Highly Sensitive Sequencing of the Sulfated Domains of Heparan Sulfate*

(Received for publication, October 28, 1998, and in revised form, January 28, 1999)

Catherine L. R. Merry, Malcolm Lyon, Jon A. Deakin, John J. Hopwood‡, and John T. Gallagher§

From the Cancer Research Campaign and University of Manchester Department of Medical Oncology, Christie Hospital NHS Trust, Manchester, M20 4BX, United Kingdom and ‡The Lymphoma Diseases Research Unit, Department of Chemical Pathology, Women’s and Children’s Hospital, North Adelaide, South Australia 5006, Australia

The heparan sulfates (HS) are hypervariable linear polysaccharides that act as membrane co-receptors for growth factors, chemokines, and extracellular matrix proteins. In most instances, the molecular basis of protein recognition by HS is poorly understood. We have sequenced 75% of the sulfated domains (S-domains) of fibroblast HS, including all of the major ones. This analysis revealed tight coupling of N- and 2-O-sulfation and a low frequency but precise positioning of 6-O-sulfates, which are required functional groups for HS-mediated activation of the fibroblast growth factors. S-domain sequencing was conducted using a novel and highly sensitive method based on a new way of reading the sequence from high performance liquid chromatography separation profiles of metabolically labeled HS-saccharides following specific chemical and enzymatic scissions. The implications of the patterns seen in the sulfated domains for better understanding of the synthesis and function of HS are discussed.

Many biological macromolecules have information encoded in their primary structure. In proteins, the amino acid sequence determines the secondary and tertiary structural characteristics of the folded protein, whereas in DNA and RNA the nucleotide sequence is the means of storing coded information that can be read by the transcriptional and translational machinery of the cell. Information for ligand binding and activation is also present within the structure of complex saccharides (1, 2). Heparan sulfate (HS) is one of a class of polysaccharides known as glycosaminoglycans, and recent research indicates that it expresses important protein recognition domains within its primary sequence (3, 4). However, unlike nucleic acids and protein, there is no established method available to read this primary information.

The near ubiquitous occurrence and strategic positioning of heparan sulfate proteoglycans, both at the surface of most mammalian cells and in the extracellular matrix, is a good indicator of their role in cell-cell recognition and cell-matrix adhesion (5, 6). Heparan sulfate proteoglycans are key sites of interaction and signaling when cells form focal adhesions on extracellular matrix substrates such as fibronectin (7). A very extensive range of growth factors and cytokines, including key angiogenic factors such as basic fibroblast growth factor (bFGF) (8, 9), hepatocyte growth factor (10), and vascular endothelial growth factor (11), have been shown to not only bind HS in vitro but to require its presence as a membrane co-receptor for optimal activation of their cognate, high affinity signaling receptors. The similarity between the co-receptor role of HS and the effect of heparin in facilitating the inactivation of thrombin by antithrombin is indicative of many biological processes being driven by HS catalysis. It may therefore be possible to exploit an understanding of HS-growth factor interactions to modulate the effects of these ligands (12) in a similar way to the current clinical use of heparin. Although there is a substantial body of information concerning the general similarities (and differences) in structure between HS from different sources (13–17), regrettably little is still known about its detailed primary structure. Furthermore, although common molecular designs in the form of spatially discrete sulfated domains (S-domains) can be recognized within the HS family, cells impose their own imprint on this design through variations in sugar sequence and patterns of sulfation, which are known to be of physiological significance.

The HS chain is initially formed in the Golgi as a polymer of alternating N-acetylgalactosamine (GlcNAc) and glucuronic acid (GlcUA) residues. A number of modification steps then follow. First, selected GlcNAc residues undergo linked N-deacetylation and N-sulfation. This N-sulfation is a prerequisite for the epimerization of adjacent GlcUA to iduronic acid (IdoUA). Following epimerization, O-sulfation takes place at various positions around these saccharides, i.e. at C-2 of IdoUA, C-6 (and rarely C-3) of N-sulfated glucosamine (GlcNS), and C-6 of GlcNAc. These modifications, which are clearly nonrandom, are both interdependent and quantitatively incomplete and generate regions of high sulfation that alternate with areas with little or no modification. The controlled variability of S-domain structure and spacing may be a method used by cells to direct which signaling molecules they respond to or to regulate their adhesion and migration in different extracellular matrix environments (14).

