THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 264, No. 10, Issue of April 5, pp. 5720-5735, 1989
Printed in U.S.A.

Novel Polyfucosylated *N*-Linked Glycopeptides with Blood Group A, H, X, and Y Determinants from Human Small Intestinal Epithelial Cells*

(Received for publication, July 5, 1988)

Jukka Finne‡§, Michael E. Breimer¶, Gunnar C. Hansson¶, Karl-Anders Karlsson¶, Hakon Leffler¶, Johannes F. G. Vliegenthart‖, and Herman van Halbeek**

*From the §Department of Medical Biochemistry, University of Turku, SF-20520 Turku, Finland and the Department of Medical Chemistry, University of Helsinki, SF-00170 Helsinki, Finland, the ¶Department of Medical Biochemistry, University of Göteborg, S-40033 Göteborg, Sweden, the ‖Department of Bio-Organic Chemistry, University of Utrecht, NL-3522 Utrecht AD, The Netherlands, and the **Complex Carbohydrate Research Center, University of Georgia, Athens, Georgia 30613*

A novel type of *N*-linked glycopeptides representing a major part of the glycans in human small intestinal epithelial cells from blood group A and O individuals were isolated by gel filtrations and affinity chromatography on concanavalin A-Sepharose and *Bandeiraea simplicifolia* lectin I-Sepharose. Sugar composition, methylation analysis, ^{1}H NMR spectroscopy of the underivatized glycopeptides and FAB-mass spectrometry and electron impact-mass spectrometry of the permethylated glycopeptides indicated a tri- and tetra-antennary structure containing an intersecting *N*-acetylglucosamine and an $\alpha(1\rightarrow6)$-linked fucose residue in the core unit for the majority of the glycans. In contrast to most glycopeptides of other sources, the intestinal glycopeptides were devoid of sialic acid, but contained 6–7 residues of fucose. The outer branches contained the following structures:

Fucα1-2Galβ1-3GlcNAcβ1- (H type 1)
Fucα1-2Galβ1-4GlcNAcβ1- (H type 2)
Galβ1-4(Fucα1-3)GlcNAcβ1- (X)
Fucα1-2Galβ1-4(Fucα1-3)GlcNAcβ1- (Y)
GalNAcα1-3(Fucα1-2)Galβ1-3GlcNAcβ1- (A type 1)
GalNAcα1-3(Fucα1-2)Galβ1-4GlcNAcβ1- (monofucosyl A type 2)
GalNAcα1-3(Fucα1-2)Galβ1-4(Fucα1-3)GlcNAcβ1- (difucosyl A type 2)
GalNAcα1-3(Fucα1-2)Galβ1-4(Fucα1-3)GlcNAcβ1-3Galβ1-4(Fucα1-3)GlcNAcβ1- (trifucosyl A type 2)

The blood group determinant structures were mainly of type 2, whereas glycolipids from the same cells contained mainly type 1 determinants. The polyfucosylated glycans represent a novel type of blood group active glycopeptides. The unique properties of the small intestinal glycopeptides as compared with glycopeptides of other tissue sources may be correlated with the specialized functional properties of the small intestinal epithelial cells.

The epithelial surface of the small intestine, consisting of a single layer of cells, has many important physiological functions, such as the selective absorption of nutrients and the formation of a barrier to chemical and infectious agents of the gastrointestinal tract. For such function, the properties of the cell surface components are anticipated to be of major importance. The small intestinal epithelium has been found to be unusually rich in glycosphingolipids (1–3). Moreover, some glycolipids, such as the blood group active glycolipids, seem to be especially characteristic of this tissue (1–3).

In view of the many structural and antigenic similarities between glycoproteins and glycolipids (4), it would be of interest to get information of the properties of protein-bound carbohydrates of the small intestinal epithelium. Such information could also give indications as to the functional properties of these glycans. Although studies have been carried out on the properties of the free mucus or soluble glycoprotein enzymes isolated from the gastrointestinal tract (5, 6), little is known of the properties of the cellular glycoproteins and their carbohydrate units. In the present work we describe the isolation and characterization of human small intestinal epithelial glycopeptides, which appear to represent a novel type of blood group active carbohydrate units of glycoproteins.

EXPERIMENTAL PROCEDURES

Isolation of Intestinal Epithelial Cells—Specimens of human small intestine were obtained during surgery. Permission was given by the Ethical Committee, the Medical Faculty, University of Göteborg. A sample of proximal ileum from a patient with the blood group status A_1Le(a−b+) secretor and a sample of proximal jejunum from a patient with the blood group status HLe(a−b−) secretor were studied. The blood group A patient was operated on because of a leiomyoma of the small intestine. The blood group O patient was resected *ad modum* Billroth II 30 years earlier and was now being treated with total gastrectomy due to a ventricular cancer. The intestinal specimens were free of tumor tissue. The blood group status of the patients was established by routine typing of the red blood cells and saliva per-

* This work was supported by the Sigrid Jusèlius Foundation, the Academy of Finland, Grant 3967 from the Swedish Medical Research Council, the Netherlands Foundation for Chemical Research (SON/ZWO), Grant UUKC 83-13 from the Netherlands Cancer Foundation (K. W. F.), and by Grant HL-38213 from the United States National Institutes of Health. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed: Dept. of Medical Biochemistry, University of Turku, Kiinamyllynkatu 10, SF-20520 Turku, Finland.

formed at the Blood Bank, Sahlgrenska Hospital, Göteborg. The total epithelial cells were isolated by incubation in EDTA-containing buffer, as previously described (7).

Extraction and Analysis of Glycolipids—The epithelial cells were extracted with methanol and chloroform/methanol 1:2 (v/v) as previously described (7). The total nonacid glycolipids were isolated from this extract (8) and analyzed by thin layer chromatography, mass spectrometry, and ^{1}H NMR spectroscopy as described elsewhere (2, 7, 9, 10). The delipidated cells were further extracted as described by Svennerholm and Fredman (11) to ensure extraction of the more polar glycolipids, but no significant amounts of glycolipids were detected in this extract as analyzed by thin layer chromatography. The delipidated cells were finally treated with absolute ethanol and dried under reduced pressure (12). Immunostaining with monoclonal antibodies on thin layer plates was performed as described in Ref. 13.

Preparation and Fractionation of Glycopeptides—Aliquots of 200 mg of the dried delipidated cells were subjected to extensive digestion with Pronase as previously described (12). After removal of the material precipitable with cetylpyridinium chloride, the total glycopeptides were purified by gel filtration on a column (2 × 45 cm) of Sephadex G-25 Fine eluted with 10 mM pyridine-acetic acid buffer, pH 5.0 (12). The glycopeptides were fractionated by gel filtration on a column (2 × 75 cm) of Sephadex G-50 Fine eluted with 0.1 M pyridine-acetic acid buffer, pH 5.0.

Prior to affinity chromatography, the glycopeptides were passed through columns of Dowex 50W-X16 in order to remove residual free peptides, and *N*-^{3}H-acetylated in their peptide moiety (12). The glycopeptides were separated on a column (4.0 × 5.2 cm) of concanavalin A-Sepharose by elution with a two-step gradient of methyl-α-D-glucoside (12). The glycopeptide fractions were desalted by gel filtration on Sepadex G-50 and subjected to affinity chromatography on a column of *Bandeiraea simplicifolia* lectin I-Sepharose under conditions described before (14).

Analytical Methods—The sugar composition of the glycopeptides was determined by gas-liquid chromatography after methanolysis (15), using *myo*-inositol as an internal standard. Detection of the glycopeptides in gel filtration was carried out by the anthrone reaction (16) or by radioactivity (12). Permethylation of the glycopeptides was carried out using potassium *t*-butoxide as a base (17). The permethylated samples were degraded by acetolysis/acid hydrolysis and analyzed as their alditol acetates by gas-liquid chromatography using selective ion monitoring as previously described (18). Defucosylation of aliquots (40–50 nmol) of the glycopeptides was carried out for methylation analysis or NMR purposes by treatment with 0.1 M HCl at 100 °C for 1 h; the excess HCl was removed under vacuum overnight (19).

^{1}H NMR Spectroscopy—Samples for NMR experiments were repeatedly dissolved in ^2H_2O (99.96 atom % ^{2}H, Aldrich) at room temperature and p^{2}H 6 with intermediate lyophilization. The deuterium-exchanged glycopeptides were subjected to ^{1}H NMR spectroscopy at 500 MHz (Bruker WM-500 instrument equipped with an Aspect-2000 computer, SON NMR facility, Dept. Biophysical Chemistry, University of Nijmegen, the Netherlands). Spectra were recorded at room and elevated probe temperatures (27 and 52 °C, respectively). For solvent peak suppression in the spectra of fraction *AS* and its defucosylated analogue, a water-eliminated Fourier transform pulse sequence (π-τ-π/2-acquisition, with composite, nonselective π-pulse) was applied (20). Further experimental details have been described (21). ^{1}H chemical shifts are expressed in ppm downfield from internal 4,4-dimethyl-4-silapentane-1-sulfonate. They were actually measured relative to internal acetone (δ 2.225) with an accuracy of 0.002 ppm.

Mass Spectrometry—100 μg of the *OS* glycopeptides and 60 μg of the *AS* glycopeptides were permethylated and $LiAlH_4$-reduced (22, 23). The samples were analyzed on a VG ZAB-HF mass spectrometer (Manchester, United Kingdom) equipped with a VG 11-250 data system. Both the electron impact-mass spectrometry (EI-MS)[1] and the fast atom bombardment-mass spectrometry techniques were applied in the positive mode. The calibration was controlled by manually counting peaks in paper spectra. The figures reproduced refer to the nominal masses obtained this way.

[1] The abbreviations used are: EI-MS, electron impact mass spectrometry; Cer, ceramide; Fuc, fucose; GalNAc-ol, *N*-acetylgalactosaminitol; Hex, hexose; HexNAc, *N*-acetylhexosamine.

RESULTS

Preparation and Fractionation of Small Intestinal Glycopeptides—After extensive proteolytic degradation, 80% of the glycopeptides prepared from delipidated human small intestinal epithelial cells were found to be not precipitated with cetylpyridinium chloride. By gel filtration on Sephadex G-50, the cetylpyridinium chloride-soluble glycopeptides were separated into two fractions, one eluting in the void volume of the column and one eluting in the included volume (Fig. 1*A*). The glycopeptides eluted in the void volume were degraded by mild alkaline borohydride treatment into fragments of different sizes (Fig. 1*B*). Analysis of the sugar composition of these fractions revealed the presence of both mannose and *N*-acetylgalactosaminitol (results not shown), indicating that they contained mixtures of both *N*- and *O*-glycosidic glycans. These glycopeptides and oligosaccharides have not been studied further, as yet.

