Glycoconjugate J (1984) 1:123—140

The Structure of a Trisialyl Di-antennary *N*-Type Glycopeptide Obtained from Rat Plasma Hemopexin

NICOLE BERNARD[1], ROBERT ENGLER[1], GERARD STRECKER[2], JEAN MONTREUIL[2], HERMAN VAN HALBEEK[3] and JOHANNES F G VLIEGENTHART[3]*

[1]*Laboratoire des Protéins de la Réaction Inflammatoire, U.E.R. Biomédicale des Saint-Pères, F-75270 Paris Cedex 06, France*
[2]*Laboratoire de Chimie Biologique, Université des Sciences et Techniques de Lille I et Laboratoire Associé au C.N.R.S. no. 217, F-59655 Villeneuve d'Ascq Cedex, France*
[3]*Department of Bio-Organic Chemistry, University of Utrecht, Croesestraat 79, NL-3522 AD Utrecht, The Netherlands*

Received July 10, 1984.

Key words: ^{1}H-NMR spectroscopy, glycopeptide, rat hemopexin

Rat hemopexin is a plasma glycoprotein that contains 18.3% carbohydrate consisting of only *N*-glycosidically-linked oligosaccharide chains. Glycopeptides obtained from hemopexin by Pronase® digestion could be separated on Concanavalin A-Sepharose into three fractions. The lectin-binding fraction has been characterized as a mixture of monosialyl and disialyl di-antennary compounds ending in *N*-acetylneuraminic acid residues α(2-6)-linked to galactose in the respective branches [Bernard N, Lombart C, Strecker G, Montreuil J, Van Halbeek H, Vliegenthart JFG (1983) Biochimie 65:185-92].

The structures of the glycans in the Concanavalin A non-binding fractions were determined by a combination of methylation analysis and 500-MHz ^{1}H-NMR spectroscopy. Some of them appeared to be tri-antennary glycans. However, the major component of these fractions possesses the following structure:

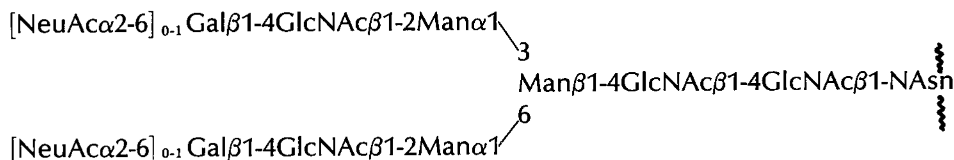
```
                    NeuAcα2
                          ↓
                          6
NeuAcα2-3Galβ1-3GlcNAcβ1-2Manα1
                                  \
                                   3
                                    Manβ1-4GlcNAcβ1-4GlcNAcβ1-NAsn
                                   6
                                  /
NeuAcα2-6Galβ1-4GlcNAcβ1-2Manα1
```


*Author for correspondence

This type of structure has been encountered before in some bovine blood coagulation factors as well as in rat α_1-acid glycoprotein, but the ^{1}H-NMR parameters for it are first reported here. Furthermore, by methylation analysis, the occurrence of the NeuAcα2-8NeuAc disaccharide element was demonstrated in a minor part of the carbohydrate moiety of rat hemopexin. This element has also been reported previously for rat brain glycopeptides.

Hemopexin is a plasma β-glycoprotein which binds heme with high affinity. Its biological function is the removal of excess heme from the bloodstream. If the hemoglobin level exceeds the binding capacity of haptoglobin, the heme formed by dissociation of hemoglobin is bound immediately to hemopexin, thereby counteracting hemolysis as much as possible. The complex formed between heme and hemopexin is then taken up by liver parenchymal cells [1-4]. Structural studies of rat hemopexin have shown that the glycoprotein contains 18.3% (by weight) asparagine-linked carbohydrate chains, which account for the molecular heterogeneity exhibited by hemopexin [2]. In order to elucidate the possible physiological role of the sugar moiety of hemopexin in the heme-binding process, in the recognition of the heme-hemopexin complex by hepatocytes, and in the uptake of asialo-hemopexin by the liver mediated by the hepatic asialoglycoprotein receptor [4], the structures of the carbohydrate chains and their microheterogeneity were studied.

Pronase digestion of rat hemopexin led to a mixture of glycopeptides which was subsequently fractionated by affinity chromatography on Con A-Sepharose. Three fractions were obtained, one of which (III) bound to Con A, while the other two (I and II) did not bind. The structures of the glycopeptides present in fraction III were determined [5] to be mono- and disialyl di-antennary chains of the *N*-acetyllactosamine type:



We report here the structure determination of the glycans present in the Con A-non-binding fractions I and II. A preliminary account of this study has been presented previously [6].

Materials and Methods

Materials

Hemopexin was prepared from rat plasma as described previously [7]. Ultrogel AcA202 was purchased from IBF (Genevilliers, France). Con A-Sepharose was obtained from Pharmacia (Uppsala, Sweden) and Bio-Gel P-2 from Bio-Rad (Richmond, CA, USA). $[^{14}C]$Acetic anhydride (5-10 mCi/mmole) was purchased from New England Nuclear (Boston, MA, USA). Pronase® was obtained from Calbiochem-Behring, (San Diego, CA, USA). Trifluoroacetic anhydride was purchased from Fluka (Buchs, Switzerland). All other reagents were of the highest grade available and were used without further purification. The oligosaccharides used as reference substances for structural studies, namely lactose (Lac), monosialyllactoses (II^6NeuAc-Lac and II^3NeuAc-Lac), disialyllactose (II^3{NeuAc}$_2$-Lac), lacto-*N*-tetraose ($LcOse_4$), sialyllacto-*N*-tetraose a (IV^3NeuAc-$LcOse_4$) and b (III^6NeuAc-$LcOse_4$) and disialyllacto-*N*-tetraose (IV^3NeuAcIII^6NeuAc-$LcOse_4$) were isolated from human milk [8-10].

Isolation and Fractionation of Hemopexin Glycopeptides

After extensive Pronase® digestion of rat hemopexin [11], the glycopeptides obtained were purified by gel filtration on an Ultrogel AcA202 column, and subsequently *N*-$[^{14}C]$ acetylated in their peptide moiety to facilitate monitoring their elution from various columns. Free radioactivity was separated from the labelled glycopeptides by gel filtration on Bio-Gel P-2 (1.5 × 62 cm), eluted with distilled water. The labelled glycopeptides were fractionated by affinity chromatography on Con A-Sepharose as described [5]. The resulting three fractions (I-III) were desalted and separated from α-methyl glucoside by Bio-Gel P-2 filtration (1.5 × 35 cm) by eluting with distilled water.

Desialylation of Hemopexin Glycopeptides

For comparative structural investigations, the asialo analogues of the glycopeptides in Con A fraction II were prepared by treatment of an aliquot of this fraction with 0.05 M trifluoroacetic acid for 30 min at 100°C. The reaction mixture was subsequently passed over Bio-Gel P-2, to remove, *inter alia*, the liberated sialic acid.

Analytical Methods

Radioactivity was determined in a Packard Tricarb liquid scintillation counter. Qualitative carbohydrate analysis of glycopeptides was carried out by classical colourimetric methods [12]. Quantitative analysis was performed by GLC after methanolysis with 0.5 M HCl/methanol for 24 h at 80°C, and trifluoroacetylation [13]. Analyses were carried out on a Varian Aerograph 2700 (glass column packed with 5% OV210; temperature 90-210°C, 2°/min).

