THE JOURNAL OF BIOLOGICAL CHEMISTRY

Vol. 263, No. 21, Issue of July 25, pp. 10168-10174, 1988
Printed in U.S.A.

Structural Studies on Sulfated Glycopeptides from the Carbohydrate-Protein Linkage Region of Chondroitin 4-Sulfate Proteoglycans of Swarm Rat Chondrosarcoma

DEMONSTRATION OF THE STRUCTURE GAL(4-*O*-SULFATE)β1-3GALβ1-4XYLβ1-*O*-SER*

(Received for publication, February 16, 1988)

Kazuyuki Sugahara‡§ and Ikuo Yamashina¶

From the ‡Departments of Pediatrics and Biochemistry, The Joseph P. Kennedy, Jr. Mental Retardation Research Center, Pritzker School of Medicine, University of Chicago, Chicago, Illinois 60637 and the ¶Department of Biological Chemistry, Faculty of Pharmaceutical Sciences, Kyoto University, Kyoto 606, Japan

Pieter De Waard‖, Herman Van Halbeek‖, and Johannes F. G. Vliegenthart‖**

From the ‖Department of Bio-organic Chemistry, Utrecht University, Utrecht, The Netherlands

Nonsulfated, monosulfated, and disulfated glycopeptides containing the entire carbohydrate sequence of the glycosaminoglycan-specific linkage region were isolated after exhaustive enzymatic digestions of Swarm rat chondrosarcoma proteoglycans with chondroitinase ABC, papain, and Pronase. Their structures were examined by 500 MHz ^{1}H NMR spectroscopy. The nonsulfated compound has the following structure with trace amounts of a few additional amino acids: $\Delta_{4,5}$-GlcAβ1-3GalNAcβ1-4GlcAβ1-3Galβ1-3Galβ1-4Xylβ1-*O*-Ser. The monosulfated compound has an ester sulfate on C-4 of the GalNAc residue and the disulfated compound has an additional hitherto unrecognized ester sulfate on C-4 of the second galactose residue which is remote from the innermost xylose. This new structure was confirmed by two-dimensional homonuclear Hartmann-Hahn spectroscopy. The molar ratio of the isolated nonsulfated, monosulfated, and disulfated compounds was 53:37:10 based on the serine contents. Biological significance of the newly found sulfated linkage structure is discussed.

Proteoglycans are macromolecular glycoproteins that contain a protein core to which are covalently attached large numbers of side chains of glycosaminoglycans and *N*- and/or *O*-linked oligosaccharides. Although the considerable complexity of proteoglycan structure and composition has been recognized, they are usually classified into several categories based on the attached glycosaminoglycan species including heparin, heparan sulfate, chondroitin sulfate, dermatan sulfate, and keratan sulfate (1). Although these glycosaminoglycans have different repetitive disaccharide units, they are, except keratan sulfate, assumed to be bound to the protein cores through a common glycosaminoglycan-specific linkage region, GlcAβ1-3Galβ1-3Galβ1-4Xylβ1-*O*-Ser, as previously reported by Rodén and co-workers (for a review see Ref. 2).

During biosynthesis of proteoglycans, the common linkage region and characteristic repeating structures of glycosaminoglycan chains are formed by sequential addition of the respective monosaccharide units to the precursors (for a review see Ref. 3). However, the sorting mechanisms in biosynthesis of different glycosaminoglycan chains after synthesis of the common linkage remain enigmatic.

The present work was undertaken to investigate additional structural peculiarities or possible modifications of the linkage region. The high field ^{1}H NMR spectra of the linkage glycopeptides isolated from rat chondrosarcoma proteoglycans revealed a previously unrecognized ester sulfate on C-4 of the second galactose residue which is remote from the innermost xylose.

EXPERIMENTAL PROCEDURES

Materials—Proteoglycan monomers were prepared from rat chondrosarcoma as reported (4) according to Oegema *et al.* (5). Enzymes were obtained from the following sources; Pronase (Calbiochem), chondroitinase ABC (Miles Laboratories), papain (23 units/mg; Worthington Biochemical), sialidase from *C. perfringens* (Sigma) and alkaline phosphatase from *E. coli* (Type III) (Sigma). Dowex 1 × 2 (AG 1 × 2, 200–400 mesh) and Dowex 50W × 2 (200–400 mesh) were purchased from Bio-Rad.

Analytical Methods—Uronic acid, hexose, and sialic acid were determined according to the carbazole method of Bitter and Muir (6) using glucuronic acid as standard, the orcinol-sulfuric acid method of Hewitt (7) with galactose as standard, and the resorcinol method of Jourdian *et al.* (8) using *N*-acetylneuraminic acid as standard, respectively. Amino acid analyses were carried out according to Spackman *et al.* (9).

Chondroitinase Digestion—Proteoglycan subunit (497 mg) was digested with chondroitinase ABC according to Saito *et al.* (10) in a total volume of 6.0 ml of 0.04 M Tris-HCl buffer, pH 8.0, containing 0.05 M sodium acetate, 0.05% bovine serum albumin, and 2.5 units of the enzyme. The digestion was carried out for 20 h at 37 °C; an additional 1 unit (0.1 ml) of the enzyme was added after 10 h. The reaction was terminated by boiling for 2 min and the digests were separated on a column (2.2 × 96 cm) of Sephadex G-50 (fine) in 0.15 M NaCl. The fractions were monitored by UV-absorption at 232 and 280 nm. The excluded materials were dialyzed against water, lyophilized, and redissolved in the chondroitinase buffer. The chondroitinase digestion was repeated twice more in the same manner as above. After the third digestion, no UV-absorbing material was detected at

* This investigation was supported in part by Grants HD-04583 and HD-09402 from the National Institute of Health and Home for Destitute and Crippled Children (U. S. A.). This investigation was supported by the Netherlands Foundation for Chemical Research (SON) with financial aid from the Netherlands Organization for the Advancement of Pure Research (ZWO) and Netherlands Foundation for Cancer Research Grant UUKC 83-13 (KWF). 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.