What is clear is that HS binds to effector proteins via S-domains with specific sugar sequences and sulfation patterns, and the importance of these interactions has been revealed in studies of development. Studies in mice indicate the key role of

---

* This work was supported by grants from the Cancer Research Campaign (to M. L., J. A. D., and J. T. G.), the National Health and Medical Research Council of Australia (to J. J. H.), the Biotechnology and Biological Sciences Research Council (to C. L. R. M.), and Glaxo Wellcome (to C. L. R. M.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed: Cancer Research Campaign and University of Manchester Department of Medical Oncology, Christie Hospital NHS Trust, Manchester, M20 4BX, UK. Tel.: 44 161 446 3201; Fax: 44 161 446 3269; E-mail:jgallagher@pier.man.ac.uk.

§ The abbreviations used are: HS, heparan sulfate; GlcNAc, N-acetylgalactosamine; GlcUA, glucuronic acid; IdoUA, iduronic acid; GlcNS, N-sulfated glucosamine; bFGF, basic fibroblast growth factor; SAX-HPLC, strong anion exchange high performance liquid chromatography; dp, degree of polymerization or number of saccharide units, e.g. dp6 = hexasaccharide; aMan, anhydromannose; 2-sulfatase, α-1-iduronate-2-sulfatase; iduronidase, α-1-iduronidase; 6-sulfatase, α-1-glucosamine-6-sulfatase; S-domain, sulfated domain; HexA, hexuronic acid.
sulfation of HS for normal kidney growth (18, 19). Also, HS and its associated core proteins are mediators in the decapentaplegic and wingless pathways in *Drosophila* (20–23).

The best known example of the involvement of HS and heparin in molecular recognition is the antithrombin III binding sequence, the elucidation of which was a landmark discovery that required the concerted use of a variety of specialized techniques and many years of work (24). More recently, high affinity binding sites have been identified for bFGF (25, 26). These sites are of relatively simple structure, being mainly composed of 4–5 repeat units of IdooUA(2S)-GlcNS, but intriguingly, these sequences fail to elicit a biological response to bFGF unless substituted by one or more 6-O-sulfate groups (27, 28). This has led to the view that HS may act as a template for cell activation, with the 6-O-sulfate groups being needed to dock an HS-bFGF complex onto the signaling receptor (29). Other members of the FGF family (27) and also unrelated growth factors, e.g. hepatocyte growth factor (10), are responsive to different HS structures than those that affect bFGF activity, but their active-site sequences are unknown. Likewise, the Hep II domain of fibronectin accommodates an extended S-domain of 14–16 sugar residues, but the sulfation pattern of the binding sequence is unknown (30). The development of a rapid sequencing method for HS is vital if we are to understand its detailed structure-function relationships and exploit this knowledge for therapeutic use.

The method we describe here allows the progressive sequencing of regions of metabolically radiolabeled HS purified from cells grown in culture. We have been able to sequence trace quantities of HS oligosaccharides rapidly and accurately. All the major S-domains falling within the size range of hexa- to octasaccharides, together with a proportion of the deacetylated oligosaccharides have been sequenced without selection by ligand binding. This approach has yielded new and intriguing information on the molecular structure of HS.

**EXPERIMENTAL PROCEDURES**

**Reagents**

Cell culture media and donor calf serum were from Life Technologies, Inc. Heparinase I (Flavobacterium heparinum; heparin lyase EC 4.2.9.7), heparinase II (F. heparinum; no EC number assigned), and heparinase III (F. heparinum; heparitin-sulfate lyase EC 4.2.2.8) were all purchased from Grampian Enzymes (Orkney, UK). Bio-Gel P-10 (fine grade) was from Bio-Rad. ProPac PA-1 analytical columns were from Bio-Rad. ProPac PA-1 (4.6 × 250 mm) column pre-equilibrated in distilled water adjusted to pH 5.5 with HCl and eluted with increasing NaCl (maximum of 1.5 M). Oligosaccharides were fractionated using a range of linear gradients optimized for the separation of each size group. Fractions of 0.5 ml were collected, and the aliquots were subjected to scintillation counting (see Fig. 1). By collection of only the central fractions from each peak, single species were obtained that were then desalted by passage over a PD-10 column eluted in H2O and lyophilized. Each peak was checked for purity by reapplication onto a ProPac PA-1 column and elution using the extended gradient conditions described below. Only those oligosaccharides observed as single peaks under these conditions were further analyzed by sequencing.

**Disaccharide Analysis**

Aliquots containing 5000 cpm of 3H were taken from each of the isolated oligosaccharides and lyophilized. These were then digested with a combination of heparinase I, II, and III enzymes and assayed for their specific disaccharide composition as detailed previously (26).