Methylation analysis was carried out as follows. Glycopeptides were methylated as described previously [14, 15] and then methanolysed in 0.5 M HCl/methanol. The partially methylated methyl glycosides were acetylated [15] (pyridine/acetic anhydride, 1/1 by vol, 0.2 ml, 20 min, 100°C). The products were analysed by GLC-MS (Riber model 10-10, Rueil-Malmaison, France) using a capillary column (0.35 mm × 60 m) coated with Cpsil-5CB (temperature 130-260°C, 5°/min).

Prior to ^{1}H-NMR spectroscopic analysis, the hemopexin glycopeptides as well as the reference oligosaccharides were repeatedly treated with $^{2}H_2O$ at p^{2}H 7 and room temperature. After each exchange treatment, the materials were lyophilized. Finally, each sample was redissolved in 0.4 ml $^{2}H_2O$ (99.96 atom% ^{2}H, Aldrich, Milwaukee, WI, USA). 500-MHz ^{1}H-NMR spectroscopy was performed on a Bruker WM-500 spectrometer (SON hf-NMR facility, Department of Biophysical Chemistry, Nijmegen University, The Netherlands) operating under control of an Aspect-2000 computer. Experimental details have been described previously [16, 17]. For solvent-peak suppression, a WEFT-pulse sequence (composite 180° pulse - delay - 90° pulse - acquisition) was applied [18]. Resolution enhancement of the spectra was achieved by Lorentzian-to-Gaussian transformation. The probe temperature was kept at 27.0 (±0.1)°C. Chemical shifts are expressed in ppm downfield from internal sodium 4,4-dimethyl-4-silapentane-1-sulphonate (DSS), but were actually measured by reference to internal acetone (δ 2.225 in $^{2}H_2O$ at 27°C), with an accuracy of 0.002 ppm.

Results

Hemopexin (80 mg) was isolated from 100 ml of pooled plasma from 20 Wistar rats. The total carbohydrate content of the glycoprotein was found to be 18.3% [1]. Galactose, mannose, *N*-acetylglucosamine and *N*-acetylneuraminic acid were present as constituent monosaccharides. The molar ratios in which these sugars occur in the intact glycoprotein are given in Table 1.

Table 1. Molar carbohydrate composition of rat plasma hemopexin and of its pronase-digest glycopeptides before and after fractionation on Con A.

Material	Molar ratio of carbohydrates[a]			
	Gal	Man	GlcNAc	NeuAc
Hemopexin	2.0	3	4.0	2.4
Pronase glycopeptides	2.1	3	4.0	2.4
Con A fraction I	2.4	3	4.5	2.9
Con A fraction II	2.0	3	3.9	3.1
Con A fraction III	2.0	3	3.8	2.0

[a] Calculated on the basis of 3 Man residues per glycan molecule.

The carbohydrate moiety of hemopexin was obtained as a mixture of glycopeptides, by Pronase digestion followed by removal of peptides and amino acids over Ultrogel AcA202. Glycopeptides were labeled in the peptide part by *N*-[^{14}C]acetylation. The recovery of carbohydrate in the total mixture of glycopeptides amounted to 90%; the molar ratio of monosaccharides was the same as that for the intact glycoprotein (see Table 1). The glycopeptides were separated by affinity chromatography on Con A-Sepharose into three fractions, in the ration 1:2:6 (see Fig. 1). Fraction I showed no interaction at all with

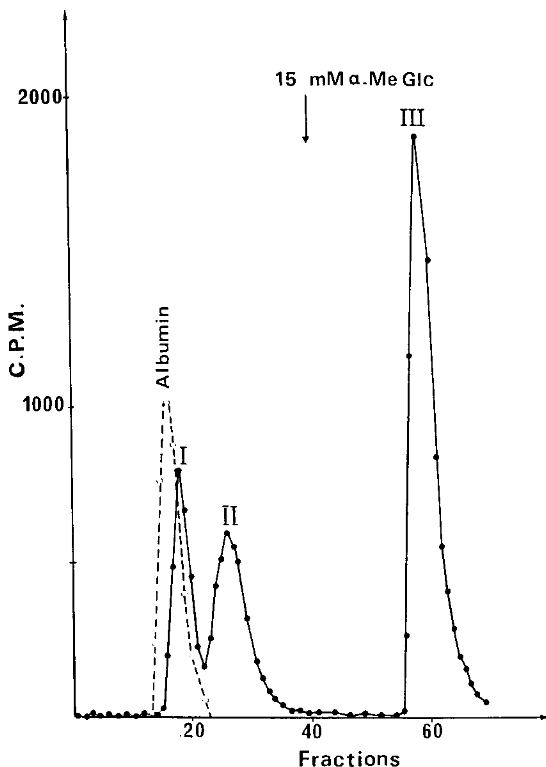


Figure 1. Fractionation of rat hemopexin glycopeptides by Con A-Sepharose. The column (1 × 24 cm) was equilibrated with 0.05 M sodium acetate buffer pH 6.0, containing 100 mM NaCl and 1 mM $CaCl_2$, $MgCl_2$ and $MnCl_2$. Elution was started with 40 ml of this buffer, followed by a 15 mM solution of α-methyl glucoside in the buffer. Fractions of 1 ml were collected and analysed for radioactivity.

Con A. Fraction II was also eluted with the starting buffer, but it was slightly but significantly retarded by Con A. Upon rechromatography, both I and II were eluted as single peaks at the same positions as in the first run, indicating that the binding capacity of the Con A column had not been exceeded. Fraction III bound tightly to the Con A column and was eluted by 15 mM α-methyl glucoside in the starting buffer. The molar ratios of the sugars in the three Con A fractions are included in Table 1.

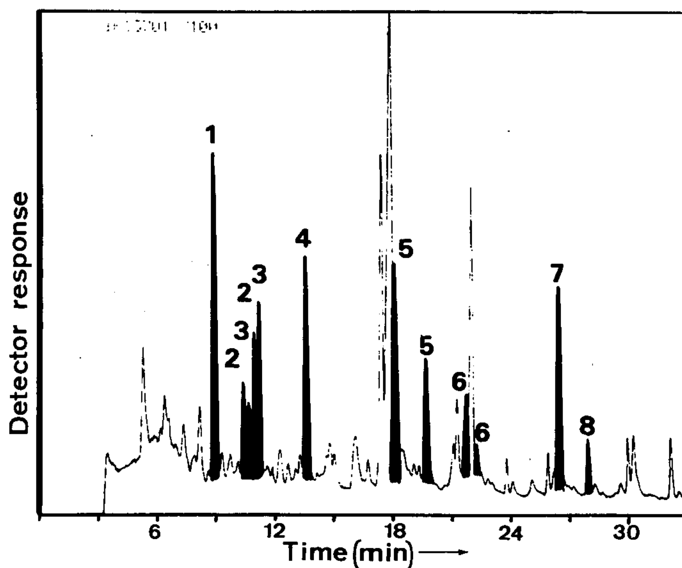


Figure 2. Gas-liquid chromatographic analysis of partially methylated monosaccharide derivatives present in the methanolysate of the permethylated glycopeptide fraction II from hemopexin.

1; $3,4,6$-Me_3-Man. **2**; $2,3,4$-Me_3-Gal. **3**; $2,4,6$-Me_3-Gal. **4**; $2,4$-Me_2-Man. **5**; $3,6$-Me_2-GlcN(Me)Ac. **6**; 4-Me-GlcN(Me) Ac. **7**; $4,7,8,9$-Me_4-Neu(Me)Ac. **8**; $4,7,9$-Me_3-Neu(Me)Ac.
Detailed conditions for the GLC analysis are given in the text.

