Eur. J. Biochem. *169*, 399–411 (1987)

Primary structure determination of seven novel N-linked carbohydrate chains derived from hemocyanin of *Lymnaea stagnalis*

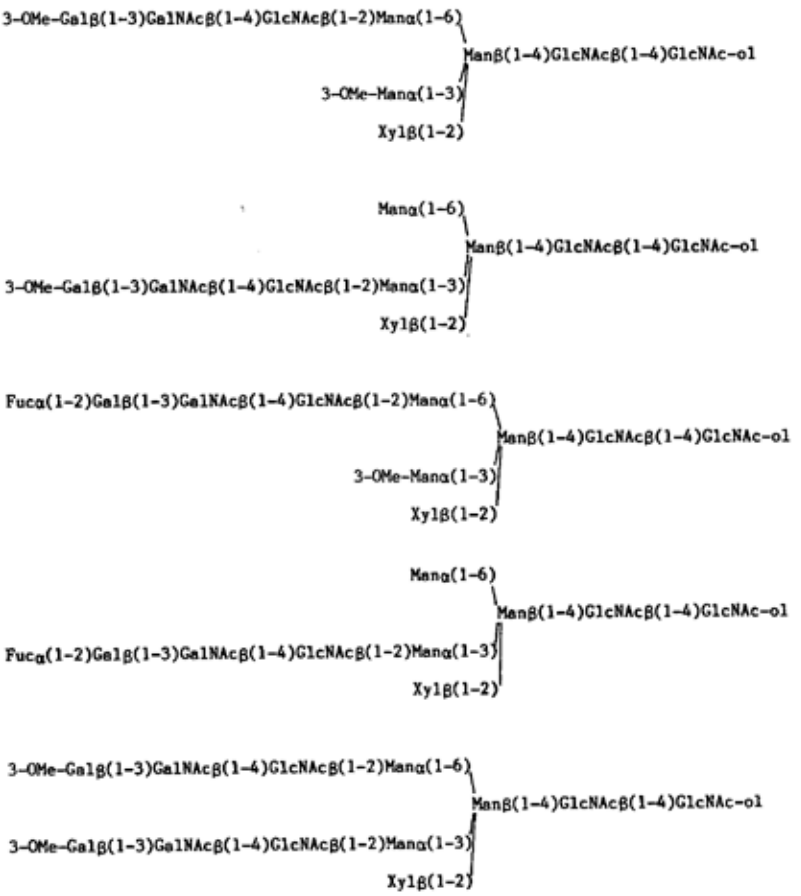
3-O-methyl-D-galactose and *N*-acetyl-D-galactosamine as constituents of xylose-containing N-linked oligosaccharides in an animal glycoprotein

J. Albert VAN KUIK[1], Rint P. SIJBESMA[1], Johannis P. KAMERLING[1], Johannes F. G. VLIEGENTHART[1] and Edward J. WOOD[2]

[1] Department of Bio-Organic Chemistry, Transitorium III, Utrecht University
[2] Department of Biochemistry, University of Leeds

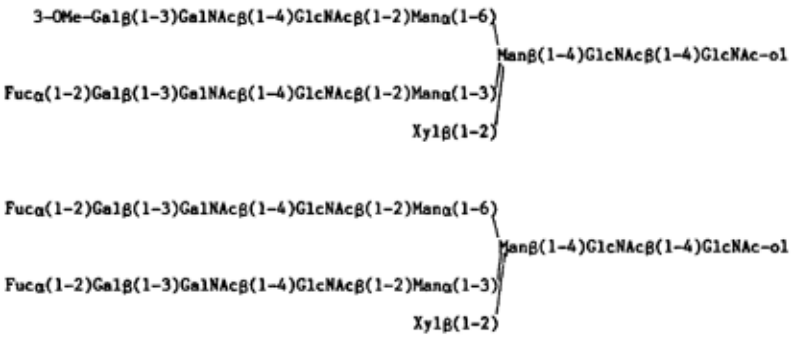
(Received May 20/July 9, 1987) – EJB 87 0583

Hemocyanin from the freshwater snail *Lymnaea stagnalis* is a high-molecular-mass copper-containing glycoprotein which functions as oxygen carrier in the hemolymph. To release the carbohydrate chains, the protein was digested by pronase followed by hydrazinolysis and reduction. The oligosaccharide-alditols were purified by gel permeation chromatography on Bio-Gel P-4, followed by HPLC on a Lichrosorb-NH_2 column. Using 500-MHz ^{1}H-NMR spectroscopy, in conjunction with sugar, methylation and deamination analysis, the following seven novel primary oligosaccharide structures could be unravelled.

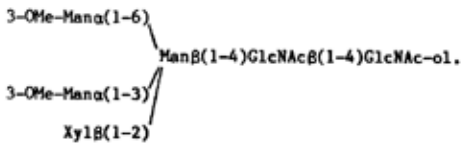


Correspondence to J. F. G. Vliegenthart, Afdeling Bio-Organische Chemie, Rijksuniversiteit te Utrecht, Transitorium III, Postbus 80.075, NL-3508 TB Utrecht, The Netherlands

Abbreviations. Xyl, D-xylose; 3-OMe-Gal, 3-*O*-methyl-D-galactose; GalNAc, *N*-acetyl-D-galactosamine; Fuc, L-fucose; 3-OMe-Man, 3-*O*-methyl-D-mannose; GlcNAc-ol, *N*-acetyl-D-glucosaminitol.



Hemocyanins are high-molecular-mass copper-containing glycoproteins, which are responsible for the oxygen transport in many arthropods and molluscs. The primary structure of the *N*-glycosidically linked low-molecular-mass carbohydrate chains of hemocyanin from the gastropods *Helix pomatia* and *Lymnaea stagnalis* have been determined [1, 2]. In *H. pomatia*, the usual pentasaccharide core structure contains an additional Xyl residue in β(1-2)-linkage to β-Man, a feature which is uncommon for animal glycoproteins. The low-molecular-mass carbohydrate chain from *L. stagnalis* hemocyanin is a variant wherein the terminal Man residues are 3-*O*-methylated. It was established to be:



Sugar analysis of native *L. stagnalis* hemocyanin revealed as additional monosaccharide constituents 3-OMe-Gal, Gal, GalNAc and Fuc [3]. The liberation of the *N*-glycosidically linked carbohydrate chains of this hemocyanin by hydrazinolysis of a pronase digest, and the fractionation of these chains as oligosaccharide-alditols on Bio-Gel P-4 have been described previously [2]. The structure determination of carbohydrate chains with higher molecular mass will be described in this article.

MATERIALS AND METHODS

Preparation of Bio-Gel P-4 fractions

The purified *L. stagnalis* hemocyanin was stripped of copper, digested with pronase, and treated with hydrazine to liberate the carbohydrate chains. After re-*N*-acetylation these chains were reduced, purified by high-voltage paper electrophoresis, and fractionated on Bio-Gel P-4. Four fractions were obtained, denoted I – IV. The analysis of fraction IV has been described previously [2].

Determination of the absolute configuration of the monosaccharides

Determination of the absolute configuration of the constituent monosaccharides from *L. stagnalis* hemocyanin was performed as described [4, 5]. Trimethylsilylated (–)-2-butyl glycosides were analyzed by GLC, on a capillary CP-sil5 CB WCOT fused silica column (0.34 mm × 25 m, Chrompack).

Alkaline borohydride treatment

Alkaline borohydride treatment, on 100 µg native glycoprotein, was carried out as described [1].

HPLC

Bio-Gel P-4 fractions I – III were further fractionated by HPLC [6] using a Perkin-Elmer series 3 liquid chromatograph, equipped with a Rheodyne injection valve. A column (4 × 250 mm) of Lichrosorb-NH_2 (5 µm, Merck) was used. The column was run isocratically with a mixture of acetonitrile/water for Bio-Gel P-4 fraction III (70:30, v/v) and II (68:32, v/v) or with a linear gradient starting with a mixture of acetonitrile/water (70:30, v/v) changing in 80 min to 60:40 (v/v) for fraction I, at a flow rate of 1 ml/min. The elution patterns were monitored at 205 nm.

500-MHz ^{1}H-NMR spectroscopy

Carbohydrates were repeatedly exchanged in ^2H_2O (99.96 atom % ^{2}H, Aldrich) with intermediate lyophilization. ^{1}H-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 in the Fourier-transform mode at a probe temperature of 27 °C. NOE difference spectroscopy was performed according to [7]. Resolution enhancement of the spectra was achieved by Lorentzian-to-Gaussian transformation [8]. Chemical shifts (δ) are given relatively to sodium 4,4-dimethyl-4-silapentane-1-sulfonate, but were actually measured indirectly to acetone (δ = 2.225 ppm) [9].

Sugar analysis

Samples containing 50 nmol carbohydrate were subjected to methanolysis (1.0 M methanolic HCl, 24 h, 85 °C) followed by GLC of the trimethylsilylated (re-*N*-acetylated) methyl glycosides on a capillary CP-sil5 CB WCOT fused silica column (0.34 mm × 25 m, Chrompack) [10].

Methylation analysis

Methylation analysis on 100-µg samples was carried out as described [2]. Because of the presence of natural occurring 3-*O*-methylated monosaccharides, C^2H_3I was used. Partially methylated alditol acetates were analyzed by GLC-MS; oven temperature program, 110 – 230 °C at 4 °C/min.