**S-domain Sequencing**

**Partial Nitrous Acid Scission—**Nitrous acid affects deaminative scission of glycosidic linkages of GlcNS-HexA, releasing inorganic 35SO4 and converting GlcNS to anhydro-mannosamine (aMan). Each oligosaccharide was subjected to partial depolymerization using dilute nitrous acid as described (34), with aliquots of the reaction stopped at a number of time points to generate a range of intermediates of the depolymerization process. Briefly, each oligosaccharide was lyophilized, then resuspended in 80 ml of H2O to which was added 10 μl each of 10 mm NaNO2 and 190 mm HCl. At each stop point (usually 30 min, 1 h, 2 h, 3 h, and 4 h), 20 μl of the reaction mixture was removed and added to a common vial containing 25 μl of 0.2 M sodium acetate, pH 5.0. All procedures and reagents were at 4°C. As can be seen in Fig. 2, cleavage of a hexasaccharide, with two internal GlcNS residues, will produce two distinct tetrasaccharides, one of which will contain a Δ4,5-unsaturated uronate at the nonreducing end (from the action of heparinase III) and one with a natural uronate at this position. Also present in the reaction mixture will be three disaccharides and some remaining intact dp6. If a hexasaccharide has one internal GlcNAc residue, then only one tetrasaccharide can be formed. If two internal GlcNAc residues are present, the entire hexasaccharide will be resistant to nitrous acid cleavage.

**Lysosomal Exonuclease Digestions—**Following nitrous acid depolymerization, samples to be digested with enzymes were desalted by passage over PD-10 size exclusion columns eluted with H2O, then lyophilized. Digests were set up in a total volume of 25 μl of 40 mm sodium acetate, pH 4.5. 2-Sulfatase and iduronidase were used either singly or combined sequentially; 6-sulfatase was only used following use of the former two enzymes. Each individual enzyme digest was incubated for 12 h at 37°C. Iduronidase was used at 0.29 milliunits/digest, 2-sulfatase was used at 0.54 milliunits/digest, and 6-sulfatase was used at 0.096 units/digest. Occasionally, glucuronidase (1.000 units/digest) was also used for sequence analysis. After treatment with the enzymes, samples were adjusted to 1 ml by addition of H2O/OHCl, pH 3.5.

**HPLC Separation Conditions for Sequence Analysis—**A single ProPac PA-1 SAX column (4.6 × 250 mm) was used with a linear gradient running from 0–0.75 M NaCl over 110 min. Samples were loaded using a 1-ml sample injection loop. Following sample loading, the loop was washed onto the column with 1 ml of H2O before application of the gradient. The column was eluted at a flow rate of 1 ml min−1, collecting 0.5-ml fractions. Radioactivity in each fraction was measured by scintillation counting.

**RESULTS**

S-domains generated by heparinase III digestion were separated by Bio-Gel P10 gel filtration and SAX-HPLC to yield a range of oligosaccharide species differing in length, sugar sequence, and sulfation pattern. The HPLC separations for dp6–10 are shown in Fig. 1.
The deduced IdoUA-GlcNAc disaccharide must be at the reducing end. The entire sequence of dp6a is therefore ΔUA-GlcNSO$_2$-IdoA(2S)-GlcNSO$_3$-IdoA-GlcNAc. The disaccharide analysis of dp6a (Table I) is compatible with this sequence.

The major octasaccharide 8a was subjected to a similar analysis, generating the profiles seen in Fig. 4. Again, by observing the ways in which each of the peaks seen in the profile of the initial partial nitrous acid scission (Fig. 4, panel A) are affected by the sequencing enzymes, it is possible to assign fragments for sequencing. A number of peaks were observed, but the key ones for sequencing are the hexa- and tetrasaccharides R6 and R4. The enzyme digestions (panels B–D) show that both R6 and R4 contain terminal IdoUA(2S)/IdoA(2S) units. Another tetrasaccharide derived from the giddle of dp8a (thus designated M4) was also present after partial nitrous acid treatment, and this fragment also contained an IdoUA(2S) terminal unit, whereas a third tetrasaccharide (U4) was enzyme-resistant and therefore from the unsaturated end. The presence of three tetrasaccharides (R4, M4, and U4) after partial nitrous acid shows that the three internal GlcN residues of dp8a are N-sulfated. An N-acetylated disaccharide is also present in dp8a, and this must therefore be at the reducing end. This is compatible with the fact that iduronidase yields free GlcNAc from the nonsulfated

For dpl0 only, no fractions were collected for the first 46 min of the gradient.

**Fig. 1.** Preparative SAX chromatography of sized HS oligosaccharides. Sized dp6, dp8, or dp10 oligosaccharide populations were applied to a single ProPac PA-1 column, eluted with a linear gradient of NaCl in MilliQ water, pH 3.5, at a flow rate of 1 ml/min, and 0.5-ml fractions were collected. The gradient used were dp6, 0–1 NaCl over 80 min (A); dp8, 0–1.2 NaCl over 90 min (B); dp10, 0–1.2 NaCl over 180 min (C). Elution profiles were monitored by scintillation counting (solid line, $^{3}$H; dashed line, $^{35}$S), and individual peaks pooled sharply (one or two central fractions only). For dp6 only, no fractions were collected for the first 46 min of the gradient.

