Biochem. J. (1991) 277, 297–303 (Printed in Great Britain)

Sequence analysis of heparan sulphate indicates defined location of *N*-sulphated glucosamine and iduronate 2-sulphate residues proximal to the protein-linkage region

Jeremy E. TURNBULL*‡ and John T. GALLAGHER†

*Department of Clinical Research and †Cancer Research Campaign, Department of Medical Oncology, University of Manchester Christie Hospital and Holt Radium Institute, Wilmslow Road, Manchester M20 9BX, U.K.

A strategy that we originally used to identify an *N*-acetylated domain adjacent to the protein-linkage sequence of heparan sulphate proteoglycan (HSPG) [Lyon, Steward, Hampson & Gallagher (1987) Biochem. J. **242**, 493–498] has been adapted for analysis of the location of $GlcNSO_3$-HexA and $GlcNSO_3(\pm 6S)$-IdoA(2S) units most proximal to the core protein. [^{3}H]Glucosamine-labelled HSPG from human skin fibroblasts was depolymerized by using HNO_2 or heparinase under conditions that allowed cleavage of all susceptible linkages. The degraded PG was coupled to Sepharose beads through the protein component, enabling specific recovery of protein-linked resistant oligosaccharides. These were released by treatment with alkaline borohydride and analysed by gel filtration and gradient PAGE. This strategy allowed investigation of the sequence of sugar residues along the chain relative to a common reference point (i.e. the reducing end of the chain). HNO_2 scission confirmed the presence of a well-defined *N*-acetylated sequence predominantly 9–12 disaccharide units in length proximal to the core protein. Heparinase scission produced two classes of oligosaccharides (M_r approx. 7000 and 15000) with the general formula:

$$IdoA(2S)\text{-}GlcNSO_3\text{-}[HexA\text{-}GlcNR]_n\text{-}HexA\text{-}GlcNSO_3\text{-}[HexA\text{-}GlcNAc]_{9-12}\text{-}GlcA\text{-}Gal\text{-}Gal\text{-}Xyl$$

in which the average value for n is 1–2 for the 7000-M_r species and approx. 22 for the 15000-M_r species. The latter oligosaccharides extend to about one-third of the total length of the HS chains (M_r approx. 45000). HNO_2 scission of these oligosaccharides enabled hypothetical models for their sequence to be proposed. The general arrangement of *N*-sulphated and *N*-acetylated disaccharides between the proximal $GlcNSO_3$ and terminal IdoA(2S) residues of the 15000-M_r fragment was similar to that in the original polysaccharide, suggesting the possibility of a tandemly repeating pattern in the sequence of HS.

INTRODUCTION

In the past few years a number of distinct structural features have been recognized in glycosaminoglycans [for reviews see Gallagher *et al.* (1986), Fransson (1989) and Gallagher (1990)]. In corneal keratan sulphate the highly sulphated disaccharides GlcNAc(6S)-β1,3Gal(6S) are located in block sequences distal to the protein-linkage region, where the disaccharide units are largely unsulphated (Oeben *et al.*, 1987). A long unsulphated sequence of disaccharide units is contiguous with the protein-linkage region of HS in fibroblast HSPGs (Lyon *et al.*, 1987) and a non-random 'periodic' distribution of GlcA residues has been detected along the dermatan sulphate chain (Fransson *et al.*, 1990). A specific pentasaccharide sequence determines the high-affinity binding of heparin to antithrombin (Lindahl *et al.*, 1984), this sequence being found only in a minority of heparin chains, often in multiple copies (Jacobsson *et al.*, 1986). It is still not clear whether the intra-chain distribution of this sequence is random or regulated (Radoff & Danishefsky, 1984; Linhardt *et al.*, 1985, 1989; Oscarsson *et al.*, 1989).

On the basis of specific enzyme and chemical scission data, we proposed that HS is an ordered polymeric structure in which the majority of the amino and ester-linked sulphate groups are clustered in a series of short domains widely separated by long oligosaccharide sequences with a low sulphate content (Gallagher *et al.*, 1990; Turnbull & Gallagher, 1990, 1991). The sulphated regions contain the linkages susceptible to the enzyme heparinase, which cleaves disaccharides of structure $GlcNSO_3(\pm 6S)$-α1,4IdoA(2S) (Linhardt *et al.*, 1990). In the present investigation our aim was to gain more information on the location of these disaccharide units by determining their proximity to the core proteins of HSPG. The proximal location of these specific disaccharide units was compared with those of general structure $GlcNSO_3$-HexA, identified by HNO_2 hydrolysis. We have used a sequencing strategy in which the protein-linked reducing end of the HS chain serves as a common reference point for determining the location of specific disaccharide units downstream (i.e. towards the non-reducing end of the chain) from the protein-linkage region (Lyon *et al.*, 1987).

EXPERIMENTAL

Materials

D-[1-^{3}H]Glucosamine (specific radioactivity 27 mCi/mg) was obtained from Amersham International. Heparinase, heparitinases I and II and chondroitin ABC lyase were obtained from Seikagaku Kogyo Co., Tokyo, Japan. Cell culture supplies,

Abbreviations used: dp, degree of polymerization (i.e. for a disaccharide dp = 2 etc.); Gal, galactose; GlcA, glucuronic acid; GlcNAc, *N*-acetylglucosamine; GlcNAc(6S), 6-*O*-sulphated *N*-acetylglucosamine; $GlcNSO_3$, *N*-sulphated glucosamine; HS, heparan sulphate; HSPG(s), heparan sulphate proteoglycan(s); HexA, hexuronic acid; IdoA, iduronic acid; IdoA(2S), iduronic acid 2-sulphate; PG(s); proteoglycan(s); Xyl, xylose; PBS, phosphate-buffered saline.

