% (v/v) Triton X-100/50 mM-NaCl/50 mM-Tris/HCl buffer, pH 8.0 (containing 10% proteinase inhibitors), and then further digested with 0.5 unit of chondroitin ABC lyase at 37 °C for 16 h. The digest was applied to a column (1 cm × 5 cm) of DEAE-cellulose equilibrated with 0.1% (v/v) Triton X-100/0.25 M-NaCl/20 mM-sodium phosphate buffer, pH 7.4. After a thorough washing with the above solution, the radiolabelled heparan sulphate proteoglycans were eluted with 0.1% (v/v) Triton X-100/1 M-NaCl/0.2 M-NaHCO₃ and then coupled directly to CNBr-activated Sepharose 4B. The combination of enzyme treatments and ion-exchange chromatography should ensure that the proteoglycans are not contaminated with sialylated oligosaccharides or other glycos-

aminoglycans that are produced by cultures of human skin fibroblasts (Gallagher *et al.*, 1983).

Degradation of heparan sulphate proteoglycan

Controlled partial nitrous acid digestion of the Sepharose-bound heparan sulphate proteoglycan was performed as described by Radoff & Danishefsky (1984). The Sepharose-coupled proteoglycan was suspended in 19 mM-HCl, and sufficient 0.1 M-NaNO₂ was added to give a final concentration of 0.25 mM. After incubation at 4 °C for various lengths of time the reaction was terminated by the addition of an excess of ammonium sulphamate. Rapid and complete nitrous acid digestion was achieved by the low-pH method of Shively & Conrad (1976).

Fig. 1. Gel filtration on Sephacryl S-300 of the Sepharose-bound oligosaccharides remaining after nitrous acid hydrolysis of heparan sulphate proteoglycan-Sepharose

Radiolabelled heparan sulphate proteoglycan-Sepharose was degraded to various degrees with nitrous acid: (a) untreated control, (b) partial degradation, and (c) complete degradation (see the Materials and methods section). After removal of all solubilized fragments, the remaining Sepharose-bound material was released by treatment with alkaline borohydride and then applied to a Sephacryl S-300 column (1 cm × 135 cm) eluted with 4 M-guanidinium chloride/50 mM-phosphate buffer, pH 6.0, at a flow rate of 10 ml/h. Fractions (0.8 ml) were collected and counted for radioactivity. —, ³H radioactivity; ----, ³⁵S radioactivity; ●, ³H/³⁵S ratio. The void volume (*V*₀) and total volume (*V*_t) were determined with Blue Dextran and Phenol Red respectively.

Intact heparan sulphate chains, or fragments remaining after nitrous acid treatment, were released from the Sepharose-bound proteoglycan by alkaline elimination with 50 mM-NaOH/1 M-NaBH₄ at 45 °C for 24 h. Solubilized material was collected by filtration of the neutralized sample over sintered glass with thorough washing.

Digestion of heparan sulphate with heparan sulphate lyase was performed in 10 mM-calcium acetate/0.1 M-sodium acetate buffer, pH 7.0, containing 50 µg of bovine serum albumin/ml at 43 °C for 7 h with 2.5 units of enzyme in the presence of 50 µg of unlabelled carrier heparan sulphate.

Gel chromatography

Sephacryl S-300 (1 cm × 135 cm) was eluted with 4 M-guanidinium chloride/50 mM-sodium phosphate buffer, pH 6.0, at a flow rate of 10 ml/h. Bio-Gel P-30 (1 cm × 113 cm) was eluted with 0.25 M-NH₄HCO₃ containing 10 µg of bovine serum albumin/ml, 1 µg of chondroitin sulphate/ml and 1 µg of heparin/ml at a flow rate of 3.2 ml/h. The radioactivity in column effluents was monitored by liquid-scintillation counting.