Deamination

For de-*N*-acetylation, 70 µg of a thoroughly dried oligosaccharide-alditol was dissolved in 150 µl anhydrous hydra-

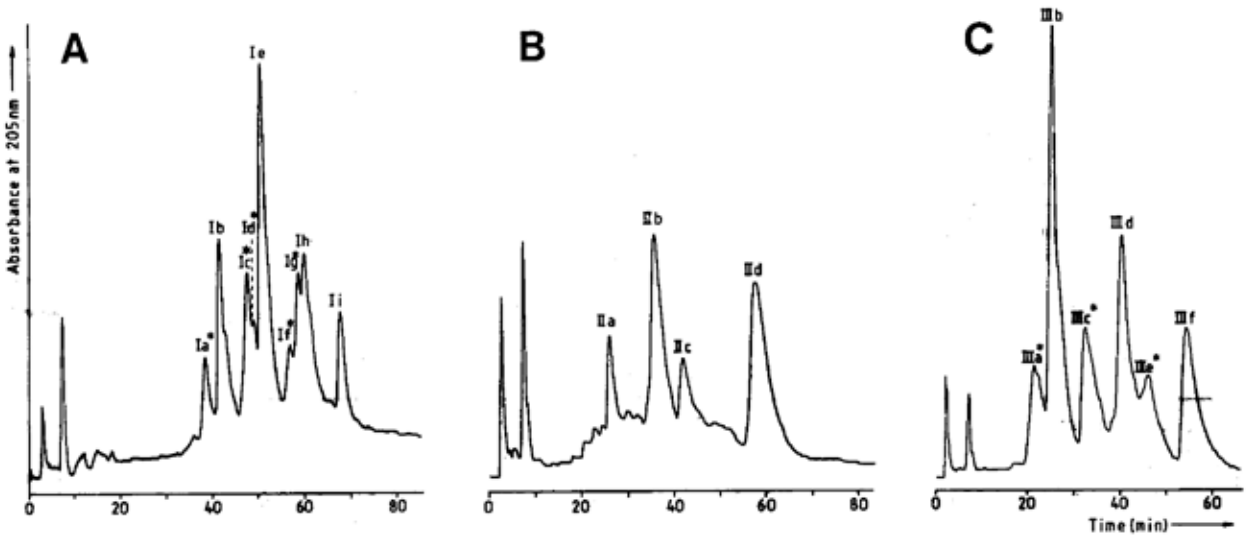


Fig. 1. *Elution profiles on Lichrosorb-NH_2* (HPLC) of oligosaccharide-[1-^{2}H]alditols from Bio-Gel P-4 fractions I–III [2]. The peaks marked with an asterix are (partly) composed of hydrazinolysis artefacts. (A) Bio-Gel P-4 fraction I, fractionated with a linear gradient starting with a mixture of acetonitrile/water (70:30, v/v) changing in 80 min to 60:40 (v/v). (B) Bio-Gel P-4 fraction II, fractionated isocratically with a mixture of acetonitrile/water (68:32, v/v). (C) Bio-Gel P-4 fraction III, fractionated isocratically with a mixture of acetonitrile/water (70:30, v/v)

zine in a 1-ml Pierce vial, and heated for 24 h at 100 °C. After evaporation of hydrazine, the sample was deaminated by 100 µl of a solution of 5 mg/ml $NaNO_2$ in 0.5 M acetic acid and subsequently reduced by 0.5 mg $NaBH_4$ [11]. Samples were trimethylsilylated and analyzed by GLC on a capillary CP-sil5 CB WCOT fused silica column (0.34 mm × 25 m, Chrompack) and by GLC-MS; oven temperature program, 120 °C during 2 min, then 120–240 °C at 4 °C/min. Trimethylsilylated 2,5-anhydro-mannitol and 2,5-anhydro-talitol were used as references.

GLC-MS

Samples were analyzed on a Carlo Erba GC/Kratos MS 80/Kratos DS55 system; electron energy, 70 eV; accelerating voltage, 2.7 kV; ionization current, 100 µA; ion-source temperature, 225 °C; BPl capillary WCOT fused silica column (0.33 mm × 25 m; Scientific Glass Engineering).

RESULTS

Alkaline borohydride treatment of the native glycoprotein did not give rise to the formation of *N*-acetylgalactosaminitol; therefore, it is suggested that the glycoprotein does not contain *O*-glycosidically bound carbohydrate chains. Determination of the absolute configuration of the constituent monosaccharides from *L. stagnalis* hemocyanin, revealed that all residues are present as D sugars, except Fuc, which has L configuration. The carbohydrate chains of *L. stagnalis* hemocyanin were released by hydrazinolysis of a pronase digest. After reduction, and purification by high-voltage paper electrophoresis, the neutral oligosaccharide-alditols were fractionated on a Bio-Gel P-4 column, yielding four fractions (I–IV) [2].

The Bio-Gel P-4 fractions I–III [2] were further fractionated by HPLC on Lichrosorb-NH_2 (Fig. 1). Sugar analysis data of the main HPLC fractions are listed in Table 1. Methylation analysis data of two HPLC fractions are presented in Table 2. Separation of Bio-Gel P-4 fraction II on HPLC yielded four fractions, denoted IIa–IId (Fig. 1B).

Table 1. *Molar carbohydrate composition of oligosaccharide-alditol HPLC fractions from* Lymnaea stagnalis *hemocyanin*
Values are based on Man taken as 2 in IIIb and IId and as 3 in IIIf, Ib, Ie and Ih

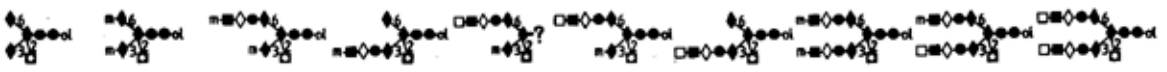
Monosaccharide	IIIb	IId	IIIf	Ib	Ie	Ih
Fuc		0.6	0.7		1.0	1.9
Xyl	0.7	0.8	1.0	1.0	0.9	1.0
3-OMe-Man	0.8	0.7				
3-OMe-Gal	0.7			1.5	0.6	
Man	2.0	2.0	3.0	3.0	3.0	3.0
Gal	0.1	0.7	0.9		1.0	2.2
GalNAc	0.8	0.6	0.9	1.6	2.0	2.2
GlcNAc	1.5	1.5	1.7	2.4	2.5	2.9
GlcNAc-ol	0.6	0.5	0.4	0.6	0.6	0.8

Table 2. *Methylation analysis of oligosaccharide-alditol HPLC fractions from* Lymnaea stagnalis *hemocyanin*
Some values of the methylation analysis data are too low in comparison to the sugar analysis data and the intensity of the anomeric signals in the ^{1}H-NMR spectra. This is probably due to the low amount of material (50–100 µg) available to this procedure. 4-Mono-*O*-[^{2}H]methyl-mannitol was taken as 1.0

Partially methylated alditol acetate	IIIb	IId
2,3,4-Tri-*O*-[^{2}H]methyl-xylitol	0.4	1.0
2,3,4-Tri-*O*-[^{2}H]methyl-fucitol		0.9
3-Mono-*O*-methyl-2,4,6-tri-*O*-[^{2}H]methyl-mannitol	0.7	0.9
3-Mono-*O*-methyl-2,4,6-tri-*O*-[^{2}H]methyl-galactitol	0.2	
3,4,6-Tri-*O*-[^{2}H]methyl-mannitol	0.9	1.1
3,4,6-Tri-*O*-[^{2}H]methyl-galactitol		0.3
1,3,5,6-Tetra-*O*-[^{2}H]methyl-2-*N*-[^{2}H]methyl-acetamido-2-deoxyglucitol	+	+
4-Mono-*O*-[^{2}H]methyl-mannitol	1.0	1.0
3,6-Di-*O*-[^{2}H]methyl-2-*N*-[^{2}H]methyl-acetamido-2-deoxyglucitol	1.1	0.6
4,6-Di-*O*-[^{2}H]methyl-2-*N*-[^{2}H]methyl-acetamido-2-deoxygalactitol	+	+

Table 3. *Relevant ^1H chemical shifts of structural-reporter groups of constituent monosaccharides for oligosaccharide-alditols from* Lymnaea stagnalis *hemocyanin and those for reference compound R [1]*

Chemical shifts are in ppm downfield from sodium 4,4-dimethyl-4-silapentane-1-sulfonate in ^2H_2O at 27°C acquired at 500 MHz (but were actually measured relative to internal acetone: δ = 2.225 ppm). For numbering of the monosaccharides and complete structures, see Fig. 2. In the table heading, the structures are represented by short-hand symbolic notation (cf. [9]): ●, GlcNAc; ◆, Man; ⊡, Xyl; ◇, GalNAc; ■, Gal; □, Fuc; m, 3-OMe, n.d. = not determined