‡ To whom correspondence should be addressed.

gel-filtration media, electrophoresis and electrotransfer equipment and materials and fluorography supplies were all as described previously (Turnbull & Gallagher, 1990). Tris, heparin (from pig intestinal mucosa; ammonium salt), chondroitin sulphate, HS (from bovine kidney), proteinase inhibitors and CNBr-activated Sepharose 4B were all supplied by Sigma Chemical Co. All other reagents and chemicals were of AnalaR or AristaR grade from BDH Chemicals.

Cell culture, radiolabelling and preparation of intact heparan sulphate proteoglycans

Confluent cultures of adult human skin fibroblasts were prepared as described previously (Turnbull & Gallagher, 1990), and were incubated for 72 h with [^{3}H]glucosamine (10 μCi/ml). The medium was removed and the cell layers were carefully washed twice with warm (37 °C) PBS (phosphate-buffered saline; Dulbecco's medium A).

The cell layers were then extracted by incubation for 24 h at 4 °C with 1% (v/v) Triton X-100 and 1 M-NaCl in PBS containing a cocktail of proteinase inhibitors (10 mM-EDTA, 5 mM-*N*-ethylmaleimide, 20 mM-benzamidine hydrochloride, 50 mM-6-aminohexanoic acid, 0.25 mM-phenylmethanesulphonyl fluoride, 0.1 mg of soya-bean trypsin inhibitor/ml and 0.1 mg of BSA/ml). The resulting extract was centrifuged (200 *g* for 10 min) to remove any insoluble material, and the supernatant was retained for further purification.

The cell-associated PGs were subjected to initial purification by ion-exchange chromatography. Samples were applied to a DEAE-Sephacel column (1 cm × 5 cm) and washed through with 0.3 M-NaCl in 20 mM-sodium phosphate buffer, pH 6.8, to elute contaminating proteins and hyaluronic acid. Remaining PGs (HS and chondroitin sulphate/dermatan sulphate) were then eluted with 1 M-NaCl in 20 mM-sodium phosphate buffer, pH 6.8, and diluted with 4 vol. of 50 mM-Tris/acetate buffer, pH 8.0, containing proteinase inhibitors at 10% of the standard concentration. Chondroitin sulphate and dermatan sulphate were then degraded by addition of 1 unit of chondroitin ABC lyase and incubation for 16 h at 37 °C. HSPGs were recovered by step elution with 1 M-NaCl from a 2 ml DEAE-Sephacel column after washing with 0.3 M-NaCl to elute chondroitin sulphate/dermatan sulphate oligosaccharides and peptide fragments. After addition of 20 μg of unlabelled chondroitin sulphate/ml of carrier and dialysis against distilled water, the radiolabelled HSPGs were freeze-dried in preparation for further analysis.

Depolymerization of HSPGs

Heparitinase (heparitinase I) and heparitinase II were used at a concentration of 20 munits/ml in 100 mM-sodium acetate/0.2 mM-calcium acetate, pH 7.0, containing 1 mg of bovine kidney HS/ml as carrier. Samples were incubated at 37 °C for 16 h. Heparinase was used at a concentration of 20–60 munits/ml in the same buffer as heparitinase but containing 1 mg of heparin/ml as carrier. Enzymes with the appropriate carrier substrates and ^{3}H-labelled HSPGs were incubated at 37 °C for 16 h. Increase in absorbance at 232 nm was measured to ensure that the end point of depolymerization was reached with each batch of digests. Deaminative cleavage was carried out with low-pH HNO_2 as described by Shively & Conrad (1976). Samples were dried down by centrifugal evaporation, reconstituted in 10 μl of 1 M-HNO_2 solution and incubated for 15 min at 20 °C. The reaction was stopped by addition of 2 μl of 1 M-Na_2CO_3.

Gel chromatography

Gel chromatography of intact chains and glycosaminoglycan oligosaccharides was performed on Sepharose CL-6B, Bio-Gel P-6 and Sephadex G-100 columns (120 cm × 1 cm) in 0.5 M-NH_4HCO_3 eluted at a flow rate of 4 ml/h. Fractions (1 ml in each case) were collected and small samples taken for scintillation counting as described previously (Turnbull & Gallagher, 1991). Estimates of the size of fragments resolved on Sepharose CL-6B and Sephadex G-100 were based on the calibrations published by Wasteson (1971) and Laurent *et al.* (1978) respectively.

Oligosaccharide mapping by gradient PAGE

Oligosaccharide mapping by gradient PAGE was performed essentially as described by Turnbull & Gallagher (1988), with minor modifications. Metabolically radiolabelled HS oligosaccharide samples were electrophoresed (500 V for 20 h) on a 20–26% gradient polyacrylamide gel (32 cm × 16 cm × 0.75 mm) and transferred by electroelution on to positively charged nylon membranes (Turnbull & Gallagher, 1990). Separated oligosaccharides were detected by fluorography. Fluorographs were scanned with a Shimadzu scanning densitometer at 550 nm (0.05 mm × 2 mm slit width).

Coupling of HSPG protein core to Sepharose

After depolymerization of the HS chains with different scission reagents, HSPGs were coupled to CNBr-activated Sepharose via their protein cores according to the manufacturer's specifications. Briefly, the dry gel was reconstituted and washed in 1 mM-HCl, then washed and suspended in coupling buffer (0.2 M-$NaHCO_3$/0.5 M-NaCl, pH 8.5). This gel suspension was then mixed immediately with HSPG dissolved in coupling buffer and rotated end-over-end for 16 h at 4 °C. Remaining active groups were blocked by mixing the gel for a further 16 h at 4 °C with 0.1 M-Tris/0.5 M-NaCl, pH 8.0. The gel suspension was then transferred to a 10 ml Econocolumn (BioRad Laboratories) and the supernatant was eluted. After being washed with three cycles of alternating pH (0.5 M-acetate/0.5 M-NaCl, pH 4, and 0.1 M-Tris/0.5 M-NaCl, pH 8.0), the gel was stored in the latter buffer. Coupling efficiencies were routinely in the range 40–50%. As described previously, the method only immobilizes those oligosaccharides that are covalently linked to the HSPG core protein (Lyon *et al.*, 1987).