RESULTS

Preparation of Sepharose-bound heparan sulphate proteoglycan

Adult human skin-fibroblast proteoglycans were radiolabelled by incubation of cultured cells with [³H]glucosamine and Na₂³⁵SO₄. After removal of the culture medium the cell-layer-associated proteoglycans were extracted with Triton X-100/NaCl. Excess radiolabel and other small radiolabelled products were removed from the extract by chromatography on DEAE-cellulose. Elution with 1 M-NaCl gave a partially purified proteoglycan fraction containing both heparan sulphate proteoglycans and chondroitin/dermatan sulphate proteoglycans. The latter were removed by treatment with testicular hyaluronidase and chondroitin ABC lyase followed by a further ion-exchange chromatography step. The purified heparan sulphate proteoglycans were then coupled to CNBr-activated Sepharose 4B via their polypeptide chains, resulting in a population of solid-phase heparan sulphate chains that, by definition, all contain an intact protein-linkage sequence at their reducing end. The coupling yield was consistently 40–50% of the radiolabel. Entire chains (or fragments derived from them) can be recovered from the gel, peptide-free, by alkaline cleavage of the xylose-serine linkage.

Analysis of gross sulphate distribution within the polymer chain

Alkaline-borohydride treatment of the Sepharose-bound heparan sulphate proteoglycan released intact heparan sulphate chains. Chromatography of these chains on Sephacryl S-300 gave a broad peak with a K_{av} of 0.22 (Fig. 1a). The ³H and ³⁵S radiolabels were co-distributed and, within experimental error, there was no discernible variation in the ³H/³⁵S ratio across the peak.

In order to investigate the relative distribution of sulphate groups along the length of the polysaccharide chain a series of Sepharose-bound fragments of various sizes was generated. Three portions of the heparan

Fig. 2. Characterization of the nitrous acid-resistant protein-linkage-region sequence of heparan sulphate by gel filtration on Bio-Gel P-30

Intact chains (a) and nitrous acid-resistant fragments (b) were released from heparan sulphate proteoglycan-Sepharose by alkaline elimination. A further sample of the released nitrous acid-resistant fragments was digested with heparan sulphate lyase (c). Samples were chromatographed on a column (1 cm × 113 cm) of Bio-Gel P-30 eluted with 0.25 M-NH₄HCO₃ containing 10 µg of bovine serum albumin/ml, 1 µg of chondroitin sulphate/ml and 1 µg of heparin/ml at a flow rate of 3.2 ml/h. Fractions (0.53 ml) were collected and counted for radioactivity. —, ³H radioactivity; ----, ³⁵S radioactivity. The void volume and total volume were determined with Blue Dextran and ³⁵SO₄²⁻ respectively. All three samples were derived from equivalent quantities of proteoglycan-Sepharose, and for comparison all the profiles are shown on a common c.p.m. scale. The peak maximum in (a) is off-scale; the actual peak values are 26.8 × 10³ and 21.9 × 10³ c.p.m. of ³H and ³⁵S respectively.

sulphate proteoglycan-Sepharose were treated for differing periods of time with the dilute nitrous acid preparation (see the Materials and methods section). Nitrous acid randomly cleaves glycosidic linkages on the reducing side of N-sulphoglucosamine residues. Incubation for 20 min, 1 h and 2 h released 8%, 22% and 45% of the ³H label respectively into solution, thus ensuring adequate representation on the Sepharose beads of all

Fig. 3. Molecular-size calibration of Bio-Gel P-30 and estimation of the size of the *N*-acetylated protein-linkage domain of heparan sulphate

Intact heparan sulphate chains released from proteoglycan–Sephacryl S-300 with alkaline borohydride were completely digested with nitrous acid. The digest was applied to a column of Bio-Gel P-30 (see Fig. 2 for details). The resulting oligosaccharide profile (a) was used to construct a molecular-size calibration curve for the column (b). The *N*-acetylated linkage fragment (elution position marked by an arrow) was estimated, by extrapolation, to be approximately ten disaccharide units in length.