The primary structures of the glycopeptides present in the major fraction (III) were established by methylation analysis in conjunction with 500-MHz ^{1}H-NMR spectroscopy [5]. The fundamental structure was found to be a di-antenna of the *N*-acetyllactosamine type. The branches are terminated by *N*-acetylneuraminic acid residues in α(2-6)-linkage to galactose. Some heterogeneity was observed in the degree of sialylation, mainly in the *N*-acetyllactosamine unit β(1-2)-linked to the α(1-6)-core mannose. The carbohydrate composition of the minor fraction I (see Table 1) and also its behaviour on TLC after hydrazinolysis and *N*-reacetylation [5] is compatible with that of a tri-antennary *N*-glycosidic structure.

The most intriguing feature of the rat hemopexin carbohydrates is the unusual sugar composition of glycopeptide fraction II. The ratio of *N*-acetylglucosamine to galactose being 4:2 (see Table 1) points to a di-antennary *N*-glycosidic structure; this fraction, however, contains three residues of *N*-acetylneuraminic acid for only two galactoses.

Structure Determination of Con A-fraction II Glycopeptides

The methyl derivatives obtained by methanolysis of the permethylated glycopeptide fraction II were identified by GLC-MS as $3,4,6$-Me_3-Man, $2,3,4$-Me_3-Gal, $2,4,6$-Me_3-Gal, $2,4$-Me_2-Man, $3,6$-Me_2-GlcN(Me)Ac, 4-Me-GlcN(Me)Ac, $4,7,8,9$-Me_4-Neu(Me)Ac and $4,7,9$-Me_3-Neu(Me)Ac, occurring in the approximate ratios of 1.8 : 0.9 : 0.9 : 1.0 : 2.7 : 0.9 : 2.8 : 0.2 (see

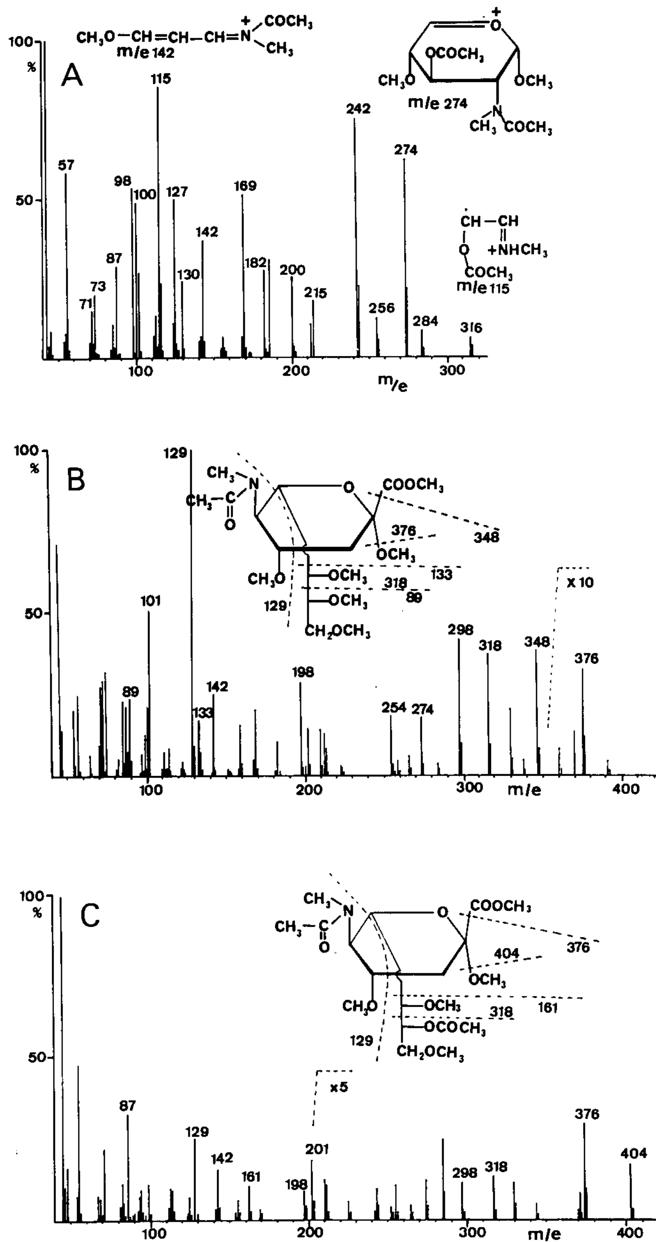


Figure 3. (A) Mass spectrum of the methyl glycoside of 4-mono-*O*-methyl-3,6-di-*O*-acetyl-*N*-methyl-*N*-acetyl-glucosamine [4-Me-GlcN(Me)Ac] (see Fig. 2, peak **6**).

(B) Mass spectrum of the methyl ester methyl glycoside of 4,7,8,9-tetra-*O*-methyl-*N*-methyl-*N*-acetylneuraminic acid [4,7,8,9-Me_4-Neu(Me)Ac] (see Fig. 2, peak **7**).

(C) Mass spectrum of the methyl ester methyl glycoside of 4,7,9-tri-*O*-methyl-8-mono-*O*-acetyl-*N*-methyl-*N*-acetylneuraminic acid [4,7,9-Me_3-Neu(Me)Ac] (see Fig. 2, peak **8**).