Residue	Reporter group	Chemical shift in compound									
											
		R	IVa	IIIb	IIIc1	IIIc2	IIId	IIIf	Ib	Ie	Ih
		ppm									
GlcNAc-1-ol	H-2	4.239	4.237	4.238	4.237		4.237	4.238	4.237	4.238	4.237
	NAc	2.057	2.056	2.058[a]	2.057		2.058	2.056	2.055[c]	2.057[d]	2.057
GlcNAc-2	H-1	4.634	4.620	4.637	4.633		4.638	4.634	4.634	4.635	4.632
	NAc	2.073	2.065	2.080	2.074	2.085	2.084	2.073[b]	2.080	2.080	2.080
Man-3	H-1	4.883	4.898	4.885	4.876		4.887	4.877	4.879	4.879	4.880
	H-2	4.270	4.281	4.272	4.258	4.270	4.273	4.255	4.251	4.254	4.255
Man-4	H-1	5.122	5.168	5.166	5.136	5.166	5.165	5.136	5.137	5.137	5.137
	H-2	4.039	4.305	4.297	4.145	4.298	4.298	4.141	4.143	4.141	4.140
	OCH_3		3.443	3.441		3.442	3.442				
Man-4′	H-1	4.913	4.961	4.896	4.915	4.897	4.895	4.916	4.893	4.894	4.895
	H-2	3.983	4.230	4.097	3.979	4.090	4.091	3.973	4.096	4.096	4.089
	OCH_3		3.415								
Xyl-X	H-1	4.449	4.453	4.450	4.433	4.454	4.451	4.440	4.436	4.436	4.436
	H-2	3.377	3.378	3.382	n.d.	n.d.	3.384	3.371	3.368	3.370	n.d.
	H-3	3.437	n.d.	n.d.	n.d.	n.d.	n.d.	3.433	n.d.	n.d.	n.d.
	$H\text{-}5_{ax}$	3.250	3.264	3.260	3.250	3.262	3.261	3.247	3.246	3.248	3.248
	$H\text{-}5_{eq}$	n.d.	4.015	4.012	4.012	4.012	4.012	4.010	4.010	4.009	4.009
GlcNAc-5	H-1				4.552			4.518	4.516	4.518	4.517
	NAc				2.041			2.040	2.042	2.045[e]	2.043
GlcNAc-5′	H-1			4.554		4.558	4.557		4.557	4.557	4.558
	NAc			2.041		2.041	2.040		2.042	2.040[e]	2.043
GalNAc-GN	H-1				4.575			4.425	4.573	4.427	4.427
	H-4				4.183			4.103	4.184	4.105	4.105
	NAc				2.048			2.070[b]	2.047	2.070	2.071
GalNAc-GN′	H-1			4.580		4.436	4.434		4.583	4.578	4.436
	H-4			4.184		4.103	4.105		4.184	4.185	4.105
	NAc			2.053[a]		2.078	2.077		2.053[c]	2.054[d]	2.077
Gal-G	H-1				4.462			4.611	4.461	4.613	4.614
	H-3				3.315			n.d.	3.317	n.d.	n.d.
	H-4				4.199			n.d.	4.198	n.d.	n.d.
	OCH_3				3.428				3.429		
Gal-G′	H-1			4.461		4.616	4.615		4.461	4.463	4.614
	H-2			3.538		n.d.	n.d.		n.d.	3.533	n.d.
	H-3			3.316		n.d.	n.d.		3.317	3.316	n.d.
	H-4			4.199		n.d.	n.d.		4.198	4.199	n.d.
	OCH_3			3.429					3.429	3.430	
Fuc-F	H-1							5.229		5.230	5.231
	H-5							4.208		4.216	4.210
	CH_3							1.205		1.208	1.208
Fuc-F′	H-1					5.230	5.229				5.231
	H-5					n.d.	4.213				4.210
	CH_3					1.210	1.209				1.208

[a–e] Assignments may have to be interchanged.

Although the fractions separated perfectly well, they contained too low an amount of sugar to determine the structure of the carbohydrate chains.

To elucidate the primary structure of the carbohydrate chains present in the various HPLC fractions obtained from the Bio-Gel P-4 fractions I and III, 500-MHz ^{1}H-NMR spectra in ^2H_2O were recorded. Relevant ^{1}H-NMR data are compiled in Table 3, together with the data of IVa [2] and reference compound R from *Helix pomatia* α-hemocyanin [1]. ^{1}H-NMR spectroscopy revealed that most of the low-intensity HPLC peaks contain 'hydrazinolysis artefacts' [12]. The latter are modified by the hydrazinolysis procedure at GlcNAc-**1** and/or GlcNAc-**2**, which results in lower retention times on HPLC than the corresponding intact oligosaccharide-alditols. The ^{1}H-NMR parameters of these fractions are not included in Table 3.

Bio-Gel P-4 fraction III

Bio-Gel P-4 fraction III will be discussed first, as it contains the structures with the lowest molecular mass. This fraction was separated into six sub-fractions denoted IIIa–IIIf (Fig. 1C). The structural-reporter-group regions of the spectrum of IIIb (for sugar analysis, see Table 1) are presented in Fig. 2A. The intensity of the anomeric proton signals points to the presence of one major component (> 90%). Comparing the chemical shifts of the signals of H-2 and NAc of GlcNAc-**1**-ol, H-1 of GlcNAc-**2**, H-1 and H-2 of Man-**3**, and H-1 of Xyl of IIIb, with those of R, leads to the conclusion that they have the Xylβ(1-2)Manβ(1-4)GlcNAcβ(1-4)GlcNAc-ol sequence in common. The presence of a terminal 3-OMe-Man (methylation analysis, Table 2) in α(1-3) linkage to Man-**3**, is evident from comparison of the H-1 (δ = 5.166 ppm) and H-2 (δ = 4.297 ppm) signals of Man-**4**, in IIIb with those in IVa. The chemical shifts of the Xyl structural-reporter groups (H-1, δ = 4.450 ppm; H-2, δ = 3.382 ppm; H-5$_{ax}$, δ = 3.260 ppm; H-5$_{eq}$, δ = 4.012 ppm) are essentially identical to those observed for IVa, wherein Xyl is also found attached to Man-**3**, next to a 3-*O*-methylated Man-**4**. The assignments of Xyl H-1 and H-5$_{eq}$ were made by selective decoupling of H-2 and H-5$_{ax}$, respectively.

According to the sugar and methylation analysis, additionally a linear tetrasaccharide is present at C-6 of Man-**3**, consisting of Man substituted at C-2, GlcNAc substituted at C-4, GalNAc substituted at C-3 and terminal 3-OMe-Gal. Comparing the H-1 and H-2 signals of Man-**4**′ of IIIb (δ = 4.896 ppm and δ = 4.097 ppm, respectively) with those of R, shows that Man-**4**′ is not terminal in IIIb. The extension of Man-**4**′ is also reflected by the shift of the NAc signal of GlcNAc-**2** [9], from δ = 2.073 ppm in R to δ = 2.080 ppm in IIIb. Since the coupling constants of the three remaining

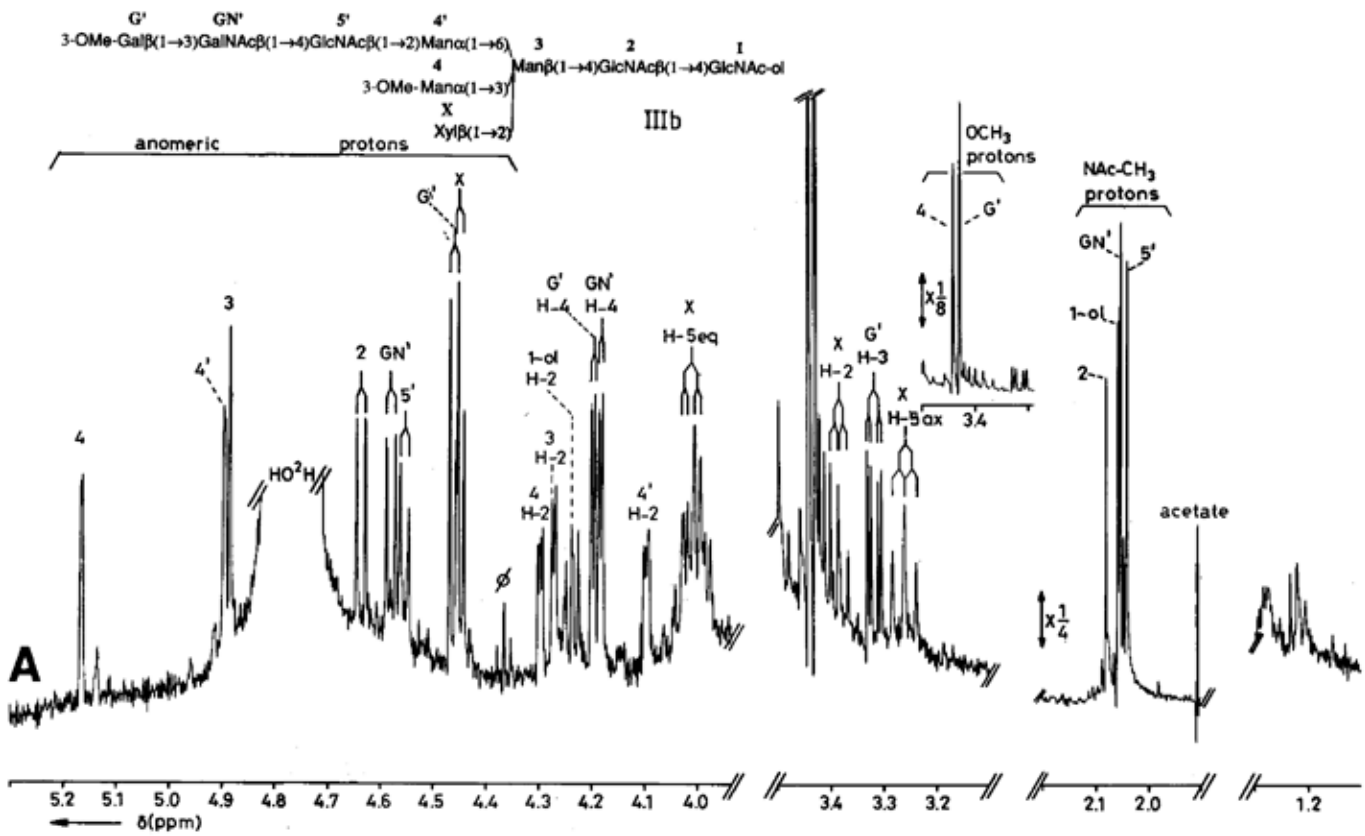
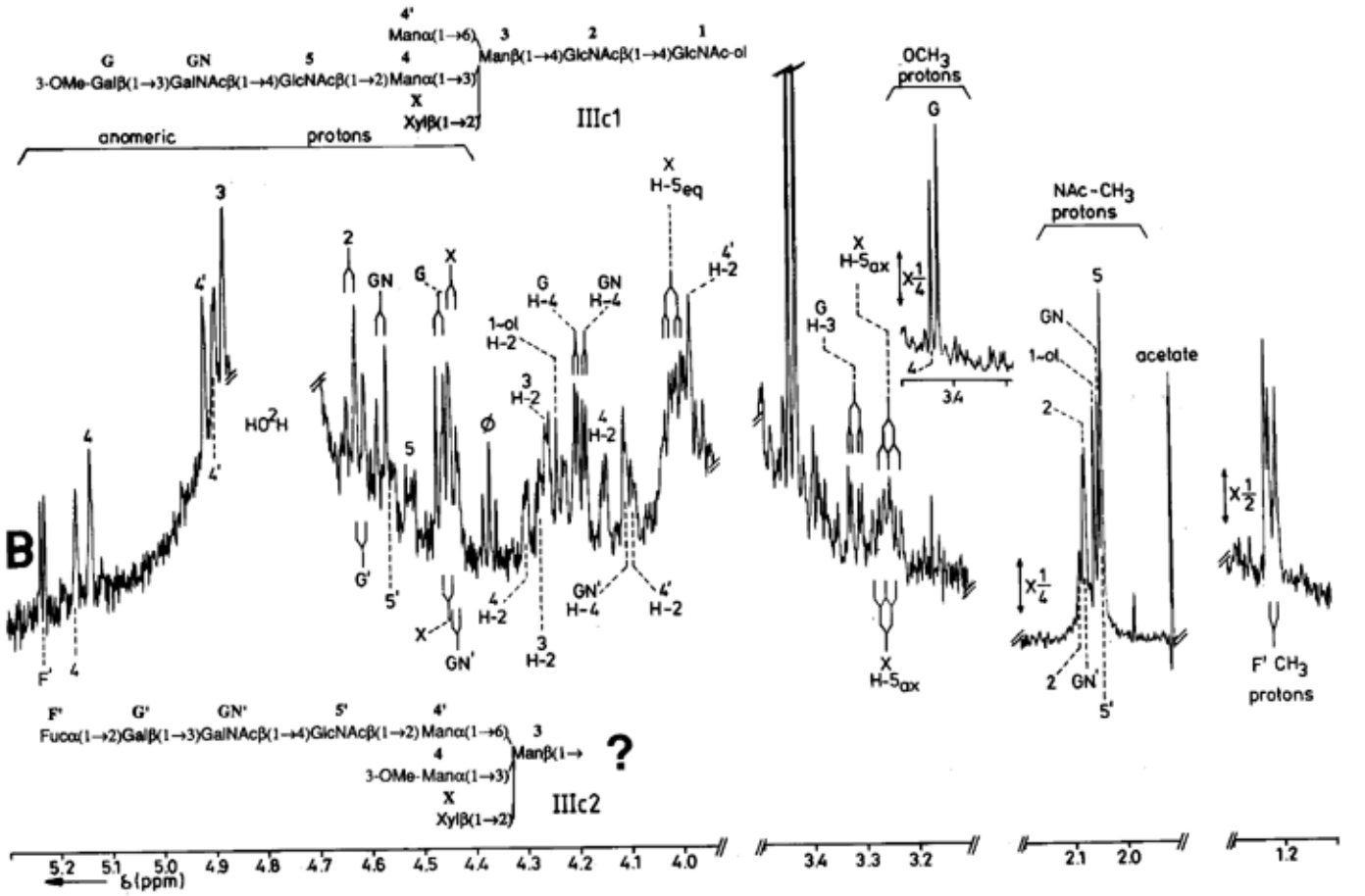