§ Present address: Dept. of Biological Chemistry, Faculty of Pharmaceutical Sciences, Kyoto University, Kyoto 606, Japan.

** Present address: Complex Carbohydrate Research Center, University of Georgia, Athens, GA 30613.

the column volume when monitored at 232 nm.

Papain Digestion—The chondroitinase ABC-digested proteoglycan (63.2 mg) was incubated with 2 mg of papain in a total volume of 2.0 ml of 0.1 M sodium acetate buffer, pH 5.5, containing 5 mM EDTA and 5 mM cysteine-HCl. The digestion was performed for 45 h at 63 °C; additional 0.8 mg of papain and 1.8 mg of cysteine-HCl were added after 22 and 34 h. Following incubation the sample was lyophilized, mixed with 1.0 ml of 5% trichloroacetic acid and centrifuged in a Beckman microcentrifuge for 10 min. The trichloroacetic acid extract was subjected to gel filtration on a column (1.2 × 106 cm) of Sephadex G-50 (fine) using water as eluent. The fractions which were positive to carbazole were pooled and further digested with Pronase as described below. The trichloroacetic acid treatment and the subsequent gel filtration here and elsewhere were carried out below 4 °C.

Pronase Digestion—Pronase digestion of the papain glycopeptides (52.5 μmol uronic acid) was performed in 1.25 ml of 0.1 M sodium borate buffer, pH 8.0, containing 0.01 M calcium acetate, using 1 mg of Pronase for 45 h; an additional 1 mg of Pronase was added after 20 and 30 h. Following incubation the sample was lyophilized, mixed with 1 ml of 5% trichloroacetic acid, and centrifuged in a Beckman microcentrifuge for 10 min. The precipitate was washed with 0.3 ml of 5% trichloroacetic acid three times and the combined supernatant fluid was chromatographed on Sephadex G-25. The carbazole positive fractions were pooled and retreated with Pronase in the same manner as above to complete the digestion.

Sialidase Digestion—The Pronase digest (30.9 μmol of uronic acid) was incubated with 1.0 unit of sialidase in a total volume of 1.3 ml of 0.04 M sodium acetate buffer, pH 5.1, at 37 °C. The reaction was monitored by determining released sialic acid according to Warren (11). After 24 h of incubation, the reaction mixture was lyophilized, mixed with 0.25 ml of 5% trichloroacetic acid, and centrifuged in a Beckman microcentrifuge for 10 min. The precipitate was washed with another 0.25 ml of 5% trichloroacetic acid and the combined supernatant fluid was applied to a Sephadex G-25 (fine) column (1.2 × 97 cm), which was eluted with 2% acetic acid. The glycopeptide fraction separated from the released sialic acid was recovered and concentrated to dryness.

Ion Exchange Chromatography—The glycopeptide fraction (corresponding to 13.6 μmol Ser and 28 μmol of uronic acid) obtained after sialidase treatment was dissolved in water, and applied to a column (0.7 × 6.5 cm) of Dowex 1 × 2 (Cl^- form). The column was washed with 15 ml of water (a flow-through fraction, not shown) and then eluted by a linear gradient with 50 ml each of water and 1.0 M NaCl. Two-ml fractions were collected and monitored by the carbazole reaction. The separated fractions (D-1–D-5) were desalted on a column (1.2 × 82 cm) of Sephadex G-25 (fine) using 2% acetic acid as eluent and lyophilized.

Paper Electrophoresis—Paper electrophoresis was performed at a potential of 50 V/cm for 2 h on a Whatman 3MM paper in a 0.06 M sodium borate buffer, pH 9.5 (12). Following electrophoresis the paper was dried and glycopeptides were detected with 0.025% fluorescamine in acetone. For preparative purposes the area corresponding to each fluorescent spot was extracted with water and the extract was chromatographed on a Sephadex G-25 column (1.2 × 82 cm) using 2% acetic acid as eluent to remove borate. Alternatively, the extract was passed through a column (0.5 × 6 cm) of Dowex 50 × 2 (H^+ form) and evaporated to dryness with methanol three times.

500-MHz ^1H NMR Spectroscopy—Glycopeptides were repeatedly exchanged in ^2H_2O (99.96% ^2H, Aldrich) with intermediate lyophilization.

^1H NMR spectra were recorded on a Bruker WM-500 spectrometer (SON hf = NMR facility, Department of Biophysical Chemistry, University of Nijmegen, The Netherlands) operating at 500 MHz at a probe temperature of 22 °C. Resolution enhancement of the spectra was achieved by Lorentzian-to-Gaussian transformation (13). The two-dimensional HOHAHA spectrum (14) was recorded on a Bruker AM-500 spectrometer (Dept. of NMR spectroscopie, University of Utrecht, The Netherlands) operating at 500 MHz at a probe temperature of 62 °C. The large HOD signal was suppressed by presaturation during 1.2 s. 469 Spectra of 2048 data points were recorded, with 48 scans/t_1 value. A MLEV-17 mixing sequence of 80 ms was used, the 90° ^1H pulse width was 27 μs. The total measuring time was 12 h. The time domain data were multiplied with a phase shifted sine-bell, phase-sensitive Fourier transformation was performed after zero filling to a 4096 × 2048 data matrix size. Chemical shifts are given relative to sodium 4,4-dimethyl-4-silapentane-1-sulfonate, but were actually measured indirectly to acetone in ^2H_2O (δ = 2.225 ppm) (15).