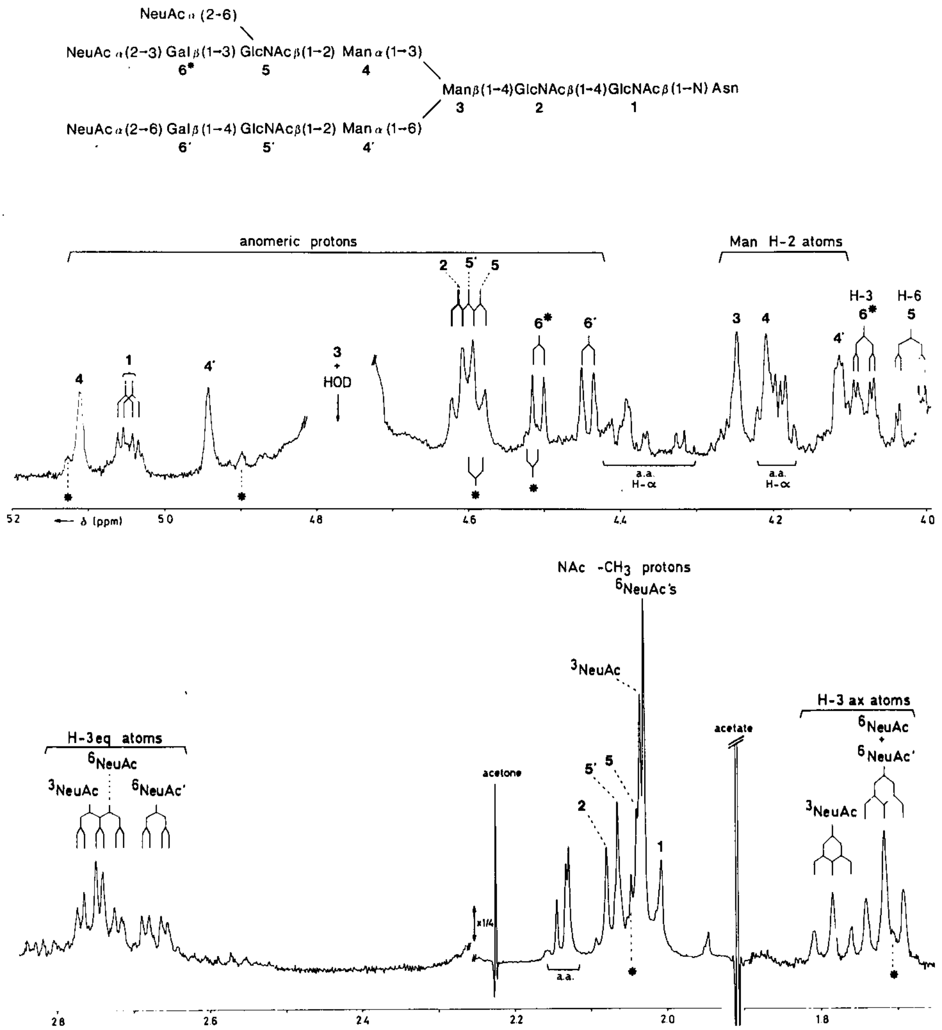


Figure 4. Structural-reporter-group regions of the 500-MHz ^{1}H-NMR spectrum (^2H_2O; p^{2}H 7; 27°C) of Con A-fraction II glycopeptides derived from rat plasma hemopexin. The bold numbers in the spectrum refer to the corresponding monosaccharide residues in the structure. Superscripts at a residue are used only if it is necessary to discriminate between equally named but differently linked monosaccharides. The relative intensity scale of the *N*-acetyl methyl proton region differs from that of the other parts of the spectrum, as indicated. Signals indicated by a.a. stem from amino acid protons. The signals marked by asterisks indicate the presence of one or more minor components in the mixture of glycopeptides which may correspond to the findings by methylation analysis (see text).

also Figs. 2 and 3). These results suggest that indeed di-antennary branching may be present since the ratio of mono(C-2)- to di(C-3 and C-6)-substituted mannose derivatives is 2:1. Furthermore, one of the galactose residues could bear *N*-acetylneuraminic acid

Table 2. ^{1}H Chemical shifts of structural-reporter groups of constituent monosaccharides for glycopeptide fraction II and its desialylated analogue derived from rat plasma hemopexin, together with those for structurally related di-antennary glycopeptides A-C, used as reference compounds.

Reporter group	Residue[a]	Chemical shift[b] in[c]				
		A	GP-II(-SA)	B	C	GP-II
H-1	GlcNAc-**1**	5.094	5.049[d] 5.041	5.088	5.068	5.052[d] 5.045
	GlcNAc-**2**	4.616	4.615	4.616	4.682[e]	4.615
	Man-**3**	4.765	4.76	4.773	4.77	4.77
	Man-**4**	5.121	5.118	5.133	5.118	5.111
	Man-**4'**	4.928	4.926	4.949	4.940	4.943
	GlcNAc-**5**	4.582	4.600	4.603	4.575	4.586
	GlcNAc-**5'**	4.582	4.580	4.603	4.606	4.601
	^{4}Gal-**6**	4.467	—	4.442	4.545	—
	^{3}Gal-**6***	—	4.445	—	—	4.505
	^{4}Gal-**6'**	4.473	4.473	4.447	4.447	4.440
H-2	Man-**3**	4.246	4.249	4.254	4.253	4.247
	Man-**4**	4.190	4.200	4.195	4.190	4.209
	Man-**4'**	4.109	4.113	4.116	4.11	4.114
H-3	Gal-**6(*)**[f]	3.67	3.7	3.67	4.11	4.079
H-3ax	^{6}NeuAc	—	—	1.716[g]	—	1.716[h]
	^{3}NeuAc	—	—	—	1.800	1.784
	^{6}NeuAc'	—	—	1.716	1.717	1.716
H-3eq	^{6}NeuAc	—	—	2.666[g]	—	2.733[h]
	^{3}NeuAc	—	—	—	2.758	2.758
	^{6}NeuAc'	—	—	2.672	2.671	2.670
NAc	GlcNAc-**1**	2.004	2.007	2.002	2.012	2.007
	GlcNAc-**2**	2.079	2.077	2.081	2.094[e]	2.079
	GlcNAc-**5**	2.050	2.045	2.067	2.048	2.039
	GlcNAc-**5'**	2.046	2.048	2.063	2.066	2.064
	^{6}NeuAc	—	—	2.029[i]	2.032[j]	2.029
	^{3}NeuAc	—	—	—	—	2.034
	^{6}NeuAc'	—	—	2.028[i]	2.030[j]	2.029

[a] For numbering of monosaccharide residues, see Fig. 4. It should be noted that Gal-**6** indicates Galβ(1-4)-linked whereas Gal-**6*** denotes Galβ(1-3)-linked to GlcNAc-**5**. A superscript preceding the name of a sugar residue indicates to which position of the adjacent monosaccharide it is linked.

[b] All data were acquired at 500 MHz, for neutral solutions of the compounds in ^2H_2O, at 27°C. Data for reference compounds A, B and C were taken from [16].

[c] In the table-heading, structures are represented by short-hand symbolic notation (compare [16]): ◆– = Man; ●– = GlcNAc; ■– = Gal; ○– = NeuAcα2-6; △– = NeuAcα2-3; □– = Fuc. The types of linkage between Gal and GlcNAc in ■——● branches are indicated by indices to be either β1-3 or β1-4. For the complete structure of hemopexin GP-II, see Fig. 4. GP-II(-SA) means: desialylated GP-II.

[d] The multiplicity of this doublet is due to heterogeneity of the peptide moiety of GP-II (compare [16]).

[e] The chemical shifts for these protons are affected by the presence of Fuc in α(1-6)-linkage to GlcNAc-**1** (see [16]).

[f] For C, Gal-**6** and for GP-II, Gal-**6***.

[g] For NeuAc α(2-6)-linked to Gal-**6**.

[h] For NeuAc α(2-6)-linked to GlcNAc-**5**.

[i j] Assignments may have to be interchanged.

at C-3, and the other galactose might be substituted by *N*-acetylneuraminic acid at C-6. The occurrence of 4-Me-GlcN(Me)Ac (for its mass spectrum, see Fig. 3A) might indicate that the third *N*-acetylneuraminic acid residue is linked to either C-3 or C-6 of an *N*-acetylglucosamine residue that simultaneously bears a second substituent at C-6 or C-3, respectively. The detection of a small amount of 4,7,9-Me_3-Neu(Me)Ac (its mass spectrum is depicted in Fig. 3C and for comparison, the mass spectrum of 4,7,8,9-Me_4-Neu(Me)Ac is shown in Fig. 3B) indicates that part of the *N*-acetylneuraminic acid residues are substituted by another sugar at C-8. According to the ratio of 4,7,8,9-Me_4-Neu (Me)Ac to 4,7,9-Me_3-Neu(Me)Ac, which is 2.8 : 0.2, the glycopeptides containing the C-8 substituted *N*-acetylneuraminic acid residue represent at most 15-20% of the apparent mixture.

In order to establish the complete primary structure of the glycopeptides present in fraction II, the authentic sample as well as its desialylated analogue were subjected to 500-MHz ^{1}H-NMR spectroscopy. The relevant parts of the resolution-enhanced ^{1}H-NMR spectrum of fraction II are depicted in Fig. 4. The chemical shifts of the structural-reporter groups [16] for the trisialyl fraction II, its desialylated counterpart and for some structurally related di-antennary glycopeptides are listed in Table 2.