The protein-linked HS oligosaccharides were released from Sepharose beads by treatment with 5 vol. of alkaline borohydride solution (50 mM-NaOH/1 M-$NaBH_4$). After incubation at 45 °C for 16 h, the mixture was neutralized by addition of acetic acid, and the released oligosaccharides were eluted from the Sepharose on an Econocolumn. After addition of 10 μg of chondroitin sulphate/ml as carrier, the samples were dialysed against distilled water with Spectrapor 6 1000-M_r cut-off membrane tubing (Pierce Chemical Co.), and freeze-dried.

RESULTS

Preparation of end-referenced oligosaccharides from HSPGs

^{3}H-radiolabelled HSPGs secreted by human skin fibroblasts were partially purified from a cell-layer extract by ion-exchange chromatography, and contaminating glycosaminoglycans were removed by treatment with chondroitin ABC lyase in the presence of proteinase inhibitors. The HS chains from these proteoglycans were eluted as a symmetrical peak from Sepharose CL-6B ($K_{av.}$ 0.34) equivalent to a mean M_r of 45 000 (Fig. 1*a*).

The ^{3}H-labelled HSPGs were treated with HNO_2, heparitinase II or heparinase, and depolymerization of all susceptible linkages was checked by comparing Bio-Gel P-6 chromatography profiles with those described previously for this HS (Turnbull & Gallagher, 1990, 1991). Material that was eluted in the void volume, containing both oligosaccharides bound to protein and

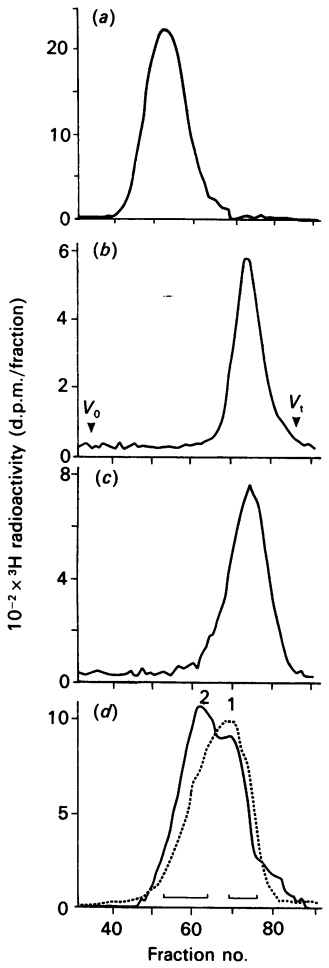

Fig. 1. Gel filtration on Sepharose CL-6B of end-referenced oligosaccharides obtained by complete depolymerization with different reagents

Radiolabelled HSPG was subjected to complete depolymerization by low-pH HNO_2, heparitinase II or heparinase. The corresponding end-referenced oligosaccharides were obtained by coupling of the degraded HSPGs to Sepharose after Bio-Gel P-6 chromatography, and release by alkaline borohydride of the fragments attached to the gel-bound protein cores. The M_r values of these fragments (see Table 1) were estimated by elution from a Sepharose CL-6B column (1 cm × 120 cm). The elution profile for intact HS chains released from the HSPG by alkaline borohydride treatment is shown in panel (*a*). The other panels show the profiles for end-referenced oligosaccharides obtained by treatment of the HSPG with (*b*) low-pH HNO_2, (*c*) heparitinase II or (*d*) heparinase. The fractions indicated by the bars in panel (*d*) (corresponding to Oligo-1 and Oligo-2) were collected for further analysis. The superimposed dotted curve in panel (*d*) shows the gel-filtration profile of oligosaccharides derived from the non-reducing ends and internal sections of HS chains by heparinase depolymerization. These oligosaccharides represent that fraction of material that failed to bind to CNBr-activated Sepharose after heparinase treatment of the HSPG. This material was treated with alkaline borohydride and subjected to gel filtration as above.

'free' oligosaccharides of size dp ⩾ 14, was pooled and protein-linked fragments were isolated by coupling to CNBr-activated Sepharose (see the Experimental section). After extensive washing to remove uncoupled material, oligosaccharides attached to Sepharose-bound protein were released by alkaline borohydride treatment. By definition, all the latter contain intact xylitol residues at their reducing ends, and can therefore be described as 'end-referenced' oligosaccharides. It follows that, for a particular scission technique, the size of these fragments defines the location

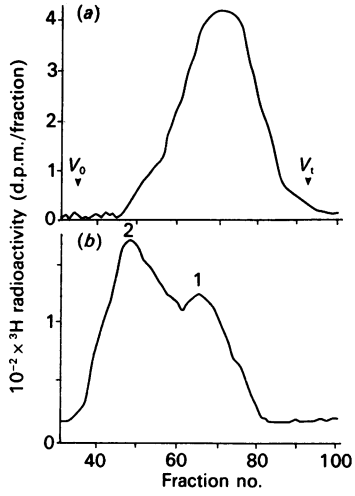

Fig. 2. Gel filtration on Sephadex G-100 of end-referenced oligosaccharides obtained by complete depolymerization with different reagents

Radiolabelled HSPG was subjected to complete depolymerization by (*a*) low-pH HNO_2 and (*b*) heparinase. The corresponding end-referenced oligosaccharides were prepared (see Fig. 1) and their M_r values (see Table 1) were estimated by elution from a Sephadex G-100 column (1 cm × 120 cm).