RESULTS

Preparation of Linkage Glycopeptides—Proteoglycan monomers (497 mg of dry weight) were repeatedly digested with chondroitinase ABC as described under "Experimental Procedures" until no further unsaturated oligosaccharide was formed. The protein core (63.2 mg dry weight) was isolated by gel filtration and digested with papain. When chromatographed on a Sephadex G-50 column, the papain digest showed a single carbazole positive peak, in which essentially all orcinol positive materials were recovered. This peak (52.5 μmol of uronic acid) was further digested with Pronase and chromatographed on a column of Sephadex G-25 (Fig. 1). The major fraction, which contained 82% of the applied uronic acid, was redigested with Pronase. Likewise, a major fraction containing 91% of the applied uronic acid was recovered from the second Pronase digest. This fraction turned out to contain a small yet significant amount of sialic acid (2.5 μmol/30.9 μmol of uronic acid), which was indicative of the contamination with mucin-type glycopeptides. In order to facilitate the separation of the linkage glycopeptides from the mucin-type glycopeptides by anion exchange chromatography, the sample was treated with sialidase. After removal of the released sialic acid on a Sephadex G-25 column, the sample was chromatographed on a Dowex 1 column and separated into six glycopeptide fractions (a flow-through fraction and D-1–D-5) as shown in Fig. 2 with quantitative recovery of uronic acid. The chemical composition of each fraction is shown in Table I. The three major fractions (D-1, D-3, and D-4) were subjected to borate paper electrophoresis at pH 9.5. The results are illustrated in Fig. 3. D-1 and D-3 gave rise to only one major spot (R_{serine} = 1.55 and 2.05, respectively) as detected with fluorescamine. D-4 gave two major spots; the faster (R_{serine} = 2.45) and slower (R_{serine} = 2.05) migrating spots were referred to as D-4F and D-4S, respectively. These glycopeptides were isolated by preparative borate paper electrophoresis followed by removal of borate as described under "Experimental Procedures." Finally, each glycopeptide was passed through a small column (0.7 × 27.5 cm) of Sephadex G-25 using double distilled water to achieve complete desalting. The yield of each glycopeptide was 2.33 μmol (D-1), 1.33 μmol (D-3), 0.42 μmol (D-4F), and 0.28 μmol (D-4S), respectively, based on the serine content.

Carbohydrate Composition—The carbohydrate composition of the isolated linkage glycopeptides is shown in Table II. The

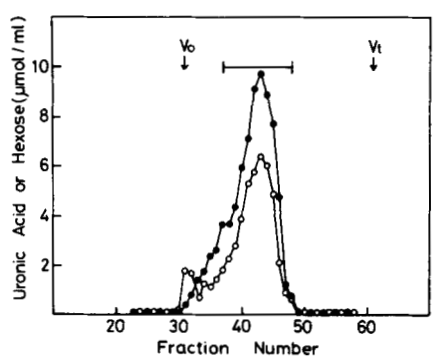


FIG. 1. **Gel filtration of the Pronase digest.** The Ponase digestion was carried out as described under "Experimental Procedures." Following trichloroacetic acid precipitation the supernatant fluid was chromatographed on a column (1.2 × 97 cm) of Sephadex G-25 (fine) with water as eluent. One-ml fractions were collected, and an aliquot was used to monitor uronic acid (○) and hexose (●). V_o and V_t were determined with blue dextran and [^3H]glucosamine, respectively.

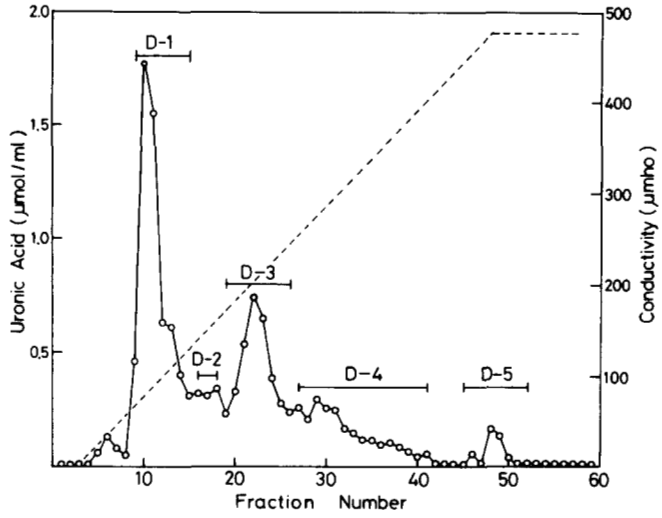


FIG. 2. **Dowex column chromatography of the glycopeptides.** The glycopeptide fraction obtained after the sialidase treatment was fractionated on a Dowex 1 column by a salt gradient elution and the fractions were monitored by the carbazole reaction as described under "Experimental Procedures."

TABLE I

Chemical composition of the glycopeptide fractions obtained by Dowex column chromatography

Serine was determined by amino acid analysis. Trace amounts of some other amino acids (Cys, Glu, Gly, Asp, Leu, Thr, and Val) were also detected. Uronic acid was quantitated by the carbazole reaction. The color formation due to hexose has not been corrected. Hexose was estimated by the orcinol-sulfuric acid method and has not been corrected for the color formation due to saturated or unsaturated disaccharide, or xylose. Numbers in parentheses represent the molar ratio normalized to Ser 1.00.

Component	D-FT[a]	D-1	D-2	D-3	D-4	D-5	Total
				μmol			
Ser	0.36 (1.00)	3.80 (1.00)	0.63 (1.00)	2.09 (1.00)	1.22 (1.00)	0.48 (1.00)	8.58
Uronic acid	1.74 (4.83)	11.75 (3.09)	1.40 (2.22)	5.52 (2.64)	3.60 (2.95)	1.37 (2.85)	25.38
Hexose	6.59 (18.30)	17.30 (4.55)	2.29 (3.63)	8.11 (3.88)	5.48 (4.49)	2.29 (4.77)	42.06

[a] D-FT, a flow-through fraction (not shown in Fig. 2).

molar ratio of Xyl[1] and Gal is 1:2, which suggests the presence of the partial structure: -Galβ1-3Galβ1-4Xylβ1-*O*-Ser, *i.e.* the usually occurring linkage region in chondroitin sulfate proteoglycans (1). Furthermore, one GlcA and one GalNAc residue are present in each glycopeptide.

The occurrence of $\Delta_{4,5}$-GlcA, possibly as a result of chondroitinase digestion, could not be established by the gas-liquid chromatography as it was not demonstrated either for a reference component $\Delta_{4,5}$-GlcAβ1-3GlcNAc. However, its occurrence was evident from the UV-absorbance and was confirmed by NMR analyses below.