**Fig. 2.** Representation of the various fragments generated by partial nitrous acid scission of a hexasaccharide containing two internal GlcNS residues. Partial depolymerization generates two tetrasaccharides and three disaccharides in addition to some uncleaved hexasaccharide. One tetrasaccharide (U4) retains the nonreducing end unsaturated uronate, and the other (R4) retains the reducing end. An aMan residue is formed at all cleaved GlcNS sites. Nitrous acid scission requires a GlcNS residue; therefore, if a hexasaccharide contained a GlcNS residue at position b, only one tetrasaccharide (sequence a–d) would be generated by nitrous acid, together with one disaccharide, IdoUA-GlcNAc (IdoA-GlcNAc). Alternatively, if the GlcNS residue at position d were changed to GlcNAc, nitrous acid would still only generate one tetrasaccharide (sequence c–f) together with one disaccharide ΔUA-aMan. The blue arrows represent exoenzyme digest steps; 1, 2-sulfatase; 2, a combination of 2-sulfatase and iduronidase. Neither of these enzymes has any effect on fragments with a ΔUA residue at position a. The exoenzyme digest steps are used to identify the nonreducing terminal sugars. For example, the tetrasaccharide R4 is digested during each enzyme step (to generate R4’ and R4”), indicating a terminal IdoUA(2S)/IdoA(2S) residue, whereas the disaccharide IdoUA-GlcNAc (IdoA-GlcNAc) is unchanged by the first enzyme step but digested by the second, indicating a terminal IdoUA residue.

Individually purified oligosaccharides from Fig. 1 were partially depolymerized using dilute nitrous acid (Fig. 2), and the scission products were separated by SAX-HPLC. A typical profile is shown for the major hexasaccharide species designated 6a (Fig. 3A). In this profile, we can recognize peaks that correspond to the nonsulfated disaccharides (fraction 18, labeled non-S; IdoUA-aMan, GlcUA-aMan, and ΔUA-aMan all co-elute as a single peak), the sulfated disaccharide IdoUA(2S)-aMan (fraction 45, labeled ISM), free $^{35}$SO$_4$ (fraction 38), two tetrasaccharides (fractions 78 and 95, designated R4 and U4 in Fig. 2), and the original hexasaccharide at fraction 118. The disaccharides were identified by comparison with known standards. R4 and U4 were confirmed as tetrasaccharides by their size elution position on a Bio-Gel P-10 column (data not shown). The appearance of two tetrasaccharides indicates that there are two internal GlcNS residues in dp6a. One of these tetrasaccharide fragments, common with the original dp6 fragment, will contain a Δ4,5-unsaturated uronate at its nonreducing end and will be resistant to lysosomal enzymes; the other will be linked to the reducing end and will be susceptible to the enzymes. It is possible to distinguish between tetrasaccharides R4 and U4 by comparing the profiles in Fig. 3, panels B–D. Digestion with 2-sulfatase causes the R4 peak (at fraction 78) to shift to fraction 45 (R4’); further digestion with iduronidase causes an additional shift to fraction 31 (R4”). This tetrasaccharide is therefore unequivocally identified as being derived from the reducing end of dp6a (thus designated R4), and it contains a terminal IdoUA(2S). The other tetrasaccharide is not affected by the enzymes and is derived from the unsaturated nonreducing end of dp6a (thus designated U4). As expected, the original dp6a is also unchanged after exoenzyme digestion. The disaccharide identified as IdoUA(2S)-aMan moves as expected after 2-sulfatase treatment but not with iduronidase alone, providing a useful internal control for the action of the enzymes. The last piece of information to be extracted from the profiles is the identity of the residues in the reducing terminal disaccharide of dp6a. ΔUA-GlcNAc was seen in the disaccharide analysis of the hexasaccharide, and iduronidase generates a peak in the position of free GlcNAc from the nonsulfated disaccharides. Nitrous acid scission showed that all the internal amino sugars were N-sulfated; therefore

**Sequence Analysis of Radiolabeled HS**
The complete sequence of dp8a is therefore IdoUA(2S)-GlcNS-IdoUA-GlcNAc, and this fact essentially solves the sequence. Other identified fragments (e.g. R4 and U4) are all compatible with the final sequence shown in Fig. 5.