The spectrum of hemopexin fraction II glycopeptides reveals the characteristic features of di-antennary asparagine-bound *N*-glycosidic carbohydrate chains consisting of the common [16] trimannosyl-*N,N'*-diacetylchitobiose core extended by *N*-acetylglucosamine residues in β(1-2)-linkage to each of the α-linked mannose residues (see Fig. 4). This can be concluded from comparison of the chemical shifts of the structural-reporter groups for GlcNAc-**1** and -**2**, and Man-**3**, -**4** and **4'** of the intact as well as the desialylated fraction II with the corresponding values for reference glycopeptides A and B, being the di-antennary glycopeptide II-6 from asialo-α_1-acid glycoprotein and its *in vitro* α(2-6)-sialylated analogue, respectively [16] (see Table 2). Similarly, it is evident that the α(1-6)-linked branch of fraction II consists of an *N*-acetyllactosamine unit that bears *N*-acetylneuraminic acid in α(2-6)-linkage to Gal-**6'** (for this NeuAc', δH-3ax 1.716 and δH-3eq 2.670; for Gal-**6'**, δH-1 4.440; for GlcNAc-**5'**, δH-1 4.601 and δNAc 2.064; and, most decisive, for Man-**4'**, δH-1 4.943; compare B and C in Table 2). However, two additional sets of *N*-acetylneuraminic acid H-3 signals are observable in Fig. 4, namely, for H-3ax at δ 1.716 and 1.784, and for H-3eq at δ 2.733 and 2.758, respectively. The remaining Gal H-1 doublet is found at δ 4.505 ($J_{1,2}$ = 7.6 Hz), the H-1 signal of GlcNAc-**5** at δ 4.586 and its *N*-acetyl signal at δ 2.039. Such an assembly of chemical shift values could not be correlated with a structural element observed before by ^{1}H-NMR in *N*- or *O*-glycosidic carbohydrate chains of glycoproteins. Combination of the results of methylation analysis and ^{1}H-NMR spectroscopy, in particular the complete characterization of the (1-6)-linked branch of the di-antenna, leaves two reasonable structural alternatives for the (1-3)-linked branch. The latter could consist of a Galβ1-3GlcNAcβ1-2 or a Galβ1-6GlcNAcβ1-2 unit, being extended by an *N*-acetylneuraminic acid residue in an α(2-3)-linkage to galactose and, in addition, by another one in either an α(2-6)- or α(2-3)-linkage to the *N*-acetylglucosamine residue. To discriminate between these possibilities, and to arrive at an unambiguous assignment of the ^{1}H-NMR features of the (1-3)-branch of the di-antenna in terms of its primary structure, use was made of the ^{1}H-NMR spectral features of lactose, lacto-*N*-tetraose, and their monosialyl and disialyl analogues. Pertinent data acquired for the two series of compounds have been compiled in Table 3; the 500-MHz ^{1}H-NMR spectra of the four lacto-*N*-tetraose oligosaccharides are depicted in Fig 5.

The spectrum of monosialyllacto-*N*-tetraose a (Fig. 5B) reveals that the NeuAc-**a** residue that is α(2-3)-linked to a Galβ1-3GlcNAcβ1- unit can be characterized by its H-3ax signal at δ 1.786, its H-3eq signal at δ 2.759 and its *N*-acetyl signal at δ 2.031 (see Table 3). The attachment of such an *N*-acetylneuraminic acid residue affects the chemical shift of H-1 of the substituted galactose residue: the latter shifts downfield from δ 4.439 to 4.508 (Table 3). Also, H-3 of ^{3}Gal emerges out of the bulk of the skeleton proton signals, towards δ 4.085. Both features are in line with the effects of similar extensions ([16, 17]; see, for example, the step from Lac to II^{3}NeuAc-Lac, Table 3). This implies that the NeuAcα2-3 Galβ1-**3**GlcNAcβ1- structural element is clearly discernible from NeuAcα2-3Galβ1-**4** GlcNAcβ1-, on the basis of the combination of the resonance positions of NeuAc H-3ax and Gal H-1 and H-3. The presence of GlcNAcβ1- in this sequence is essential, since if *N*-acetylglucosamine is replaced by *N*-acetylgalactosaminitol, the H-3ax and H-3eq of *N*-acetylneuraminic acid resonate at δ 1.800 and δ 2.774, respectively [17]. Furthermore, it should be noted that in the case of a β(1-3)-linkage between galactose and *N*-acetylglucosamine, the chemical shifts of GlcNAc H-1 and NAc are hardly influenced by α(2-3)-sialylation of galactose (see Table 3).

From the spectrum of monosialyllacto-*N*-tetraose b (Fig. 5C) it can be inferred that the NeuAc-**b** residue which is α(2-6)-linked to a GlcNAcβ1- residue substituted with galactose in a β(1-3)-linkage, shows an H-3ax triplet at δ 1.689, an H-3eq doublet of doublets at δ 2.744 and an *N*-acetyl singlet at δ 2.033. This set of chemical shift values is so far unique for this type of *N*-acetylneuraminic acid environment (compare [16, 17]). Introduction of such a branching *N*-acetylneuraminic acid at GlcNAc C-6 of a Galβ1-3GlcNAcβ1- element results in an upfield shift for H-1 (Δδ -0.03 ppm) and for NAc (Δδ -0.006 ppm) of the *N*-acetylglucosamine, whereas the chemical shifts of the ^{3}Gal structural-reporter groups are not affected (see Table 3).

The spectrum of disialyllacto-*N*-tetraose (Fig. 5D) shows that the sets of chemical shift values proposed above to be characteristic for the NeuAcα2-3Galβ1-3GlcNAcβ1- and Galβ1-3[NeuAcα2-6]GlcNAcβ1- sequences, respectively, are found essentially unaltered, as compared to Figs. 5B and C. Apparently, the *N*-acetylneuraminic acid residues **a** and **b** do not interfere with their mutual environments and the effects which they exert on the chemical shifts of structural reporters of neighbouring residues are independent and additive (see Table 3).

Taking into consideration the results of methylation analysis, and extrapolating the aforementioned ^{1}H-NMR findings for the lacto-*N*-tetraose series to the spectral features of hemopexin glycopeptide fraction II, the set of H-3ax and H-3eq signals at δ 1.784 and δ 2.758 can be attributed to *N*-acetylneuraminic acid in α(2-3)-linkage to Galβ1-3GlcNAc, β(1-2)-linked to Man-**4** in the di-antenna. The resonance position of H-1 of ^{3}Gal (designated Gal-**6***) is in accordance with this proposal (δ 4.505). Moreover, the Gal-**6*** H-3 signal is found at δ 4.079. The upfield shift of Gal-**6*** H-1 upon desialylation (from δ 4.505 to 4.445, see Table 2) confirms the linkage between Gal-**6*** and GlcNAc-**5** to be β1-3, since the latter value corresponds to that for Gal-IV H-1 in lacto-*N*-tetraose (Table 3). Concomitantly, due to the complete removal of sialic acid from glycopeptide II, the H-1 and NAc signals of GlcNAc-**5** shifted downfield to δ 4.600 and δ 2.045, respectively. The chemical shift values for Gal-**6'** and GlcNAc-**5'** H-1 and NAc after desialylation of fraction II are the usual ones for an asialo *N*-acetyllactosamine branch (see Table 2, compound A). The shift effects shown by these reporters and by H-1 of Man-**4'** corroborate the α(2-6)-linkage of

Table 3. ^{1}H Chemical shifts of structural-reporter groups of constituent monosaccharides for lactose, lacto-*N*-tetraose, and several of their sialyl analogues, isolated from human milk.