500-MHz ^1H NMR Spectroscopy—The ^1H NMR spectrum of fraction D-1 is presented in Fig. 4*A* and the ^1H NMR data in Table III. Comparison of the spectral data with those of the reference compound Galβ1-3Galβ1-4Xylβ1-*O*-Ser (17) allows a straightforward assignment of the signals of the Xyl-*1*

[1] The abbreviations used are: Xyl, D-xylose; $\Delta_{4,5}$-GlcA, D-gluco-4-enepyranosyluronic acid; GalNAc, 2-deoxy-2-*N*-acetylamino-D-galactose; HOHAHA, homonuclear Hartmann-Hahn.

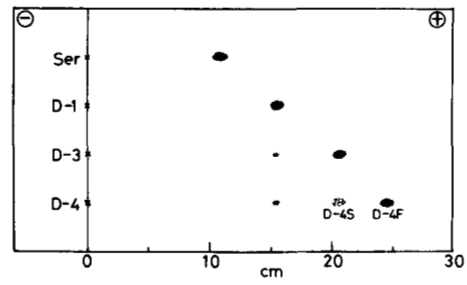


FIG. 3. **Paper electrophoresis of the glycopeptides.** Each glycopeptide fraction (corresponding to approximately 10 nmol of Ser) obtained by the Dowex column chromatography was subjected to paper electrophoresis and the separated components were detected with fluorescamine as described under "Materials and Methods." The results are illustrated.

TABLE II

Carbohydrate composition of linkage glycopeptides from rat chondrosarcoma proteoglycans

Glycopeptides were subjected to methanolysis (1.0 M methanolic HCl, 24 h, 85 °C) followed by gas-liquid chromatography of the trimethylsilylated (re-*N*-acetylated) methyl glycosides on a capillary CP sil CB WCOT fused silica column (0.34 mm × 25 m; Chrompack) (16).

Residue	D-1	D-3	D-4F
		molar ratio	
Xyl	1.0	1.0	1.0
Gal	1.9	1.8	1.9
GlcA	0.6	0.7	0.7
GalNAc	0.8	0.8	0.5

H-1, H-2, and H-5_{ax} atoms, together with those of Gal-*2* H-1, and H-4. From the observation that the chemical shifts of the Gal-*2* structural reporter groups are identical to those of the reference compound, it follows that Gal-*2* is substituted at position 3 by βGal-*3*. The resonances at δ = 4.664 and δ = 4.155 ppm are assigned to Gal-*3* H-1 and H-4, respectively. They are shifted downfield 0.049 and 0.228 ppm, respectively, when compared to the reference compound. These shift increments correspond very well with those observed for the Gal-*2* signals in Galβ1-4Xylβ1-*O*-Ser upon extension with Gal-*3* in β1-3 linkage. In conjunction with the sugar analysis data, this suggests that Gal-*3* is extended with a GalNAcβ1-4GlcAβ1-3 unit. In turn this unit is further elongated with $\Delta_{4,5}$-GlcA as can be deduced from the signals at δ = 5.184 and δ = 5.896 ppm which are characteristic for the H-1 and H-4 atoms of $\Delta_{4,5}$-GlcA, respectively (18, 19). This leaves two anomeric signals still to be interpreted. One signal coincides with that of Gal-*3* H-1 at δ = 4.664 ppm. Irradiation of the last signal reveals coupling to the Gal-*3* H-2 at δ = 3.746 as well as to the signal at δ = 3.454 ppm typical for βGlcA H-2. Therefore the signal at δ = 4.664 belongs to the GlcA-*4* and Gal-*3* anomeric protons. Irradiation of the remaining H-4 signal at δ = 4.098 ppm, belonging to GalNAc-*5*, reveals the H-3 atom of this residue at δ = 3.901 ppm. In turn by irradiation also H-2 at δ = 4.002 ppm and the H-1 atom at δ = 4.538 ppm could be assigned. The resonances of Xyl-*1* H-5_{eq} and the Ser protons are hidden under the H-4 signals and under the H-3 signal of GalNAc-*5*. By irradiation of $\Delta_{4,5}$-GlcA-*6* H-4, the H-3 of this residue is found at δ = 4.093 ppm, almost coinciding with the GalNAc-*5* H-4. Therefore, the structure of the compound present in fraction D-1 is:

$\Delta_{4,5}$-GlcAβ1-3GalNAcβ1-4GlcAβ1-3Galβ1-3Galβ1-4Xy1β1-*O*-Ser
6 *5* *4* *3* *2* *1*