The glycopeptides eluted in the included volume of the Sephadex G-50 column (Fig. 1*A*), containing 64 and 60% of the carbohydrate of blood group A and O small intestinal glycopeptides, respectively, were *N*-^{3}H-acetylated in their peptide moiety and subjected to fractionation by lectin affinity chromatography.

In contrast to glycopeptides of several other tissues, only minor amounts of the small intestinal glycopeptides bound to concanavalin A-Sepharose (Fig. 2). This result was confirmed by quantitative sugar analysis which indicated that only 2.6–3.4% and 3.3–6.6% of the total sugars were found in the two fractions eluted from the column with 20 and 200 mM methyl α-D-glucoside, respectively.

The glycopeptides not bound to concanavalin A-Sepharose were desalted by gel filtration and subjected to chromatography on a column containing the immobilized *B. simplicifolia* lectin I, specific for terminal α-galactose and α-*N*-acetylgalactosamine (Fig. 3). Glycopeptides from blood group O small intestine did not bind to the column, whereas a major part of

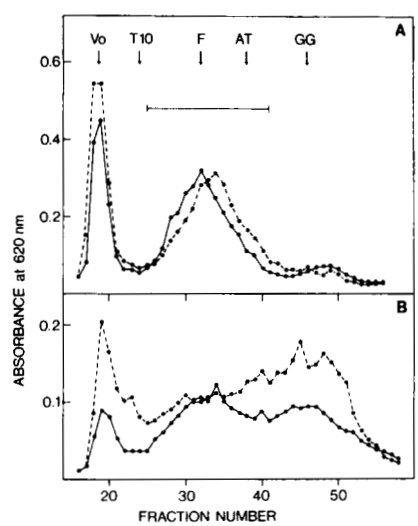


FIG. 1. **Gel filtration of glycopeptides from human small intestinal epithelial cells.** The column (2 × 75 cm) of Sephadex G-50 Fine was eluted with 0.1 M pyridine-acetic acid buffer, pH 5.0. Fractions of 4.5 ml were collected and analyzed for neutral sugar by the anthrone method (absorbance at 620 nm). *A*, total cetylpyridinium chloride-soluble glycopeptides; *B*, the material in the void volume in A after alkaline borohydride treatment. V_o, void volume; *T10*, dextran M_r 10,000; *F*, fetuin glycopeptide; *AT*, asialotransferrin glycopeptide; *GG*, Galβ(1→3)GalNAc-ol. ●——●, glycopeptides from blood group A individual; ●- - -●, glycopeptides from blood group O individual. The fractions indicated by the bars were used for further purification of the glycopeptides.

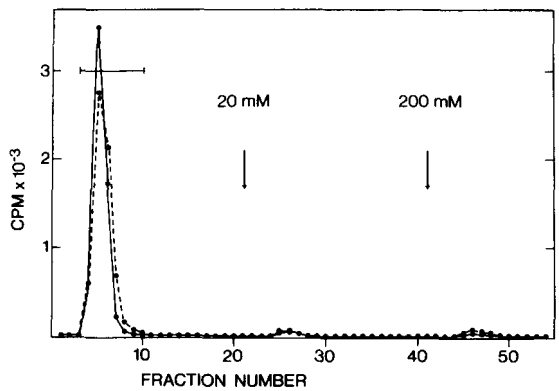


FIG. 2. **Affinity chromatography of glycopeptides from human small intestinal epithelial cells on concanavalin A-Sepharose.** The affinity column (4.0 × 5.2 cm) was eluted with the starting buffer and 20 and 200 mM methyl-α-D-glucoside was included in the buffer at the positions indicated by *arrows*. Fractions of 15 ml were collected and analyzed for radioactivity of the *N*-^{3}H-acetylated glycopeptides. ●——●, glycopeptides from blood group A individual; ●- - -●, glycopeptides from blood group O individual. The fractions indicated by the bars were used for further purification of the glycopeptides.

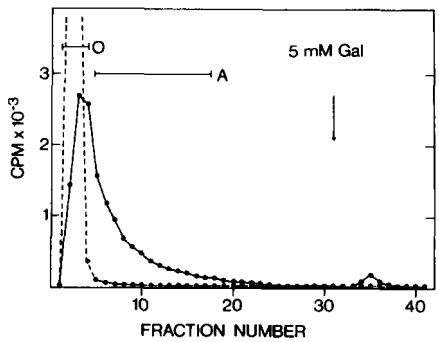


FIG. 3. **Affinity chromatography of glycopeptides from human small intestinal epithelial cells on *B. simplicifolia* lectin I-Sepharose.** The affinity column (1.4 × 10 cm) was eluted with the starting buffer, and 5 mM galactose was included in the buffer at the *arrow*. Fractions of 6 ml were collected and analyzed for radioactivity of the *N*-^{3}H-acetylated glycopeptides. ●——●, glycopeptides from the blood group A individual; ●- - -●, glycopeptides from the blood group O individual. The fractions indicated by the bars were used for further purification of the glycopeptides.

the blood group A glycopeptides were retarded and required several column volumes of buffer for elution. In contrast to the blood group A active glycopeptides from human erythrocytes chromatographed under the same conditions (14), only minor amounts of the blood group A small intestinal glycopeptides were bound to the column and eluted in the elution buffer, when galactose was included (Fig. 3).

The blood group O glycopeptides and the retarded fraction of blood group A glycopeptides were finally subjected to gel filtration (Fig. 4). The middle portions of each peak, fractions *AS* and *OS*, were subjected to structural characterizations.

Sugar Composition and Methylation Analysis—The sugar composition of the glycopeptides isolated is shown in Table I. The values are expressed assuming 3 mannose residues/molecule, which is in accordance with the molecular size of the glycopeptides in gel filtration (Fig. 4). The proportions of mannose, galactose and *N*-acetylglucosamine were found to be typical of *N*-glycosidic glycopeptides of various sources. However, the glycopeptides differed from previously studied glycopeptides in their high content of fucose and in the absence of sialic acid residues. In addition, the glycopeptides

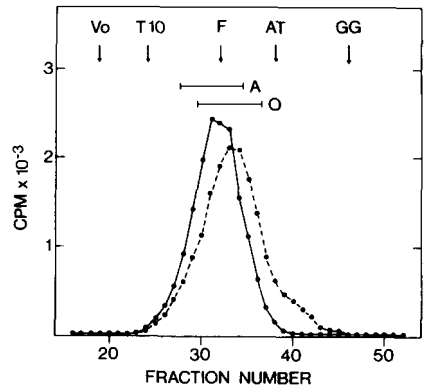


FIG. 4. **Gel filtration of the glycopeptides isolated by affinity chromatography.** The glycopeptides were subjected to gel filtration as described in the legend of Fig. 1. The middle part of each peak (*A* and *O*), as shown by the bars, was used for structural characterizations.

TABLE I

Monosaccharide composition of glycopeptides from human small intestinal epithelial cells

The glycopeptides from individuals of blood group A and O were isolated as shown in Figs. 1–4. The values are given assuming 3 mol of mannose/mol of glycopeptide.

Constituent	Blood group A (*AS*)	Blood group O (*OS*)
	mol/mol	
Fuc	7.4	5.7
Man	3.0	3.0
Gal	4.0	4.0
GalNAc	2.9	
GlcNAc	6.8	6.7
NeuAc		

from the blood group A individual contained *N*-acetylgalactosamine, a rare constituent of *N*-glycosidic carbohydrate units.

Methylation analysis (Table II) revealed a substitution pattern of mannose residues suggesting tetra- and triantennary core structures for the glycopeptides. The presence of approximately 1 residue of 3,4,6-tri-substituted mannose concomitant with 1 residue of terminal *N*-acetylglucosamine suggested that the chitobiose-linked mannose residue of the core contained an intersecting *N*-acetylglucosamine residue.

In *AS* glycopeptides the presence of terminal *N*-acetylgalactosamine and 2,3-di-substituted galactose, which were not found in *OS* glycopeptides (Table II), suggested the presence of the blood group A determinant structure (Scheme 1). In accordance with this, defucosylation with mild acid converted most of the 2,3-di-substituted galactose into 3-mono-substituted galactose. In *OS* glycopeptides, the 2-mono-substituted galactose residues were converted into terminal galactose residues by defucosylation, which suggested the presence of the H determinant structure (Scheme I) in these glycopeptides.

An unusually high proportion of the *N*-acetylglucosamine residues was found to be 3,4-di-substituted in both *AS* and *OS* glycopeptides (Table II). After defucosylation most of these residues were found to be replaced by 4-mono-substituted *N*-acetylglucosamine, which suggested the possible presence of the X antigen structure (Scheme 1) in the glycopeptides. The presence of some 3-mono-substituted *N*-acetylglucosamine after defucosylation suggested that a small proportion of the blood group determinants could be of type 1.

^{1}*H NMR Spectroscopy*—To arrive at a more definite eval-

uation of the heterogeneity that appeared to occur in the small intestinal glycopeptide fractions *OS* and *AS* from the data of sugar and methylation analyses (see Tables I and II), 500 MHz ^{1}H NMR spectroscopy was applied, both before and after defucosylation. The relevant parts of the ^{1}H NMR spectra of fraction *OS*, recorded before and after defucosylation, are compared in Fig. 5. The characteristic features of the ^{1}H NMR spectra of glycopeptides *OS* and *AS* and their defucosylated analogues (*df-OS* and *df-AS*, respectively) are summarized in Table III. The structures discussed and the coding of residues are indicated in Scheme 2.

TABLE II

Methylation analysis of native and defucosylated small intestinal glycopeptides

Defucosylation of the glycopeptides was carried out by mild acid treatment. The relative proportions of the hexosamine derivatives were taken from the mass fragmentography scans and combined using the data of the total sugar composition. The proportion of the fucose derivative was not determined due to interfering peaks near the solvent front.

