# Determination of the Primary Structures of 16 Asialo-Carbohydrate Units Derived from Human Plasma $\alpha_1$-Acid Glycoprotein by 360-MHz $^1$H NMR Spectroscopy and Permethylation Analysis†

Bernard Fournet, Jean Montreuil, Gérard Strecker, Lambertus Dorland, Johan Haverkamp, Johannes F. G. Vliegenthart, J. Paul Binette, and Karl Schmid*

ABSTRACT: The present study was initiated in order to determine the primary structures of the carbohydrate units of desialyzed human plasma $\alpha_1$-acid glycoprotein. In an earlier investigation 16 glycopeptides which originated from four of the five glycosylation sites of this protein were prepared in the homogeneous state. Since the elucidation of the structures of such a large number of carbohydrate units by the conventional chemical and enzymatic procedures would require a very long period of time and considerable amounts of material, a new method—360-MHz $^1$H NMR spectroscopy—in combination with methylation analysis and partial acetolysis was introduced in the present study. This highly sensitive method allowed the elucidation of the structures of the carbohydrate units of the mentioned glycopeptides to be carried out in a very short period of time. The present investigation revealed that the structures of the 16 heteroglycans can be grouped into five classes, and, except for those of two classes, the structures are new. Two classes designated A and B possess the known bi- and triantennary structures, respectively. A third class designated C proved to be tetraantennary in its structure, whereas the other classes designated BF and CF are, in comparison with the structures of classes B and C, characterized by an additional fucose residue in a hitherto unknown position and linkage. The latter monosaccharide residue is attached in an $\alpha(1\rightarrow3)$ bond, vicinal to Gal$\beta(1\rightarrow4)$, to the GlcNAc residue **7**. The present study also revealed that each glycosylation site of pooled $\alpha_1$-acid glycoprotein possesses carbohydrate units with different structures.

The chemical heterogeneity of the carbohydrate moiety of human plasma $\alpha_1$-acid glycoprotein was firmly established over a decade ago. However, structural studies on this moiety have afforded contradictory results with regard to both the structure of the core and the degree of peripheral branching of the carbohydrate units (Bayard & Fournet, 1976; Hatcher & Jeanloz, 1974; Jeanloz, 1972; Montreuil, 1975; Schmid, 1975; Schwick et al., 1977; Spiro, 1973; Yoshizawa & Yamashina, 1978). Recently, Schmid et al. (1973) have elucidated the complete amino acid sequence of this protein and thereby demonstrated that $\alpha_1$-acid glycoprotein possesses five glycosylation sites. Furthermore, these investigators have more recently purified a large number of glycopeptides derived from this protein (Schmid et al., 1977). The availability of these glycopeptides in homogeneous state afforded the possibility to determine the primary sequence of the carbohydrate units and thereby explain the heterogeneity of the carbohydrate moiety of $\alpha_1$-acid glycoprotein.

The aim of the present study was to elucidate the primary carbohydrate structures of the previously isolated 16 glycopeptides of $\alpha_1$-acid glycoprotein. To minimize the time required for such sequence determinations, a new technique, namely, 360-MHz $^1$H NMR spectroscopy, was successfully introduced into this area of investigation.

## Materials and Methods

*$\alpha_1$-Acid Glycoprotein.* This protein was isolated from pooled normal human plasma by a method described by Bürgi & Schmid (1961). The homogeneity of this globulin with regard to its polypeptide chain was established by several criteria of purity (Schmid, 1975). Asialo glycopeptides[1] were prepared in the homogeneous state and designated as described previously (Schmid et al., 1977). It should be noted that, subsequently, these asialo glycopeptides are referred as glycopeptides.

*The Carbohydrate Composition of the Glycopeptides.* The carbohydrate compositions of the glycopeptides were determined by gas–liquid chromatography (GLC[2]) after treatment of the glycopeptides with 0.5 M methanol–HCl for 24 h at 80 °C and pertrifluoroacetylation (Fournet et al., 1974; Zanetta et al., 1972). This procedure allows simultaneous determination of both the neutral and amino sugars and was carried out with the aid of an Aerograph 1200 (Varian, Orsay, France) equipped with a flame ionization detector and a glass column (0.3 × 300 cm, 3% Carbowax 6000 on Chromosorb W-HMDS, 60–80 mesh, nitrogen as carrier gas at a flow rate of 30 mL/min, and temperature programmed from 110 to 200 °C at 2 °C/min).

† From the Laboratoire de Chimie Biologique, Université des Sciences et Techniques de Lille I, Villeneuve D'Ascq, France (B.F., J.M., and G.S.), Laboratory of Organic Chemistry, University of Utrecht, Croesestraat 79, Utrecht, The Netherlands (L.D., J.H., and J.F.G.V.), and the Department of Biochemistry, Boston University School of Medicine, Boston, Massachusetts (J.P.B. and K.S.). *Received May 22, 1978.* This investigation was supported by the Centre National de la Recherche Scientifique (Laboratoire Associé No. 217: Biologie physico-chimique et moléculaire des glucides libres et conjugués), the Netherlands Foundation for Chemical Research (SON) with financial aid from the Netherlands Organization for the Advancement of Pure Research (ZWO), Hoffmann-La Roche, Nutley, N.J., and the National Institute of General Medical Sciences (GM-10374), U.S. Public Health Service, and from the Veterans Administration Hospital (0958-01-V.A.), Buffalo, N.Y. 14215.

[1] $\alpha_1$-Acid glycoprotein was desialyzed prior to chymotryptic digestion in order to render possible the separation of the formed glycopeptides (Schmid et al., 1977). The presence of sialyl residues in the glycopeptides would give rise to significant shift increments in the $^1$H NMR spectra for a number of protons in the sialylated antenna which, however, could readily be interpreted (Dorland et al., 1978; Montreuil & Vliegenthart, 1978).

[2] Abbreviations used are: GLC, gas–liquid chromatography; Fuc, fucose; Gal, galactose; GlcNAc, *N*-acetylglucosamine; Man, mannose; OMe, *O*-methyl.

0006-2960/78/0417-5206$01.00/0 

TABLE I: The Carbohydrate Composition of Glycopeptides Derived from Asialo-$\alpha_1$-Acid Glycoprotein.[a]

| glycopeptide | | monosaccharide residues (expressed in mol per mol of glycopeptide) | | | | total no. of residues |
|---|---|---|---|---|---|---|
| | | Fuc | Gal | Man | GlcN | |
| class A | GP[b]II-6 | 0 | 1.92 (2) | 3.00 | 3.97 (4) | 9 |
| class B | GPII-5 | 0 | 2.71 (3) | 3.00 | 5.17 (5) | 11 |
| | GPIII-3 | 0 | 2.64 (3) | 3.00 | 4.93 (5) | 11 |
| | GPIII-7 | 0 | 2.77 (3) | 3.00 | 4.94 (5) | 11 |
| | GPIV-7 | 0 | 2.60 (3) | 3.00 | 4.73 (5) | 11 |
| | GPV-5 | 0 | 2.99 (3) | 3.00 | 4.99 (5) | 11 |
| class BF | GPII-4 | 0.71 (1) | 2.82 (3) | 3.00 | 5.10 (5) | 12 |
| class C | GPII-3 | 0 | 3.68 (4) | 3.00 | 5.94 (6) | 13 |
| | GPIII-2 | 0 | 3.72 (4) | 3.00 | 5.57 (6) | 13 |
| | GPIII-6 | 0 | 3.82 (4) | 3.00 | 5.99 (6) | 13 |
| | GPIV-6 | 0 | 3.75 (4) | 3.00 | 5.96 (6) | 13 |
| | GPV-4 | 0 | 3.54 (4) | 3.00 | 5.73 (6) | 13 |
| class CF | GPIII-5 | 0.85 (1) | 4.09 (4) | 3.00 | 5.64 (6) | 14 |
| | GPIV-5 | 0.94 (1) | 3.97 (4) | 3.00 | 6.20 (6) | 14 |
| | GPV-2 | 0.91 (1) | 3.52 (4) | 3.00 | 6.21 (6) | 14 |
| | GPV-3 | 1.12 (1) | 4.06 (4) | 3.00 | 5.82 (6) | 14 |

[a] The number of Man residues per carbohydrate chain was assumed to be 3.00. [b] For designation, see Schmid et al. (1977).