The ^1H NMR spectrum of fraction D-3 is given in Fig. 4*B*

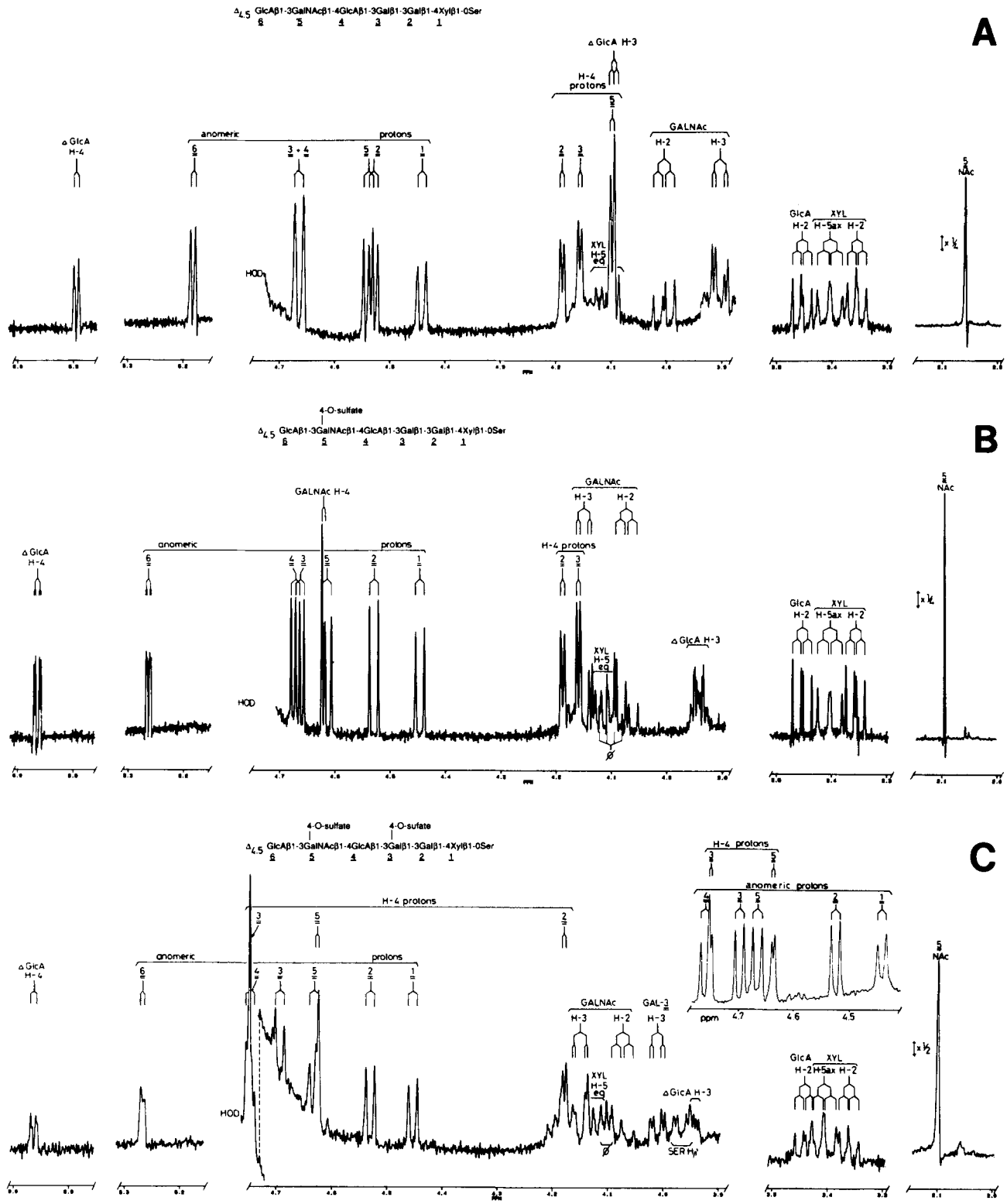


FIG. 4. **Structural reporter-group regions of the resolution-enhanced 500 MHz ^{1}H NMR spectra of fractions D-1 (*A*), D-3 (*B*) and D-4F (*C*) recorded in ^2H_2O at 22 °C.** The inset in *C* is the spectrum of fraction D-4F recorded at 70 °C. The *numbers* and *letters* in the spectra refer to the corresponding residues in the structures.

and the corresponding NMR data in Table III. Comparison with the data of fraction D-1 shows that the signals characteristic for the -Galβ1-3Galβ1-4Xylβ1-*O*-Ser partial structure are present. The carbohydrate analysis suggests that this element is elongated in the same way as in fraction D-1. By ^{1}H-decoupling experiments it is demonstrated that the GalNAc-5 H-1, H-2, H-3, and H-4 signals are shifted downfield with $\Delta\delta$ = 0.077, 0.069, 0.248, and 0.523 ppm, respectively. A downfield shift of 0.523 ppm for an H-4 signal indicates the presence of *O*-sulfate at C-4 of GalNAc-5 (20). This *O*-sulfation has a slight, but significant effect on the signals of GlcA-4 H-1 ($\Delta\delta$ = 0.008 ppm), making it separated

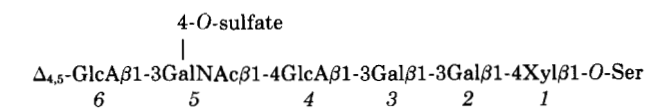
TABLE III

^1H chemical shifts of structural-reporter-groups of the constituent monosaccharides of linkage glycopeptide fractions together with those for a reference compound R (17)

Chemical shifts are given in ppm downfield from internal sodium 4,4-dimethyl-4-silapentane-1-sulfonate but were actually measured indirectly to acetone in ^2H_2O (δ = 2.225 ppm) at 22 °C (reference compound at 27 °C) or at 62 °C (two-dimensional HOHAHA of D-4F).

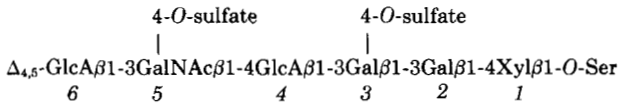
Residue	Reporter group	Chemical shift in				
		R	D-1	D-3	D-4F	D-4F 62 °C
Ser	H-α	4.260				4.191
	H-β	3.995				3.940
	H-β'	4.038			3.965	4.052
Xyl-*1*	H-1	4.470	4.443	4.445	4.453	4.469
	H-2	3.381	3.356	3.357	3.362	3.386
	H-3	3.623	3.608	3.619		3.627
	H-4	3.874		3.866		3.859
	H-5 Ax	3.416	3.403	3.403	3.406	3.422
	H-5 eq	4.124	4.12	4.12	4.126	4.130
Gal-*2*	H-1	4.534	4.529	4.530	4.530	4.522
	H-2	3.789	3.670	3.675	3.680	3.699
	H-3	3.821	3.826	3.826	3.831	3.822
	H-4	4.194	4.188	4.188	4.177	4.167
Gal-*3*	H-1	4.615	4.664	4.664	4.694	4.699
	H-2	3.606	3.746	3.744	3.777	3.798
	H-3	3.668	3.797	3.800	4.010	4.022
	H-4	3.927	4.155	4.160	4.749	4.757
GlcA-*4*	H-1		4.664	4.672	4.747	4.767
	H-2		3.454	3.454	3.444	3.470
	H-3		3.616	3.635	3.614	3.619
	H-4					3.823
	H-5					3.710
GalNAc-*5*	H-1		4.538	4.615	4.632	4.679
	H-2		4.002	4.071	4.074	4.083
	H-3		3.901	4.149	4.149	4.158
	H-4		4.098	4.621	4.626	4.642
	NAc		2.057	2.095	2.098	2.091
ΔGlcA-*6*	H-1		5.184	5.265	5.265	5.258
	H-2		3.793	3.833	3.784	3.831
	H-3		4.093	3.943	3.942	3.950
	H-4		5.896	5.966	5.963	5.938