Fig. 2. *Structural-reporter-group regions of the resolution-enhanced 500-MHz ^{1}H-NMR spectra of oligosaccharide-[1-^{2}H]alditols derived from* Lymnaea stagnalis *hemocyanin recorded in ^2H_2O at 27°C.* (A) Spectrum of fraction IIIb. (B) Spectrum of fraction IIIc. The structural-reporter groups of IIIc1 are given above the spectrum, and the structural-reporter groups of IIIc2 are marked below the spectrum. (C) Spectrum of fraction IIId. (D) Spectrum of fraction IIIf. (E) Spectrum of fraction Ib. (F) Spectrum of fraction Ie. (G) Spectrum of fraction Ih. The numbers and letters in the spectra refer to the corresponding residues in the structures. The relative intensity scale of the *N*-acetyl and *O*-methyl regions differs from that of the other parts of the spectra as indicated. ∅ indicates impurity


4'
Manα(1→6)
3
2
1
G
GN
5
4
Manβ(1→4)GlcNAcβ(1→4)GlcNAc-ol
3-OMe-Galβ(1→3)GalNAcβ(1→4)GlcNAcβ(1→2)Manα(1→3)
X
Xylβ(1→2)
IIIc1
anomeric
protons
OCH3
protons
NAc-CH3
protons
B
HO2H
acetate
F' CH3
protons
F'
G'
GN'
5'
4'
Fucα(1→2)Galβ(1→3)GalNAcβ(1→4)GlcNAcβ(1→2)Manα(1→6)
Manβ(1→
?
IIIc2
δ(ppm)

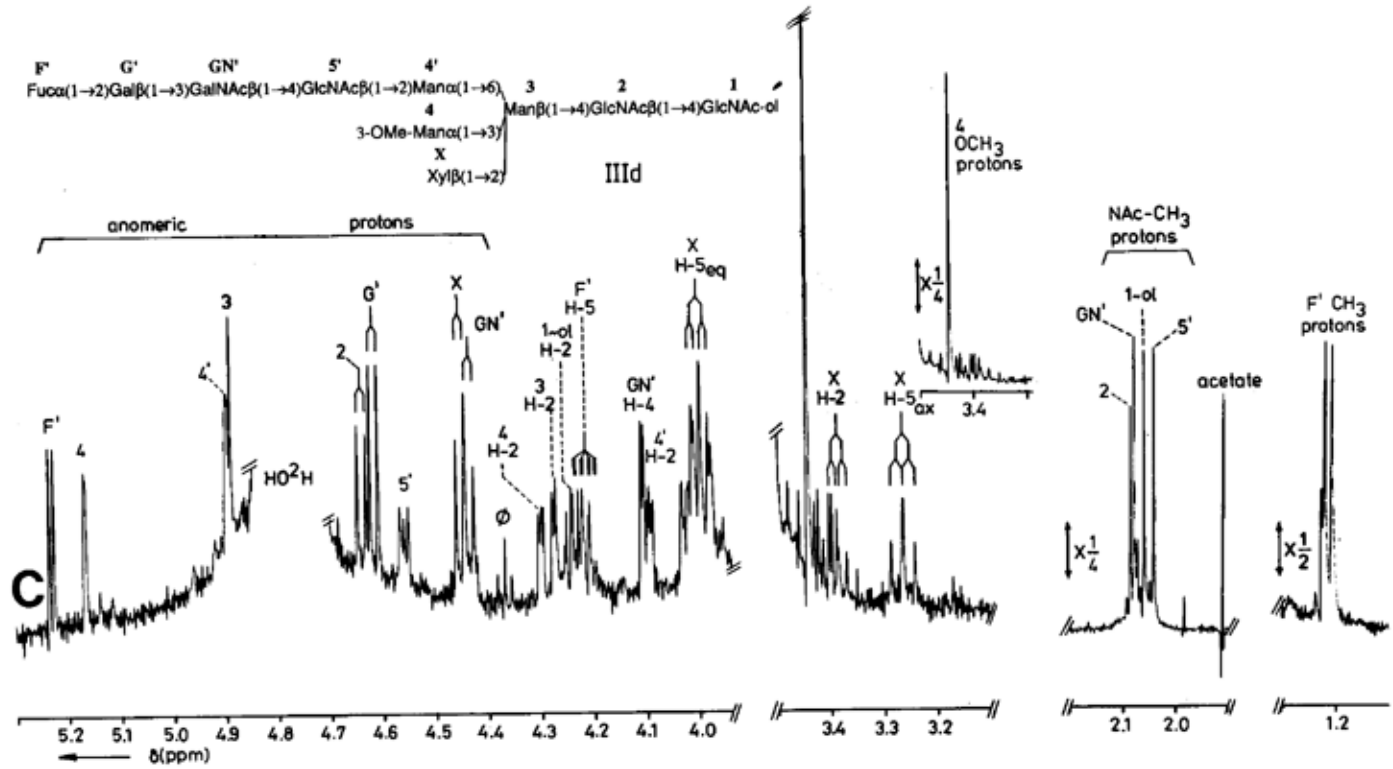

F'
G'
GN'
5'
4'
Fucα(1→2)Galβ(1→3)GalNAcβ(1→4)GlcNAcβ(1→2)Manα(1→6)
3
2
1
Manβ(1→4)GlcNAcβ(1→4)GlcNAc-ol
4
3-OMe-Manα(1→3)
X
Xylβ(1→2)
IIId
anomeric
protons
4
OCH3
protons
NAc-CH3
protons
C
HO2H
acetate
F' CH3
protons
δ(ppm)