Table 1. M_r estimates for end-referenced oligosaccharides derived by complete depolymerization with HNO_2, heparitinase II and heparinase

End-referenced oligosaccharides were derived by complete depolymerization with the reagents shown as described in the text. In the case of heparinase these oligosaccharides produce a bimodal distribution on both Sepharose CL-6B and Sephadex G-100 (Figs. 1*d* and 2*b*), designated as Oligo-1 and Oligo-2. M_r estimates are based on the calibrations published by Wasteson (1971) and Laurent *et al.* (1978). Conditions used were not identical with these published calibrations, and the estimates are therefore only approximate.

				Heparinase	
		HNO_2	Heparitinase II	Oligo-1	Oligo-2
Sepharose CL-6B	$K_{av.}$	0.77	0.77	0.70	0.56
	M_r	4800	4800	7500	15000
Sephadex G-100	$K_{av.}$	0.60	–	0.52	0.25
	M_r	5000	–	6500	15000

of the first susceptible linkages downstream from the xylitol moiety.

Estimation of the size of end-referenced oligosaccharides derived by different depolymerization techniques

Initial characterization of end-referenced oligosaccharides was achieved by gel filtration on Sepharose CL-6B (Fig. 1) and Sephadex G-100 (Fig. 2). The M_r estimates derived from these separations are summarized in Table 1.

HNO_2-resistant oligosaccharides

The end-referenced glycans obtained by HNO_2 hydrolysis were eluted as a symmetrical peak after gel filtration on Sepharose CL-6B (Fig. 1*b*) or Sephadex G-100 (Fig. 2*a*), and had an M_r of approx. 4800–5000. Since HNO_2 hydrolysis specifically breaks the HS chain at *N*-sulphated disaccharides, these fragments will

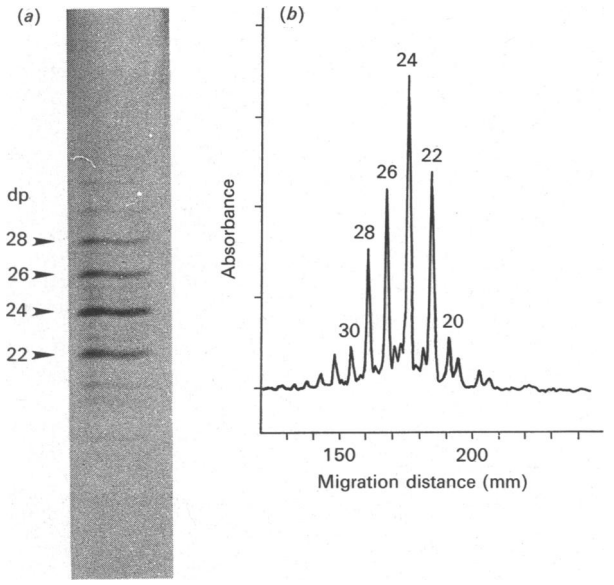

Fig. 3. Oligosaccharide mapping of the location of the first $GlcNSO_3$ residue downstream from the protein core

(*a*) Oligosaccharide mapping of reducing-end oligosaccharides resistant to low-pH HNO_2 treatment. End-referenced oligosaccharides were prepared from radiolabelled HSPG as described in Fig. 1, resolved on a 20–60% gradient PAGE gel (10000 V·h; 27 cm resolving gel, 5 cm stacking gel) and electrotransferred to nylon membrane (10 V for 4 h) for fluorography. The dp sizes shown were estimated from a ladder series of ^3H-labelled hyaluronic acid oligosaccharide standards (derived by partial hyaluronidase scission) resolved on the same gel. (*b*) Densitometric scan of the oligosaccharide map shown in (*a*), with the dp size of the major peaks indicated.

be devoid of *N*-sulphate groups and their molecular size defines the position of the first downstream $GlcNSO_3$ unit. The HNO_2-resistant fragments have the general structure:

$$HexA \rightarrow GlcNAc \rightarrow [GlcA \rightarrow GlcNAc]_n \rightarrow GlcA \rightarrow Gal \rightarrow Gal \rightarrow Xyl$$

Assuming an M_r of 395 for the *N*-acetylated disaccharide units (i.e. including an NH_4^+ counterion), the M_r corresponds to fragments 12–13 disaccharide units in length, containing an extended sequence of 10–11 *N*-acetylated disaccharide units ($n = 9$–10) adjoining the protein-linkage tetrasaccharide. End-referenced oligosaccharides obtained by complete depolymerization with heparitinase II gave a profile on Sepharose CL-6B identical with those derived by deaminitive hydrolysis (Fig. 1*c*). Since unsulphated *N*-acetylated sequences are largely resistant to this enzyme (Linhardt *et al.*, 1990), this result confirms the above conclusions regarding the size of the extended *N*-acetylated sequence proximal to the protein core.

The end-referenced fragments obtained by complete HNO_2 hydrolysis were further characterized by oligosaccharide mapping on 20–26% gradient PAGE gels. They resolved as a series of discrete bands ranging from 9 to 16 disaccharide units in size (Figs. 3*a* and 3*b*), with the bands corresponding to 11–14 disaccharide units (dp 22–28) being the major constituents (Fig. 3*b*). Minor bands were present between each of the major bands, these possibly being due to variants of the basic extended *N*-acetylated structures differentially substituted with ester sulphate or phosphate groups, the latter being linked to the terminal xylose residues (Fransson *et al.*, 1985). Further treatment of the oligosaccharides with low-pH HNO_2 did not result in any change in the banding pattern, whereas heparitinase treatment eliminated all the bands (results not shown). These results indicate clearly that, although there is some variation in position of the most proximal *N*-sulphate groups, they are predominantly separated from the protein-linkage tetrasaccharide of HS by a sequence of 9–12 *N*-acetylated disaccharide units.