*Permethylation.* Glycopeptides (1 mg of a preparation) were methylated according to the method of Hakomori (1964) as modified by Björndal et al. (1970). After extraction with chloroform, the methylated products were purified on Sephadex LH-20 (Pharmacia Fine Chemicals, Uppsala, Sweden; column 1 × 20 cm, chloroform–ethanol 1:1, 2-mL fractions; detection of carbohydrate by the phenol–sulfuric acid method of Dubois et al. (1956)). The methylated compounds were then treated with 0.5 M methanol–HCl for 24 h at 80 °C and the formed mixtures of methylglycosides analyzed by GLC before (neutral monosaccharides only) and after peracetylation (neutral monosaccharides and hexosamines) in pyridine–acetic anhydride (1:1, 1 mL, 100 °C, 30 min).

*Partial Acetolysis.* Partial acetolysis of the glycopeptides was carried out according to Bayard & Montreuil (1972). The obtained oligosaccharides were desalted by passage through a column of Dowex 50, X8 ($H^+$ form) and then reduced in 0.2 M $NaBD_4$ at room temperature overnight. After acidification, boric acid was removed by evaporation with methanol. The resulting compounds were then purified by passage through a Dowex 50, X8 ($H^+$ form), permethylated, and subsequently analyzed by GLC–mass spectrometry (Riber, Model 10-10, Rueil-Malmaison, France; capillary glass column coated with SE-30, temperature program from 150 to 300 °C at 4 °C/min; detection by mass fragmentography at *m/e* values 236, 260, and 277; Fournet et al., unpublished results). Appropriate authentic oligosaccharides were included in this investigation as references.

*The 360-MHz $^1$H NMR Spectroscopy.* For $^1$H NMR spectroscopic analysis the neutralized glycopeptides were repeatedly treated with $D_2O$ at room temperature and intermediate lyophilization. The samples were then dissolved in 0.5 mL of $D_2O$ to give a concentration between 2 and 30 mM and subjected for 30 min to $^1$H NMR spectroscopy (Bruker HX-360 spectrometer, Bruker Spectrospin, A.G., Zürich, Switzerland; operated in the Fourier transform mode). Probe temperatures of 25 and 60 °C were selected so that the signals near the HDO resonance could be detected (Strecker et al., 1977). The observed chemical shifts ($\delta$) are reported relative to the position of sodium 2,2-dimethyl-2-silapentane-5-sulfonate (indirectly to acetone in $D_2O$: $\delta$ = 2.225 ppm). The accuracy of the chemical shift values observed on high resolution $^1$H NMR spectroscopy is about 0.001 ppm, whereas these values are reproducible within 0.002 ppm. It is important to note that integration of the area of the signals directly yielded the number of each monosaccharide and acetyl residues.

## Results

### A. *Chemical and Physicochemical Studies of the 16 Glycopeptides*

1. *Carbohydrate Composition.* According to the molar distribution of their monosaccharides, the 16 glycopeptides (Table I) were divided into five classes (A, B, BF, C, and CF).

2. *Permethylation Studies.* The investigation of the methylated glycopeptides (Table II) indicated that the compounds of each class yielded the same methylated neutral and amino sugars in identical molar ratios. Typical gas–liquid chromatograms of the carbohydrate units of each class of glycopeptides are presented in Figure 1. The substitution patterns of the Man and GlcNAc residues and the occurrence of Fuc residues, indicated by the data presented in Table II, suggested that the heteroglycans possess structures with increasing complexity when the structures of class A to class CF were compared with each other.

3. *The 360-MHz $^1$H NMR Studies.* The interpretation of the obtained 360-MHz $^1$H NMR spectra of the 16 glycopeptides was carried out as described by Dorland et al. (1977a) who took advantage of the corresponding spectra of appropriate reference compounds. For the interpretation of these spectra in terms of structural assignments, the resonances of the anomeric protons, the H-2 protons of the Man residues, the H-5 and H-6 protons of the Fuc residues, and the *N*-acetyl protons were used (Dorland et al., 1978; Strecker et al., 1978). Moreover, the shift increments due to the increasing degree of branching and the addition of Fuc in the heteroglycan structures (Tables III and IV) were found to be far greater than 0.002 ppm, the reproducibility of these values. It is of particular interest to note that the chemical shift data of the Man H-1 and H-2 protons (Table IV) were used for the rec-

TABLE II: Molar Ratios of Monosaccharide Methyl Ethers Present in the Methanolysates of the Permethylated Glycopeptides Derived from Asialo-$\alpha_1$-Acid Glycoprotein.

| glycopeptide | | methylated monosaccharides (mol/mol of glycopeptide) | | | | | | | | total no. of residues |
|---|---|---|---|---|---|---|---|---|---|---|
| | | 2,3,4-tri-OMe[a]-Fuc | 2,3,4,6-tetra-OMe-Gal | 3,4,6-tri-OMe-Man | 3,4-di-OMe-Man | 2,4-di-OMe-Man | 3,6-diOMe-Man | 3,6-diO-Me-Glc-NAcNMe[b] | 6-OMe-GlcNAc-NMe | |
| class A | GPII-6 | 0 | 1.88 (2) | 1.88 (2) | 0 | 1.00 (1) | 0 | 4.21 (4) | 0 | 9 |
| class B | GPII-5 | 0 | 3.06 (3) | 1.02 (1) | 0 | 0.97 (1) | 0.86 (1) | 4.55 (5) | 0 | 11 |
| | GPIII-3 | 0 | 2.91 (3) | 1.32 (1) | 0 | 0.90 (1) | 0.76 (1) | 4.76 (5) | 0 | 11 |
| | GPIII-7 | 0 | 2.85 (3) | 1.07 (1) | 0 | 1.05 (1) | 1.09 (1) | 4.57 (5) | 0 | 11 |
| | GPIV-7 | 0 | 3.08 (3) | 1.07 (1) | 0 | 0.98 (1) | 0.87 (1) | 4.70 (5) | 0 | 11 |
| | GPV-5 | 0 | 3.12 (3) | 1.20 (1) | 0 | 1.00 (1) | 0.88 (1) | 4.99 (5) | 0 | 11 |
| class BF | GPII-4 | 0.45 (1) | 3.00 (3) | 1.20 (1) | 0 | 1.00 (1) | 0.95 (1) | 3.90 (4) | 0.65 (1) | 12 |
| class C | GPII-3 | 0 | 3.81 (4) | 0 | 0.64 (1) | 0.90 (1) | 0.87 (1) | 5.74 (6) | 0 | 13 |
| | GPIII-2 | 0 | 4.12 (4) | 0 | 0.72 (1) | 1.00 (1) | 0.96 (1) | 5.81 (6) | 0 | 13 |
| | GPIII-6 | 0 | 3.76 (4) | 0 | 0.95 (1) | 1.00 (1) | 0.99 (1) | 5.54 (6) | 0 | 13 |
| | GPIV-6 | 0 | 3.71 (4) | 0 | 1.00 (1) | 1.10 (1) | 0.94 (1) | 5.60 (6) | 0 | 13 |
| | GPV-4 | 0 | 3.85 (4) | 0 | 0.97 (1) | 1.00 (1) | 0.81 (1) | 5.55 (6) | 0 | 13 |
| class CF | GPIII-5 | 0.84 (1) | 3.83 (4) | 0 | 0.88 (1) | 1.00 (1) | 0.96 (1) | 4.94 (5) | 0.80 (1) | 14 |
| | GPIV-5 | 0.80 (1) | 3.89 (4) | 0 | 0.87 (1) | 1.00 (1) | 0.91 (1) | 4.92 (5) | 0.78 (1) | 14 |
| | GPV-2 | 0.68 (1) | 3.81 (4) | 0 | 1.01 (1) | 1.00 (1) | 0.91 (1) | 5.21 (5) | 0.61 (1) | 14 |
| | GPV-3 | 0.82 (1) | 3.75 (4) | 0 | 1.04 (1) | 0.88 (1) | 0.77 (1) | 4.86 (5) | 0.52 (1) | 14 |