Residue[a]	Reporter group	Anomer of oligosaccharide[b]	Chemical shift[c] in[d]							
			Lac	II⁶NeuAc-Lac	II³NeuAc-Lac	II³(NeuAc)₂Lac	LcOse₄	IV³NeuAc-LcOse₄	III⁶NeuAc-LcOse₄	IV³NeuAc,III⁶NeuAc-LcOse₄
Glc-I	H-1	α	5.221	5.225	5.220	5.220	5.220	5.221	5.219	5.220
		β	4.663	4.667	4.661	4.659	4.661	4.663	4.661	4.662
	H-2	β	3.284	3.308	3.281	3.287	3.280	3.286	3.281	3.284
⁴Gal-II	H-1	α	4.448	4.429	4.530	4.527	4.439	4.443	4.437	4.440
		β			4.528	4.523				
	H-3	α	3.64	n.d.[e]	4.114	4.095	n.d.	n.d.	n.d.	n.d.
		β			4.110	4.090				
	H-4	α,β	3.910	3.936	3.959	3.969	4.150	4.150	4.175	4.171
³GlcNAc-III	H-1	α	—	—	—	—	4.734	4.741	4.701	4.706
		β	—	—	—	—	4.730	4.737	4.697	4.703
	NAc	α,β	—	—	—	—	2.027	2.030[f]	2.021	2.021
³Gal-IV	H-1	α,β	—	—	—	—	4.439	4.508	4.437	4.500
	H-3	α,β	—	—	—	—	n.d.	4.085	n.d.	4.078
	H-4	α,β	—	—	—	—	3.912	3.943	3.964	3.940
³NeuAc-**a**	H-3ax	α,β	—	—	1.799	1.739	—	1.786	—	1.782
	H-3eq	α	—	—	2.757	2.682	—	2.759	—	2.756
		β				2.679				
	NAc	α,β	—	—	2.030	2.068	—	2.031[f]	—	2.028[g]
⁶NeuAc-**b**	H-3ax	α,β	—	1.739	—	—	—	—	1.689	1.688
	H-3eq	α,β	—	2.715	—	—	—	—	2.744	2.740
	NAc	α,β	—	2.030	—	—	—	—	2.033	2.031[g]
⁸NeuAc-**c**	H-3ax	α,β	—	—	—	1.741	—	—	—	—
	H-3eq	α,β	—	—	—	2.779	—	—	—	—
	NAc	α,β	—	—	—	2.031	—	—	—	—

NeuAc′ to Gal-**6′**. It should be noted that the differences in chemical shifts observed between the corresponding reporter groups of the Gal-**6***-GlcNAc-**5** and the Gal-**6′**-GlcNAc-**5′** branches in the desialylated fraction II (see Table 2) are essentially in accordance with those existing between Galβ1-3GlcNAcβ1-3GalNAcOL (for GlcNAc, δH-1 4.653; δNAc 2.073) and Galβ1-4GlcNAcβ1-3GalNAcOL (for GlcNAc, δH-1 4.631; δNAc 2.083) obtained from cystic fibrosis bronchial mucin glycoproteins [19], and similarly between Galβ1-3GlcNAcβ1-3Gal (for ^{3}Gal, δH-1 4.438) and Galβ1-4GlcNAcβ1-3Gal (for ^{4}Gal, δH-1 4.482), isolated from human meconium [20].

Returning to the spectrum of hemopexin fraction II (Fig. 4), the third set of H-3ax (δ 1.716) and H-3eq (δ 2.733) signals must be ascribed to the *N*-acetylneuraminic acid that is α(2-6)-linked to GlcNAc-**5**. This assignment is primarily based on the aforementioned difference in chemical shift observed for the H-1 and *N*-acetyl signals of GlcNAc-**5** for the sialyl and the desialylated fraction II (Δδ -0.02 ppm and Δδ -0.006 ppm, respectively: see Table 2). These shift decrements are in accordance with the effects of α(2-6)-sialylation of a β(1-3)-substituted *N*-acetylglucosamine residue, described above. Obviously, the precise chemical shifts of GlcNAc H-1 and NAc, and even those of *N*-acetylneuraminic acid α(2-6)-linked to this *N*-acetylglucosamine, are considerably influenced by the type of linkage in which *N*-acetylglucosamine is involved and by the residue to which it is attached (compare glycopeptide II with disialyllacto-*N*-tetraose, Tables 2 and 3).

On the basis of the results of methylation analysis and 500-MHz ^{1}H-NMR spectroscopy, the major component of hemopexin Con A-fraction II possesses the following structure:

```
                NeuAcα2
                      |
                      6
NeuAcα2-3Galβ1-3GlcNAcβ1-2Manα1\
         6*     5          4    \3
                                 Manβ1-4GlcNAcβ1-4GlcNAcβ1-NAsn
                                /6 3         2            1
NeuAc′α2-6Galβ1-4GlcNAcβ1-2Manα1/
          6′    5′         4′
```

---

[a] For coding of monosaccharide residues, see Fig. 5. A superscript preceding the name of a sugar residue indicates to which position of the adjacent monosaccharide it is linked.

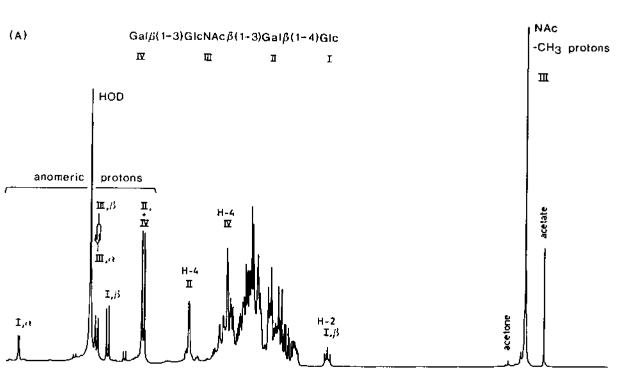
[b] Each of these oligosaccharides ending in the reducing Galβ1-4Glc unit occurs, under the applied measuring conditions, as a mixture of two anomers having Glc in the α- and β-pyranose form, in the ratio of 7:10, respectively (compare [17]).

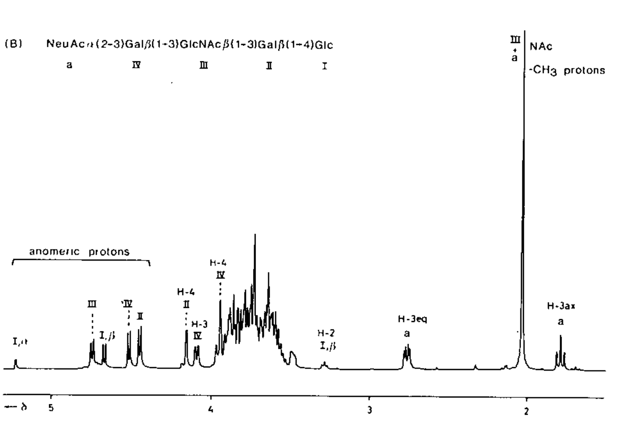
[c] Data were acquired at 500 MHz, for neutral solutions in $^{2}H_2O$ at 27°C. Chemical shifts are in ppm relative to internal DSS (using internal acetone at δ 2.225).