Heparinase-resistant oligosaccharides

In order to locate the first downstream IdoA(2S) residues, heparinase was employed as the scission reagent (Linhardt *et al.*, 1990). The end-referenced oligosaccharides derived by prolonged incubation with heparinase produced a bimodal distribution on both Sepharose CL-6B (Fig. 1*d*) and Sephadex G-100 (Fig. 2*b*). They comprised two partially resolved major oligosaccharide peaks, designated Oligo-1 and Oligo-2, with estimated M_r of 7000 and 15000 respectively (Table 1). Allowing for the bias of $[^3H]$glucosamine radiolabel (which is directly proportional to oligosaccharide size), the molar ratio of Oligo-1 to Oligo-2 was estimated to be approx. 2:1. The smaller peak (Oligo-1) is thus the major species. The distribution of these oligosaccharides on Sepharose CL-6B was unchanged after prolonged retreatment with further heparinase (result not shown). The data indicate that the first downstream IdoA(2S) units are predominantly clustered in two well-defined locations in two different species of HS chains from skin fibroblasts. These heparinase-resistant oligosaccharides have the general structure:

$$IdoA(2S)\text{-}GlcNSO_3\text{-}[HexA\text{-}GlcNR]_n\text{-}GlcA\text{-}Gal\text{-}Gal\text{-}Xyl$$

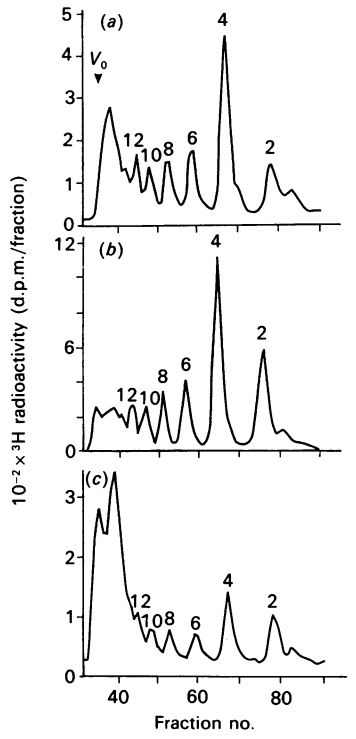

Fig. 4. Gel filtration on Bio-Gel P-6 of the products from HNO_2 hydrolysis of reducing-end heparinase-resistant domains

End-referenced heparinase-resistant domains were prepared and partially resolved by gel filtration on Sepharose CL-6B (see Fig. 1*d*). Fractions corresponding to Oligo-1 and Oligo-2 as shown by the bars were pooled, freeze-dried and depolymerized with low-pH HNO_2. The resulting oligosaccharide mixtures were separated by gel filtration on a Bio-Gel P-6 column (1 cm × 120 cm). (*a*) Oligo-2 (15000-M_r domain); (*b*) intact HS chains; (*c*) Oligo-1 (7000-M_r domain).

As expected the IdoA(2S) moieties are more remote from the protein core than the most proximal $GlcNSO_3$ residues. In the HS chains that yield the 15000-M_r heparinase-resistant fragment (Oligo-2 in Fig. 1*d*), the distance of the most proximal sulphated iduronate moieties from the protein is equivalent to one-third of the total length of the polysaccharide chain (M_r approx. 45000).

The end-referenced oligosaccharides obtained by complete heparinase depolymerization were also subjected to gradient PAGE oligosaccharide mapping. The fragments ranged in size from approx. dp 26–28 upwards, but distinct banding was not evident and the sample spread unresolved up the gel (results not shown). The complexity of these oligosaccharide mixtures apparently prevents resolution under the separation conditions used.

Estimation of the size of heparinase-resistant domains located internally and at the non-reducing ends of HS chains

After heparinase treatment of HSPG, the oligosaccharides that failed to bind to CNBr-activated Sepharose would be enriched in fragments derived from central and peripheral (i.e. non-reducing end) locations in the HS chains. Analyses of these fragments on Sepharose CL-6B indicated that they were similar in size to or slightly larger than the Oligo-1 reducing-end fragments (Fig. 1*d*; $K_{av.}$ 0.68, M_r approx. 7000–10000) and distinctly smaller than the 15000-M_r Oligo-2 fragment. This indicates that heparinase-resistant domains as large as Oligo-2 are principally confined to the reducing terminal, whereas other sections of the chain mainly comprise smaller resistant domains (M_r approx. 7000–10000).

Further characterization of heparinase-resistant end-referenced oligosaccharides

Additional information on the structure of Oligo-1 and Oligo-2 was obtained by recovering appropriate fractions from Sepharose CL-6B chromatography (Fig. 1*d*) and subjecting the concentrated preparations to HNO_2 hydrolysis to cleave at *N*-sulphated disaccharides. The breakdown products were separated on Bio Gel P-6 (Fig. 4*a* and 4*c*) and the elution profiles compared with the corresponding profile for the parent HS (Fig. 4*b*). The profiles revealed that both Oligo-1 and Oligo-2 were enriched in large *N*-acetylated sequences (dp $\geqslant$ 16). This was expected because of the presence of the reducing-end *N*-acetylated sequence, which comprises a greater proportion of these oligosaccharides than of the total HS polysaccharide. In addition, there was a depletion of disaccharide units (representing 41% and 43% of susceptible linkages in Oligo-1 and Oligo-2 respectively compared with 49% in intact chains), indicating a lower content of contiguous *N*-sulphated sequences. Oligo-1 contained only minor quantities of intermediate-sized fragments (dp 6–14), whereas Oligo-2 fragments in this size range comprised a significant proportion of the total breakdown products, with tetrasaccharides (dp 4) being the most prominent component (Fig. 4*a*). Indeed, apart from the over-representation of the long *N*-acetylated sequences and a small deficit in disaccharide products (see the Discussion section), Oligo-2 yielded a depolymerization profile that was remarkably similar to that obtained with the original HS chains (compare Figs. 4*a* and 4*b*).