from the Gal-*3* H-1, and on the GlcA-*4* H-3 ($\Delta\delta$ = 0.019 ppm). The effect on $\Delta_{4,5}$-GlcA-*6* is larger: downfield shifts of $\Delta\delta$ = 0.081, 0.040 and 0.070 ppm are observed for the H-1, H-2, and H-4 signals, respectively. Furthermore, there is an upfield shift of $\Delta\delta$ = 0.150 for the H-3 signal of $\Delta_{4,5}$-GlcA-*6*. A long range (W-type) coupling with J = 1.5 Hz is present between H-1 and H-3 and between H-2 and H-4 of $\Delta_{4,5}$-GlcA-*6* (19). Apparently, the conformation of $\Delta_{4,5}$-GlcA-*6* is affected by the 4-sulfate group of GalNAc-*5*. Therefore the structure of the compound in fraction D-3 is:



The ^1H NMR spectrum of fraction D-4S is identical (data not shown) with that of D-3, consistent with the fact that they have the same R_{serine} value on paper electrophoresis.

The ^1H NMR spectrum of fraction D-4F is given in Fig. 4*C* and the corresponding ^1H NMR data are summarized in Table III. Comparison with the data for the fractions D-1 and D-3 suggests that the same sequence of saccharides is present. In the anomeric region two H-4 signals occur at δ = 4.626 and δ = 4.749 ppm. This is more evident in a spectrum recorded at 70 °C, wherein the HOD line is not disturbing this region (see inset, Fig. 4*C*). As discussed for fraction D-3, the former belongs to GalNAc-*5*, bearing an *O*-sulfate group at C-4. Since the chemical shift of Gal-*2* H-1 is unaltered, while the Gal-*3* H-1 undergoes a downfield shift of $\Delta\delta$ = 0.030 ppm in comparison to fraction D-3, it is concluded that the signal at δ = 4.749 ppm has to be assigned to Gal-*3* H-4. The downfield shift of $\Delta\delta$ = 0.589 ppm is a result from the presence of an *O*-sulfate group at C-4 of Gal-*3*. The presence of a phosphate can be excluded because in the ^1H NMR spectrum no P-H coupling can be observed. The Gal-*3* H-3 signal is now visible at Δ = 4.010 ppm indicating a downfield shift of $\Delta\delta$ = 0.210 ppm. Again the effect on the neighboring sugar residues shows a similar tendency as discussed for C-4 sulfation of GalNAc-*5*. The anomeric proton of GlcA-*4* and GalNAc-*5* are shifted downfield with $\Delta\delta$ = 0.075 and $\Delta\delta$ = 0.017 ppm, respectively, while for Gal-*2* H-4 a minor upfield shift of $\Delta\delta$ = 0.010 ppm is observed. Therefore, the structure of the compound in fraction D-4F is:



To prove the assignment of this interesting structure we decided to run a two-dimensional HOHAHA spectrum (14). This technique has the advantage that all protons within a scalar coupling network can have cross-peaks in the two-dimensional spectrum, depending on the mixing time. This results in subspectra of the individual monosaccharides on cross-sections of the two-dimensional spectrum. The two-dimensional HOHAHA spectrum of fraction D-4F is given in Fig. 5. The spectrum is recorded at 62 °C to avoid disturbance of the anomeric region by the HOD signal. The applied mixing time of 80 ms is not long enough to allow magnetization transfer through the small couplings between H-4 and H-5 of both Gal and GalNAc. Mixing times which do allow this transfer would have resulted in a poor signal-to-noise ratio for this low amount of material. However, the important connectivities, especially between the H-1 and H-4 of both Gal and GalNAc are prominent in Fig. 5. A close inspection of this figure reveals that the cross-peaks of the coinciding GlcA-*4* H-1 and Gal-*3* H-4 signals are separated in two slightly different lines as is indicated in the figure. The connectivities between Gal-*3* H-1 and H-4 are clear from a cross-peak close to the diagonal and from the common cross-peaks with the H-2 and H-3 atoms. In this way a full assignment of the Ser, Xyl-*1*, GlcA-*4*, and $\Delta_{4,5}$-GlcA-*6* residues and a partial assignment of the Gal-*2*, Gal-*3*, and GalNAc-*5* residues can be made. The slightly different chemical shift values due to higher temperature are given in a separate column in Table III.

DISCUSSION

In the present study we have isolated nonsulfated (D-1), monosulfated (D-3/D-4S), and disulfated (D-4F) glycopeptides with a common carbohydrate backbone containing the entire carbohydrate sequence characteristic of the glycosaminoglycan-specific linkage region. D-3 (and D-4S) is sulfated on the GalNAc residue, while D-4F is sulfated, unexpectedly, on the second Gal in addition to the GalNAc. The molar ratio of D-1, D-3/D-4S, and D-4F is 53:37:10 based on the serine contents of the purified fractions. The original proteoglycan, however, could have rather different ratios when the other fractions and any biases in recoveries are taken into account.