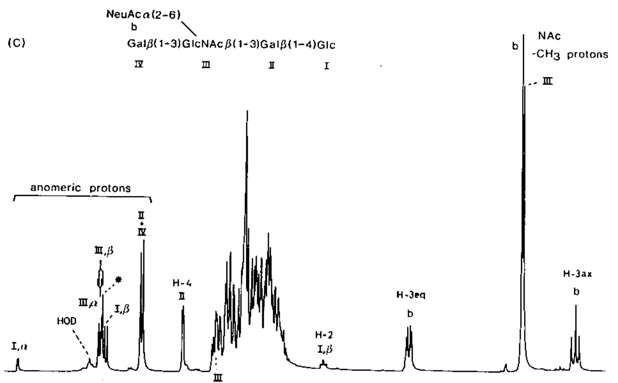
[d] For complete structures of the lacto-*N*-tetraose series, see Fig. 5. In the table-heading, structures are represented by short-hand symbolic notation (compare [16]): ⊞ = Glc; ■– = Gal; ●– = GlcNAc; ○– = NeuAcα2-6; △– = NeuAcα2-3; ▽– = NeuAcα2-8.

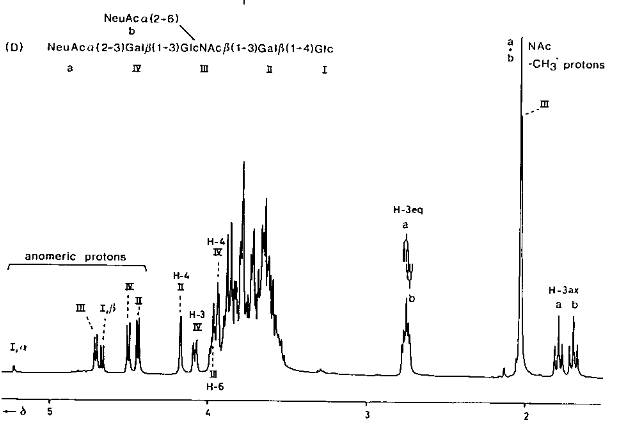
[e] n.d., value could not be determined by inspection.

[f,g] Assignments may have to be interchanged.


(A)
Galβ(1-3)GlcNAcβ(1-3)Galβ(1-4)Glc
IV III II I
HOD
anomeric protons
NAc
-CH3 protons
III
acetate
acetone
III,β
III,α
II + IV
H-4 IV
H-4 II
I,β
H-2 I,β
I,α
δ
5
4
3
2


(B)
NeuAcα(2-3)Galβ(1-3)GlcNAcβ(1-3)Galβ(1-4)Glc
a IV III II I
III + a
NAc
-CH3 protons
anomeric protons
III
I,β
IV
II
H-4 II
H-3 IV
H-4 IV
H-2 I,β
H-3eq a
H-3ax a
I,α
δ
5
4
3
2


(C)
NeuAcα(2-6)
b
Galβ(1-3)GlcNAcβ(1-3)Galβ(1-4)Glc
IV III II I
b
NAc
-CH3 protons
III
anomeric protons
II + IV
III,β
III,α
I,β
HOD
H-4 II
III H-6
H-2 I,β
H-3eq b
H-3ax b
I,α
δ
5
4
3
2


(D)
NeuAcα(2-6)
b
NeuAcα(2-3)Galβ(1-3)GlcNAcβ(1-3)Galβ(1-4)Glc
a IV III II I
a + b
NAc
-CH3 protons
III
anomeric protons
III
I,β
IV
II
H-4 II
H-3 IV
H-4 IV
III H-6
H-3eq a
b
H-3ax a b
I,α
δ
5
4
3
2

Methylation analysis pointed to the occurrence of a minor (less than 20%) glycopeptide constituent in hemopexin fraction II, containing a C-8 substituted *N*-acetylneuraminic acid residue. Careful inspection of the 500-MHz ^{1}H-NMR spectrum of the hemopexin fraction II (Fig. 4) revealed that a small amount of contaminating glycopeptide(s) is indeed present in the mixture. This is deduced from the presence of low-intensity signals marked by asterisks in Fig. 4, in particular from the NeuAc H-3ax signal at δ 1.707, the *N*-acetyl signal at δ 2.046, the Gal H-1 doublet at δ 4.516, the Man-**4'** H-1 signal at δ 4.898 and the Man **4** H-1 signal at δ 5.127. Comparison of these data with those listed for disialyllactose (see Table 3) shows that it is unlikely that a NeuAcα2-8NeuAcα2-3Galβ1-4GlcNAcβ1- element occurs in the minor constituent(s) of fraction II, since the characteristic H-3ax and H-3eq signals of this kind of NeuAcα2-8 are missing. The lack of suitable reference data, for example, for NeuAcα2-8NeuAc linked α2-6 to Galβ1-3/4, α2-3 to Galβ1-3 or α2-6 to GlcNAcβ1-, makes it impossible to locate or to define further the NeuAcα2-8NeuAc element, the occurrence of which was established by methylation analysis.

Now that it has been demonstrated that all *N*-glycosidic carbohydrate chains of rat hemopexin share the common mannotriose branching element, it can be estimated from the molecular weight of the glycoprotein (M_r ~ 62 000 [2]) together with the mannose content (3.7% [1]), that rat hemopexin must have four glycosylation sites. This means at least four out of ten [1] Asp residues as determined per molecule by amino acid analysis, stem from glycosylated Asn.

Discussion

Rat plasma hemopexin is composed of a single polypeptide chain with M_r ~ 62 000. The protein contains as much as 18.3% carbohydrate [1], which is linked to four asparagine residues. Although the physiological role of the sugar chains of hemopexin has not yet been elucidated, it is tempting to suggest that they play a role in the recognition of the heme-hemopexin complex by hepatocytes. In the present study, the structures of the carbohydrate chains of hemopexin were determined after their release from the protein as glycopeptides and subsequent fractionation by affinity chromatography. In accordance with the known binding affinities of Con A [21], the major glycopeptides, namely, the mono- and disialyl di-antennae

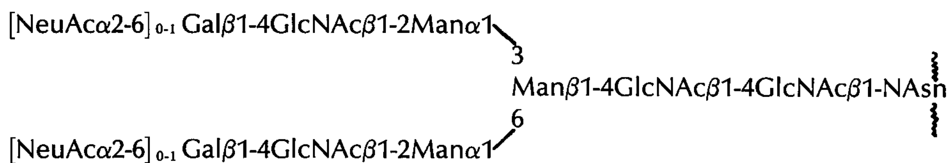


Figure 5. 500-MHz ^{1}H-NMR spectra (^2H_2O; p^{2}H 7; 27°C) of (A) lacto-*N*-tetraose (LcOse$_4$); (B) monosialyllacto-*N*-tetraose a (IV^{3}NeuAc-LcOse$_4$); (C) monosialyllacto-*N*-tetraose b (III^{6}NeuAc-LcOse$_4$); and (D) disialyllacto-*N*-tetraose (IV^{3}NeuAc, III^{6}NeuAc-LcOse$_4$). The HO^{2}H signals in spectra (B), (C) and (D) were suppressed by a WEFT-pulse sequence. Roman numbers in the spectra refer to the corresponding residues in the structures; α and β refer to the respective anomers of the oligosaccharides.

bound tightly to the lectin, whereas the tri-antennary compounds did not [5].