DISCUSSION

The sequence analysis of linear polymers requires a method for generating a common reference point or reading frame from which the sequence can be read. Radiochemical or fluorescent end-labelling procedures have been described that are particularly useful for studies where metabolic incorporation of isotopically labelled precursors is impractical (Fransson *et al.*, 1990; Jackson, 1990). In the present study we have used the core protein of HSPG to enable specific recovery of biosynthetically radiolabelled HS oligosaccharides originating at the point of attachment to the protein component. This approach allows the position of defined downstream linkages most proximal to the core protein to be identified. In principle, partial scission should allow the location of linkages further downstream to be obtained. However, the separation techniques in current use do not yet provide the adequate resolution of large HS oligosaccharides required for complete sequence analysis. Nevertheless, the present approach afforded some revealing insights into the sequence of HS and gave evidence for a possible repeating pattern in the chain structure.

In confirmation of our initial study on the sequence of HS (Lyon *et al.*, 1987), the first position of *N*-sulphation was separated by an average of ten *N*-acetylated disaccharide units from the protein-linkage tetrasaccharide (Figs. 1 and 2). Although gradient PAGE analyses showed minor variation in the location of these proximal *N*-sulphates (Fig. 3), an extended sequence of *N*-acetylated disaccharide units in this position was a highly conserved feature of the HS preparation and corresponds to the longest sequence of this type in the polysaccharide chains (Fig. 4*b*; see also Lyon *et al.*, 1987). Gradient PAGE studies on the reducing-end *N*-acetylated sequence in HS from 210C embryonic mouse fibroblasts (Winterbourne & Mora, 1981) gave very similar results (J. E. Turnbull, M. Lyon and J. T. Gallagher, unpublished work). *N*-Acetyl-rich oligosaccharides adjacent to the protein-linkage sequence have been detected in other species of HS (Cifonelli, 1968; Parthasarathy & Spiro, 1984; Trescony *et al.*, 1989) and may be a general characteristic.

A more complex picture emerged when heparinase was used to locate the first downstream positions of the specific disaccharide substrate $GlcNSO_3(\pm 6S)$-IdoA(2S). Heparinase scission identified two protein-linked oligosaccharides (Figs. 1*d* and 2*b*) with average M_r values of 7000 (Oligo-1) and 15000 (Oligo-2), reflecting a differential clustered location of the enzyme-cleavage sites. Taking into account the *N*-acetylated sequence adjacent to the protein-linkage tetrasaccharide, these heparinase-resistant oligosaccharides have the general structure:

IdoA(2S)-$GlcNSO_3$-[HexA-GlcNR]$_n$-HexA-$GlcNSO_3$-[HexA-GlcNAc]$_{9-12}$-GlcA-Gal-Gal-Xyl

The average values for n are approx. 1–2 and 22 for Oligo-1 and Oligo-2 respectively. Although the polydispersity of the M_r profiles (Figs. 1*d* and 2*b*) suggests some variation in the location of the proximal $GlcNSO_3(\pm 6S)$-IdoA(2S) units, their distribution is clearly non-random and reveals an unexpected polymorphism in HS structure. Several different types of HSPG are synthesized by human fibroblasts in culture (Lories *et al.*, 1989; Schmidtchen *et al.*, 1990), and it is possible that polysaccharide chains with different internal sequences may be attached to separate core proteins.

Oligo-2 (M_r 15000) was the largest heparinase-resistant fragment in skin fibroblast HS and comprised about one-third of the length of the polymer chain. A 15000-M_r oligosaccharide fraction was identified in a previous study of heparinase depolymerization of HS (Turnbull & Gallagher, 1991) but its location in the chain was not established. The results presented here indicate that these large oligosaccharides are mainly confined to the reducing end of HS, linked directly to the protein core. M_r values of large heparinase-resistant oligosaccharides released by heparinase from central and distal regions of the polysaccharide were mainly in the range 7000–10000 (Fig. 1*d*). These fragments represent the

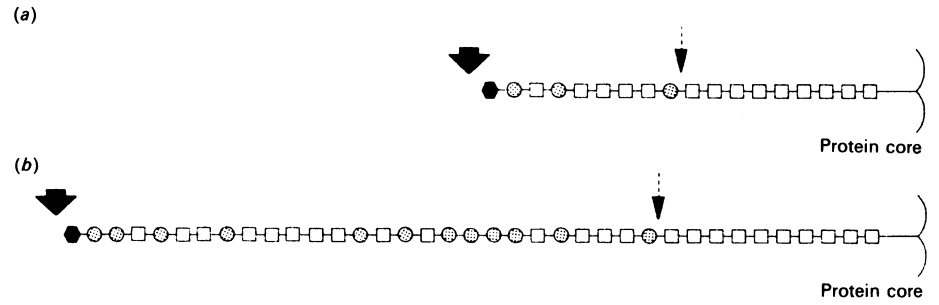

Fig. 5. Model of the hypothetical sequences of Oligo-1 (M_r 7000) and Oligo-2 (M_r 15000) reducing-end heparinase-resistant domains