More complex species are also easily solved using this technique. When the products of partial nitrous acid scission of octasaccharide 8d were separated by HPLC, the presence of 6-O-sulfation was indicated by a peak corresponding to the disaccharide IdoUA(2S)-aMan(6S) (Fig. 6). Therefore an additional enzyme digestion step was included in this analysis, combining 6-sulfatase, 2-sulfatase, and iduronidase. Once again, it is easy to identify fragments that are not affected by any of the enzyme treatments (i.e. U6 and the intact dp8) as well as one tetrasaccharide (R4, the same as was seen in 6a and 8a), which shifts with the 2-sulfatase alone and further with 2-sulfatase plus iduronidase. An additional peak is also observed that elutes at fraction 190 (R6) and that shifts with 2-sulfatase to fraction 150 (R'). This latter peak displays only a small shift when subsequently treated with iduronidase (fraction 148, R6''), but the combination of these two enzymes must be removing the hexuronate as the subsequent addition of 6-sulfatase causes a further significant shift to fraction 98 (R6''''). Fragment R6 therefore contains a nonreducing terminal sequence of IdoUA(2S)-GlcNS(6S) and is a reducing end fragment of dp8d. The presence of R6 after nitrous acid scission shows that the unsaturated disaccharide of dp8d is N-sulfated. Confirmation that the unsaturated uronic acid lacks 2-sulfation comes from the presence of the non-sulfated disaccharide peak in all of the profiles (fraction 18); specifically this peak is still observed after the combined 2-sulfatase and iduronidase digest. Combining the sequence data of R6 with the presence of R4, the sequence of dp8d is ΔUA-GlcNSO₃-IdoUA(2S)-GlcNSO₃-IdoUA-ΔUA-GlcNAc. This same principle of observing the results of digesting partial nitrous acid-generated fragments with exoenzymes was therefore used to deduce the sequences of all the oligosaccharides labeled in Fig. 1, panels A and B, as well as two decasaccharides from Fig. 1, panel C (see Table II). In each case, the ability to match common fragments between analyses (e.g. R4 in 6a, 8a, and 10a) and to independently ascertain the disaccharide composition by multienzyme digestion helped to confirm the sequences.

DISCUSSION

The present study describes a method by which trace quantities of metabolically radiolabeled HS oligosaccharides can be

---

**TABLE I**

Disaccharide compositions of HS oligosaccharides

| Oligosaccharide | ΔUA-GlcNAc | ΔUA-GlcNS | ΔUA(2S)-GlcNS | ΔUA(2S)-GlcNS(6S) |
|----------------|------------|----------|--------------|------------------|
| 6a             | 1.17       | 1.00     | 1.09         |                  |
| 8a             | 0.99       | 1.00     | 2.13         |                  |
| 8d             | 0.97       | 1.00     | 1.14         | 0.88             |
| 10a            | 0.94       | 1.00     | 2.92         |                  |

**FIG. 3.** Sequence analysis of hexasaccharide 6a. A, fragment profile generated by partial nitrous acid scission resolved on a ProPac PA-1 column eluted with a linear 0–0.75 M NaCl gradient over 110 min. Aliquots of these fragments (≥5,000 cpm of 3H/run) were subsequently digested with 2-sulfatase (B), iduronidase (C), or both enzymes sequentially (D). Fractions (0.5 min) were collected and counted for radioactivity (solid line, 3H; dashed line, 35S). Identified peaks are nonsulfated disaccharide pool (non-S, fraction 18), free 35SO₄ (fraction 38), free GlcNAc (Nac, fraction 5), disaccharide IdoUA(2S)-aMan (IdoA(2S)-aMan (ISM), fraction 45), tetrasaccharide R4 (green line), tetrasaccharide U4 (blue line), and intact hexasaccharide 6a (fraction 118).
directly sequenced without the need to chemically end-label samples. The idea of using the inherent specificities of the lysosomal exoenzymes to identify the nonreducing terminal residues within HS is nothing new in itself (35). The key to their specific use in the present method is the ability of the controlled partial cleavage with nitrous acid to open up the internal sequence of the oligosaccharide for analysis and the method of reading the sequence from the HPLC separation profiles. The positions of N-sulfation are read from the partial nitrous acid profiles, as illustrated in Fig. 3 for the hexasaccharide 6a. Other sulfation positions and sugar residues are identified by the use of the specific exoenzymes. Although this investigation has so far been limited to dp6–dp10 oligosaccharides, the method could in principle be extrapolated to the analysis of longer complex sequences within the range of resolution of the HPLC column.