By Con A-fractionation of rat hemopexin glycopeptides, another mixture (II) of glycopeptides was obtained that bound only weakly to the lectin. These were characterized as di-antennae possessing a relatively high content of sialic acid. The latter is probably responsible for the deviating affinity properties with respect to Con A. The main component is composed of a trisialyl di-antennary oligosaccharide linked to a short peptide:

```
                    NeuAcα2
                          ↓
                          6
 NeuAcα2-3Galβ1-3GlcNAcβ1-2Manα1
                                 \3                                    |
                                  Manβ1-4GlcNAcβ1-4GlcNAcβ1-NAsn
                                 /6                                    |
 NeuAc α2-6Galβ1-4GlcNAcβ1-2Manα1
```

A minor component of this fraction contains the NeuAcα2-8NeuAc disaccharide element, which has been described to occur in rat brain glycopeptides [22], but it has not previously been encountered in rat circulating glycoproteins. The structure of this minor constituent of the mixture could not be completely unraveled by the combination of methylation analysis and ^{1}H-NMR spectroscopy. However, at present, FAB-MS studies are being carried out to solve this problem.

Di- and tri-antennary *N*-glycosidically-linked glycans, containing one or more NeuAcα2-3Galβ1-3[NeuAcα2-6]GlcNAcβ1- sequences directly linked to the core, have been found in bovine prothrombin [23], bovine blood coagulation factors X (Stuart factor) [24], IX (Christmas factor) [25] and II [26]. A structural variant, containing *N*-acetylneuraminic acid α(2-4)- instead of α(2-3)-linked to galactose, was reported for bovine cold-insoluble globulin [27]. Therefore, it was suggested [28] that such sialylated Galβ1-3GlcNAc structures might be characteristic for *bovine* glycoproteins. After the discovery of type 1 [Galβ1-3GlcNAc] chains in α_1-acid glycoprotein from *rat* plasma [29], this proposal can now be definitely rejected, since the same type of sialylated type 1 structure appears to occur in the *rat* plasma glycoprotein, hemopexin.

It should be mentioned that previous structural elucidation of the bovine oligosaccharides possessing such sialylated Galβ1-3GlcNAc (type 1) chains was carried out by a combination of methylation analysis and sequential exoglycosidase digestion, performed on the reduced carbohydrate chains liberated from the protein by hydrazinolysis [23-27]. In the present paper, evidence for the trisialyl di-antennary glycopeptide from rat plasma hemopexin is based on methylation analysis, and independently on high-resolution (500-MHz) ^{1}H-NMR spectroscopy. The latter technique can rapidly locate such a type 1 unit in a specific branch of a di-antenna, which also contains a normal *N*-acetyllactosamine (type 2) branch. This result is difficult to achieve by other routes of structural analysis.

Acknowledgements

This investigation was supported by Grant No. 7911753 (CL) from the Institut National de la Santé et de la Recherche Médicale, by the Faculté Broussais, Hôtel-Dieu, Université Paris VI, by the Centre National de la Recherche Scientifique (LA 217, Relation structure-fonction des constituants membranaires and RCP 529, Glucides et glycoconjugués), by the Netherlands Foundation for Chemical Research (SON/ZWO), and by Grant UUKC 83-13 from the Netherlands Foundation for Cancer Research (KWF). The authors are indebted to Mrs. Y. Rondeau (Paris) for technical assistance, to Dr. J-M. Wieruszeski (Lille) and Dr. J.H.G.M. Mutsaers (Utrecht) for assistance in recording the mass and NMR spectra, respectively, and to Ms. C.L.E. de Lint (Utrecht) for typing the manuscript.

References

1 Bernard N, Lombart C, Jayle MF (1975) Biochimie 57:551-57.
2 Bernard N, Lombart C, Waks M (1980) Eur J Biochem 103:271-76.
3 Smith A, Morgan WT (1981) J Biol Chem 256:10902-9.
4 Bernard N, Dodeur M, Dumont JP, Durand G, Engler R, Feger J (1982) Biol Cell 43:173-78.
5 Bernard N, Lombart C, Strecker G, Montreuil J, Van Halbeek H, Vliegenthart JFG (1983) Biochimie 65:185-92.
6 Bernard N, Engler R, Strecker G, Montreuil J, Van Halbeek H, Vliegenthart JFG (1983) in Proc 7th Int Symp Glycoconjugates, eds. Chester MA, Heinegård D, Lundblad A, Svensson S, Secretariat, Lund, p 200.
7 Bernard N, Lombart C, Jayle MF (1976) Biochimie 58:1429-31.
8 Kuhn R, Baer HH (1956) Chem Ber 89:504-11.
9 Kuhn R, Gauhe A (1965) Chem Ber 98:395-413.
10 Grimmonprez L, Montreuil J (1968) Bull Soc Chim Biol 50:843-55.
11 Zinn AB, Marshall JS, Carlson DM (1978) J Biol Chem 253:6761-67.
12 Montreuil J, Spik G (1963) in Méthodes colorimétriques de dosage des glucides totaux, Lille, France.
13 Zanetta JP, Breckenridge WC, Vincendon G (1972) J Chromatogr 69:291-304.
14 Finne J, Krusius T, Rauvala H (1980) Carbohydr Res 80:336-39.
15 Fournet B, Strecker G, Leroy Y, Montreuil J (1981) Anal Biochem 116:489-502.
16 Vliegenthart JFG, Dorland L, Van Halbeek H (1983) Adv Carbohydr Chem Biochem 41:209-374.
17 Vliegenthart JFG, Dorland L, Van Halbeek H, Haverkamp J (1982) Cell Biol Monogr 10:127-72.
18 Van Halbeek H, Vliegenthart JFG, Winterwerp H, Blanken WM, Van den Eijnden DH (1983) Biochem Biophys Res Commun 110:124-31.
19 Van Halbeek H, Dorland L, Vliegenthart JFG, Hull WE, Lamblin G, Lhermitte M, Boersma A, Roussel P (1982) Eur J Biochem 127:7-20.
20 Herlant-Peers M-C, Montreuil J, Strecker G, Dorland L, Van Halbeek H, Veldink GA, Vliegenthart JFG (1981) Eur J Biochem 117:291-300.
21 Debray H, Montreuil J (1981) in Lectins - Biology, Biochemistry, Clinical Biochemistry, ed. Bøg-Hansen TC, Vol. 1, De Gruyter, New York, p 221-30.

22 Finne J (1982) J Biol Chem 257:11966-70.
23 Mizuochi T, Yamashita K, Fujikawa K, Kisiel W, Kobata A (1979) J Biol Chem 254:6419-25.
24 Mizuochi T, Yamashita K, Fujikawa K, Titani K, Kobata A (1980) J Biol Chem 255:3526-31.
25 Mizuochi T, Taniguchi T, Fujikawa K, Titani K, Kobata A (1983) J Biol Chem 258: 6020-24.
26 Berger EG, Buddecke E, Kamerling JP, Kobata A, Paulson JC, Vliegenthart JFG (1982) Experientia 38:1129-62.
27 Takasaki S, Yamashita K, Suzuki K, Iwanaga S, Kobata A (1979) J Biol Chem 254:8548-53.
28 Kobata A (1979) Cell Struct Funct 4:169-81.
29 Yoshima H, Matsumoto A, Mizuochi T, Kawasaki T, Kobata A J Biol Chem 256:8476-84.