lengths of oligo- and poly-saccharides corresponding to the distribution of *N*-sulphoglucosamine residues. After being washed on a sintered-glass filter to remove all the solubilized fragments (derived from the distal regions of the chains), the three portions of gel were pooled. The gel was then treated with alkaline borohydride to release the proximal fragments all containing an intact reducing terminal sequence of Gal–Gal–Xyl. These were chromatographed on Sephacryl S-300 and gave a broad profile (Fig. 1b) covering a wide range of molecular sizes. The $^3\text{H}/^{35}\text{S}$ ratio increased with decreasing molecular size of the fragments (Fig. 1b), reaching the highest values in the smallest fragments present with K_{av} values of 0.6–0.7. An identical profile for the $^3\text{H}/^{35}\text{S}$ ratio was obtained when corresponding fragments were generated by a series of partial digestions of the proteoglycan–Sephacryl S-300 with heparan sulphate lyase (results not shown). Together these findings indicate a decreasing degree of sulphation with increasing proximity to the protein core.

To investigate the position(s) of the *N*-sulphoglucosamine units that are in closest proximity to the protein core, a further sample of heparan sulphate proteoglycan–Sephacryl S-300 was subjected to complete nitrous acid hydrolysis under the low-pH conditions of Shively & Conrad (1976). The fragments that remain attached to the Sepharose beads will be devoid of *N*-sulphate groups and will correspond to the general structure:

These fragments were recovered from the gel by alkaline elimination and chromatographed on Sephacryl S-300 (Fig. 1c). These gave a single relatively sharp peak with a surprisingly low K_{av} value of 0.67. The elution position of this material corresponded to that of the smallest fragments (with the high $^3\text{H}/^{35}\text{S}$ ratio) generated by the partial nitrous acid digest. There is thus evidence for a relatively extended non-sulphated domain adjacent to the protein-linkage sequence of this heparan sulphate.

Further analysis of *N*-sulphation in the proximity of the protein-linkage region

In order to be able to characterize the above large non-sulphated domain in greater detail, further experiments were performed with the use of Bio-Gel P-30 for improved size resolution. Intact heparan sulphate chains, prepared as described in the Materials and methods section, were chromatographed, as a control, on a Bio-Gel P-30 column (Fig. 2a). These were eluted in the void volume of the gel, but also present was a small amount of smaller and predominantly ^3H -labelled material (for example a peak with a K_{av} of 0.66). This latter small material presumably consists of alkali-labile *O*-linked oligosaccharides derived either from the proteoglycan or from some contaminating glycoprotein.

Complete nitrous acid treatment of heparan sulphate proteoglycan–Sephacryl S-300 released 87% of the ^3H label. Subsequent alkaline elimination of the nitrous acid-resistant Sepharose-bound fragments released the remaining 13% of the ^3H label (a negligible amount of ^3H label remained associated with the Sepharose beads). Chromatography of the latter fragments on Bio-Gel P-30 gave the profile shown in Fig. 2(b). A single relatively sharp peak of ^3H -labelled material was eluted with a K_{av} of 0.21 (cf. Fig. 1c). Small amounts of lower-molecular-mass species were also present, but these were almost identical in profile and quantity with those present in the undegraded control (cf. Fig. 2a), which would be consistent with them being nitrous acid-stable *O*-linked oligosaccharides not derived from heparan sulphate chains. The major heparan sulphate-derived peak should be susceptible to degradation by bacterial heparan sulphate lyase, which shows specificity for the GlcNAc α 1 $\rightarrow$ 4GlcA linkage (Silverberg *et al.*, 1985). Treatment of a parallel sample of the nitrous acid-resistant fragments (i.e. Fig. 2b) with heparan sulphate lyase resulted in complete elimination of the major ^3H -labelled peak (Fig. 2c). The major product of digestion was eluted in the total volume of the gel confirming that the nitrous acid-resistant fragment comprises an *N*-acetylated repeat sequence. A smaller peak with an anomalously retarded elution position was also present. This is probably a small quantity of free *N*-acetyl[^3H]-glucosamine produced by a β -glucuronidase contaminant in the heparan sulphate lyase preparation. Prolonged incubation with heparan sulphate lyase leads to a gradual

accumulation of radioactivity in this peak. It is again noticeable that a small amount of ^3H -labelled material remained across the whole of the included volume, representing the background level of contaminating oligosaccharides (cf. Figs. 2a and 2b).