Constituent	Substitution	Blood group A (*AS*)		Blood group O (*OS*)	
		Native	Defucosylated	Native	Defucosylated
		mol/mol			
Gal	Terminal	0.4	0.9	0.8	3.1
	3	0.9	2.2	1.1	0.8
	2	0.2	0.1	2.2	0.1
	6				
	2,3	2.5	0.9		
	3,6				
Man	2	0.3	0.5	0.3	0.7
	2,4	0.9	0.8	1.0	0.6
	3,6	0.1	0.2	0.1	0.1
	2,6	0.7	0.5	0.5	0.4
	3,4,6	1.0	1.1	1.1	1.1
GlcNAc	Terminal	0.9	1.5	0.9	1.4
	4	2.1	4.4	2.4	4.3
	3	0.2	0.7	0.8	0.6
	3,4	3.3	0.1	2.3	0.2
	4,6	0.3	0.1	0.3	0.1
GalNAc	Terminal	2.3	2.7		

The NMR-spectroscopic data confirm that both fractions *OS* and *AS* contain *N*-glycosidic glycopeptides of the *N*-acetyllactosamine type. The *N*-type is concluded from the presence of signals that are typical (21) for the structural element (·→4)GlcNAcβ(1→4)GlcNAcβ(1→*N*)Asn, namely, the H-1 doublet at $\delta \approx 5.05$ ($J_{1,2} > 9$ Hz) and the NAc singlets at δ 2.011 and $\delta \approx 2.094$. The singularity of the signal at δ 2.011 in the spectra of both (*df-*)*OS* and (*df-*)*AS* indicates that the residual peptide moiety (resulting after Pronase digestion) of the fractions is homogeneous in composition (21). A Fuc residue was found to occur in α(1→6)-linkage to the Asn-bound GlcNAc residue (GlcNAc-1) of all glycans in fractions *OS* and *AS*. This is revealed (21) by the position of the NAc signal of GlcNAc-2, namely, δ 2.095 (for *OS*) or δ 2.093 (for *AS*). The α(1→6)-linked Fuc has its H-1 signal at δ 4.87, its H-5 signal at δ 4.12, and its CH_3 signal at δ 1.200 (see Table III). Only the CH_3 doublet at δ 1.200 could be observed individually (for *OS*, see Fig. 5*C*); the H-1 signal of Fucα(1→6) coincides with a Manα(1→6) H-1 signal and with some of the Fucα(1→3) H-5 signals, whereas the H-5 signal of Fucα(1→6) is found among Man H-2 and Fucα(1→2) H-5 signals (see below). The core portion of the glycopeptides in *OS* and *AS* is completed by 3 Man residues in the assembly that is common to all *N*-type glycan chains; evidence for the latter statement is outlined below.

It became evident from methylation analysis (see Table II) that the compounds in fractions *OS* and *AS* were likely to contain highly branched $Man_3GlcNAc_2$ core units. The number of β-linked GlcNAc residues attached to the core, and their linkage positions, are in general expected (21) to become available from inspection of the Man H-2 region of the ^{1}H NMR spectrum ($4.0 < \delta < 4.3$ ppm), in conjunction with counting the number of NAc methyl singlets ($2.0 < \delta < 2.1$ ppm). In the spectrum of *OS* (see Fig. 5*A*), the NAc methyl proton region shows signals equivalent to a total of about 7.5 NAc groups (based on the unit intensity of the signals at δ 2.011 and δ 2.095); thus, 5–6 peripheral GlcNAc residues occur in the glycans of fraction *OS*. From integration of the corresponding region of the spectrum of *AS*, we concluded that a

SCHEME 1

Blood-group and differentiation antigenic determinants

Coding	Structure
H type-1	Fucα(1→2)Galβ(1→3)GlcNAcβ(1→•)
H type-2	Fucα(1→2)Galβ(1→4)GlcNAcβ(1→•)
X	Galβ(1→4)GlcNAcβ(1→•) Fucα(1→3)/
Y	Fucα(1→2)Galβ(1→4)GlcNAcβ(1→•) Fucα(1→3)/
B type-1/2	Fucα(1→2)Galβ(1→3/4)GlcNAcβ(1→•) Galα(1→3)/
Difucosyl B	Fucα(1→2)Galβ(1→3/4)GlcNAcβ(1→•) Galα(1→3)/ Fucα(1→4/3)/
A type 1/2	Fucα(1→2)Galβ(1→3/4)GlcNAcβ(1→•) GalNAcα(1→3)/
Difucosyl A	Fucα(1→2)Galβ(1→3/4)GlcNAcβ(1→•) GalNAcα(1→3)/ Fucα(1→4/3)/
Trifucosyl A	Fucα(1→2)Galβ(1→3/4)GlcNAcβ(1→3)Galβ(1→4)GlcNAcβ(1→•) GalNAcα(1→3)/ Fucα(1→4/3)/ Fucα(1→3)/

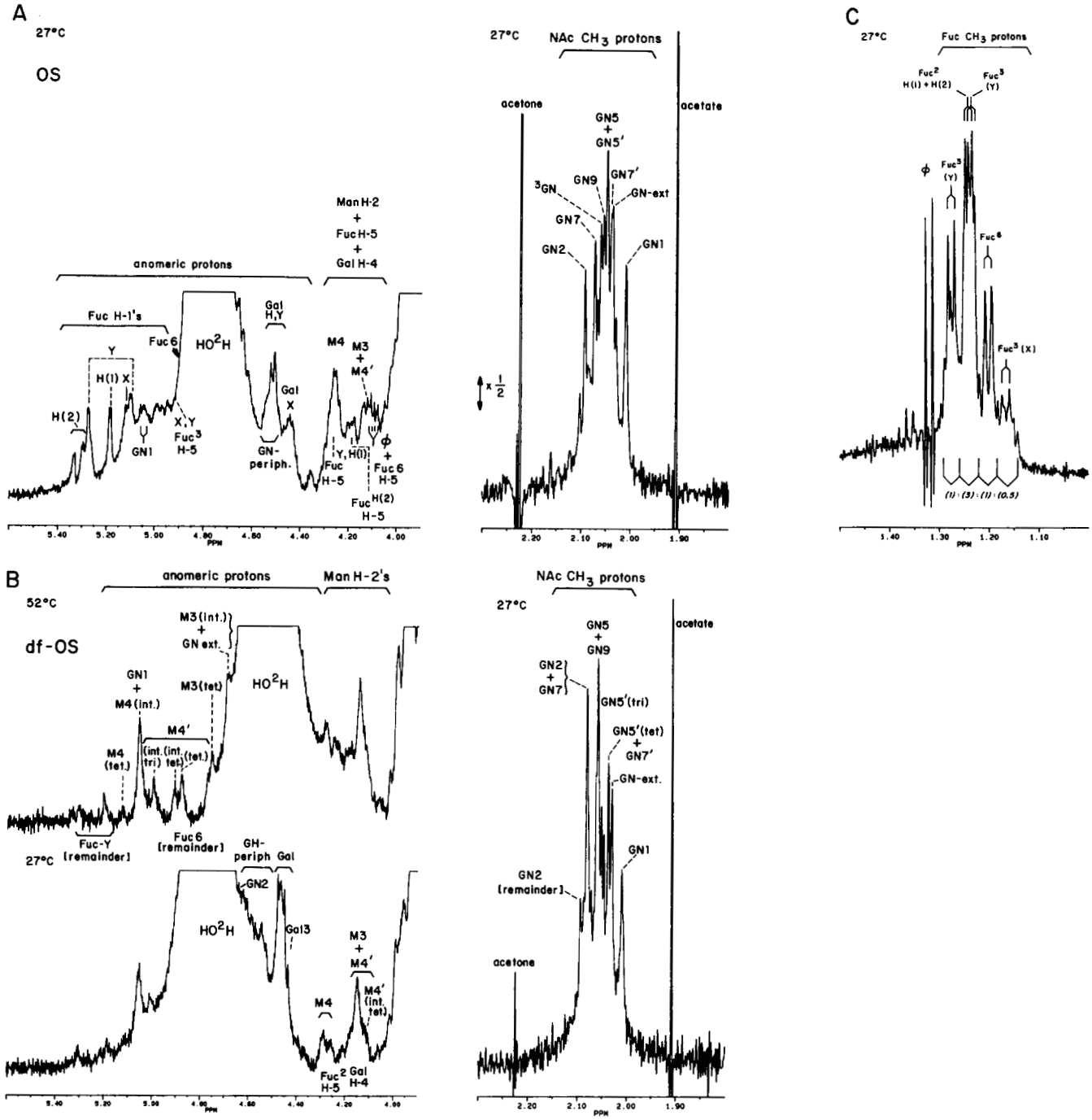


FIG. 5. **Structural-reporter group regions of the resolution-enhanced 500 MHz ^{1}H NMR spectra of glycopeptide fraction *OS*, recorded before (*trace A*) and after (*trace B*) defucosylation.** The anomeric-proton region ($3.9 < \delta < 5.6$ ppm) of the spectrum of *df-OS*, recorded at 27 °C is shown on the *bottom trace* of *B*; the corresponding region of the spectrum recorded at 52 °C is shown on the *top trace* of *B*. *Trace C* shows the Fuc CH_3 region of the spectrum of *OS* in *trace A*. For the comprehensive structure of glycopeptides *OS*, see Scheme 2. The numbering of the core and backbone residues is as indicated in Table III and Scheme 2. Branching patterns are indicated as *tri* (for tri-antennary), *tet* (for tetra-antennary), and *int* (for intersected). *GN-ext* stands for extra (*i.e.* additional) GlcNAc residue. Superscripts as in Fuc^2, Fuc^6, Gal^3 indicate the position of the glycosidic linkage in which the residue is involved; ^{3}GN indicates that the GlcNAc residue is substituted at the C-3 position. The key to the coding used for blood group and differentiation antigenic determinants is Scheme 1. *Remainder* in *trace B* refers to the fact that defucosylation was not complete (see text). Signals marked by ϕ stem from a frequently occurring noncarbohydrate, nonprotein contaminant.

total of 9.5 NAc groups is present in compound *AS*. Sugar analysis (Table I) indicated that 2.5 of these belong to the GalNAc residues that give rise to the blood group A activity of the native *AS* glycopeptides. Thus, also taking into account that two signals belong to GlcNAc-1 and -2, we arrived at a number of 5 peripheral GlcNAc residues in *AS*. The spectra of both *OS* (Fig. 5*A*) and *AS*, however, showed highly crowded patterns of partially overlapping resonances in the $4.0 < \delta < 4.3$ regions; among those, the Man H-2 signals could not readily be located. The signals obscuring this spectral region

TABLE III

^1H chemical shifts of pertinent structural-reporter groups of the blood group active glycopeptides (and their defucosylated analogues) from human small intestinal epithelial cells from individuals with blood group O and A