[a] *O*-Methyl is abbreviated as OMe. [b] Acetyl-*N*-methyl is abbreviated as NAcNMe.

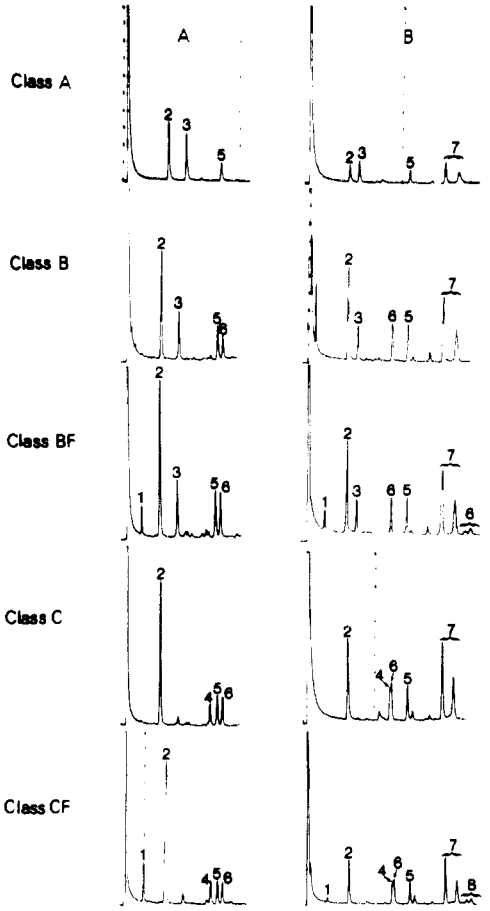



FIGURE 1: Gas-liquid chromatograms of methylated methylglycosides before (A) and after (B) peracetylation. The designations of the peaks (1–8) refer to the following compounds: methyl 2,3,4-tri-*O*-methylfucoside (peak 1), methyl 2,3,4,6-tetra-*O*-methylgalactoside (peak 2), methyl 3,4,6-tri-*O*-methylmannoside (peak 3), methyl 3,4-di-*O*-methylmannoside (peak 4), methyl 2,4-di-*O*-methylmannoside (peak 5), methyl 3,6-di-*O*-methylmannoside (peak 6), methyl 3,6-di-*O*-methyl-*N*,*N*-acetylmethylglucosaminide (peak 7), and methyl 6-mono-*O*-methyl-*N*,*N*-acetylmethylglucosaminide (peak 8).

ognition of the substitution patterns of the trimannosidobranching core. It is also noteworthy that the amino acid residues which are peptide bond linked to the asparagine residues of the glycopeptides (Schmid et al., 1977) have negligible effects on the spectra of the glycan moieties except on the H-1 of GlcNAc **1**.

The results of the $^1$H NMR and methylation analyses (Tables II and III) demonstrate that the carbohydrate moieties of all peptides have in common the same structural unit as found in class A compounds (see below). In addition, the $^1$H NMR spectra of the 16 glycopeptides were found to be characteristic for each class of oligosaccharides (Figure 2). The total number of anomeric protons (i.e., number of monosaccharide residues) of each of the five structures, as obtained by integration of the resonances listed in Table III, was found to be identical with that determined by chemical analysis. However, it should be noted that the $^1$H NMR spectroscopy yields more accurate values.

### B. *Primary Carbohydrate Structures*

1. *Class A Glycopeptides.* The results of the methylation analysis of glycopeptide GPII-6 (Table II) showed that the two Gal residues occupy terminal positions. Further, the two Man derivatives indicated the presence of a Man residue that is substituted in positions 3 and 6 by additional Man residues forming a trimannoside that is characteristic of carbohydrate units linked N-glycosidically to asparagine (Montreuil, 1975; Kornfeld & Kornfeld, 1976). The $^1$H NMR spectra revealed the complete sequence of the monosaccharides, the C–C bonds and the anomeric linkages, of the carbohydrate unit of this glycopeptide. It should also be noted that the spectrum of glycopeptide GPII-6 was found to be identical with that of the asialoglycopeptide of human serotransferrin (Dorland et al., 1975). Thus, these data demonstrated the presence in this glycopeptide of the biantennary structure shown in Chart I.

2. *Class B Glycopeptides.* The methylation studies of the class B glycopeptides (Table II), when compared with the corresponding data of the class A compound, demonstrated that the three Gal residues occupy terminal positions and that

CHART I: Structure of the Class A Oligosaccharide.

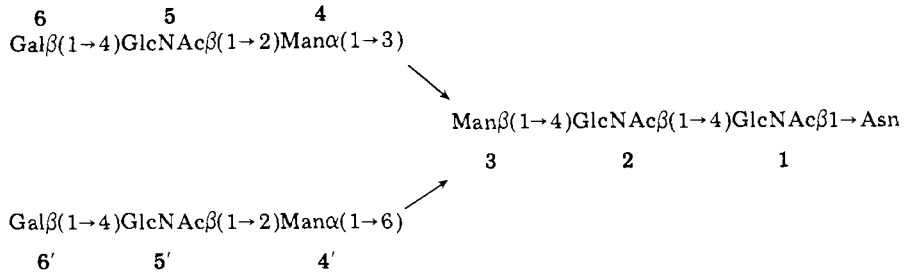



CHART II: Structure of the Class B Oligosaccharide.