Recently, there have been a limited number of advances in the development of methodologies for routinely sequencing small quantities of HS. The use of mass spectrometry, although sensitive (36), suffers from a significant drawback in that it cannot by itself discriminate between isomeric mass forms within the oligosaccharides, such as IdUA/GlcUA stereoisomers, or 3S/6S positional isomers of GlcN. Another approach has been to specifically end-label oligosaccharides with a fluorescent tag. Partial depolymerization then generates a range of fragments similar to those used in this study but simplified in their analysis by the exclusion of all those unlabeled fragments that are not co-terminal with the end-label (37). Tagging methods require additional handling procedures and μg quantities of material for sequencing. The present method is applicable to

\[
\text{[Sequence Analysis of Radiolabeled HS]}
\]

**Fig. 4.** Sequence analysis of octasaccharide 8a. A, fragment profile generated by partial nitrous acid resolved on a ProPac PA-1 column eluted with a linear 0–0.75 m NaCl gradient over 110 min. Aliquots of these fragments (±5,000 kcpm of \(^{3}\text{H}\)/run) were subsequently digested with 2-sulfatase (B), iduronidase (C), or both enzymes sequentially (D). Fractions (0.5 min) were collected and counted for radioactivity (\(^{3}\text{H}\), solid line). Identified peaks are nonsulfated disaccharide pool (fraction 18), free \(^{35}\text{SO}_{4}\) (fraction 38), disaccharide IdUA(2S)-aMan (IdoA(2S)-aMan, fraction 45), tetrasaccharide R4 (green line), tetrasaccharide U4 (blue line), tetrasaccharide M4 (pink line), hexasaccharide R6 (red line), and intact octasaccharide 8a (fraction 165).

**Fig. 5.** Sequence analysis of decasaccharide 10a. A, fragment profile generated by partial nitrous acid resolved on a ProPac PA-1 column eluted with a linear 0–0.75 m NaCl gradient over 110 min. Aliquots of these fragments (±5,000 cpm of \(^{3}\text{H}\)/run) were subsequently digested with 2-sulfatase (B), iduronidase (C), or both enzymes sequentially (D). Fractions (0.5 min) were collected and counted for radioactivity (\(^{3}\text{H}\), solid line; \(^{35}\text{S}\), dashed line). Identified peaks are nonsulfated disaccharide pool (fraction 18), free \(^{35}\text{SO}_{4}\) (fraction 38), disaccharide IdUA(2S)-aMan (IdoA(2S)-aMan, fraction 45), tetrasaccharide R4 (green line), tetrasaccharide U4 (blue line), tetrasaccharide M4 (pink line), hexasaccharide R6 (red line), hexasaccharide U6 (brown line), and intact decasaccharide 10a (fraction 198).
the analysis of trace quantities of HS produced in cultured cells, the main requirement being the availability of approximately 30,000–40,000 cpm of each $^3$H-oligosaccharide to be sequenced (which includes a supportive disaccharide analysis). Also, this is realized without recourse to additional chemical modification, with the increased handling and losses this entails, and with access to only standard laboratory equipment (i.e. HPLC and a scintillation counter). Fluorescent tagging is valuable, however, where the oligosaccharide for sequencing is not available metabolically radiolabeled.

Nitrous acid is a specific scission agent that is reported to act with equal probability at all GlcNS residues within an oligosaccharide (38). It would therefore be theoretically possible to obtain an evenly proportioned mixture of fragments from a partial nitrous acid depolymerization. However, when used to cleave the substrates described above, we observed a degree of unexpected selectivity. With all the substrates, the reducing end tetrasaccharide proved to be more resistant than other regions, with fragment R4 always being more abundant than U4, irrespective of the nature of the oligosaccharide. The presence of the specific R4 peak (IdoA(2S)-GlcNAc) was in fact a significant aid to sequencing, as can be seen in the profiles of Figs. 3–6. It is possible that the local chemical environment of the reducing end makes these tetrasaccharides relatively nitrous acid-resistant. It was also noted that the longer or more complex oligosaccharides were more susceptible to nitrous acid scission, with the reaction times needing to be adjusted accordingly.

A minority of the sequenced S-domains (i.e. 6b, 8b, and 8e) contain a single glucuronate, whose position in the reducing end disaccharide was determined by the additional use of glucuronidase to release free GlcNAc from the nonsulfated disaccharide peak (data not shown). Although not strictly necessary,
it is also possible to positively confirm the unsaturated nature of a completely enzyme-resistant fragment by either specific chemical (mercuric acetate treatment (39)) or enzymatic (Δ-glycuronidase digestion) removal of the terminal unsaturated glycuronate residue.