Residue[a]	Reporter group	Chemical shift[b] in *df-OS* 27 °C	*df-OS* 52 °C	*OS* 27 °C	*df-AS*[c] 27 °C	*AS*[c] 27 °C
GlcNAc-**1**	H-1	5.055	5.055	5.056	5.050	5.06
	NAc	2.011	2.014	2.011	2.010	2.011
GlcNAc-**2**	H-1	4.629	ND[d]	ND	4.63	ND
	NAc	2.081[e]	2.079[e]	2.095	2.083	2.093
Fucα(1→6)	H-1			4.875		4.871
	H-5			4.12		ND
	CH_3	—[e]	—[e]	1.200		1.201
Man-**3**	H-1	ND	4.686[int/]	ND	4.695[int]	4.695[int]
		ND	4.752[tet]	ND	ND	ND
	H-2	4.147	4.145	4.140	4.148	ND
Man-**4**	H-1	5.055[int]	5.055[int]	5.052[int]	5.079[int]	5.039[int]
		5.12[tet]	5.122[tet]	ND	ND	ND
	H-2	4.286	4.283	ND	4.286	ND
Man-**4′**	H-1	5.003[inttri]	4.998[inttri]	4.992[inttri]	5.05[inttri]	5.06[inttri]
		ND	4.877[tet]	4.875[tet]	ND	ND
		4.916[inttet]	4.906[inttet]	4.950[inttet]	4.963[inttet]	5.039[inttet]
	H-2	4.147[inttri]	4.145[inttri]	4.140[inttri]	4.148	ND
		4.11[inttet]	4.11[inttet]	4.11[inttet]	4.105	ND
GlcNAc-**5***	H-1	4.58*	ND	4.52	4.54	4.576
	NAc	2.059[t]	2.060[t]	2.051[t]	2.061[t]	2.054[t]
GlcNAc-**5′**	H-1	4.56	ND	4.52	4.54	4.576
	NAc	2.052[inttri]		2.051[inttri]	2.040[t]	2.036/2.024
		2.039[inttet]	2.042[inttet]	2.042[inttet]	2.041[inttet]	
GlcNAc-**7**	H-1	4.52	ND	4.52	4.54	4.576
	NAc	2.081	2.079	2.075	2.083	2.074
GlcNAc-**7′**	H-1	4.53	ND	4.52	4.54	4.576[t]
	NAc	2.039[t]	2.042[t]	2.040[t]	2.040[t]	2.036/2.024
GlcNAc-**9**	H-1	4.47	ND	4.46	4.47	4.45
	H-4	3.283	3.285	3.277	3.28	3.287
	NAc	2.059[t]	2.060[t]	2.062[t]	2.061[t]	2.070[t]
Gal[3] (**6***)	H-1	4.438*	ND	4.56	4.45 (*1*)[c]	4.436 (*1*)[c]
Gal-**6′/8/8′**	H-1	4.46	ND	4.443	4.609 (*3*)	4.689 (*3*)
				4.518		
	H-4[#]	4.147	4.145	ND	4.215 (*3*)	4.198 (*3*)
GlcNAc-ext β(1→3)	H-1	4.70	4.686	ND	ND	ND
	NAc	2.032	2.034	2.057	ND	ND
				(2.032)		
Gal-ext β(1→4)	H-1	4.46	ND	4.45	ND	ND
Fucα(1→2)[I]	H-1			5.188		5.216
	H-5			4.26		4.26
	CH_3			1.235		1.29/1.242
Fucα(1→2)[2]	H-1			5.300		5.300 (*3*)
				(5.334)		
	H-5			4.21		4.32 (*3*)
	CH_3			1.235		1.293 (*3*)
Fucα(1→2)[Y]	H-1			5.278		5.279 (<*1*)
	H-5			4.25		4.25
	CH_3			1.240		1.293 (<*1*)
Fucα(1→3)[X]	H-1			5.120		
	H-5			4.87		
	CH_3			1.166		
Fucα(1→3)[Y]	H-1			5.102		5.119[g] (<*1*)
	H-5			4.87		4.87
	CH_3			1.277		1.279
GalNAc	H-1				5.174	5.200
	H-2				4.215	4.235
	H-5				ND	4.16
	NAc				2.040	2.036

[a] For comprehensive structures of *OS* and *AS* and coding of residues, see Scheme 2.

[b] Chemical shifts were acquired at 500 MHz for neutral ^2H_2O solutions at 27 °C, and for *df-OS* also at 52 °C (δ HO^2H 4.509).

[c] Italicized numbers in parentheses refer to relative-intensity ratios of the signals in the NMR spectra.

[d] ND, value could not be determined, because signal overlapped with HO^2H resonance at the applied probe temperature, or with other signals.

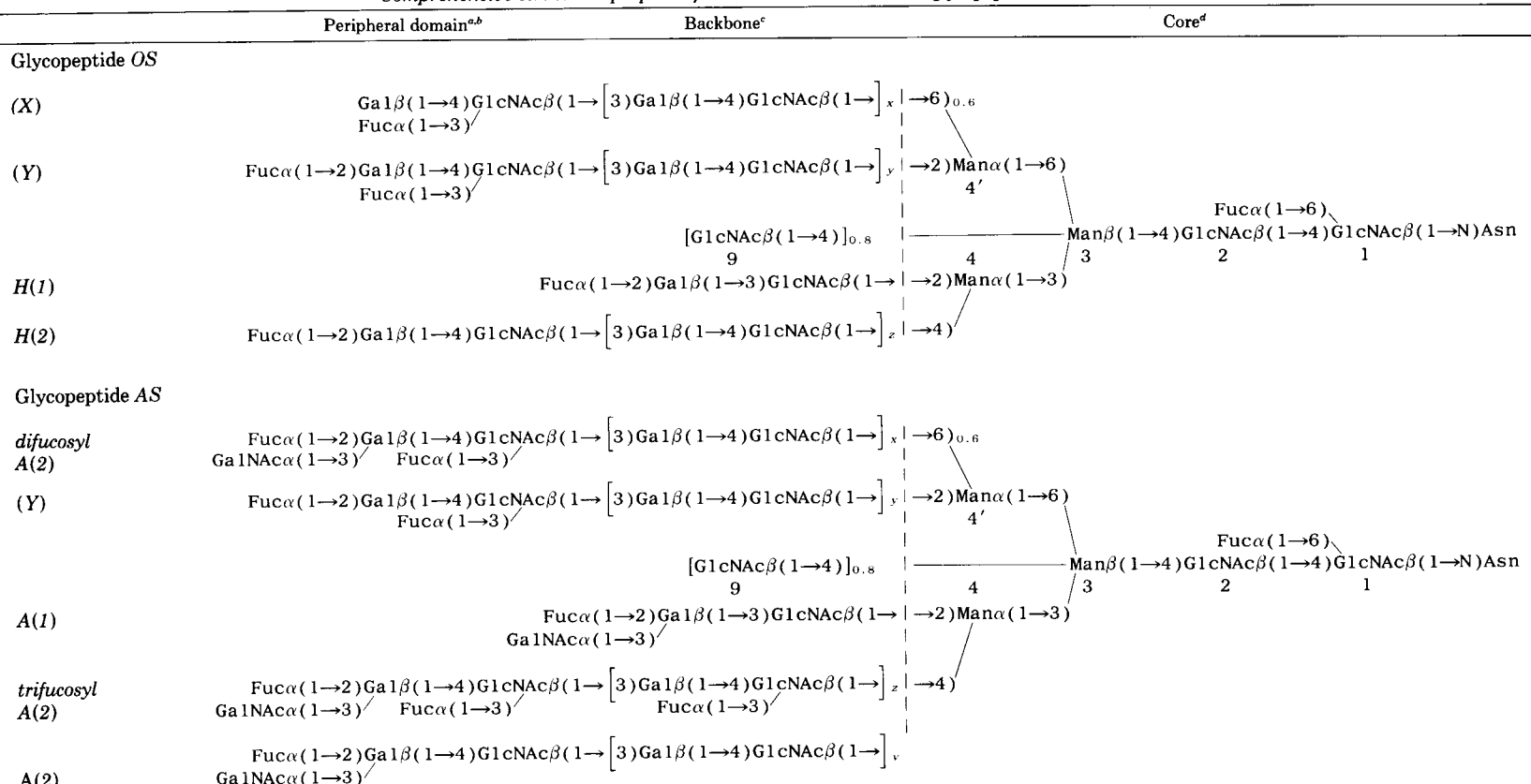
[e] Defucosylation under the applied conditions (see text) was not complete. For example, the Fucα(1→6) residue linked to GlcNAc-1 remained present in at least 35% of *OS* and *AS*. Two signals were observed in the spectra of *df-OS* and *df-AS* for the NAc methyl protons of GlcNAc-2, at δ 2.081 and δ 2.094, in the intensity ratio 5:1.

[f] Additional symbols used are: [t], tentative assignment; [1,2], refer to the corresponding blood-group (H for *OS*, A for *AS*) antigenic determinants (see Scheme 1). [int], stands for intersected compounds; [inttri], for intersected tri-antennary compounds; [inttet], stands for intersected tetra-antennary compounds; [tet] for nonintersected tetra-antennary compounds; *, assuming arbitrarily that the β(1→3)-linked Gal is Gal-6* (see text and legend to Scheme 2); [#], substituted by the "additional" *N*-acetyllactosamine unit in β(1→3)-linkage.

[g] Alternatively, this set of Fucα(1→3) reporter group signals may be assigned to the third Fuc residue in the trifucosyl A determinant (see text).

SCHEME 2

Comprehensive structures proposed for human small-intestinal glycopeptides OS and AS



[a] The type-1 blood-group determinants are linked directly to the core. Here, they are arbitrarily shown β(1→2)-linked to Man-4; however, their precise location is not known (see text).

[b] The average number of GalNAc residues per molecule *AS* is 2.5; this is in agreement with the occurrence of the β(1→6)-linked branch in 60% of the molecules.

[c] The sum of the indices $x + y + z$ (for *OS*) and $+ v$ (for *AS*) = 1, indicating that the glycopeptides contain, on average, one "additional" *N*-acetyllactosamine unit, which varies in type of terminal antigenic structure.

[d] The branch β(1→6)-linked to Man-4′ is present in 60% of the glycopeptides, the intersecting GlcNAc-9 residue occurs in 80% of the molecules.

were thought to be either Fuc H-5 signals or signals of structural-reporter groups of other monosaccharides that had emerged out of the bulk resonance of skeleton protons ($3.4 < \delta < 4.0$) due to the presence of the Fuc residues. To enable the determination of the sites of substitution of the peripheral GlcNAc residues to the $Man_3GlcNAc_2$ core unit, the *OS* and *AS* glycopeptides were subjected to mild acid treatment resulting in the removal of most (but not all; see footnote *e* to Table III) of the Fuc residues.