Fig. 2B, C

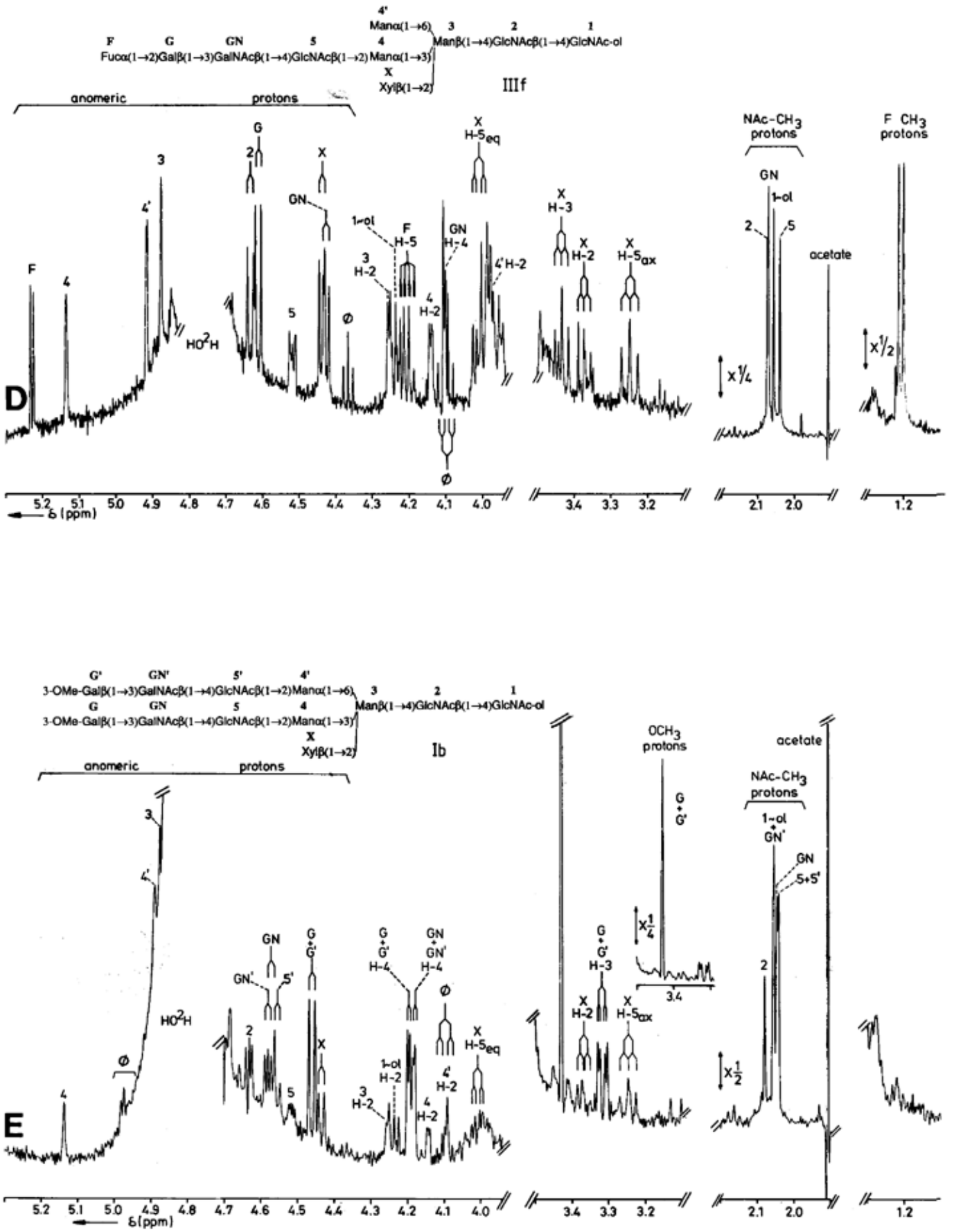

F Fucα(1→2)
G Galβ(1→3)
GN GalNAcβ(1→4)
5 GlcNAcβ(1→2)
4' Manα(1→6)
4 Manα(1→3)
X Xylβ(1→2)
3 Manβ(1→4)
2 GlcNAcβ(1→4)
1 GlcNAc-ol
IIIf
anomeric protons
D
NAc-CH3 protons
F CH3 protons
acetate
δ (ppm)
3-OMe-Galβ(1→3)
Ib
OCH3 protons
E
δ(ppm)

Fig. 2D, E

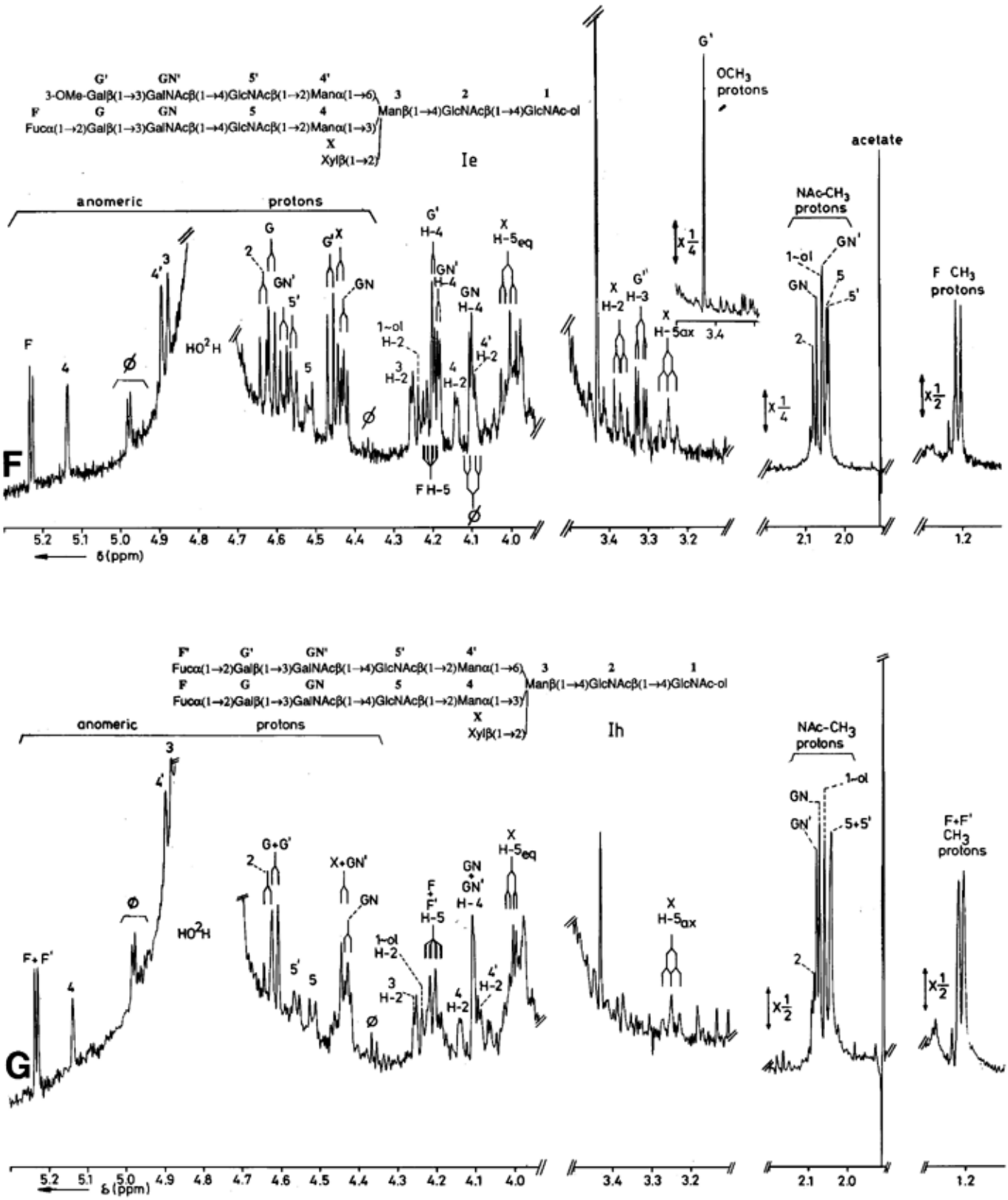


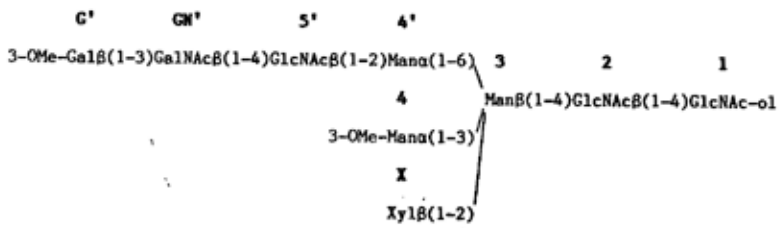
Fig. 2F, G

unassigned anomeric proton signals are about 8 Hz, it can be concluded that 3-OMe-Gal, GalNAc and GlcNAc all have β configuration. The chemical shifts of the 3-OMe-Gal structural-reporter groups (H-1, H-3, H-4 and OCH_3) in III b could be found by studying first the shift effects going from β-D-galactopyranose to 3-*O*-methyl-β-D-galactopyranose (Table 4). It can be seen that 3-*O*-methylation of Gal results in an upfield shift of H-3 ($\Delta\delta = -0.303$ ppm) and a

Table 4. *^{1}H-NMR data for the β-D-glycopyranoses of galactose and 3-O-methyl galactose*
Chemical shifts are in ppm downfield from sodium 4,4-dimethyl-4-silapentane-1-sulfonate in ^2H_2O at 27°C acquired at 500 MHz (but were actually measured relative to internal acetone: δ = 2.225 ppm)