The models depict hypothetical sequences for each reducing-end heparinase-resistant domain (*a*, Oligo-1; *b*, Oligo-2), and correspond approximately to their average size (M_r 7000 and 15000 respectively). The internal sequences are shown in terms of the location of *N*-sulphated and *N*-acetylated disaccharides, derived from the low-pH HNO_2 scission data (Figs. 3 and 4). The relative proportions and distributions of these disaccharide units in the model sequences for Oligo-1 and Oligo-2 are rationalized from their HNO_2 scission profiles (Figs. 4*a* and 4*c*). Each fragment will terminate with an IdoA(2S) residue at the non-reducing end, and will contain an extended *N*-acetylated sequence at the reducing end. In addition, the tendency of IdoA(2S) residues to be located centrally within *N*-sulphated IdoA-repeat sequences (Turnbull & Gallagher, 1991) strongly suggests that sequences of this type will be located at the non-reducing end of these fragments. The remainder of the actual sequences are speculative. For example, the hypothetical sequence for the Oligo-2 contains a sequence of *N*-sulphated disaccharide units located in a position relative to the protein core similar to that shown at the non-reducing end of Oligo-1. However, since it does not contain an IdoA(2S) residue, it is therefore not susceptible to heparinase cleavage. This type of arrangement could explain the presence of the larger heparinase-resistant oligosaccharides in approximately one-third of the chains. Whether the mixed *N*-sulphated/*N*-acetylated tetrasaccharides occur in contiguous sequences remains unknown, but in view of their abundance (Fig. 4*a*) it is likely that such arrangements occur to some extent. □, *N*-Acetylated disaccharide; ⊙, *N*-sulphated disaccharide; ●, IdoA(2S)→$GlcNSO_3$; ↓, HNO_2-cleavage site proximal to core protein; ▼, heparinase-cleavage site proximal to core protein.

average spacing between heparinase-cleavage sites in HS, as described in an earlier study (Turnbull & Gallagher, 1991).

HNO_2 hydrolysis suggested that the basic arrangement of *N*-acetylated and *N*-sulphated disaccharides in Oligo-2 was very similar to that found in the original HS chain (Figs. 4*a* and 4*b*). The similarity does not of course extend to the large *N*-acetylated sequences (dp ≥ 16 in Fig. 4*a*) contiguous with the protein-linkage region, which are enriched in the oligosaccharide fraction. The data are consistent with the possibility that the sequence pattern of *N*-acetyl and *N*-sulpho derivatives in Oligo-2 is tandemly repeated in the HS chain. Hypothetical sequences for Oligo-2 and Oligo-1 are described in Fig. 5. The sequences are designed to accommodate the average size of the oligosaccharides (Table 1), the HNO_2-scission data (Fig. 4), the extended reducing-end *N*-acetylated domain, the non-reducing terminal sugar produced by heparinase cleavage [IdoA(2S)] and the central location of these cleavage sites within *N*-sulphated sequences (Turnbull & Gallagher, 1991).

In the hypothetical sequence for Oligo-2, an *N*-sulphated region of four disaccharide units is located around a position equivalent to the terminal IdoA(2S) in Oligo-1. We suggest that this sequence may have escaped modification by an iduronate sulphotransferase, thus resulting in these larger heparinase-resistant oligosaccharides. Part of a second *N*-sulphated domain is located at the reducing side of the heparinase-cleavage site in Oligo-2, and the continuing sequence immediately beyond this point is likely to consist of a few additional *N*-sulphated disaccharide units and may also contain a second IdoA(2S) residue. This view is supported by the observation that heparitinase treatment of these oligosaccharides (results not shown) generates IdoA-repeat sequences much smaller than those derived from intact chains (Turnbull & Gallagher, 1990, 1991). In addition, the average length of the $GLcNSO_3$-IdoA-repeat domains in this HS is 4–6 disaccharide units (Turnbull & Gallagher, 1991). The presence of additional *N*-sulphated disaccharides in this position would make up the deficit in relative disaccharide content between Oligo-2 and the HS chain (Figs. 4*a* and 4*b*). Whether sequences composed of alternate *N*-acetylated and *N*-sulphated disaccharides are contiguous (as suggested in Fig. 5) remains unknown, but in view of their abundance in the oligosaccharides (Fig. 4*a*) it is likely that such arrangements occur to some extent. Variations in the lengths of the *N*-acetylated regions dp ≥ 6 (Fig. 4*a*) are likely to be a major factor contributing to the size polydispersity of Oligo-2.

The current findings have important implications for the biosynthesis of HS, a complex process carried out by membrane-bound enzymes in the Golgi system. The polysaccharide is believed to be assembled directly on to the protein core as a non-sulphated precursor called heparan or *N*-acetylheparosan, composed of GlcA-GlcNAc repeats (Lindahl *et al.*, 1986). Conversion into HS begins with the deacetylation and *N*-sulphation of a proportion of the GlcNAc residues, analyses of different HS species suggesting that about 50% of the amino sugars are converted into $GlcNSO_3$ (Gallagher & Walker, 1985). As discussed above, the *N*-sulphate groups tend to be clustered in discrete regions of the polymer chain. Epimerization of GlcA to IdoA and an extensive series of ester-linked sulphations complete HS biosynthesis. The mechanisms that determine the final structure of the polysaccharide are unknown. Proximal *N*-sulphation and particularly iduronate 2-*O*-sulphation occur at an appreciable distance from the core protein (Figs. 1, 2 and 3). This could be due to a restricted yet consistent access of the relevant polymer-modifying enzymes to the internal region of the polysaccharide chain. Alternatively, the sequences may be the products of enzymic processes that accurately and reproducibly target modification events to regions of defined distance from the protein core. Although there is no direct evidence in favour of either possibility, comparisons with the proximal sequences in heparin are quite revealing. In heparin, *N*-sulphation appears to occur within 1–2 disaccharide units of the tetrasaccharide linkage sequence and IdoA(2S) residues are located only 6 disaccharide units downstream (Rosenfeld & Danishefsky, 1988). These internal modifications occur despite a very high density of heparin chains along the central polypeptide of the PG. In HSPGs, the glycosaminoglycans are quite sparsely distributed and would appear to be much more accessible to enzymic processing.