Backbone Structure of the Components of Fractions OS and AS—The chemical shift values of the predominant Man H-1 and H-2 signals in the spectrum of *df-OS* have been included in Table III. At least three different sets of values were observed, pointing to the occurrence of three different types of branching in *OS*. The set of values δ 5.055, 4.998, and 4.686 (measured at 52 °C) for H-1 of Man-4, -4′, and -3, respectively, suggested the presence of intersected triantennary type glycopeptides (see Table IV; compare Ref. 21). The second set of values, namely, δ 5.122, 4.877, and 4.752, pointed to the presence of nonintersected tetra-antennary type of compounds (see Table IV; compare Ref. 21). The third set of values, namely, δ 5.055, 4.906, and 4.686 for H-1 of Man-4, -4′, and -3, respectively, had not been previously observed. In view of the results of methylation analysis (Table II), this third set of values is attributed to glycopeptides with an intersected tetra-antennary type of branching. The NMR line of reasoning for this assignment is based on the analogy between, on one hand, the chemical shift differences between the new set of Man H-1 (and H-2) chemical shifts and the one previously established for nonintersected tetra-antenna, and, on the other hand, the chemical shift effects observed for the Man H-1 atoms in the step from intersected tri- to nonintersected triantenna (see Table IV). The relative intensity ratio of these three sets of signals (taking into account that also at δ 5.055 the H-1 signal of GlcNAc-1 is located, that the signal at δ 4.877 at least in part stems from the remainder of Fucα(1→6), and that the signal at δ 4.686 partly belongs to the H-1 doublet of one of the peripheral GlcNAc residues, see later) points to the occurrence of intersected tri-, intersected tetra-, and nonintersected tetra-antennary compounds in the ratio 2:2:1. That is to say, Man-4 is substituted by GlcNAc-5 and -7 at C-2 and C-4, respectively, in all components of (*df-*)*OS*. Man-4′ bears GlcNAc-5′ and -7′ at C-2 and C-6, respectively, in about 60% of the constituents of (*df-*)*OS*; in the remaining 40% of the molecules, Man-4′ is substituted only at C-2 by GlcNAc-5′. Finally, Man-3 is substituted at C-4 by a β-linked (so-called intersecting) GlcNAc residue (usually referred to as GlcNAc-9) in 80% of the molecules in (*df-*)*OS* (see Scheme 2). Relatively intense signals in the Man H-2 region of the spectrum of *df-OS* are seen at δ 4.28 and 4.14-4.12 (Man-4 H-2 signal, and coinciding H-2 signals of Man-3 and -4′, respectively) (see Fig. 5*B*); these values are in agreement with the predominance of intersected type of glycopeptides in (*df-*)*OS* (compare Table IV).

The branching patterns of the glycopeptides in *df-AS* are proposed to be the same as those in *df-OS*, although the ratios in which they occur might be slightly different. Apparently, the exact chemical shift values for H-1 and H-2 of Man-3, -4, and -4′ for *df-AS* (Table III) are modified by the presence of the GalNAcα(1→3)Galβ(1→4) arms, compared with those for *df-OS* and the reference compounds (Table IV).

The presence of the intersecting GlcNAc-9 residue in the majority (≈80%) of glycopeptide molecules, both in *OS* and *AS*, is corroborated by the occurrence of the H-4 signal of this residue outside the bulk of skeleton resonances, at $\delta \approx$ 3.28 (see Table III; *cf.* Ref. 24). These findings on the occurrence of the various branching patterns, and the ratio in which they occur, from the NMR data correlate rather well with the results of methylation analysis (Table II).

Each of the peripheral GlcNAc residues (denoted 5, 5′, 7,

TABLE IV

Chemical shift data for the mannose H-1 and H-2 atoms in intersected di-, tri-, tetra-, and penta-antennary Asn-linked oligosaccharides of the N-acetyllactosamine type[a]

The majority of these data have been taken from Refs. 21, 25–28.

Type of branching	Chemical shift of					
	Man-3 H-1	Man-4 H-1	Man-4′ H-1	Man-3 H-2	Man-4 H-2	Man-4′ H-2
—5′—4′ \ 3—2— —5—4— /	4.77	5.12	4.93	4.25	4.19	4.11
—5′—4′ \ 9—3—2— —5—4— /	4.70	5.06	5.00	4.18	4.25	4.15
—5′—4′ \ 3—2— —5—4— / —7— /	4.76	5.12	4.92	4.21	4.22	4.11
—5′—4′ \ 9—3—2— —5—4— / —7— /	4.69	5.06	5.00	4.15	4.28	4.14
—7′ \ —5′—4′ \ 3—2— —5—4— / —7— /	4.76	5.13	4.87	4.21	4.22	4.09
—7′ \ —5′—4′ \ 9—3—2—[a] —5—4— / —7— /	4.69	5.05	4.92	4.15	4.28	4.12
—7′ 10′ \ —5′—4′ \ 9—3—2— —5—4— / —7— /	4.71	5.06	4.89	4.15	4.28	4.16

[a] The values for the intersected tetra-antennary type of branching were inferred from this study.

and 7′) linked to the α-Man residues is substituted by a Gal residue. From methylation analysis (Table II), it was deduced that one of these Gal residues is present in (1→3)-linkage to GlcNAc, whereas the other Gal residues are (1→4)-linked, thus forming part of *N*-acetyllactosamine units. The β(1→3)-linked Gal residue in *df-OS* has its H-1 signal at δ 4.438 (*cf.* Refs. 29–33); the other Gal residues have their H-1 signals at δ ≈ 4.47. The determination of the site of attachment of the Galβ(1→3)GlcNAcβ(1→·) residue (so-called type-1 unit; see Scheme 1) to the core is, at least in principle, feasible by NMR, for example by comparison with the data published for the biantennary glycopeptide from rat plasma hemopexin (29) having Galβ(1→3)GlcNAcβ(1→·) attached to C-2 of Man-4 and those for the triantennary glycopeptide obtained from calf fetuin (30, 31) having the Galβ(1→3)GlcNAcβ(1→·) unit attached to C-4 of Man-4. However, the complexity of the β-anomeric proton regions as well as of the NAc methyl proton regions of the spectra of *df-OS* and *df-AS* did not permit the deduction of the position of the type-1 unit unambiguously. (In the data listed in Table III, we tentatively assume, for reasons of convenience in tabulating the chemical shifts, that the type-1 unit is linked to C-2 of Man-4.)

The fifth (for tri-) or sixth (for tetra-antennary compounds) peripheral GlcNAc residue is not linked to a Man residue but is rather present in β(1→3)-linkage to one of the aforementioned Gal residues. In turn, this GlcNAc bears a Gal residue in β(1→4)-linkage. Typical ^{1}H NMR spectroscopic features of the β(1→3)-linked GlcNAc are its H-1 doublet at δ ≈ 4.69 and its NAc methyl singlet at δ 2.032 (*cf.* Refs. 21, 25, 33, 34). The H-4 signal of the Gal residue substituted by this β(1→3)-linked GlcNAc residue is found at δ ≈ 4.147 (21, 33, 34), thereby coinciding in the spectrum of (*df-*)*OS* with some Man H-2 signals (see Table III). The NMR characteristics of such an extension of a fundamental *N*-acetyllactosamine branch by an additional *N*-acetyllactosamine unit have been described before for an intersected biantennary *N*-type glycopeptide (25) and for some smaller size, mucin-type *O*-glycosidic oligosaccharide-alditols (33). Owing to the lack of reliable NMR data obtained on higher branched, *N*-type compounds with repeating *N*-acetyllactosamine units, it was not possible to determine by ^{1}H NMR spectroscopy the identity of the branch(es) to which this extra *N*-acetyllactosamine unit may be attached in *OS*. Due to the complexity of the ^{1}H NMR spectra of *AS* and *df-AS*, the occurrence of such an extra *N*-acetyllactosamine unit could neither be proven nor denied. In methylation analysis, both *AS* and *OS* contain a Gal residue monosubstituted at C-3 (Table II), which leads us to assume that the extra *N*-acetyllactosamine unit is present in *AS*, too. Mass spectrometry of the permethylated glycopeptides supports this assumption (see below).

Outer Region Structures of the Glycopeptides in Fractions OS and AS—With respect to the Fuc residues occurring in the outer regions of the *OS* glycopeptides, the following details could be elucidated from the ^{1}H NMR spectrum (Fig. 5*A*). In addition to the core-linked Fucα(1→6) residue mentioned above, Fuc residues occur in *OS* in two other types of linkages, namely α(1→2) to Gal and α(1→3) to peripheral GlcNAc. Thus, Fuc forms part of either one of the following determinants: H, X, and Y (see Scheme 1). ^{1}H NMR spectroscopic data for Fuc residues forming part of various types of blood group determinants, collected from the literature, have been compiled in Tables V and VI.

Evidence for the presence of the X-antigenic determinant is afforded by the Fuc H-1 signal at δ 5.12, the H-5 signal at δ 4.87, and the CH_3 doublet at δ 1.17 (see Table III) (21, 25, 32, 34–40). A possible reason for the somewhat diffuse shape of the CH_3 doublet (Fig. 5*A*) is the simultaneous occurrence of peptide signals in this region (*cf.* Ref. 21). The occurrence

TABLE V

^{1}H chemical shifts of pertinent structural-reporter groups of blood group and differentiation antigenic determinants H, X, and Y in oligosaccharides with type-2 backbones

Residue[a]	Reporter group	Chemical shift[b] in[c]								
		$H(2)^R$	X^R	Y^R	$H(2)^3$	X^3	Y^3	$H(2)^6$	X^6	Y^6
$GlcNAc^R$	H-1 α	5.210	5.106	5.090						
	β	4.715	4.729	4.725						
	NAc α	2.048	2.033	2.034						
	β	2.045	2.031	2.032						
$GlcNAc^3$	H-1				4.685	4.692	4.68			
	NAc				2.043	2.032	2.028			
$GlcNAc^6$	H-1							4.541	4.557	4.558
	NAc							2.064	2.056	2.057
Gal^4	H-1	4.541^β	4.455^β	4.507^β	4.550	4.45	4.476	4.539	4.442	4.498
	H-4	ND	ND	ND	3.89	3.90	3.89	3.89	3.92	ND
Fuc^2	H-1	5.308		5.277	5.310		5.276	5.305		5.279
	H-5	4.226		4.249	4.226		4.258	4.225		4.26
	CH_3	1.224		1.235	1.237		1.235	1.235		1.235
Fuc^3	H-1		5.098^R	5.086^R		5.138	5.118		5.106	5.093
	H-5		4.839	4.882		4.832	4.873		4.83	4.869
	CH_3		1.172	1.263		1.177	1.272		1.171	1.270

[a] Superscripts indicate the position of the glycosidic linkage in which the residue is involved.