Protons	Chemical shift of	
	β-Gal	3-OMe-β-Gal
	ppm	
H-1	4.575	4.583
H-2	3.484	3.501
H-3	3.638	3.335
H-4	3.921	4.204
H-5	3.698	3.672
H-6a	3.764	3.782
H-6b	3.733	3.746
OCH_3		3.439
Coupling		
	Hz	
$J_{1,2}$	7.9	7.9
$J_{2,3}$	10.0	9.9
$J_{3,4}$	3.5	3.4
$J_{4,5}$	<1.0	<1.0
$J_{5,6a}$	7.8	7.8
$J_{5,6b}$	4.2	5.0
$J_{6a,6b}$	−11.5	−11.3

to Man-4′. The H-1 signal of GlcNAc-5′ was identified by one dimensional NOE-difference spectroscopy. Presaturation of the Man-4′ H-2 signal at δ = 4.097 ppm gave rise to an NOE effect at the doublet at δ = 4.554 ppm, which is interpreted as an interglycosidic effect on the H-1 of GlcNAc-5′. A similar chemical shift was found for the H-1 signal of GlcNAc-5′ for laccase from sycamore cells (*Acer pseudoplanianus* L.) (δ = 4.553 ppm; with acetone at δ = 2.225 ppm) [13], wherein GlcNAc-5′ is linked in the same way to a Xyl-containing core structure. Furthermore, the H-1 signal of GlcNAc-5′ resonates at the same position as GlcNAc H-1 in the compound GalNAcβ(1-4)GlcNAcβ(1-2)Manα-O$(CH_2)_8$-$COOCH_3$ (δ = 4.559 ppm) (O. Hindsgaul, personal communication). It is noteworthy that a virtual coupling exists for the GlcNAc-5′ H-1 signal in the spectrum of this compound as well as in the spectrum of III b. Consequently, the remaining anomeric proton signal at δ = 4.580 ppm in the spectrum of III b is attributed to GalNAc. Two Gal or GalNAc H-4 signals, easily recognized from the typical coupling pattern ($J_{3,4}$ = 3.5 Hz, $J_{4,5}$ < 1 Hz), resonate outside the bulk of skeleton protons. As the signal resonating at δ = 4.199 ppm belongs to the H-4 signal of 3-OMe-Gal, the signal at δ = 4.184 ppm is attributed to GalNAc H-4. A similar value (δ = 4.157 ppm) was found for the H-4 signal of GalNAc in Galβ(1-3)GalNAcβ(1-4)Galβ(1-4)Glc [14]. The assignment of the two remaining NAc signals (δ = 2.041 ppm and δ = 2.053 ppm) will be described later. Based on ^{1}H-NMR spectroscopy, in combination with sugar, methylation, and deamination analysis, the main component of III b has the following structure:

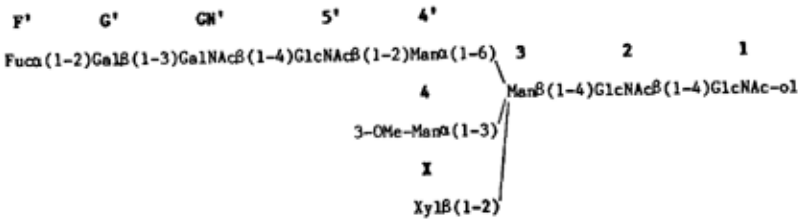


downfield shift of H-4 ($\Delta\delta$ = 0.283 ppm). In consequence, these protons will resonate outside the bulk of skeleton protons (3.9 > δ > 3.5 ppm) in the spectrum of III b. By selective ^{1}H decoupling of the H-3 signal of 3-OMe-Gal (δ = 3.316 ppm) in III b, H-2 (δ = 3.538 ppm) and H-4 (δ = 4.199 ppm) are found by difference spectroscopy. To assign the two doublets at δ = 4.46 ppm selective irradiation was carried out, yielding H-2 of Xyl and H-2 of 3-OMe-Gal in a difference spectrum. Since the signal at δ = 4.450 ppm belongs to Xyl H-1, the anomeric signal at δ = 4.461 ppm can be assigned to 3-OMe-Gal. Concerning the assignment of the OCH_3 signals in the spectrum of III b, the OCH_3 signal at δ = 3.441 ppm, which is also present in the spectrum of IV a [2], was attributed to Man-4. The other OCH_3 signal (δ = 3.429 ppm), which is not present in the spectrum of IV a, is assigned to 3-OMe-Gal.

The sugar residue attached to Man-4′, which has to be either GlcNAc or GalNAc, was traced by degradation via deamination of III b. This method yields, after reduction, free 2,5-anhydro-mannitol in the case of a -GalNAc-GlcNAc- sequence and free 2,5-anhydro-talitol when the sequence is -GlcNAc-GalNAc-. Because only 2,5-anhydro-mannitol and no trace of 2,5-anhydro-talitol was detected by GLC-MS, it can be concluded that GlcNAc, and not GalNAc, is attached

For fraction III d (sugar analysis, Table 1), the structural-reporter-group regions of the ^{1}H-NMR spectrum are presented in Fig. 2C. The equal intensity of the anomeric proton signals points to the presence of a single component. Comparison of the spectra of III d and III b shows that the structural element -4)GlcNAcβ(1-2)Manα(1-6)[3-OMe-Manα(1-3)][Xylβ(1-2)]Manβ(1-4)GlcNAcβ(1-4)GlcNAc-ol is present in both compounds. Sugar and methylation analysis (Tables 1 and 2) reveal that a linear trisaccharide has to be present as extension of the core, consisting of a terminal Fuc residue, a Gal residue substituted at C-2 and a GalNAc residue substituted at C-3. In the spectrum of III d a specific set of structural-reporter groups for Fuc is present (H-1, δ = 5.229 ppm; H-5, δ = 4.213 ppm; CH_3, δ = 1.209 ppm), which is indicative for Fuc in α(1−2)-linkage to Gal [9, 15]. This implies that the sequence of the extending trisaccharide has to be Fucα(1-2)Galβ(1-3)GalNAc. Two anomeric proton signals still have to be assigned in the spectrum of III d. Fucosylation of Gal causes a downfield shift of its anomeric proton signal [9, 16], so the sharp H-1 signal at δ = 4.615 ppm in the spectrum of III d is assigned to Gal. The H-1 and H-4 signals of GalNAc are shifted upfield ($\Delta\delta$ = −0.146 and −0.079 ppm, respectively) going from III b to III d. For the assignment of the *N*-acetyl signals, comparison of the

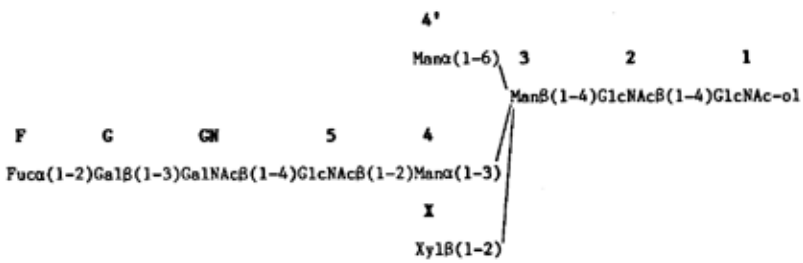
spectrum of IIIb with that of IIId reveals that only one NAc signal is shifted. This signal (δ = 2.053 ppm in IIIb and δ = 2.077 ppm in IIId) is assigned to GalNAc, since this residue is closest to either the 3-*O*-methylated or fucosylated Gal. As the *N*-acetyl signals of GlcNAc-**1**-ol and GlcNAc-**2** resonate at δ = 2.058 ppm and δ = 2.084 ppm, respectively in IIId, the *N*-acetyl signal at δ = 2.040 ppm is assigned to GlcNAc-**5**′. So the structure of IIId is



The spectrum of IIIc (Fig. 2B) reveals the presence of a mixture of two components in approximately equal amounts. This is judged from the occurrence of two Man-**4** and two Man-**4**′ H-1 signals, which all have the same intensity.

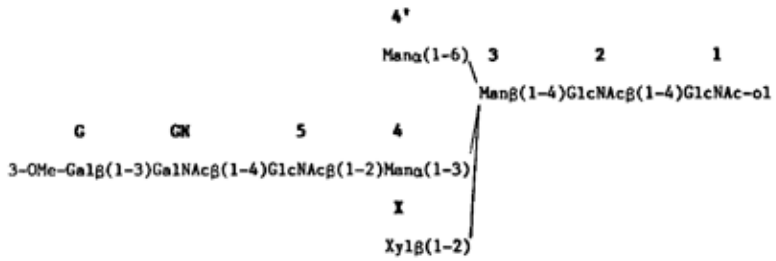
Comparison of IIIc with R, reveals a terminal Man-**4**′ (H-1, δ = 4.915 ppm; H-2, δ = 3.979 ppm), and comparing IIIc with IIIb and IIId, shows a substituted Man-**4**′ (H-1, δ = 4.897 ppm; H-2, δ = 4.090 ppm). Comparing IIIc with IVa, reveals a set of structural-reporter groups which are indicative for a terminal 3-*O*-methylated Man-**4** (H-1, δ = 5.166 ppm; H-2, δ = 4.298 ppm; OCH_3, δ = 3.442 ppm). The presence of a substituted Man-**4** in IIIc can be deduced from the Man-**4** H-1 (δ = 5.136 ppm) and H-2 (δ = 4.145 ppm) signals (see IIIf). From the equal retention time on HPLC, it can be concluded that both components have approximately the same size. So, the terminal Man-**4**′ residue and the substituted Man-**4** residue belong to one component, designated IIIc1, and the substituted Man-**4**′ together with the terminal 3-*O*-methylated Man-**4** belong to the other component, designated IIIc2.