We have elucidated the structures of all the main hexa- and octasaccharides as well as the two most abundant decasaccharides (Fig. 1), which are compiled in Fig. 7. In total, these represent approximately 75% of all the S-domains (≥dp6) derived from 3T3-fibroblast HS. Although it has been known for some time that the biosynthetic mechanisms that generate HS impose certain constraints upon the resulting structures (14, 40–42), it is only when a number of S-domains of comparable size are actually sequenced that the true pattern of regulation of individual modifications is fully revealed. In the present examples, all sequences terminate in GlcNAc. Internal IdoUA residues flanked by GlcNS are consistently modified by 2-O-sulfate; variation is seen at the reducing end, where the hexuronate positioned between GlcNS and the terminal GlcNAc is either IdoUA, GlcUA, or GlcUA(2S) at the reducing end (8a, b, and c, respectively) can be modified by the addition of a single 6-O-sulfate (giving 8d, e, and f). However, only a minor although variable fraction of each variant was modified, although in each case the central positioning of the 6-O-sulfate was conserved. Tightly controlled, nonrandom patterns such as these may well be highly significant, particularly in relation to basic and acidic FGFs, which require 6-O-sulfates in their activation sequences (27, 28), and also hepatocyte growth factor, where 6-O-sulfate groups appear to be the only sulfates required for binding (10). Because of this regulation, small alterations at the HS biosynthetic level can cause changes that stand out from a background pattern and may be more easily and rapidly recognized by HS-binding proteins. It is becoming apparent that the precise positioning of modifications within the HS chain is important not only for directly which ligand can bind (such as the differences between bFGF and hepatocyte growth factor binding sites) but also for determining whether a bound ligand is activated or held in an inactive state (28, 43, 44). Although we have not as yet attempted to sequence the longer and less abundant S-domains (≥dp12), which will also be involved in the activation of HS binding factor growth factors, it would seem likely that, for at least the 3T3-fibroblast HS, the patterns observed within the smaller S-domains will be mirrored in the larger ones but perhaps with more complex patterns of 6-O-sulfation.

The use of specific exo-glycosidases has, for a number of years, formed the basis of a method for sequencing branched N-linked oligosaccharides. The reagent array method initiated a flood of interest in the field of glycochemistry, as investigators began to realize its potential to open up the information held in these complex branched sugars (45, 46). As a result of the use of this technique, many biological problems have been addressed; for example the importance of the glycosylation of IgG in rheumatoid arthritis (47, 48). Glycosaminoglycan sequence analysis, on the other hand, still retains the reputation of being a somewhat arcane art, with detailed, accurate sequence information only available in very few cases. The technique described here can be used to routinely and reliably sequence HS from cultured cells, the most commonly studied source of HS. It is hoped that this technique will make HS sequencing a practical proposition for any reasonably equipped laboratory interested in the structure and function of these hypervariable regulatory polymers. With only a few modifications, the same method can be applied to saccharides released from HS by other scission methods such as heparinase I or partial nitrous acid. Depending upon the availability of a partial scission method and an appropriate array of specific exoenzymes, this basic protocol could in theory also be adapted to the sequencing of oligosaccharides derived from other complex glycosaminoglycans, e.g. dermatan sulfate.

The data presented in this paper reveals new and important characteristics of S-domain sequence and diversity in 3T3 fibroblast HS. In addition to helping to elucidate specific protein recognition domains in HS, this approach should form the basis for the future elucidation of HS sequence patterns that define different cell types, stages in embryonic development, and disease processes (e.g. tumorigenesis) and may also contribute toward realizing a true whole chain analysis of HS.

Acknowledgments—We thank S. Stringer, D. Pye, and R. Vivés for their helpful advice and K. Beckman and C. Boulter for production of the recombinant enzymes.