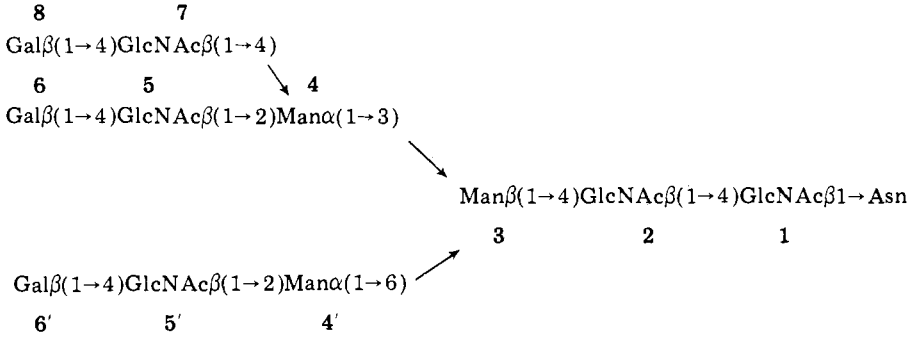
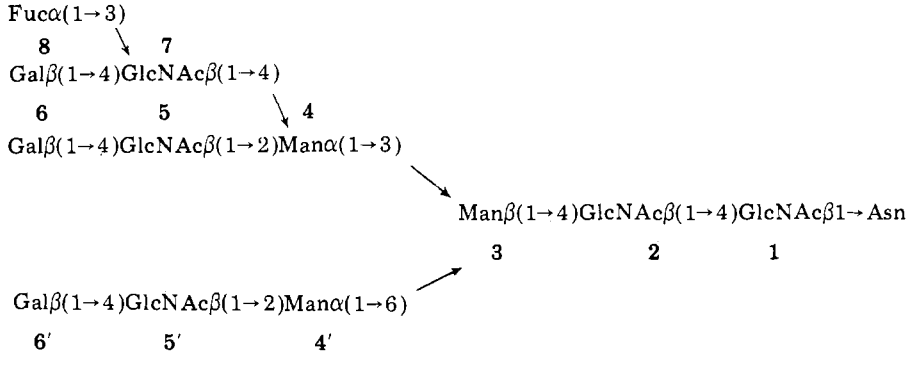



CHART III: Structure of the Class BF Oligosaccharide.




one of the two α-Man residues is substituted in positions 2 and 4, indicating the presence of an additional substituent on the double-substituted Man residue. On partial acetolysis of glycopeptide GPV-5 three oligosaccharides were formed in relatively large amounts: Gal(1→4)GlcNAc, GlcNAcβ(1→2)-Man, and GlcNAcβ(1→4)Manα(1→3)Man. These compounds after reduction and permethylation were characterized by GLC–mass spectroscopy. The fragment ions observed at *m/e* 236, 260, and 277 were indicative of terminal neutral hexitols, hexosamines, and hexosaminitols, respectively.

The mass spectrum of the trisaccharide-alditol (Figure 3) revealed that a GlcNAc residue is linked to position 4 of Man residue **4**. The ion observed at *m/e* 260 was formed by the cleavage of a hexosamine residue from the oligosaccharide, while subsequent elimination of methanol gave rise to *m/e* 228. The fragment at *m/e* 236 originated by splitting off the alditol residue from this oligosaccharide. The ions observed at *m/e* 464, 432, 400, and 440, 408, 376 established the sequence of the monomeric units in the trisaccharide as being GlcNAc-Man-Man. Further, the ion noted at *m/e* 133 demonstrated that the alditol unit was glycosylated in position 3, and the ion noted at *m/e* 296 proved that the linkage of hexosamine to the hexose unit is not (1→3). Since the sum of the intensities of the ions at *m/e* 440, 408, and 376 was larger than that of the ions observed at *m/e* 464, 432, and 400, the type of linkage is a (1→4) or a (1→6) rather than a (1→2) bond (Kärkkäinen, 1971). Moreover, the relative abundance of the ions at *m/e* 440, 408, and 376 decreased from high to low *m/e* values suggesting a (1→4) linkage of the hexosamine to the hexose residue.

In the $^1$H NMR spectra of the class B compounds, one can recognize the presence of the biantennary structure on the basis of the resonances of the anomeric protons, of the H-2 protons of Man, and of the protons of the *N*-acetyl groups (Table III). From these $^1$H NMR results and the data of the chemical analyses described above, it was concluded that the class B structure possesses a third chain consisting of Gal→GlcNAc (*N*-acetyllactosamine, i.e., residues **7** and **8**). The assignment of the position of attachment of this *N*-acetyllactosamine unit to position 4 of Man **4** (and not Man **4′**) is based on (a) the shift increments of the H-2 protons of Man residues **3** and **4** with regard to those of the corresponding increments of the class A structure (see Table IV) and (b) on the identification by GLC–mass spectrometry of the trisaccharide GlcNAcβ-(1→4)Manα(1→3)Man obtained by acetolysis. As a result of the extra singlet due to the *N*-acetyl group of GlcNAc **7** (2.078 ppm), the NAc region of the spectra of class B glyco-

TABLE III: $^1$H Chemical Shifts of Characteristic Protons of Constituent Monosaccharides for Glycopeptides Derived from $\alpha_1$-Acid Glycoprotein.