In agreement with the present findings, a study on endothelial

HSPG indicated that heparinase-cleavage sites were separated from the protein core by a relatively large resistant region of undefined size (Nader *et al.*, 1987). It seems likely that a targeted mechanism modifies *N*-acetylheparosan in the vicinity of the protein-linkage region and possibly elsewhere in the chain. In this connection we should mention an alternative possibility for localizing primary *N*-sulphation in HS. In a mastocytoma microsomal preparation it was found that the formation of full-length *N*-acetylheparosan precursors of heparin is a multi-step process, requiring a degree of concurrent *N*-sulphation (Lidholt *et al.*, 1989). Likewise, some *N*-sulphate groups may be introduced at discrete stages of the polymerization of *N*-acetylheparosan precursors in cells that synthesize HS. This mechanism could thus effectively space the *N*-sulphate groups along the polysaccharide and may provide an ordered focus for further modification steps.

In summary, the proximities of $GlcNSO_3$ and IdoA(2S) residues to the core protein of human skin fibroblast HSPGs have been determined. These constituents occupy well-defined repeating positions, $GlcNSO_3$ first appearing about 10 disaccharide units from the protein core and IdoA(2S) appearing at an average of 16 or 35 disaccharide units from the core in different populations of HS. The differential capacity for locating the proximal IdoA(2S) may correlate with different biological properties of the polysaccharide chains. The data are consistent with the HS polysaccharides being ordered polymeric structures formed by tightly regulated biosynthetic mechanisms (Turnbull & Gallagher, 1991). Furthermore, they provide evidence that a rigidly defined primary sequence of sugars is not conserved in all HS chains; rather, they appear to possess recognizable sequence patterns. Analysis of these sequence patterns still presents a formidable analytical problem. However, the strategy described here should lead to further insights concerning the primary structure of HS, and may prove particularly revealing when applied to polysaccharide from a single PG species.

We thank Dr. K. Yoshida for provision of heparitinases, Dr. M. Lyon for helpful discussion and careful reading of the manuscript and Mrs. P. Jones for secretarial assistance. We also thank the Cancer Research Campaign and the Christie Hospital Endowment Fund for financial support.

REFERENCES

Cifonelli, J. A. (1968) Carbohydrate Res. **8**, 233–242
Fransson, L.-Å. (1989) in Heparin (Lane, D. & Lindahl, U., eds.), pp. 115–134, Edward Arnold, London
Fransson, L.-Å., Silverberg, I. & Carlstedt, I. (1985) J. Biol. Chem. **260**, 14722–14726
Fransson, L.-Å., Havsmark, B. & Silverberg, I. (1990) Biochem. J. **269**, 381–388
Gallagher, J. T. (1990) Curr. Opinion Cell Biol. **1**, 1201–1218
Gallagher, J. T. & Walker, A. (1985) Biochem. J. **230**, 665–674
Gallagher, J. T., Lyon, M. & Steward, W. P. (1986) Biochem. J. **236**, 313–325
Gallagher, J. T., Turnbull, J. E. & Lyon, M. (1990) Biochem. Soc. Trans. **18**, 207–209
Jackson, P. (1990) Biochem. J. **270**, 705–713
Jacobsson, K.-J., Lindahl, U. & Horner, A. A. (1986) Biochem. J. **240**, 625–632
Laurent, T. C., Tengblad, A., Thunberg, L., Höök, M. & Lindahl, U. (1978) Biochem. J. **175**, 691–701
Lidholt, K., Kjellén, L. & Lindahl, U. (1989) Biochem. J. **261**, 999–1007
Lindahl, U., Thunberg, L., Backström, G., Riesenfield, J., Nordling, K. & Björk, I. (1984) J. Biol. Chem. **159**, 12363–12376
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
Linhardt, R. J., Cohen, D. M. & Rice, K. G. (1989) Biochemistry **28**, 2888–2894
Linhardt, R. J., Turnbull, J. E., Wang, H. M., Loganathan, D. & Gallagher, J. T. (1990) Biochemistry **29**, 2611–2617
Lories, V., Cassiman, J. & Van Den Berghe, H. & David, G. (1989) J. Biol. Chem. **264**, 7009–7016
Lyon, M., Steward, W. P., Hampson, I. N. & Gallagher, J. T. (1987) Biochem. J. **242**, 493–498
Nader, H. B., Dietrich, C. P., Buonassisi, V. & Colburn, P. (1987) Proc. Natl. Acad. Sci. U.S.A. **84**, 3565–3569
Oeben, M., Keller, R., Stuhlsatz, W. & Greiling, H. (1987) Biochem. J. **248**, 85–93
Oscarsson, L.-G., Pejler, G. & Lindahl, U. (1989) J. Biol. Chem. **264**, 296–304
Parthasarathy, N. & Spiro, R. G. (1984) J. Biol. Chem. **259**, 12749–12755
Radoff, S. & Danishefsky, I. (1984) J. Biol. Chem. **259**, 166–172
Rosenfeld, L. & Danishefsky, I. (1988) J. Biol. Chem. **263**, 262–266
Schmidtchen, A., Carlstedt, I., Malmström, A. & Fransson, L.-A. (1990) Biochem. J. **265**, 289–300
Shively, J. E. & Conrad, H. E. (1976) Biochemistry **15**, 3932–3942
Trescony, P. V., Oegema, T. R., Farnam, B. J. & Deloria, L. B. (1989) Connect. Tissue Res. **19**, 219–242
Turnbull, J. E. & Gallagher, J. T. (1988) Biochem. J. **251**, 597–608
Turnbull, J. E. & Gallagher, J. T. (1990) Biochem. J. **265**, 715–724
Turnbull, J. E. & Gallagher, J. T. (1991) Biochem. J. **273**, 553–559
Wasteson, Å. (1971) J. Chromatogr. **59**, 87–97
Winterbourne, D. J. & Mora, P. T. (1981) J. Biol. Chem. **256**, 4310–4320

Received 26 November 1990/4 January 1991; accepted 15 January 1991