[b] Chemical shifts were acquired at 500 MHz for neutral solutions at 27 °C.

[c] For complete structures of determinants, see Scheme 1. Data were compiled from Refs. 32–43 and 48–50. Here, superscripts indicate the position of the linkage in which GlcNAc is involved; thus, ^{3}, Galβ(1→4)GlcNAcβ(1→3) (usually to Galβ(1→•)); ^{6}, Galβ(1→4)GlcNAcβ(1→6) (usually to GalNAc-ol). It should be noted that the chemical shifts of the GlcNAc H-1 and NAc protons (and also those of the Gal and Fuc reporter groups) are more or less strongly dependent on the nature of the aglycone. ^{R}, reducing GlcNAc residue (or for other residues than GlcNAc: attached to reducing GlcNAc). $^\beta$, refers to the chemical shift of the reporter group in the β-anomer of the (reducing) oligosaccharide.

TABLE VI

^1H chemical shifts of pertinent structural-reporter groups of blood group antigenic determinants H, Le^a, Le^b, B, and A, in oligosaccharides with type-1 backbones

Residue[a]	Reporter group	Chemical[b] shift in[c]								
		H(1)[3]	H(1)	Le^a	Le^b	afucosyl B[d]	B(1)	difucosyl B(1)	A(1)	difucosyl A(1)[P]
GlcNAc[3]	H-1	4.654	—[c]	4.70	4.665	4.727[Rβ]	—[c]	4.598	—[c]	
	NAc	2.111		2.036	ND	2.043[R]		2.072		
Gal[3]	H-1	4.577	4.583	4.43	4.582	4.546[d]	4.697	4.725	4.676	*4.70*
	H-4	3.889	3.916	3.92	3.90	4.182	4.278	4.234	4.214	*4.17*
Fuc[2]	H-1	5.208	5.268		5.153		5.349	5.232	5.378	*5.26*
	H-5	4.273	4.269		4.354		4.317	4.391	4.322	*4.40*
	CH_3	1.230	1.235		1.252		1.227	1.274	1.232	*1.29*
Fuc[4]	H-1			5.030	5.026			5.041		*5.04*
	H-5			4.88	4.880			4.834		*4.83*
	CH_3			1.183	1.276			1.287		*1.29*
Gal[α3]	H-1					5.145	5.259	5.165		
	H-4					4.021	4.015	ND		
	H-5					4.194	4.119	4.262		
GalNAc[α3]	H-1								5.185	*5.18*
	H-2								4.251	*4.25*
	H-4								4.020	
	H-5								4.160	
	NAc								2.049	*2.05*

[a] Superscripts indicate the position of the glycosidic linkage in which the residue is involved.

[b] Chemical shifts were acquired at 300 or 500 MHz for neutral solutions at 22 or 27 °C.

[c] For complete structures of determinants, see Scheme 1. Data were compiled from Refs. 20, 36, 37, and 44–47. Here, the superscript 3 indicates the position of the linkage in which GlcNAc is involved: Galβ(1→4)GlcNAcβ(1→3) (usually to Galβ(1→•)). When no values are listed for the GlcNAc[3] reporter groups for a certain determinant, the Gal residue is β(1→3)-linked to GalNAc-ol. The two columns of chemical shifts for H(1) illustrate that the chemical shifts of the Gal and Fuc reporter groups are more or less strongly dependent on the nature of the aglycone.

[d] In the afuco B-determinant, β-Gal is actually (1→4)-linked to reducing GlcNAc, so the values refer to the afuco B(2)-structure. Additional symbols used are: [R], reducing GlcNAc residue. [β], refers to the chemical shift of the reporter-group in the β-anomer of the (reducing) oligosaccharide. [P], values predicted from the shift effects observed in the step from B(1) to difucosyl B(1).

of one of the Fuc residues in α(1→2)-linkage to Galβ(1→4) (constituting the H type-2 determinant) is revealed by the Fuc H-1 signal at δ ≈ 5.30, in conjunction with the CH_3 doublet at δ 1.235 (Table III) (21, 34–39, 41). The Galβ(1→3)GlcNAcβ(1→·) branch (type-1 unit) in *OS* appeared to be terminated also by a Fuc residue in α(1→2)-linkage to Gal; together these residues constitute the H type-1 determinant (see Scheme 1). Typical NMR parameters for this blood group determinant are the Fuc H-1 signal at δ 5.188 together with the CH_3 doublet at δ 1.235 (Table III; compare Refs. 32, 34–37). Finally, the combination of Fuc H-1 signals at δ 5.28 (Fucα(1→2)) and δ 5.10 (Fucα(1→3)) together with the CH_3 signals at δ 1.24 (Fucα(1→2)) and 1.28 (Fucα(1→3)) is highly characteristic of the occurrence of the Y determinant (35–38, 42).

The intensity ratio of the various Fuc H-1 signals, along with that of the Fuc CH_3 doublets (1:3:1:0.5) (see Fig. 5*C*), indicates that, on average, there are 5–6 Fuc residues present/molecule in *OS*, divided over 6 different positions: Fucα(1→6) to GlcNAc-1; Fucα(1→3) to GlcNAc in X determinant; Fucα(1→2) to Galβ(1→4) in H type-2 unit, probably in two different branches; Fucα(1→2) to Galβ(1→3) in H type-1 unit; and two types of Fuc in the Y determinant.

As has been outlined previously (43), the chemical shift of a Fuc H-1 signal is not only indicative of the type of linkage in which Fuc itself is involved, and the nature of the residue to which it is attached, but also depends on a number of other structural details in the branch to which it is attached (*e.g.* the type-1 or type-2 character of the Gal-GlcNAc unit and the type of linkage in which GlcNAc is involved). Thus, the chemical shift of Fuc H-1 is a most sensitive marker of the complete microenvironment of the Fuc residue, but, *mutatis mutandis*, is also the parameter whose value is the most difficult to predict among the chemical shifts of all ^1H NMR structural-reporter groups. Therefore, it is quite possible that the signal at δ 5.334 in the spectrum of *OS*, like the more commonly observed one at δ 5.300, belongs to a Fuc residue in an H type-2 determinant. However, that H type-2 unit should then be attached to Man in a type of linkage not observed before (neither β(1→2) to Man (21), nor β(1→3) or β(1→6) to Gal (41)), leading us to propose that this H type-2 unit may be β(1→4)-linked (to Man 4). Alternatively, the presence of the intersecting GlcNAc residue might modify ("finetune") the chemical shift of H-1 of a Fuc residue in the H type-2 determinant at the end of a β(1→2)-linked branch, compared with the value established for nonintersected type of structures. Therefore, the chemical shifts of the Fuc H-1 signals observed for fraction *OS* could not easily be used for branch localization of the various blood group determinants.

The alternative and usually complementary way of localizing terminal Fuc residues, *e.g.* blood group determinants, in different branches of a complex carbohydrate chain is (43) based on the chemical shift effects brought about by removal of the various Fuc residues, in particular on the positions of the NAc signals at δ ≈ 2.05. However, the complexity of the mixture of *OS* glycopeptides with respect to the heterogeneity

of the core branching pattern, the impossibility of even determining the position of the core to which the type-1 unit is linked, the nature of the Gal residue to which the extra *N*-acetyllactosamine unit is attached, etc., prevented determination of the precise distribution of the various blood group determinants over the branches of the various *OS* components, except that the H type-1 unit seems to be directly attached to one of the core α-mannose residues. If this unit were the extra Gal-GlcNAc residue linked β(1→3) to a fundamental *N*-acetyllactosamine branch, its NAc signal would be expected to be observable at δ ≈ 2.11 (34). A similar chemical shift (δ ≈ 2.10–2.11) would be predicted for the NAc signal of the H type-1 unit if it were attached β(1→4) to Man-4 (39). Since we do not observe a NAc singlet at that position, we conclude that the H type-1 unit is directly attached to the core but not to the C-4 position of Man-4 (see Scheme 2).

As to the occurrence of blood group A and other determinants in the various glycopeptide components of fraction *AS*, the ^{1}H NMR spectra of *AS* and *df-AS* permitted the following conclusions. The major difference between the spectra of preparations of *df-AS* and *df-OS* is in the occurrence of a large signal at δ 5.174 ($J_{1,2}$ 4 Hz), dominating the anomeric proton region of the spectrum of *df-AS* (not shown). This signal is attributed to H-1 of α(1→3)-linked GalNAc attached to Galβ(1→·) residues (*cf.* Refs. 36, 37, 44–47). The relative intensity of the signal is in agreement (compare Tables I and II) with the occurrence of 2 to 3 GalNAc residues/molecule of *AS* glycopeptides. The other structural-reporter group resonances of α-linked GalNAc are its H-2 (at δ ≈ 4.22) and H-5 (at δ ≈ 4.16, both contributing to the "Man H-2 region" of the spectrum, and therefore hampering deduction of the core branching patterns of *AS* components even more than the situation with *OS*). The dominant NAc methyl singlet in the spectrum of *df-AS* at δ 2.040 (Table III) is also ascribed to the terminal GalNAc residues.

In the spectrum of native *AS*, the chemical shifts of the above mentioned GalNAc protons are found to be slightly altered (see Table III). These alterations are obviously due to the presence of Fuc residues in α(1→2)-linkage to the same Gal residue as to which GalNAc is attached, completing the blood group A sequence (see Scheme 1) (compare Refs. 36, 37, 44, 47). The reporter-group resonances of the Fucα(1→2) residues involved in these A determinants were assigned on the basis of their relatively large intensity (nearly equal to that of the GalNAc reporter group signals); they are found at δ 5.30 (H-1), δ 4.32 (H-5), and δ 1.293 (CH_3) (see Table III). In particular the chemical shift of the CH_3 protons differs drastically from that known to be characteristic for blood group A determinants (δ ≈ 1.23, see Table VI). A possible explanation for the apparent downfield shift might be the presence of an additional Fuc residue in α(1→3)-linkage to GlcNAc of the *N*-acetyllactosamine unit to which the A determinant is attached (together comprising the so-called difucosyl A type-2 determinant, see Scheme 1). NMR evidence for the presence of this type of determinant (in at least two branches of the *AS* molecules, by virtue of the relative intensity of the CH_3 doublet at δ 1.29) stems from the fact that, going from the H type-2 to the Y determinant by means of a similar attachment of Fucα(1→3) to the GlcNAc residue, a similar downfield shift of the CH_3 group of Fucα(1→2) is observed (Δδ ≈ 0.04 ppm; see Table V). A similar effect on the chemical shift of the Fuc^2 CH_3 doublet in the step from a type-2 monofucosyl A to a difucosyl A (so-called AY) determinant has been observed for glycolipids containing these structural elements (51). In spite of the fact that the NMR data on the glycolipids were obtained in Me_2SO/D_2O solutions, drawing such an analogy to the NMR data reported here for glycopeptides in D_2O is meaningful because it has been shown (52) that the conformations of such blood group determinants do not change significantly from one solvent system to another. Furthermore, the shift effects on the Fuc structural-reporter groups observed in the step from *df-AS* to *AS* (Table III) are essentially analogous to those observed going from an afuco-B to a difucosyl-B determinant (Scheme 1; Table VI).[2] In addition to these two difucosyl A type-2 (and/or perhaps one difucosyl A type-1) determinants, some branches in lower amount glycopeptide components of *AS* might terminate in H or Y determinants. However, the complexity of the mixture of *AS* glycopeptides, both with respect to heterogeneity of branching pattern and to the occurrence of various outer determinants, did not permit us to prove the latter statement, let alone to localize any of the determinants in any particular branch. The tentative comprehensive structure of glycopeptides *AS* has been included in Scheme 2.