| compound | | chemical shift of H-1 of residue[a] | | | | | | | | | | chemical shift of H-2 of residue[a] | | | chemical shift of the *N*-acetyl-$CH_3$ of residue[a] | | | | | chemical shift of fucose protons | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 4′ | 5 | 5′ | 6 | 6′ | 7 | 7′ | 8 | 8′ | 3 | 4 | 4′ | 1 | 2 | 5 | 5′ | 7 | 7′ | H-1 | H-5 | H-6 |
| class A | GPII-6 | 5.095 | 4.616 | 4.765 | 5.121 | 4.928 | 4.580 | 4.580 | 4.469 | 4.469 | | | | | 4.250 | 4.191 | 4.112 | 2.004 | 2.079 | 2.050 | 2.046 | | | | | |
| class B | GPII-5 | 5.092 | 4.620 | 4.755 | 5.120 | 4.927 | 4.595 | 4.583 | 4.468 | 4.468 | 4.548 | | 4.468 | | 4.214 | 4.214 | 4.109 | 2.004 | 2.078 | 2.048 | 2.048 | 2.078 | | | | |
| | GPII-3 | 5.058 | 4.614 | 4.761 | 5.119 | 4.922 | 4.590 | 4.579 | 4.466 | 4.466 | 4.548 | | 4.446 | | 4.212 | 4.212 | 4.110 | 2.011 | 2.076 | 2.049 | 2.049 | 2.076 | | | | |
| | GPII-7 | 5.064 | 4.618 | 4.758 | 5.120 | 4.925 | 4.598 | 4.586 | 4.471 | 4.471 | 4.547 | | 4.471 | | 4.215 | 4.215 | 4.113 | 2.011 | 2.078 | 2.049 | 2.049 | 2.078 | | | | |
| | GPIV-7 | 5.089 | 4.616 | 4.755 | 5.118 | 4.922 | 4.591 | 4.580 | 4.471 | 4.471 | 4.548 | | 4.448 | | 4.213 | 4.213 | 4.109 | 2.006 | 2.077 | 2.048 | 2.048 | 2.077 | | | | |
| | GPV-5 | 5.086 | 4.617 | 4.759 | 5.122 | 4.924 | 4.594 | 4.584 | 4.475 | 4.475 | 4.546 | | 4.475 | | 4.221 | 4.221 | 4.105 | 2.011 | 2.079 | 2.052 | 2.052 | 2.079 | | | | |
| class BF | GPII-4 | 5.090 | 4.612 | 4.753 | 5.117 | 4.919 | 4.590 | 4.579 | 4.470 | 4.470 | 4.547 | | 4.446 | | 4.205 | 4.205 | 4.104 | 2.006 | 2.077 | 2.047 | 2.047 | 2.064 | | 5.110 | 4.838 | 1.179 |
| class C | GPII-3 | 5.061 | 4.612 | 4.755 | 5.129 | 4.870 | 4.592 | 4.573 | 4.471 | 4.471 | 4.549 | 4.549 | 4.471 | 4.471 | 4.214 | 4.214 | 4.092 | 2.011 | 2.081 | 2.050 | 2.040 | 2.081 | 2.040 | | | |
| | GPII-2 | 5.060 | 4.612 | 4.755 | 5.122 | 4.865 | 4.588 | 4.570 | 4.468 | 4.468 | 4.549 | 4.549 | 4.485 | 4.485 | 4.214 | 4.214 | 4.092 | 2.013 | 2.076 | 2.050 | 2.039 | 2.076 | 2.039 | | | |
| | GPII-6 | 5.063 | 4.613 | 4.760 | 5.127 | 4.865 | 4.590 | 4.572 | 4.474 | 4.474 | 4.545 | 4.545 | 4.474 | 4.474 | 4.215 | 4.215 | 4.089 | 2.010 | 2.078 | 2.051 | 2.039 | 2.078 | 2.039 | | | |
| | GPIV-6 | 5.090 | 4.614 | 4.758 | 5.129 | 4.870 | 4.594 | 4.573 | 4.466 | 4.466 | 4.551 | 4.551 | 4.446 | 4.487 | 4.215 | 4.215 | 4.093 | 2.006 | 2.077 | 2.052 | 2.040 | 2.077 | 2.040 | | | |
| | GPV-4 | 5.078 | 4.613 | 4.755 | 5.128 | 4.867 | 4.589 | 4.572 | 4.469 | 4.469 | 4.543 | 4.543 | 4.447 | 4.485 | 4.218 | 4.218 | 4.093 | 2.005 | 2.077 | 2.052 | 2.039 | 2.077 | 2.039 | | | |
| class CF | GPIII-5 | 5.057 | 4.611 | 4.755 | 5.125 | 4.866 | 4.590 | 4.576 | 4.465 | 4.465 | 4.547 | 4.547 | 4.442 | 4.487 | 4.218 | 4.218 | 4.091 | 2.014 | 2.078 | 2.053 | 2.040 | 2.069 | 2.040 | 5.113 | 4.824 | 1.176 |
| | GPIV-5 | 5.077 | 4.611 | 4.749 | 5.124 | 4.863 | 4.590 | 4.575 | 4.461 | 4.461 | 4.549 | 4.549 | 4.441 | 4.484 | 4.207 | 4.207 | 4.087 | 2.006 | 2.076 | 2.052 | 2.037 | 2.064 | 2.037 | 5.107 | 4.816 | 1.172 |
| | GPV-2 | 5.050 | 4.616 | 4.760 | 5.125 | 4.864 | 4.590 | 4.573 | 4.463 | 4.463 | 4.544 | 4.544 | 4.444 | 4.483 | 4.217 | 4.217 | 4.090 | 2.007 | 2.077 | 2.052 | 2.038 | 2.068 | 2.038 | 5.115 | 4.831 | 1.172 |
| | GPV-3 | 5.071 | 4.610 | 4.759 | 5.122 | 4.862 | 4.591 | 4.580 | 4.463 | 4.463 | 4.547 | 4.547 | 4.445 | 4.481 | 4.216 | 4.216 | 4.087 | 2.005 | 2.075 | 2.050 | 2.039 | 2.068 | 2.039 | 5.111 | 4.836 | 1.172 |

[a] The number of each monosaccharide residue is seen in Figure 2.

peptides differs significantly from that of the biantennary glycopeptide (Figure 4). Further, these additional Gal and GlcNAc residues both have $\beta$ configuration as demonstrated by the chemical shifts of the H-1 protons and the coupling constants $J_{1,2}$ (~7.5 Hz). Based on the above described evidence, class B glycopeptides have the triantennary structure shown in Chart II.

3. *Class BF Glycopeptides.* The carbohydrate composition of glycopeptide GPII-4 shows the presence of an additional Fuc residue when compared with that of the class B compounds. Methylation studies provided strong evidence that the Fuc residue is in terminal position and attached to a GlcNAc residue. From the set of resonance signals of the Man H-1 and H-2 protons, it can be concluded that this glycopeptide possesses the class B type triantennary structure. Judging from the coupling constant $J_{1,2}$ (~3.5 Hz) the Fuc residue is $\alpha$-glycosidically linked. Moreover, comparison of the *N*-acetyl signals of the class B structure with those of the class BF shows that only the corresponding signals of GlcNAc **7** have shifted. Therefore, the Fuc residue is attached to this GlcNAc residue. Further, it is important to note that the Fuc H-5 proton resonates at an unusually low field ($\delta_{H\text{-}5}$ = 4.838 ppm; Table III). A comparable observation was made with certain milk oligosaccharides (Kobata, 1977; Watkins, 1972) that contain the structural unit Fuc$\alpha$(1→4)[Gal$\beta$(1→3)]GlcNAc and thus possess the determinant for Le$^a$ blood group activity. It is important to note again that the H-5 of Fuc of these compounds also resonates at an unusually low field ($\delta$ = 4.87 ppm; Dorland et al., unpublished results). Earlier, Lemieux et al. (1975) reported the 100-MHz $^1$H NMR spectrum of a trisaccharide with this structural unit and offered, as explanation for the exceptional downfield shift of the H-5 proton, the deshielding of H-5 by the three oxygen atoms in positions 1, 4, and 5 of the vicinal Gal residue. On the basis of space-filling molecular models a similar deshielding of H-5 can be predicted for the Fuc residue if the linkages of Fuc and Gal to the GlcNAc residue are interchanged: Fuc$\alpha$(1→3)[Gal$\beta$(1→4)]GlcNAc. In this structure the same oxygen atoms are in close proximity to Fuc H-5. From our $^1$H NMR results in combination with the finding that GPII-4 is completely devoid of Le$^a$ activity, we concluded that the latter structural element is present in the carbohydrate unit of this glycopeptide. Thus, this class of heteroglycans possesses the structure shown in Chart III.

4. *Class C Glycopeptides.* From their carbohydrate compositions (Table I) and the results of the methylation studies (Table II), it was deduced that the class C glycopeptides possess a tetraantennary structure. However, the major evidence in support of this proposed structure was derived from the $^1$H NMR spectroscopic analyses (Table III). The resonance positions of the Man H-1 and H-2 protons indicated that the class B triantenna is the structural element of the class C compounds. The addition of a fourth *N*-acetyllactosamine linked in a $\beta$(1→6) bond to Man **4′** was demonstrated by the shift increments for H-1 and H-2 of this Man residue (Table IV). Other characteristic features of the $^1$H NMR spectra of class C compounds are the changed resonance positions of the *N*-acetyl protons of GlcNAc **5′** (Figure 4), in comparison with those of the triantennary structure, due to the attachment of the fourth antenna. Based on the above described data the class C glycopeptides have the tetraantennary structure shown in Chart IV.

5. *Class CF Glycopeptides.* Comparison of the carbohydrate compositions (Table I), the results of methylation analysis (Table II), and the data of the $^1$H NMR spectroscopy (Table III) of the class CF glycopeptides with those of the class C glycopeptides showed that both classes CF and C compounds

CHART IV: Structure of the Class C Oligosaccharide.