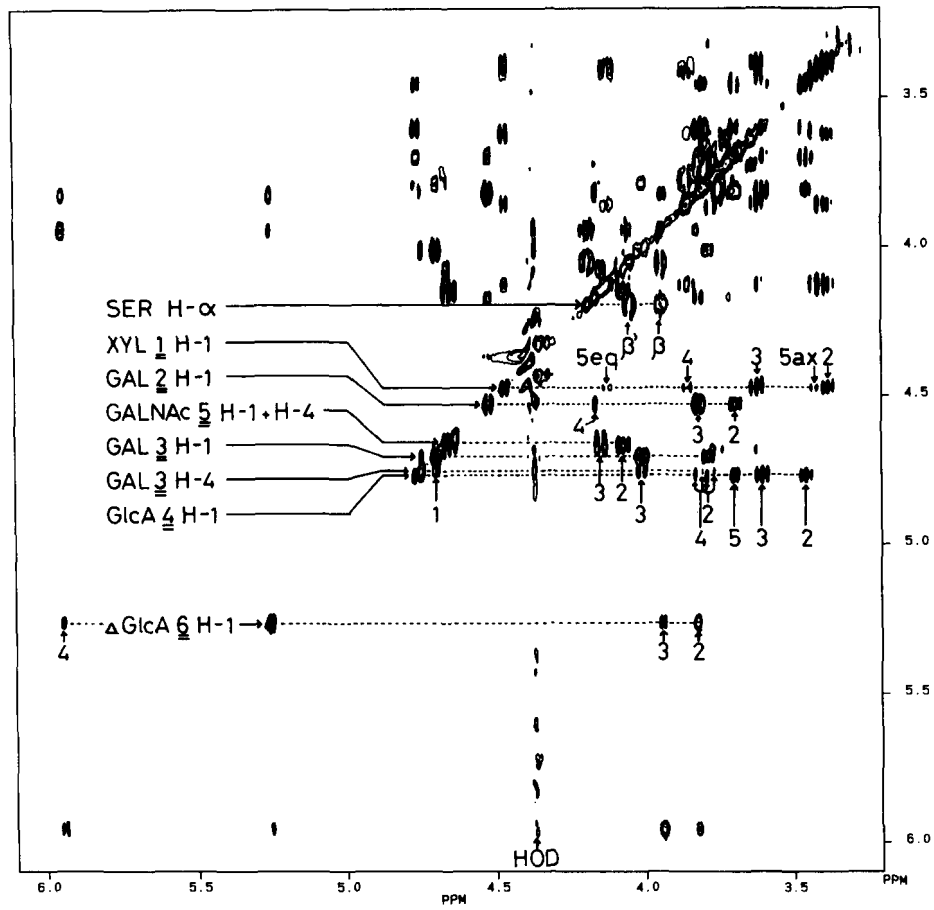


FIG. 5. **Two-dimensional HOHAHA spectrum of fraction D-4F recorded at 500 MHz in ^2H_2O at 62 °C.** Mixing time 80 ms. In the figure the diagonal peaks of the protons in the anomeric region together with the Ser α proton are indicated. The *numbers* near cross-peaks in the figure refer to the protons of the scalar coupling network belonging to a diagonal peak. The numbers of Gal-*3* H-1 and H-4 are combined. The subspectra can thus be found on the *dashed lines*.

In addition partial desulfation might have taken place during the purification. The glycopeptides retain the structure, $\Delta_{4,5}$-GlcAβ1-3GalNAc, reflecting the specificity of chondroitinase ABC employed, leaving the innermost disaccharide unit attached (21, 22). The behavior of the glycopeptides on Dowex 1 chromatography or on paper electrophoresis correlates well with their negative charges assumed from their sulfate and uronic acid contents.

In the pioneering work by Rodén and co-workers (2), the structure of the common linkage region was deduced from the chemical analyses of the hydrolytic fragments of linkage glycopeptides. The preparative procedures of the fragments included mild acid hydrolysis and Dowex 1 treatment which could cause desulfation and removal of sulfated fragments, respectively. Lindahl and Rodén (23) isolated from chondroitin 4-sulfate proteoglycans of bovine nasal septa a sulfated glycopeptide with a carbohydrate composition corresponding to the linkage structure, GlcA-GalNAc-GlcA-Gal-Gal-Xyl-Ser but did not report the location of the sulfate group. Partial acid hydrolysis (pH 1.5 at 100 °C for 4 h) of this glycopeptide yielded fragments including two acidic yet nonsulfated glycopeptides which had carbohydrate compositions corresponding to the parent compound and GlcA-Gal-Gal-Xyl-Ser, respectively. Thus, desulfation took place indeed during mild acid hydrolysis under the above conditions. It cannot be excluded that a monosaccharide other than GalNAc had been sulfated.

During the preparation of enzymatically deglycosylated protein cores from embryonic proteochondroitin sulfate, Oike *et al.* (22) have observed sulfated oligosaccharide-like stubs which are refractory to the actions of chondroitinases ABC, AC-II, and keratanase, but releasable by alkali. These stubs may contain a 4-*O*-sulfated galactose.

The linkage region of chondroitin 4-sulfate proteoglycan of Swarm rat chondrosarcoma has also been studied recently by Oegema *et al.* (24), who demonstrated an ester phosphate on C-2 of the xylose residue in the common linkage region. Xylose-2-phosphate was found on one of 3–4 chondroitin sulfate chains. Although these investigators detected sulfate in the linkage fragments, they did not quantitate or localize it. More recently Fransson *et al.* (25) isolated a negatively charged linkage fragment, expected to have 2 galactoses, 1 xylose, and 1 serine, from bovine lung heparan sulfate, and demonstrated that most of the heparan sulfate chains carry xylose-2-phosphate. A possibility of sulfation of the fragment, however, was not investigated. None of the glycopeptides isolated in the present study was phosphorylated on xylose. In addition to the NMR data, the good recovery of xylose in the methanolysis products (Table II) is also indicative of the absence of phosphate (24, 33). Alkaline phosphatase treatments did not change the electrophoretic patterns of the glycopeptides (data not shown). It may be that dephosphorylation took place during the enzymatic digestions of the isolation procedure. A significant alkaline phosphatase activity was detected at least in the Pronase preparation as measured according to Garen and Levinthal (34). Likewise, D-1 and D-3/D-4S might be desulfation products resulting from D-4F during the manipulation.

Although to our knowledge a 4-*O*-sulfated galactose has never been reported for glycoproteins from animal sources, it occurs in carrageenin (26, 27), a polysaccharide obtained from red algae.