REFERENCES

1. Dwek, R. A. (1995) Biochem. Soc. Trans. 23, 1–25
2. Lowe, J. B. (1994) in Molecular Glycobiology (Fukuda, M., and Hindsaul, O., eds) pp. 165–205, IRL Press at Oxford University Press, Oxford
3. Conrad, H. E. (1998) Heparin Binding Proteins, Academic Press, Inc., San Diego, CA
4. Gallagher, J. T. (1998) Trends Glycosci. Glycotechnol. 10, 137–144
5. Bernfield, M., Kokenyesi, R., Kato, M., Hinkes, M. T., Spring, J., Gallo, R. L., and Lose, E. J. (1992) Annu. Rev. Cell Biol. 8, 365–393
6. Kielder, L., and Lindahl, U. (1991) Annu. Rev. Biochem. 60, 443–475
7. Woods, A., McCarthy, J. B., Furcht, L. T., and Couchman, J. R. (1993) Mol. Biol. Cell 4, 605–613
8. Rappaport, A. C., Krutka, A., and Obwin, B. B. (1991) Science 252, 1705–1708
9. Yayon, A., Klagsbrun, M., Esko, J. D., Leder, P., and Ornitz, D. M. (1991) Cell 64, 841–848
10. Lyon, M., Deakin, J. A., Mizuno, K., Nakamura, T., and Gallagher, J. T. (1994) J. Biol. Chem. 269, 11216–11223
11. Tessler, S., Rockwell, P., Hicklin, D., Cohen, T., Levi, B. Z., Witte, L., Lemiachez, I. R., and Neufeld, G. F. (1994) J. Biol. Chem. 269, 12456–12461
12. David, G. P. (1993) FASEB J. 7, 283–303
13. Fransson, L.-A. (1989) in Heparin (Lane, D. A., and Lindahl, U., eds) pp. 115–134, Edward Arnold, London
14. Gallagher, J. T., Turnbull, J. E., and Lyon, M. (1992) Int. J. Biochem. 24, 563–569
15. Lyon, M., Deakin, J. A., and Gallagher, J. T. (1994) J. Biol. Chem. 269, 11208–11215
16. Macarana, M., Sakya, Y., Tawada, Y., Yoshida, K., and Lindahl, U. (1996) J. Biol. Chem. 271, 17804–17810
17. Gallagher, J. T., and Walker, A. S. (1985) Biochem. J. 230, 665–674
18. Bollee, S. L., Fletcher, J. M., Beddington, R. S., and Wilson, P. J. (1998) Gene Des. 12, 1894–1906
19. Davies, J., Lyon, M., Gallagher, J. T., and Garrod, D. (1995) Development 121, 1507–1517
20. Binnari, R. C., Staveley, E. B., Johnson, W. A., Godavarti, R., Sasisekharan, R., and Manoukian, A. S. (1997) Development 124, 2623–2632
21. Hacker, U., Lin, X., and Perrimon, N. (1997) Development 124, 3565–3573
22. Jean, S. M., Nakate, H., Nakamura, J., Anzani, A., Oakes, R., Kalusa, V., Golden, C., and Selleck, S. B. (1997) Development 124, 4115–4120
23. Reichsman, F., Smith, L., and Cumberledge, S. (1996) J. Cell Biol. 135, 821–827
24. Lindahl, U., Thumbreg, L., Backström, G., Riesenfeld, J., Nordling, K., and Björk, I. (1984) J. Biol. Chem. 259, 12368–12376
25. Halsbich, H., Suzuki, S., Saito, T., Tamura, T., Harada, T., Yoshida, K., and Kimata, K. (1992) Biochem. J. 285, 805–811
26. Turnbull, J. E., Fornig, D. G., Wright, L. K., and Wilkinson, M. C., and Gallagher, J. T. (1992) J. Biol. Chem. 267, 10337–10341
27. Galas, R., Macarana, M., Obwin, B. B., Lindahl, U., and Rappaport, A. C. (1993) J. Biol. Chem. 268, 23906–23914
28. Pye, D. A., Vives, R. R., Turnbull, J. E., Hyde, P., and Gallagher, J. T. (1998) J. Biol. Chem. 273, 22936–22943
29. Kan, M., Wang, F., Xu, J., Crabich, J. W., Hou, M., and Koechean, W. L. (1995) Science 259, 1918–1921
30. Walker, A., and Gallagher, J. T. (1996) Biochem. J. 317, 871–877
31. Bielecki, J., Hopwood, J. D., Wilson, P. J., and Anson, D. S. (1993) Biochem. J.
Sequence Analysis of Radiolabeled HS

289, 241–246
32. Unger, E. G., Durrant, J., Anson, D. S., and Hopwood, J. J. (1994) Biochem. J. 304, 43–49
33. Litjens, T., Bielicki, J., Anson, D. S., Friderici, K., Jones, M. Z., and Hopwood, J. J. (1997) Biochem. J. 327, 89–94
34. Radoff, S., and Danishefsky, I. (1984) J. Biol. Chem. 259, 166–172
35. Linker, A. (1979) Biochem. J. 183, 711–720
36. Rhoemberg, A., Ernst, S., Sasisekharan, R., and Biemann, K. (1998) Proc. Natl. Acad. Sci. U. S. A. 95, 4176–4181
37. Turnbull, J. E., Hopwood, J. J., and Gallagher, J. T. (1997) in A Laboratory Guide to Glycoconjugate Analysis (Jackson, P., and Gallagher, J. T., eds) Vol. 9, pp. 261–278, Birkhaeuser Verlag, Basel, Switzerland
38. Bienkowski, M., and Conrad, H. E. (1985) J. Biol. Chem. 260, 356–365
39. Maccarana, M., Casu, B., and Lindahl, U. (1993) J. Biol. Chem. 268, 23898–23905
40. Riesenfeld, J., Hoek, M., and Lindahl, U. (1982) J. Biol. Chem. 257, 7050–7055