Mass Spectrometry—The permethylated glycopeptide fractions *OS* and *AS* were analyzed by EI and fast atom bombardment (FAB) mass spectrometry, both in the positive mode. No molecular ions could be recorded by the FAB technique. Ions up to about 2000 mass units were recorded by the EI technique. They correspond to the outer regions of the *OS* and *AS* glycopeptides, with cleavage preferentially occurring at glycosidic linkages involving C-1 of HexNAc residues.

Fig. 6 shows the mass spectrum of the permethylated *OS* glycopeptides recorded at 360 °C. The intersecting GlcNAc-9 was found as a terminal HexNAc at *m/z* 260. Terminal trisaccharides Fuc-Hex-HexNAc were found at *m/z* 638 (and 606 = 638–32) corresponding to a blood group H sequence (both type-1 and -2 according to NMR), and an X determinant (Hex-[Fuc]HexNAc) (supported by the peak at *m/z* 432). The difucosyl Y determinant found in the NMR spectra is corroborated by the peaks at *m/z* 812 Fuc-Hex-[Fuc]HexNAc and 780 (812–32). The extension of this structure by a Hex-HexNAc unit is verified by the peaks at *m/z* 1261 and 1229. An extension of the H or X terminals by a Hex-HexNAc unit can be seen at *m/z* 1087 and 1055. The mass spectrum thus supports the interpretation of NMR spectra regarding outer region structures and once again shows the extent of heterogeneity. The relative distribution of these terminals given in Scheme 2 is also in rough accordance with the intensities of the different ions in the mass spectra. Higher molecular weight ions were also recorded up to *m/z* 2433, but of low intensity. This is only half-way up since one of the isomeric sequences has a calculated nominal mass of 4981. The interpretation of these is difficult as they probably are the result of both pyrolysis and electron-induced fragmentation.

The mass spectrum of the permethylated *AS* glycopeptides is shown in Fig. 7. Only fragments corresponding to terminal sequences could be recorded. Terminal fragments similar to the ones obtained for the *OS* glycopeptides are seen at *m/z* 638 and 812. The ions at *m/z* 812 are an indication that not all branches were terminated by HexNAc. Blood group A terminal sequences with one Fuc are found at *m/z* 883 and 851, and at 1332 and 1300 for a structure with an additional *N*-acetyllactosamine unit. Corresponding structures with one additional Fuc can be seen at *m/z* 1057 and 1506, respectively. Significant peaks at *m/z* 1680 and 1648 must be due to a structure within 3 fucose residues, the additional one probably located at the second internal HexNAc. The amount of this latter structure is probably almost as large as that of the difucosyl compound. The presence of the mono- and trifucosyl A sequences in addition to the difucosyl sequence is in agree-

[2] H. van Halbeek, unpublished results.

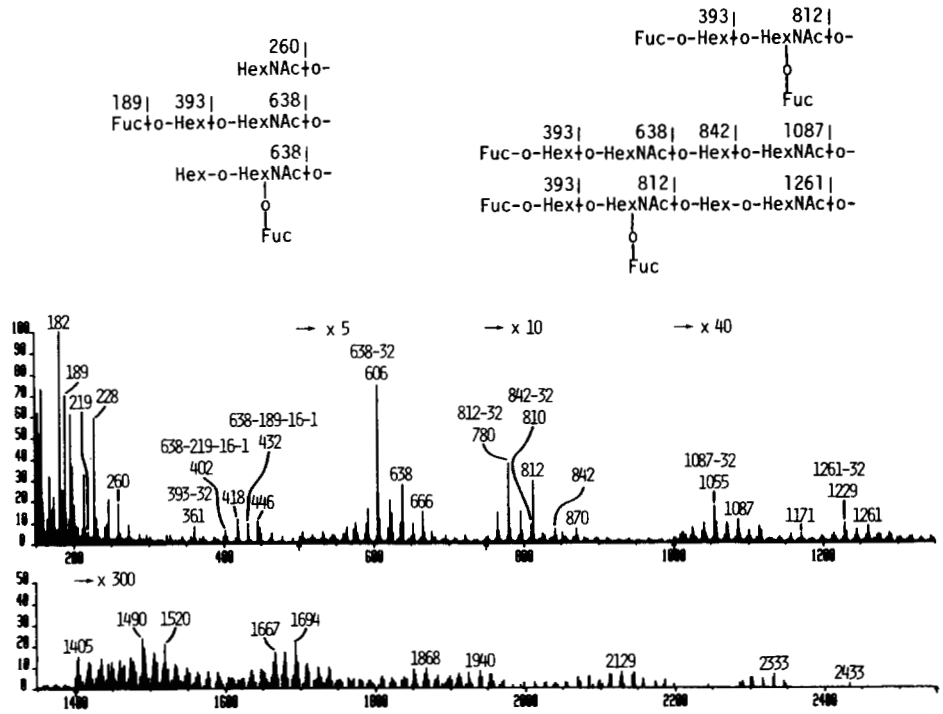


FIG. 6. **Electron impact mass spectrum of permethylated glycopeptides *OS* (2 nmol) from the blood group O individual.** The ion source was held at 360 °C, the electron energy was 40 eV, trap current 500 μA, and accelerating voltage 8 kV. The figures given refer to nominal masses.

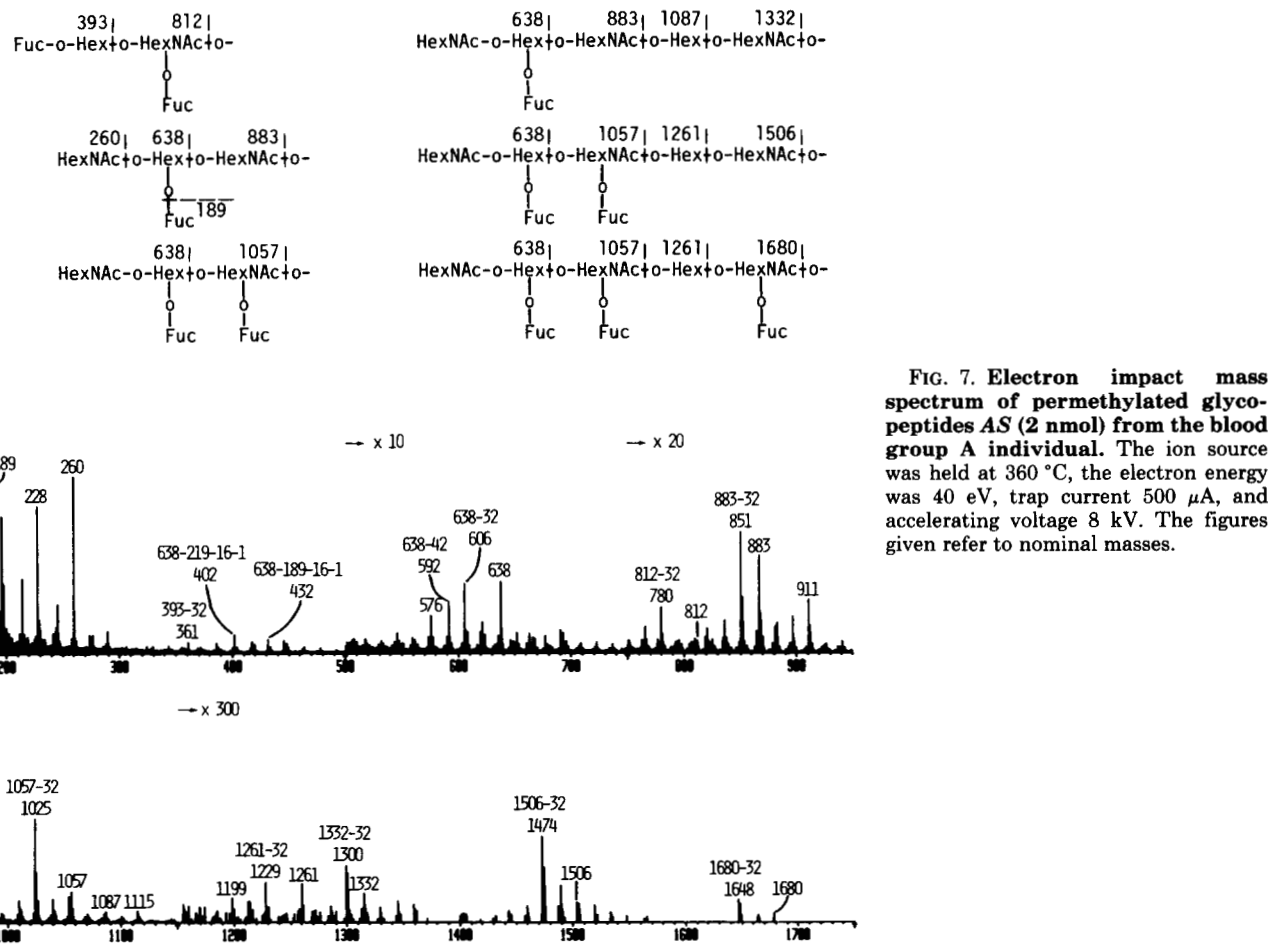


FIG. 7. **Electron impact mass spectrum of permethylated glycopeptides *AS* (2 nmol) from the blood group A individual.** The ion source was held at 360 °C, the electron energy was 40 eV, trap current 500 μA, and accelerating voltage 8 kV. The figures given refer to nominal masses.

ment with the NMR data, although the presence of the entire sequences could not be specifically pointed out with NMR due to the restrictions of this methodology. For example, the innermost fucose residue in the trifucosyl A structure would be ranked among the NMR-determined sequences as an X determinant. The peak at *m/z* 182 is the base peak in the *OS* and *AS* mass spectra. This ion is found in samples containing type-2 sequences (53) and is also in accordance with the