Molecular-size calibration of the Bio-Gel P-30 was undertaken to allow estimation of the size of the major *N*-acetylated domain in Fig. 2(b). Intact heparan sulphate chains (as in Fig. 2a) were limit-digested with nitrous acid, giving rise to a series of oligosaccharides of general structure:

(omitting the possible presence of *O*-sulphate groups on both the non-reducing terminal disaccharide and the reducing anhydromannose residue, whose contributions decrease with increasing oligosaccharide length). These fragments were chromatographed on Bio-Gel P-30 (Fig. 3a), and the peaks corresponding to $n = 0-5$ were used to generate the molecular-size calibration shown in Fig. 3(b). Extrapolation of this plot leads to a size estimation of ten disaccharide units for the linkage region domain, i.e. eight *N*-acetylated disaccharide units plus the linkage tetrasaccharide unit.

DISCUSSION

A population of solid-phase heparan sulphate chains all containing an intact protein linkage were prepared by coupling the fibroblast proteoglycan via its core protein to CNBr-activated Sepharose. Partial chain degradation can be performed *in situ* with retention of the proximal fragments on the gel and easy removal of the solubilized distally derived fragments. Alkaline cleavage of the protein linkage then generates chains/fragments all terminating in xylose at the reducing end irrespective of overall size polydispersity. This approach has similarities to that of Radoff & Danishefsky (1982) for preparing solid-phase heparin, although the latter involved chemical modification of the reducing termini of the free heparin chains before coupling them to a specially modified Sepharose gel. The latter method gave very low yields of coupled material (1–2%) compared with our 40–50%. Both procedures allow the investigation of variation in chemical composition with distance along the polymer chain, relative to a defined reference point. In the present study this reference point is the unique xylose residue common to the reducing termini of all protein-linked heparan sulphate chains. In the case of commercial heparin the majority of the chains have already been excised from a native proteoglycan, and therefore the reference points are not necessarily unique but serve mainly as a method of identifying the reducing termini and thereby orienting the chains.

Our aim was to investigate sulphate distribution within the chains derived from a specific heparan sulphate proteoglycan, namely the detergent-extractable proteoglycan from confluent adult human skin fibroblasts. The $^3\text{H}/^{35}\text{S}$ ratio of the protein-bound oligosaccharide fragments produced by partial nitrous acid hydrolysis and then released by alkaline elimination gives an indication of the distribution of sulphate groups along the original polysaccharide chain. If the ratio were constant throughout the range of oligosaccharide sizes, this would imply a relatively uniform distribution of sulphate groups. However, the sharp increase in $^3\text{H}/^{35}\text{S}$

ratio towards the lower-molecular-size range of fragments indicates the presence of sulphate-deficient domains towards the protein-linkage region of heparan sulphate. This was confirmed by determining the molecular size and sulphate content of the carbohydrate fragment(s) that remained on the Sepharose-bound proteoglycan after complete nitrous acid hydrolysis. The predominant component that was released from such proteoglycan by alkaline elimination was a non-sulphated oligosaccharide (Fig. 2b) that was completely degraded by heparan sulphate lyase (Fig. 2c), and that corresponded in molecular size to an extended *N*-acetylated sequence of eight disaccharide units attached to the protein linkage of the polymer chain, i.e.:

The identity of the non-reducing terminal uronic acid is not known and may be variable. The xylose residue may also be phosphorylated at C-2, as has been demonstrated for bovine lung heparan sulphate (Fransson *et al.*, 1985). Considering that the reducing terminal sequence differs slightly from that of the oligosaccharide series used as molecular-size calibrants of Bio-Gel P-30 (i.e. Gal-Gal-Xyl as compared with GlcNAc-GlcA-aMan), there is still a slight uncertainty in size attribution, probably of the order of $\pm$ one disaccharide unit. The nitrous acid-resistant component did not give a perfectly symmetrical peak (Fig. 2b), as some 'tailing' was evident on the lower-molecular-size side. This suggests that a small proportion of chains have slightly shorter *N*-acetylated domains proximal to the protein core. However, the important features are, firstly, the extended length of the *N*-acetylated domain, and, secondly, the apparently strong conservation of this domain among the vast majority of the chains derived from this proteoglycan species. Interestingly, early chemical analyses of heparan sulphate fractions derived from human aorta (Knecht *et al.*, 1967) and bovine lung (Cifonelli, 1968) had indicated an association of serine and neutral sugars with long nitrous acid-resistant oligosaccharides. These were estimated to contain an average of six to seven *N*-acetylated disaccharide units (Cifonelli, 1968). However, the heterogeneity of these structures is not clear, and, being tissue-derived, they will be of mixed cellular origin(s). An enrichment of non-sulphated disaccharides has also been found in the vicinity of the protein core in the glomerular basement-membrane heparan sulphate proteoglycan (Parthasarathy & Spiro, 1984).

Comparison of the size of the fibroblast domain with the complete spectrum of *N*-acetylated fragments generated by limit nitrous acid digestion of free chains indicates that this particular domain may be the largest *N*-acetylated sequence commonly present within the polymer. It can also be estimated from gel-filtration profiles (Figs. 2b and 3a) that the *N*-acetylated domain of eight disaccharide units represents approx. 14% of the ^3H label in the heparan sulphate chain. If, as inferred, every chain possesses one such *N*-acetylated domain, then the average M_r value for the intact chains (using the sulphate content quoted by Gallagher & Walker, 1985) would be approx. 23000. This is in good agreement with the value of 20000 estimated by gel chromatography of intact chains from embryonic skin fibroblasts (Carlstedt *et al.*, 1983).

The existence of a conserved *N*-acetylated domain of

specific length adjacent to the protein-linkage sequences poses important questions concerning the biosynthesis of mature heparan sulphate. It is possible that such a sequence is generated by a highly ordered enzyme mechanism that accurately and reproducibly positions an *N*-sulphate group at a defined distance from the protein core. Alternatively, this sequence may be the result of a restricted access, albeit a consistent one, of the *N*-deacetylase and/or *N*-sulphotransferase to the proximal region of the polymer. Although there are as yet no clear data supporting either possibility, structural comparisons can be made with heparin proteoglycan that may be of relevance. It has long been known from chemical analyses of mucosal heparin, in which *N*-acetylated disaccharides are comparatively rare, that between one and three *N*-acetylated hexosamine residues exist immediately adjacent to the GlcA-Gal-Gal-Xyl-Ser linkage sequence (Lindahl, 1966a,b; Cifonelli & King, 1972). This suggests that the proximity of *N*-sulphation is not determined by the protein-linkage sequence as such. Considerable steric problems may, however, exist as a result of a high frequency and close spacing of substitution of the polypeptide with glycosaminoglycan chains. There is some evidence for the clustering of glycosaminoglycan chains within proteoglycans, including the embryonic skin-fibroblast (Carlstedt *et al.*, 1983) and glomerular basement-membrane (Parthasarathy & Spiro, 1981) heparan sulphate proteoglycans. However, the density of substitution in such heparan sulphate proteoglycans cannot be significantly greater than that found in the heparin proteoglycan, where it is proposed that approximately two-thirds of the serine residues of a predominantly Ser-Gly polypeptide repeat sequence are substituted with heparin chains (Robinson *et al.*, 1978; Metcalfe *et al.*, 1980). Conversely, it may be possible that the structural differences between these heparan sulphate and heparin chains may reflect differences in access imposed by the physical organization of the membrane-bound polymer-modifying enzymes (see Lindahl *et al.*, 1986) within the relevant cell type, rather than by the conformations of the substrates.