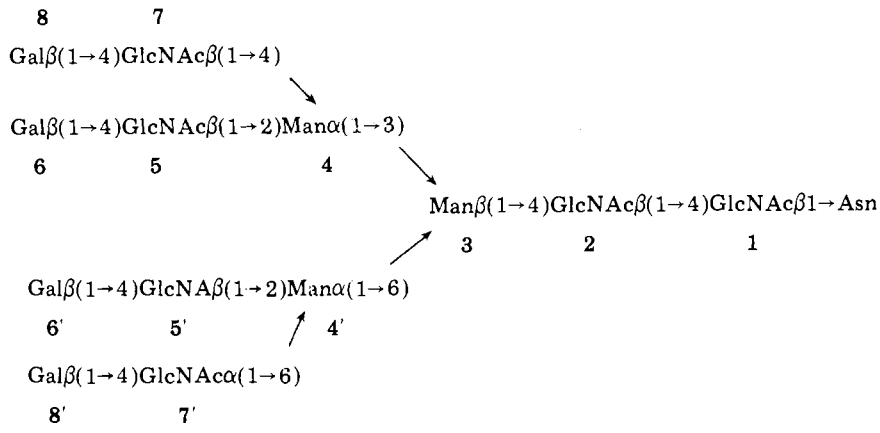
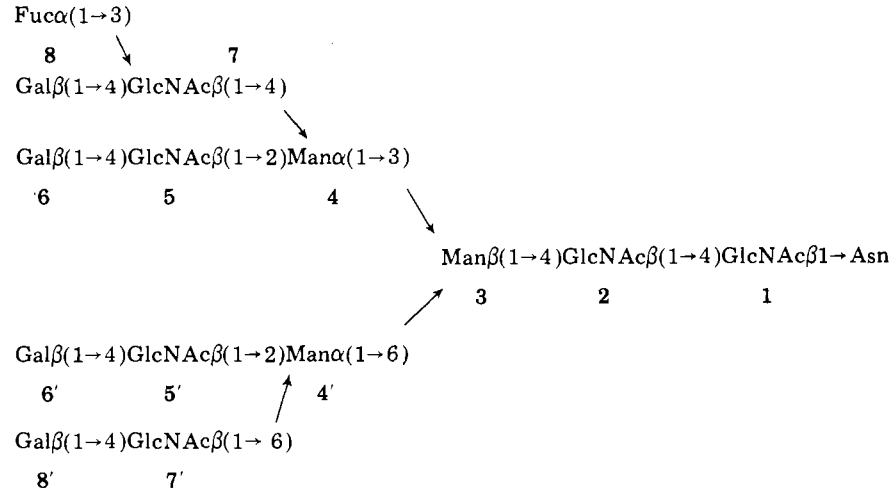



CHART V: Structure of the Class CF Oligosaccharide.




possess the same tetraantennary structure. However, the class CF peptides contained an additional terminal Fuc linked to a GlcNAc residue. The $^1$H NMR parameters for H-1, H-5, and H-6 of this Fuc and for the *N*-acetyl proton of GlcNAc 7 (Figure 4) were in complete agreement with those found for the Fuc residue of class BF glycopeptides. These data, therefore, demonstrated that this Fuc residue occupies an $\alpha(1\rightarrow3)$ position vicinal to a $\beta(1\rightarrow4)$ Gal residue both being linked to GlcNAc 7 (Figure 4). In support of this structure the class CF compounds were found to be devoid of Le$^a$ activity. Based on this evidence, the carbohydrate structure of the class CF glycopeptides is as shown in Chart V.

It should be noted that, while the above described five carbohydrate structures are devoid of Sia residues, each Gal residue of the carbohydrate units present in the native protein probably carries a Sia residue. The exact nature of these Sia→Gal bonds has, however, not yet been established.

## Discussion

The present paper describes several new findings regarding

TABLE IV: Chemical Shifts of Mannose H-1 and H-2 Protons for Glycopeptides with Bi- (Class A), Tri- (Classes B, BF),[a] and Tetraantennary (Classes C, CF)[a] Carbohydrate Chains.

| class | $\delta_{H\text{-}1}$ of residue | | | $\delta_{H\text{-}2}$ of residue | | |
|---|---|---|---|---|---|---|
| | 3 | 4 | 4' | 3 | 4 | 4' |
| A (biantenna) $n$ = 1 | 4.765 | 5.121 | 4.928 | 4.250 | 4.191 | 4.112 |
| B (triantenna) $n$ = 5 | 4.758 | 5.120 | 4.924 | *4.215* | *4.215* | 4.109 |
| | (0.003) | (0.002) | (0.003) | (0.004) | (0.004) | (0.003) |
| BF (triantenna + Fuc) $n$ = 1 | 4.753 | 5.117 | 4.919 | *4.205* | *4.205* | 4.104 |
| C (tetraantenna) $n$ = 5 | 4.757 | 5.127 | *4.867* | *4.215* | *4.215* | *4.092* |
| | (0.003) | (0.003) | (0.003) | (0.002) | (0.002) | (0.002) |
| CF (tetraantenna + Fuc) $n$ = 4 | 4.756 | 5.124 | *4.864* | *4.215* | *4.215* | *4.089* |
| | (0.005) | (0.002) | (0.002) | (0.005) | (0.005) | (0.003) |

[a] Values which differ significantly from those of the biantennary structure (class A) are italicized. The letter $n$ indicates the number of compounds analyzed of each class and $\delta$ indicates chemical shifts in ppm. The numbers in parentheses indicate mean errors of the chemical shifts observed by analyzing different ($n$) compounds of the same class.