For fraction IIIf the structural-reporter-group regions of the ^{1}H-NMR spectrum are presented in Fig. 2D. Sugar analysis (Table 1) indicates for IIIf a similar composition to IIId, except for the 3-*O*-methylation of Man. Comparing the spectra of IIIf and R, shows that the structural element Manα(1-6)[Xylβ(1-2)]Manβ(1-4)GlcNAcβ(1-4)GlcNAc-ol is present in both compounds. Man-**4**′ occurs in a terminal position, as is evident from the chemical shifts of its H-1 and H-2 signals and also from that of the NAc signal of GlcNAc-**2**. The presence of terminal Man-**4**′ implies that Man-**4** has to be substituted. The H-1 and H-2 signals of Man-**4** and also the H-1 and H-2 signals of Man-**3**, and the chemical shifts of the Xyl structural-reporter groups are shifted upfield, going from IIId to IIIf, due to the extension of Man-**4**. H-1 of GlcNAc-**5**, easily recognized by the virtual coupling, is shifted to δ = 4.518 ppm, which is comparable with the chemical shift found for GlcNAc-**5** H-1 in laccase (δ = 4.522 ppm; with acetone at δ = 2.225 ppm) [13]. As in laccase the chemical shift of the *N*-acetyl group is not affected, going from GlcNAc-**5**′ to GlcNAc-**5**. The presence of the Fucα(1-2)Galβ(1-3)GalNAcβ(1-4) structural element can be concluded from comparing the spectrum of IIId with that of IIIf. The H-1, H-5 and CH_3 signals of Fuc are essentially identical for IIIf and IIId. This implies that the structural element Fucα(1-2)Gal is present, as the Fuc parameters are very sensitive for the type of sugar chain to which Fuc is attached [9, 15]. Furthermore, Gal H-1 and GalNAc H-4 resonate at essentially the same positions for IIIf and IIId. However, GalNAc H-1 (δ = 4.425 ppm) and NAc (δ = 2.070 ppm) are found at slightly upfield positions, going from IIId to IIIf. This is due to the attachment of GalNAc to GlcNAc-**5** (IIIf) instead of to GlcNAc-**5**′ (IIId). So the structure of IIIf has to be



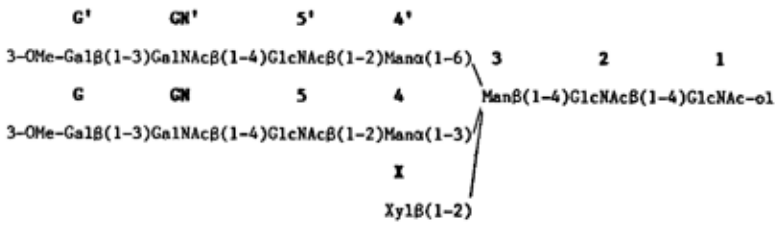
The presence of the structural element 3-OMe-Galβ(1-3)GalNAc is revealed by comparing IIIc with IIIb. The chemical shifts of H-1, H-3, H-4 and OCH_3 of Gal, and H-1 and H-4 of GalNAc are essentially identical for IIIc and IIIb. The *N*-acetyl signal of GalNAc is sensitive to the attachment of GalNAc to either GlcNAc-**5** or -**5**′ (compare IIId with IIIf). GalNAc NAc resonates at δ = 2.048 ppm for IIIc, which is profoundly different from δ = 2.053 ppm found in the spectrum of IIIb. This is indicative for linkage of GalNAc to GlcNAc-**5**, instead of to GlcNAc-**5**′, so a 3-OMe-Galβ(1-3)GalNAcβ(1-4)GlcNAcβ(1-2)Manα(1-3) structural element is present for IIIc1. The H-1 and NAc signals of GlcNAc-**5** resonate at δ = 4.522 ppm and δ = 2.041 ppm, respectively. The presence of the core-structural element Xylβ(1-2)Manβ(1-4)GlcNAcβ(1-4)GlcNAc-ol in IIIc is concluded from GlcNAc-**1**-ol H-2 (δ = 4.237 ppm) and NAc (δ = 2.057 ppm), GlcNAc-**2** H-1 (δ = 4.633 ppm) and NAc (δ = 2.074 ppm), Man-**3** H-1 (δ = 4.876 ppm) and H-2 (δ = 4.258 ppm), and Xyl H-1 (δ = 4.433 ppm), H-5_{ax} (δ = 3.250 ppm) and H-5_{eq} (δ = 4.012 ppm), when IIIc is compared with IIIf. From the chemical shift of NAc of

GlcNAc-**2** can be concluded that a terminal Man-**4′** is attached to this core structure (compare IIIc with R and IIIf). This is corroborated by the H-1 and H-5_{ax} signals of Xyl. Comparing the H-1 and H-2 signals of Man-**3** in IIIc and IIIf reveals that Man-**4** is substituted. So, the structure of IIIcl is a core-structural element consisting of Manα(1-6)[Manα(1-3)][Xylβ(1-2)]Manβ(1-4)GlcNAcβ(1-4)GlcNAc-ol is deduced from comparing Ib with R, IVa, IIIb, IIIc, IIId and IIIf. Comparing Ib with IIIb shows the presence of the 3-OMe-Galβ(1-3)GalNAcβ(1-4)GlcNAcβ(1-2)Manα(1-6) structural



The presence of the structural element Fucα(1-2)Galβ(1-3)GalNAcβ(1-4) in IIIc can be concluded from comparison of IIIc with IIId and IIIf. The chemical shift values of the H-1 and CH_3 signals of Fuc, the H-1 signal of Gal, and the H-4 signal of GalNAc are essentially identical for all three compounds. The attachment of this element to GlcNAc-**5′** can be seen from H-1 (δ = 4.436 ppm) and NAc (δ = 2.078 ppm) of GalNAc in IIIc, which are essentially identical with the values found for IIId, but differ from those presented for IIIf. The H-1 signal of GlcNAc-**5′** is found at δ = 4.558 ppm, and its NAc signal at δ = 2.041 ppm. The presence of a small NAc signal of GlcNAc-**2** at δ = 2.085 ppm, a H-2 signal of Man-**3** at δ = 4.270 ppm, in combination with signals for Xyl H-1 at δ = 4.454 ppm and H-5_{ax} at δ = 3.262 ppm, is indicative for another, incomplete, core-structural element. The virtual absence of the H-1 signal of Man-**3**, and the absence or low intensity of the H-2 and NAc signals of GlcNAc-**1**-ol, and the H-1 and NAc signals of GlcNAc-**2** are typical for the presence of a hydrazinolysis artefact. The same effects are observed in the spectra of the hydrazinolysis artefacts IIIa and IIIe (data not shown). The H-2 signal of Man-**3** and the H-1 and H-5_{ax} signals of Xyl are indicative for the presence of a 3-*O*-methylated terminal Man-**4** in this structure (compare IIIc with IVa, IIIb and IIId). The NAc signal of GlcNAc-**2** shows that Man-**4′** is substituted (compare IIIc with IIIb and IIId). This implies that IIIc2 is a hydrazinolysis artefact of IIId.

element in Ib, while comparing Ib with IIIc1 reveals the presence of the 3-OMe-Galβ(1-3)GalNAcβ(1-4)GlcNAcβ(1-2)Manα(1-3) branch. The presence of two 3-OMe-Galβ(1-3)GalNAcβ(1-4)GlcNAcβ(1-2)Man elements can be concluded from the doubling of the intensities of the 3-OMe-Gal structural-reporter groups (H-1, H-3, H-4 and OCH_3) and the H-4 signal of GalNAc, compared to the other H-1 and H-2 signals. Furthermore GalNAc H-1 signals are found at δ = 4.573 ppm and δ = 4.583 ppm, which correspond with the chemical shifts found in IIIc1 and IIIb, respectively. Both H-4 atoms of GalNAc resonate at δ = 4.184 ppm in the spectrum of Ib. The assignment of the GalNAc NAc signals is based on comparison of the spectrum of Ib with the spectra of IIIc1 and IIIb. The chemical shifts of the structural-reporter groups of GlcNAc-**5** (H-1 and NAc) in Ib are essentially identical with those found for IIIc1, while the shifts of GlcNAc-**5′** in Ib correspond with those determined in IIIb. The substitution of Man-**4** can be concluded from comparing the H-1 and H-2 signals of Man-**4** in Ib with those signals in IIIc1 and IIIf. This is evidenced by the H-1 and H-2 signals of Man-**3** and by the H-1 and H-5_{ax} signals of Xyl, which are resonating at essentially the same positions in the spectra of Ib, IIIc1 and IIIf. Man-**4′** H-1 and H-2 are found at δ = 4.893 ppm and δ = 4.096 ppm, respectively, which is indicative for a substituted Man-**4′** (compare Ib with IIIb and IIId). This is corroborated by GlcNAc-**2** NAc, resonating at δ = 2.080 ppm. This implies that a diantennary structure is present in Ib with the structure

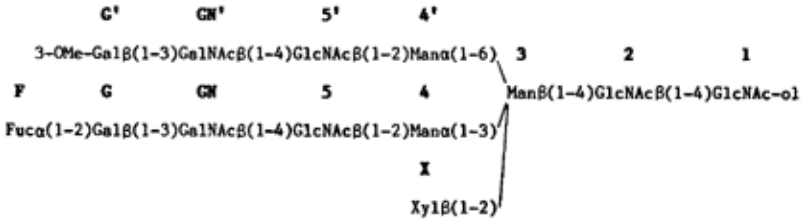


Bio-Gel P-4 fraction I

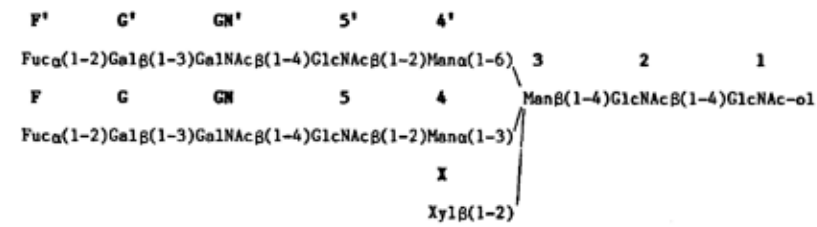
Bio-Gel P-4 fraction I was separated on HPLC into nine fractions, denoted Ia–Ii (Fig. 1A). The ^{1}H-NMR spectrum of Ib (Fig. 2E; sugar analysis, Table 1) shows the presence of a single component, which is concluded from the presence of a single set of H-1 and H-2 signals of Man, and from the intensity of the anomeric proton signals. The presence of