In conclusion, an extended *N*-acetylated domain sited directly adjacent to the protein linkage and comprising an estimated eight disaccharide unit repeats is a highly conserved feature of virtually all the heparan sulphate chains derived from the cell-associated fibroblast proteoglycan. It will be of interest to ascertain whether this is the sole non-random structural domain. It has been suggested that heparinase cleavage sites [GlcNSO₃(6-OSO₃) α 1 $\rightarrow$ 4IdoA(2-OSO₃)] in heparin are randomly distributed (Linhardt *et al.*, 1985). However, analyses of the less-commonly occurring heparin chains with strong anti-coagulant activity have indicated the existence of certain 'ordered' structural features. Glucuronate content appears to increase with distance from the reducing

terminal of these heparin molecules (Radoff & Danishefsky, 1985), and the antithrombin-III-binding sites are located at, or near to, the non-reducing termini (Radoff & Danishefsky, 1984). The latter binding sites appear to occur in multiple copies on only a minority of the chains within the heparin proteoglycan (Horner & Young, 1982). Such observations suggest that there may be mechanisms capable of generating non-random distributions of sugar sequences within these glycosaminoglycans.

The secretarial assistance of Mrs. Kathryn Eastwood is gratefully acknowledged.

REFERENCES

- Carlstedt, I., Cöster, L., Malmström, A. & Fransson, L.-Å. (1983) *J. Biol. Chem.* **258**, 11629–11635
- Cifonelli, J. A. (1968) *Carbohydr. Res.* **8**, 233–242
- Cifonelli, J. A. & King, J. (1972) *Carbohydr. Res.* **21**, 173–186
- Fransson, L.-Å., Silverberg, I. & Carlstedt, I. (1985) *J. Biol. Chem.* **260**, 14722–14726
- Gallagher, J. T. & Walker, A. (1985) *Biochem. J.* **230**, 665–674
- Gallagher, J. T., Gasiunas, N. & Schor, S. L. (1983) *Biochem. J.* **215**, 107–116
- Gallagher, J. T., Lyon, M. & Steward, W. P. (1986) *Biochem. J.* **236**, 313–325
- Horner, A. A. & Young, E. (1982) *J. Biol. Chem.* **257**, 8749–8754
- Knecht, J., Cifonelli, J. A. & Dorfman, A. (1967) *J. Biol. Chem.* **242**, 4652–4661
- Lindahl, U. (1966a) *Biochim. Biophys. Acta* **130**, 361–367
- Lindahl, U. (1966b) *Biochim. Biophys. Acta* **130**, 368–382
- Lindahl, U. & Kjellén, L. (1987) in *The Biology of the Extracellular Matrix: Biology of Proteoglycans* (Wight, T. N. & Mecham, R. P., eds.), Academic Press, New York, in the press
- Lindahl, U., Feingold, D. S. & Rodén, L. (1986) *Trends Biochem. Sci.* **11**, 221–225
- Linhardt, R. J., Merchant, Z. M., Rice, K. G., Kim, Y. S., Fitzgerald, G. L., Grant, A. C. & Langer, R. (1985) *Biochemistry* **24**, 7805–7810
- Metcalfe, D. D., Smith, J. A., Austen, K. F. & Silbert, J. E. (1980) *J. Biol. Chem.* **255**, 11753–11758
- Parthasarathy, N. & Spiro, R. G. (1981) *J. Biol. Chem.* **256**, 507–513
- Parthasarathy, N. & Spiro, R. G. (1984) *J. Biol. Chem.* **259**, 12749–12755
- Radoff, S. & Danishefsky, I. (1982) *Anal. Biochem.* **120**, 373–378
- Radoff, S. & Danishefsky, I. (1984) *J. Biol. Chem.* **259**, 166–172
- Radoff, S. & Danishefsky, I. (1985) *J. Biol. Chem.* **260**, 15106–15111
- Robinson, H. C., Horner, A. A., Höök, M., Ögren, S. & Lindahl, U. (1978) *J. Biol. Chem.* **253**, 6687–6693
- Shively, J. E. & Conrad, H. E. (1976) *Biochemistry* **15**, 3943–3950
- Silverberg, I., Havsmark, B. & Fransson, L.-Å. (1985) *Carbohydr. Res.* **137**, 227–238