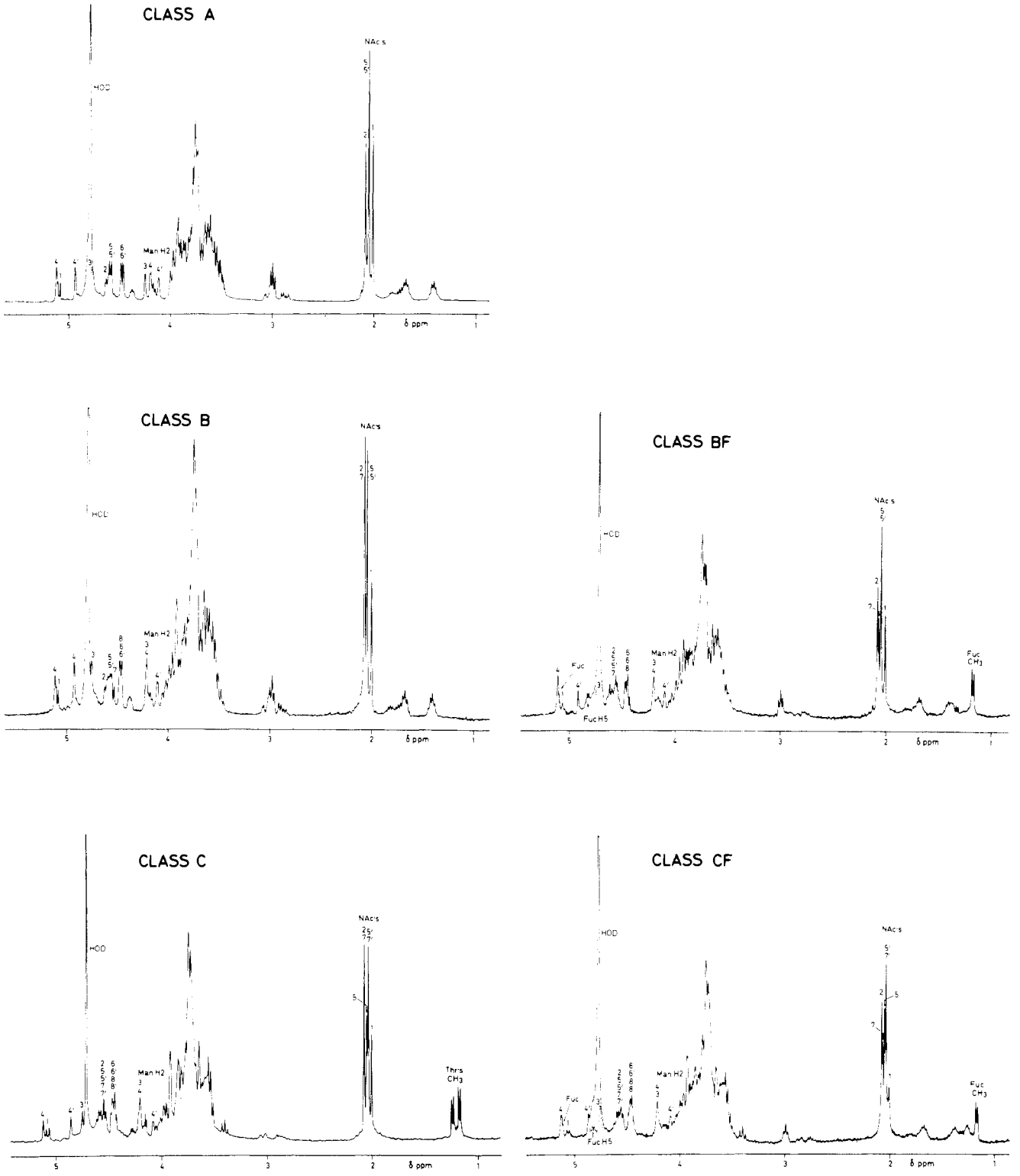



FIGURE 2: 360-MHz $^1$H NMR spectra of glycopeptides. GPII-6 (class A); GPII-5 (class B); GPII-4 (class BF); GPIV-6 (class C); GPIII-5 (class CF). It should be noted that, as judged by the results presented in Table III, these resonances were clearly separated from each other (for further information, see text).

the carbohydrate moiety of $\alpha_1$-acid glycoprotein. (1) The complete structures, including the monosaccharide sequences, the type and configuration of the glycosidic bonds, of the 16 previously isolated asialo-carbohydrate units of this protein were established. (2) $\alpha_1$-Acid glycoprotein possesses heteroglycans with five different structures.[3] Although the biantennary chain (class A), which is known to occur frequently in other glycoproteins (Kornfeld & Kornfeld, 1976; Montreuil & Vliegenthart, 1978), was found in only one of the 16 investigated glycopeptides of $\alpha_1$-acid glycoprotein, the triantennary structure (class B) has been reported to be present in only a few glycoproteins (Kondo et al., 1977; Kornfeld & Kornfeld, 1976). However, the structures of the class BF (triantennary

[3] A striking correlation could now be established between the separation of the glycopeptides on Dowex 50 (Schmid et al., 1977) and the structures of their heteroglycans: in each group of glycopeptides, these peptides were eluted in order of decreasing complexity of their structures.

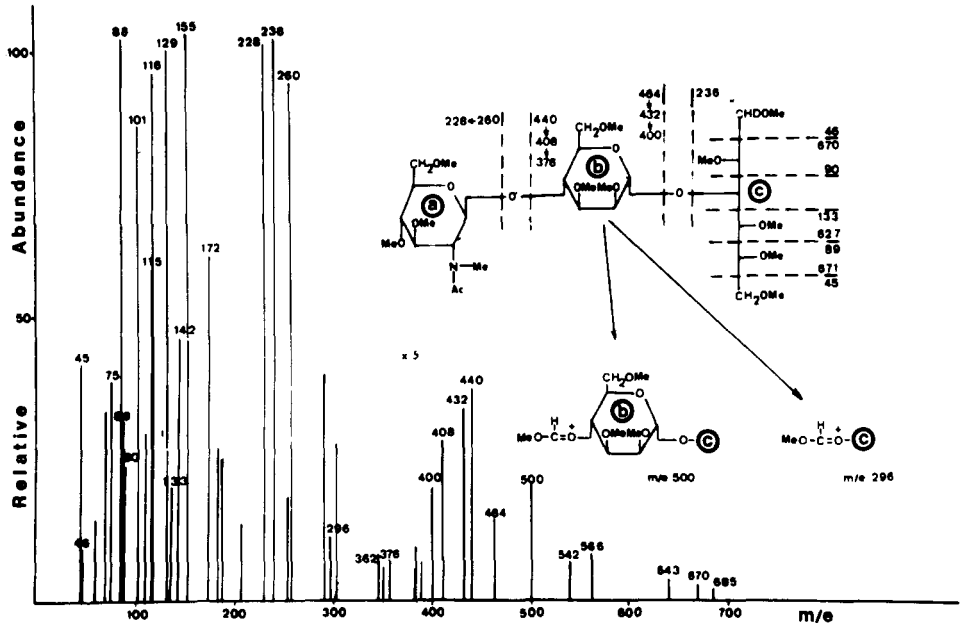



FIGURE 3: Mass spectrum of permethylated GlcNAc$\beta$(1→4)Man$\alpha$(1→3)mannitol (for further details, see text).

+ Fuc), class C (tetraantennary), and class CF (tetraantennary + Fuc) heteroglycans had not yet been described. (3) The occurrence and the position of the Fuc residue in classes BF and CF are unique. In both structures this residue is attached via an $\alpha$(1→3) bond to GlcNAc **7**. This new Fuc linkage[4] differs from the well-known Fuc$\alpha$(1→6) linkage to GlcNAc **1** reported for many glycoproteins (Kornfeld & Kornfeld, 1976; Montreuil & Vliegenthart, 1978). (4) The present data suggest that during the biosynthesis of the carbohydrate units of at least this protein, the antennae are linked, not randomly, but to highly specific positions of certain monosaccharides of the glycan units. (5) Each glycosylation site of the pooled protein bears glycan chains of different structures. The heteroglycans derived from glycosylation site II were shown to possess structures of classes A, B, BF, and C, whereas the carbohydrate units linked to glycosylation sites III, IV, and V possess classes B, C, and CF structures.

The heterogeneity of the carbohydrate moiety of pooled $\alpha_1$-acid glycoprotein pointed out in the introductory section can now be explained by the occurrence at each of the glycosylation sites of carbohydrate units with different structures. The reason for this observation, however, is still unknown, but it can be speculated that this heterogeneity may be due to individual protein molecules with different carbohydrate units at the same glycosylation site and may reflect differences in the biosynthesis and/or degradation of the carbohydrate moiety.

As to the elucidation of the three different antennary structures (classes A, B, and C), it should be noted that each of these structures can easily be established on the basis of the sets of the $^1$H NMR parameters of the Man H-1 and H-2 protons (Table IV) and the *N*-acetyl protons of the various GlcNAc residues (Figure 4). The presence of the Fuc residue in $\alpha$(1→3) linkage to GlcNAc **7** in the classes BF and CF structures can be inferred from the resonance positions of the Fuc H-1, H-5, and H-6 protons and the *N*-acetyl protons of GlcNAc **7** (Figure 4).