The ^{1}H-NMR spectrum of Ie (Fig. 2F; sugar analysis, Table 1) consists of a single compound, as can be seen from the equal intensity of the anomeric proton signals. Comparing the spectra of Ie and Ib reveals that both compounds have the same Manα(1-6)[Manα(1-3)][Xylβ(1-2)]Manβ(1-4)-GlcNAcβ(1-4)GlcNAc-ol core structure, substituted at Man-**4** and Man-**4′**. When the ^{1}H-NMR data of Ie are compared with those of IIIb and IIIc1 it can be seen that a 3-OMe-

Galβ(1-3)GalNAcβ(1-4)GlcNAcβ(1-2) branch is present in Ie. The attachment of this branch to Man-**4**′ is based on the NAc signal of GalNAc, which is found at $\delta = 2.054$ ppm (compare Ie with IIIb). Comparing the spectra of Ie, IIId and IIIf shows a Fucα(1-2)Galβ(1-3)GalNAcβ(1-4)GlcNAcβ(1-2) structural element in all three compounds. This branch is attached to Man-**4** as can be seen from the GalNAc structural-reporter groups H-1 ($\delta = 4.427$ ppm) and NAc ($\delta = 2.070$ ppm). It has to be noted that to distinguish between the H-1 signal of GalNAc and the H-1 signal of Xyl, which are both resonating at $\delta = 4.43$ ppm, selective decoupling of the H-2 signal of Xyl was utilized. The H-1 signal of Xyl was found at $\delta = 4.436$ ppm, so the chemical shift at $\delta = 4.427$ ppm is assigned to the H-1 signal of GalNAc. The positions of the H-1 and NAc signals of GalNAc are essentially the same in Ie and IIIf. So, in conclusion Ie has the following structure



The ^{1}H-NMR spectrum of Ih (Fig. 2G; sugar analysis, Table 1) reveals a single component, as can be derived from the anomeric structural-reporter groups. This compound has a diantennary structure, which has the same Manα(1-6)-[Manα(1-3)][Xylβ(1-2)]Manβ(1-4)GlcNAcβ(1-4)GlcNAc-ol core structure, substituted at Man-**4** and Man-**4**′, as already reported for Ib and Ie. The presence of two Fucα(1-2)Galβ(1-3)GalNAcβ(1-4)GlcNAcβ(1-2) structural elements is deduced from comparing the structural-reporter groups of Ih with those of IIId and IIIf. Comparing of Ih with IIId shows that one element is attached to Man-**4**′, and comparing Ih with IIIf reveals the linkage of the other element to Man-**4**. The doubling of intensity of the H-1, H-5 and CH_3 signals of Fuc, the H-1 signal of Gal and the H-4 signal of GalNAc, and the occurrence of two different sets of H-1 and NAc signals of GalNAc, and H-1 signals of GlcNAc-**5**/-**5**′, is in agreement with the fact that in Ih this structural element is indeed present in both branches. The structure of Ih is the following:



The remaining fractions contain too little carbohydrate material to analyze, but consist in part of hydrazinolysis artefacts.

DISCUSSION

Hydrazinolysis is an adequate method to liberate all types of Asn-linked carbohydrate chains. Unfortunately, it inevitable leads to chemical modifications in the *N*,*N*′-diacetylchitobiose moiety for a part of the molecules [16]. With the sophistication of separation methods, this will lead to additional fractions, giving a misleading impression of the number of different carbohydrate chains. Furthermore, the presence of hydrazinolysis artefacts can hamper the structural elucidation.

Hemocyanin from *L. stagnalis* contains a large variety of carbohydrate chains, which are extensions of the Xyl-containing core structure described for *H. pomatia* and its variant *L. stagnalis* hemocyanin [1, 2]. A Xyl-containing core structure has been described for the carbohydrate chains of the plant glycoprotein laccase [13], which has also a binuclear copper site, like the hemocyanins. The extended carbohydrate chains, described above, are all present as mono- or diantennary structures, which contain GalNAc in an *N*-glycosidic carbohydrate chain. GalNAc is substituted with 3-OMe-Galβ(1-3) or Fucα(1-2)Galβ(1-3). The latter structural elements may have a terminating effect on the elongation of the carbohydrate chain. In Ic there is some evidence for the presence of a structure, which is the reversed form of Ie, having the Fucα(1-2)Galβ(1-3)GalNAcβ(1-4)GlcNAcβ(1-2) branch attached to Man-**4**′ and the 3-OMe-Galβ(1-3)GalNAcβ(1-4)-GlcNAcβ(1-2) branch attached to Man-**4**. With this structure present, all four possible diantennary structures would occur.

The presence of the GalNAc-GlcNAc sequence in an *N*-glycosidic chain was reported for lutropin. The carbohydrate chain is demonstrated to contain sulfated GalNAc in β(1-4) linkage to GlcNAc [17]. In this case, sulfate may function as biosynthetic stop signal. The terminating function of 3-*O*-methylation is also observed for 3-OMe-Man, as this residue, is only present in terminal position ([18] and this study). Although Man-**4** as well as Man-**4**′ can be 3-*O*-methylated, in the extended structures only 3-*O*-methylated Man-**4** has been found. It might be that the 3-*O*-methylation of Man-**4**′ is biosynthetically more difficult. Remarkably, 3-*O*-methylated mannose is not present in hemocyanin from *H. pomatia*. It has to be noted that 3-OMe-Man, in conjunction with GalNAc, have also been found as sugar constituents of hemoglobin from the gastropod *Planorbis corneus* [19]. So it seems that 3-*O*-methylation in animal glycoproteins is not restricted to hemocyanin. With respect to the Fuc-containing carbohydrate structures, it is worth noting that the blood group H type 4 structure Fucα(1-2)Galβ(1-3)GalNAcβ is present [20].

The possession of these unusual carbohydrate structures may explain in part why hemocyanins have a high antigenic potency when injected into vertebrates, although the lack of terminal sialic acid residues may also give a contribution to their striking ability to stimulate the vertebrate immune system.

We thank Mr F. Roerink for the determination of the absolute configuration of the constituent monosaccharides, and Mrs C. E. M. Kolsteeg for carrying out several sugar analyses and the alkaline borohydride experiment. 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).

REFERENCES

1. Van Kuik, J. A., Van Halbeek, H., Kamerling, J. P. & Vliegenthart, J. F. G. (1985) *J. Biol. Chem. 260*, 13984–13988.
2. Van Kuik, J. A., Sijbesma, R. P., Kamerling, J. P., Vliegenthart, J. F. G. & Wood, E. J. (1986) *Eur. J. Biochem. 160*, 621–625.
3. Hall, R. L., Wood, E. J., Kamerling, J. P., Gerwig, G. J. & Vliegenthart, J. F. G. (1977) *Biochem. J. 165*, 173–176.
4. Gerwig, G. J., Kamerling, J. P. & Vliegenthart, J. F. G. (1978) *Carbohydr. Res. 62*, 349–357.
5. Gerwig, G. J., Kamerling, J. P. & Vliegenthart, J. F. G. (1979) *Carbohydr. Res. 77*, 1–7.
6. Turco, S. J. (1981) *Anal. Biochem. 118*, 278–283.
7. Wagner, G. & Wüthrich, K. (1979) *J. Magn. Reson. 33*, 675–680.
8. Ernst, R. R. (1966) *Adv. Magn. Res. 2*, 1–135.
9. Vliegenthart, J. F. G., Dorland, L. & Van Halbeek, H. (1983) *Adv. Carbohydr. Chem. Biochem. 41*, 209–374.
10. Kamerling, J. P. & Vliegenthart, J. F. G. (1982) *Cell Biol. Monogr. 10*, 95–125.
11. Strecker, G., Pierce-Cretel, A., Fournet, B., Spik, G. & Montreuil, J. (1981) *Anal. Biochem. 111*, 17–26.
12. Michalski, J. C., Peter-Katalinic, J., Egge, H., Paz-Parente, J., Montreuil, J. & Strecker, G. (1984) *Carbohydr. Res. 134*, 177–189.
13. Takahashi, N., Hotta, T., Ishihara, H., Mori, M., Tejima, S., Bligny, R., Akazawa, T., Endo, S. & Arata, Y. (1986) *Biochemistry 25*, 388–395.
14. Dorland, L., Van Halbeek, H., Vliegenthart, J. F. G., Schauer, R. & Wiegandt, H. (1986) *Carbohydr. Res. 151*, 233–245.
15. Van Halbeek, H. (1984) *Biochem. Soc. Trans. 12*, 601–605.
16. Van Halbeek, H., Dorland, L., Vliegenthart, J. F. G., Hull, W. E., Lamblin, G., Lhermitte, M., Boersma, A. & Roussel, P. (1982) *Eur. J. Biochem. 127*, 7–20.
17. Green, E. D., Van Halbeek, H., Boime, I. & Baenziger, J. U. (1985) *J. Biol. Chem. 260*, 15623–15630.
18. Lederkremer, G. Z. & Parodi, A. J. (1984) *J. Biol. Chem. 259*, 12514–12518.
19. Wood, E. J. & Chaplin, M. F. (1983) *Life Chem. Rep. Suppl. 1*, 141–142.
20. Le Pendu, J., Lambert, F., Samuelsson, B., Breimer, M. E., Seitz, R. C., Urdaniz, M. P., Suesa, N., Ratcliffe, M., Francois, A., Poschmann, A., Vinas, J. & Oriol, R. (1986) *Glycoconjugate J. 3*, 255–271.