It remains to be elucidated whether a 4-*O*-sulfated galactose is common to other glycosaminoglycans or whether it is

characteristic of chondroitin 4-sulfate or galactosaminoglycans. Sorting mechanisms in biosynthesis of different glycosaminoglycans with the same linkage region are largely unknown. Multienzyme complexes have been suggested (3), while the presence of hybrid proteoglycans with both chondroitin (or dermatan) and heparan sulfate (28–31) has complicated the problem. A presumptive galactose 4-*O*-sulfotransferase might be a key enzyme for such complexes if any and/or the galactose 4-*O*-sulfate might act as a recognition signal for sorting.

A sulfotransferase activity responsible for the formation of the galactose 4-*O*-sulfate found in the present study also remains to be demonstrated in tissues of rat chondrosarcoma. Recent evidence (32, 33) suggests that during biosynthesis of glycosaminoglycans, not only chain elongation and modification including sulfation, but also formation of the linkage region takes place in the Golgi apparatus. It is most likely that the galactose 4-*O*-sulfotransferase occurs in the Golgi apparatus.

Acknowledgments—We are deeply indebted to the late Dr. Albert Dorfman (University of Chicago) for his continuous support and encouragement and wish to thank Dr. Robert Heinrikson and Pamela Keim (University of Chicago) for the amino acid analyses.

REFERENCES

1. Poole, A. R. (1986) *Biochem. J.* **236,** 1–14
2. Lindahl, U., and Rodén, L. (1972) *Glycoproteins* (Gottschalk, A., ed) pp. 491–517, Elsevier Scientific Publishing Co., New York
3. Rodén, L. (1980) *The Biochemistry of Glycoproteins and Proteoglycans* (Lennarz, W. J., ed) pp. 267–371, Plenum Publishing Co., New York
4. Sugahara, K., and Dorfman, A. (1979) *Biochem. Biophys. Res. Commun.* **89,** 1193–1199
5. Oegema, T. R., Jr., Hascall, V. C., and Dziewiatkowski, D. D. (1975) *J. Biol. Chem.* **250,** 6151–6159
6. Bitter, T., and Muir, H. (1962) *Anal. Biochem.* **4,** 330–334
7. Hewitt, L. F. (1937) *Biochem. J.* **31,** 360–366
8. Jourdian, G. W., Dean, L., and Roseman, S. (1971) *J. Biol. Chem.* **246,** 430–435
9. Spackman, D. H., Stein, W. H., and Moore, S. (1958) *Anal. Chem.* **30,** 1190–1206
10. Saito, H., Yamagata, T., and Suzuki, S. (1968) *J. Biol. Chem.* **243,** 1536–1542
11. Warren, L. (1959) *J. Biol. Chem.* **234,** 1971–1975
12. Takasaki, S., and Kobata, A. (1978) *Methods Enzymol.* **50,** 50–54
13. Ernst, R. R. (1966) *Adv. Magn. Res.* **2,** 1–135
14. Bax, A., and Davis, D. G. (1985) *J. Magn. Res.* **65,** 355–360
15. Vliegenthart, J. F. G., Dorland, L., and Van Halbeek, H. (1983) *Adv. Carbohydr. Chem. Biochem.* **41,** 209–374
16. Kamerling, J. P., and Vliegenthart, J. F. G. (1982) *Cell Biol. Monogr.* **10,** 95–125
17. Van Halbeek, H., Dorland, L., Veldink, G. A., Vliegenthart, J. F. G., Garegg, P. Y., Norberg, T., and Lindberg, B. (1982) *Eur. J. Biochem.* **127,** 1–6
18. Hirano, S. (1970) *Org. Magn. Res.* **2,** 577–580
19. Perlin, A. S., Mackie, D. M., and Dietrich, C. P. (1971) *Carbohydr. Res.* **18,** 185–194
20. Harris, M. J., and Turvey, J. R. (1970) *Carbohydr. Res.* **15,** 57–63
21. Hascall, V. C., Riolo, R. L., Hayward, J., Jr., and Reynolds, C. C. (1972) *J. Biol. Chem.* **247,** 4521–4528
22. Oike, Y., Kimata, K., Shinomura, T., Nakazawa, K., and Suzuki, S. (1980) *Biochem. J.* **191,** 193–207
23. Lindahl, U., and Rodén, L. (1966) *J. Biol. Chem.* **241,** 2113–2119
24. Oegema, T. R., Jr., Kraft, E. L., Jourdian, G. W., and Van Valen, T. R. (1984) *J. Biol. Chem.* **259,** 1720–1726
25. Fransson, L.-Å., Silverberg, I., and Carlstedt, I. (1985) *J. Biol. Chem.* **260,** 14722–14726
26. Johnston, R., and Percival, E. G. V. (1950) *J. Chem. Soc.* 1994–1998
27. Turvey, J. R., and Williams, T. P. (1962) *J. Chem. Soc.* 2119–2122
28. Rapraeger, A., Jalkanen, M., Endo, E., Koda, J., and Bernfield, M. (1985) *J. Biol. Chem.* **260,** 11046–11052
29. David, G., and Van den Berghe, H. (1985) *J. Biol. Chem.* **260,** 11067–11074
30. Kato, M., Koike, Y., Ito, Y., Suzuki, S., and Kimata, K. (1987) *J. Biol. Chem.* **262,** 7180–7188
31. Isemura, M., Sato, N., Yamaguchi, Y., Aikawa, J., Munakata, H., Hayashi, N., Yoshizawa, Z., Nakamura, T., Kubota, A., Arakawa, M., and Hsu, C.-C. (1987) *J. Biol. Chem.* **262,** 8926–8933
32. Nuwayhid, N., Glaser, J. H., Johnson, J. C., Conrad, H. E., Hauser, S. C., and Hirschberg, C. B. (1986) *J. Biol. Chem.* **261,** 12936–12941
33. Lohmander, L. S., Hascall, V. C., Yanagishita, M., Keuttner, K. E., and Kimura, J. H. (1986) *Arch. Biochem. Biophys.* **250,** 211–227
34. Garen, A., and Levinthal, C. (1960) *Biochim. Biophys. Acta* **38,** 470–483