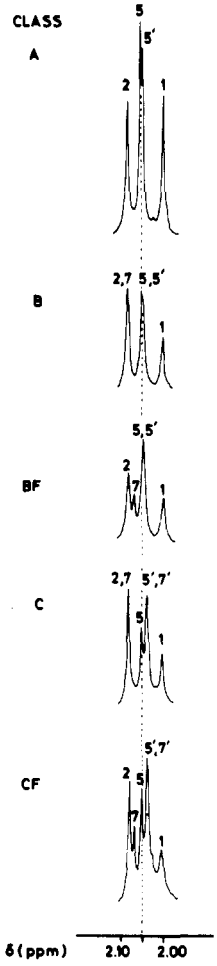



FIGURE 4: 360-MHz $^1$H NMR resonance patterns of the *N*-acetyl groups characteristic for the classes A, B, BF, C, and CF carbohydrate chains. The numbers refer to the corresponding GlcNAc residues.

The present study thus confirms the applicability of 360-MHz $^1$H NMR in conjunction with methylation analysis for the elucidation of complex carbohydrate structures as proposed in earlier investigations on related heteroglycans (Dorland et

[4] A preliminary note on this new linkage was presented earlier (Fournet et al., 1978; Montreuil & Vliegenthart, 1978). Recently, Krusius & Finne (1978) reported this linkage to be present in a glycoprotein derived from rat brain.

al., 1977a,b, 1978; Montreuil & Vliegenthart, 1978; Strecker et al., 1977). While with this $^1$H NMR spectroscopic method such a determination can be completed within 30 min and with as little as 2 mg of a homogeneous oligosaccharide, it should be realized that, if the elucidation of these carbohydrate structures were to be carried out by the conventional chemical and enzymatic techniques, not only much larger amounts of material but above all an enormous commitment of time would be required. High resolution $^1$H NMR spectroscopy, therefore, is excellently suited for the rapid elucidation of complex glycan structures.

Acknowledgments

The authors are very much obliged to Mr. L. Larsen, Institute of Laboratories, Massachusetts Department of Health, Jamaica Plain, Mass., for the generous supply of supernatant solution of Cohn fraction V of pooled normal human plasma from which $\alpha_1$-acid glycoprotein was isolated. They are grateful to Mr. Y. Leroy, CNRS technician, for his skillful assistance and to the Délégation Générale à la Recherche Scientific et Technique, the Etablissement Public Régional Nord-Pas de Calais, and the Institut Pasteur of Lille (B.F., G.S., and J.M.) for the contribution toward the purchase of the Riber mass spectrometer.

References

Bayard, B., & Fournet, B. (1976) *Carbohydr. Res. 46,* 75.

Bayard, B., & Montreuil, J. (1972) *Carbohydr. Res. 24,* 427.

Björndal, H., Hellerquist, C. G., Lindberg, B., & Svensson, S. (1970) *Angew. Chem., Int. Ed. Engl. 9,* 610.

Bürgi, W., & Schmid, K. (1961) *J. Biol. Chem. 236,* 1066.

Dorland, L., Haverkamp, J., Schut, B. L., Vliegenthart, J. F. G., Spik, G., Strecker, G., Fournet, B., & Montreuil, J. (1977a) *FEBS Lett. 77,* 15.

Dorland, L., Schut, B. L., Vliegenthart, J. F. G., Strecker, G., Fournet, B., Spik, G., & Montreuil, J. (1977b) *Eur. J. Biochem. 73,* 93.

Dorland, L., Haverkamp, J., Vliegenthart, J. F. G., Strecker, G., Michalski, J. C., Fournet, B., Spik, G., & Montreuil, J. (1978) *Eur. J. Biochem. 87,* 323.

Dubois, M., Gilles, K. A., Hamilton, K. J., Rebers, P. A., & Smith, F. (1956) *Anal. Chem. 28,* 350.

Fournet, B., Leroy, Y., & Montreuil, J. (1974) in *Méthodologie de la Structure et du Métabolisme Glycoconjugués* (Montreuil, J., Ed.) p 111, Editions du CNRS, Paris.

Fournet, B., Strecker, G., Spik, G., Montreuil, J., Schmid, K., Dorland, L., Haverkamp, J., Schut, B. L., & Vliegenthart, J. F. G. (1978) *Proc. IVth Int. Symp. Glycoconjugates, 1977* (in press).

Hakomori, S. I. (1964) *J. Biochem. (Tokyo) 55,* 205.

Hatcher, V. B., & Jeanloz, R. W. (1974) in *Méthodologie de la Structure et du Métabolisme des Glycoconjugués* (Montreuil, J., Ed.) p 329, Editions du CNRS, Paris.

Jeanloz, R. W. (1972) in *Glycoproteins* (Gottschalk, A., Ed.) p 565, Elsevier Publishing Co., Amsterdam.

Kärkkäinen, J. (1971) *Carbohydr. Res. 17,* 11.

Kobata, A. (1977) in *The Glycoconjugates* (Horowitz, M. I., & Pigman, W., Eds.) Vol. I, p 423, Academic Press, New York, N.Y.

Kondo, T., Fukuda, M., & Osawa, T. (1977) *Carbohydr. Res. 58,* 405.

Kornfeld, R., & Kornfeld, S. (1976) *Annu. Rev. Biochem. 45,* 217.

Krusius, T., & Finne, J. (1978) *Eur. J. Biochem. 84,* 395.

Lemieux, R. U., Bundle, D. R., & Baker, D. A. (1975) *J. Am. Chem. Soc. 97,* 4076.

Montreuil, J. (1975) *Pure Appl. Chem. 42,* 431.

Montreuil, J., & Vliegenthart, J. F. G. (1978) *Proc. IVth Int. Symp. Glycoconjugates, 1977* (in press).

Schmid, K. (1975), in *The Plasma Proteins* (Putnam, F. W., Ed.) Vol. I, p 183, Academic Press, New York, N.Y.

Schmid, K., Kaufmann, H., Isemura, S., Bauer, F., Emura, J., Motoyama, T., Ishiguro, M., & Nanno, S. (1973) *Biochemistry 12,* 2711.

Schmid, K., Nimberg, R. B., Kimura, A., Yamaguchi, H., & Binette, J. P. (1977) *Biochim. Biophys. Acta 492,* 291.

Schwick, H. G., Heide, K., & Haupt, H. (1977) in *The Glycoconjugates* (Horowitz, M. I., & Pigman, W., Eds.) Vol. I, p 268, Academic Press, New York, N.Y.

Spik, G., Bayard, B., Fournet, B., Strecker, G., Bouquelet, S., & Montreuil, J. (1975) *FEBS Lett. 50,* 296.

Spiro, R. G. (1973) *Adv. Protein Chem. 27,* 343.

Strecker, G., Herlant-Peers, M.-C., Fournet, B., Montreuil, J., Dorland, L., Haverkamp, J., Vliegenthart, J. F. G., & Farriaux, J.-P. (1977) *Eur. J. Biochem. 81,* 165.

Strecker, G., Fournet, B., Montreuil, J., Dorland, L., Haverkamp, J., Vliegenthart, J. F. G., & Dubesset, D. (1978) *Biochimie* (in press).

Watkins, W. M. (1972) in *Glycoproteins* (Gottschalk, A., Ed.) p 830, Elsevier Publishing Co., Amsterdam.

Yoshizawa, Z., & Yamashina, I. (1978) *Structure and Function of Glycoproteins,* Nankodo Publishing Co., Tokyo.

Zanetta, J. P., Beckenridge, W. C., & Vincendon, G. (1972) *J. Chromatogr. 69,* 291.