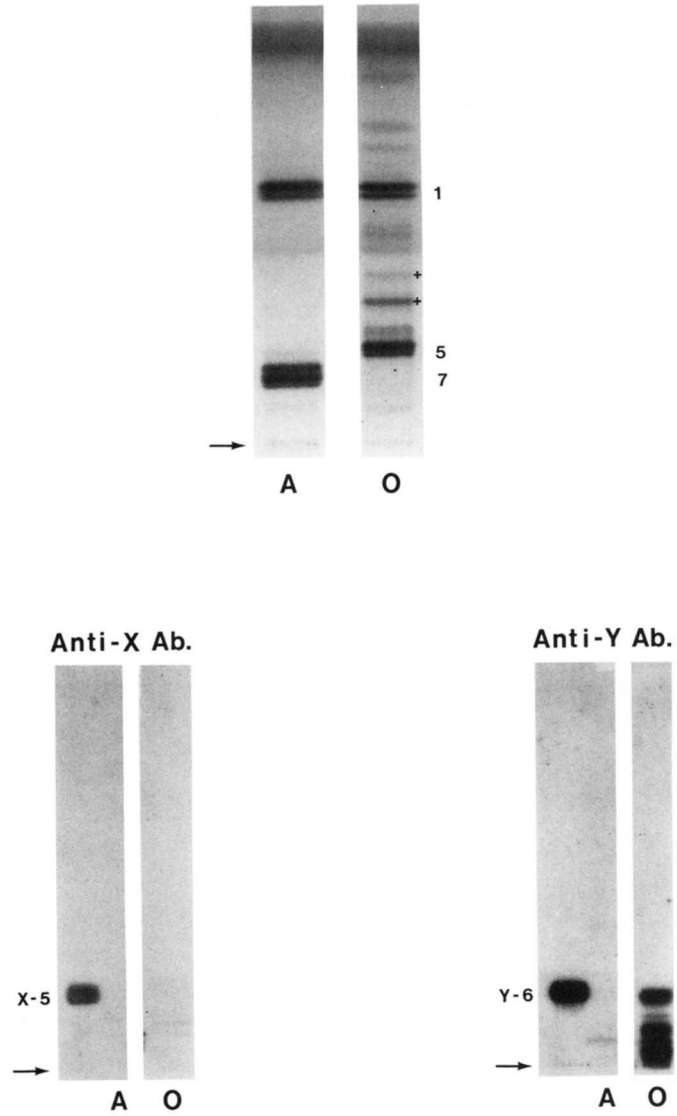


FIG. 8. **Thin layer chromatogram of total nonacid glycosphingolipid fractions from small intestinal epithelial cells of human blood group A_1Le(a−b+) secretor (*lane A*) and blood group OLe(a−b−) secretor (*lane O*) individuals.** Detection was done with a chemical reagent, anisaldehyde (*top chromatogram*) and by autoradiography after labeling with monoclonal anti-X antibodies (*bottom chromatogram, left*), and anti-Y antibodies (*bottom chromatogram, right*). In the top chromatogram, bands marked + in *lane O* are nonglycolipid contaminants and the designation to the *right* indicates number of sugar residues in each glycolipid band. A reference glycolipid, reacting with the antibody used, was applied to the *left* of the bottom chromatograms and the short hand designations, *X-5* and *Y-6*, stand for the type-2 chain isomers of the Le^a and Le^b penta- and hexaglycosylceramides, respectively. *Arrow* indicates sample origin. The amount of glycolipid applied was 20 μg (*top chromatogram*) and 10 μg (*bottom chromatograms*) for *lanes A* and *O* and 2 μg for the reference compounds. Monoclonal antibodies used were anti-X, $D_1$56-22 (13) and anti-Y, F-3 (55). Solvent was chloroform/methanol/water, 60:35:8 (by volume). Further details about the thin layer chromatography and the binding of antibodies to the plates are described in detail elsewhere (10, 13).

interpretation of the NMR spectra. No peaks giving evidence for the peptide portion were found, a result in accordance with EI-mass spectra of glycopeptides derived from human transferrin (54).

Glycolipid Composition of the Small Intestinal Epithelial Cells—The nonacid and acid glycosphingolipid fractions isolated from the epithelial cells were structurally characterized and the results are described in detail elsewhere (3). The total nonacid glycolipids of the two cases (about 8 μg of glycolipid/mg cell protein) were analyzed by thin layer chromatography including overlay with monoclonal antibodies (Fig. 8). For structural information the total fractions were separated on silicic acid chromatography into two subfractions containing monoglycosylceramides and more polar glycolipids, respectively. The polar fractions were converted to permethylated and $LiAlH_4$-reduced permethylated derivatives, and these were analyzed in detail by NMR spectroscopy and by mass spectrometry using the direct inlet and a fractional distillation from the probe. The sequences concluded and their relative abundances are summarized in Table VII. Both samples contained monoglycosylceramides and blood group type glycolipids with 5–7 sugar residues as major glycolipids. In the blood group A sample, Le^b hexaglycosylceramides and type 1 A (ALe^b) heptaglycosylceramides were present in about equal amounts. The blood group H sample had type-1 H pentaglycosylceramides as major complex glycolipid. Globotriaosylceramides and globotetraosylceramides, which are major glycolipids of the nonepithelial part of the small intestine, were absent from the epithelial cells. In both cases traces of more complex glycolipids were detected by mass spectrometry (Table VII) and by overlay with antibodies on the thin layer plate (Fig. 8). Blood group X determinants were absent from both glycolipid samples, while the blood group H sample contained a series of bands reacting with the anti-Y antibody (Fig. 8). Thus, the glycolipids had mainly type-1 chains and the glycopeptides mainly type-2 chains.

DISCUSSION

Previous studies on human intestinal glycoproteins have focused on the properties of soluble mucins (5, 56). These molecules contain carbohydrate chains mainly of the *O*-glycosidic type. In contrast, the bulk of the carbohydrate in the small intestinal cells was found to be of the *N*-glycosidic type. The polyfucosylated glycopeptides isolated from human small intestinal epithelium display structural features distinct from glycopeptides isolated from several other sources. In contrast to glycopeptides of human erythrocytes (14, 57), human brain (58), rat plasma (59), rat kidney, liver and brain (60), or different types of cells grown in culture (19, 61, 62), the small intestinal glycopeptides are almost devoid of sialic acid, which is usually the most common terminal sugar in the carbohydrate units of glycoproteins. Sialic acid is present in the minor fraction of glycopeptides precipitable with cetylpyridinium chloride, but the sugar composition of these glycopeptides suggests that it consists mainly of *O*-glycosidic sugar chains.[3] The main glycopeptide fraction, not precipitable with cetylpyridinium chloride, consists for the major part of *N*-glycosidic chains and is characterized by an unusually high content of fucose. Asparagine-linked glycopeptides with several fucose residues have previously been described from human granulocytes (63, 64). They differ from the small intestinal glycopeptides in that they are of the poly-*N*-acetyllactosamine type, contain fewer fucose residues, and also contain sialic acid. A striking feature of the intestinal glycopeptides is that the majority of potential fucosylation sites are in fact occupied by fucose residues.

A tentative overall model for the structural properties of the glycopeptides is shown in Scheme 2. In accordance with the results of fractionation on concanavalin A-Sepharose, the relative size in gel filtration, sugar composition, methylation analysis, mass spectrometry and NMR data, the glycopeptides

[3] J. Finne, unpublished results.

TABLE VII

Glycolipid structures identified in small intestinal epithelial cells

The glycolipids of the two human individuals were characterized by thin layer chromatography, mass spectrometry, ^{1}H NMR spectroscopy and antibody binding. Note that several glycolipid components are present in the fractions as seen by antibody binding in Fig. 8. The amount of substance given (+++++ to + and trace (+)) is a semiquantitation from the thin layer plate and the ion intensity of mass spectra.

Glycolipid structure	A_1Le(a−b+) secretor	HLe(a−b−) secretor
Galβ1-1'Cer,Glcβ1-1'Cer	+++++	+++++
Hex-Hex-Cer		+
Fucα1-2Gal-Hex-Cer		+
Galβ1-3GlcNAcβ1-3Galβ1-4Glcβ1-1'Cer		++
Fucα1-2Galβ1-3GlcNAcβ1-3Galβ1-4Glcβ1-1'Cer		++++
Galβ1-3(Fucα1-4)GlcNAcβ1-3Galβ1-4Glcβ1-1'Cer	+	
GalNAcα1-3(Fucα1-2)Gal-HexNAc-Hex-Hex-Cer	(+)	
Fucα1-2Galβ1-3(Fucα1-4)GlcNAcβ1-3Galβ1-4Glcβ1-1'Cer	++++	
GalNAcα1-3(Fucα1-2)Galβ1-3(Fucα1-4)GlcNAcβ1-3Galβ1-4Glcβ1-1'Cer	++++	
$(Fuc)_2$, HexNAc, $(Hex)_3$, Cer		(+)
$(HexNAc)_3$, $(Hex)_4$, Cer		(+)
$(Fuc)_2$, $(HexNAc)_2$, $(Hex)_4$, Cer		(+)
$(Fuc)_3$, $(HexNAc)_2$, $(Hex)_4$, Cer	(+)	
$(Fuc)_2$, $(HexNAc)_3$, $(Hex)_4$, Cer	(+)	

are of the tri- and tetra-antennary type, and most of them contain an intersecting *N*-acetylglucosamine residue as well as an α(1→6)-linked fucose residue attached to the core unit. The Gal-GlcNAc backbone units are mainly type-2 chains, but on average one type-1 chain/core unit is present. Glycopeptides from the blood group A individual contain the A determinant structure, whereas those from the O individual contain the corresponding H structure. In addition α(1→3)-linked fucose residues are present in the X (or SSEA-1), Y and di-, and trifucosylated A determinant structures. The exact localization of the different terminals to the various antennae was not possible with the methods available. Sequential digestions with exo- and endoglycosidases (65) could potentially be used to study this question, but the limited amounts of glycopeptides available, as well as the inhibitory effects of the multiple fucose residues on among others endo-β-galactosidase cleavability prevented the use of this approach for the structural analyses.

The type-2 blood group A, H, and X determinant structures are known to occur in asparagine-linked sugar chains of human glycoproteins (14, 57, 63, 64, 66). The occurrence of the Y determinant structure has been reported recently (42, 67), whereas the type-1 H and the difucosylated A structures seem to be novel for human *N*-linked glycans. All these structures are known constituents of glycolipids and *O*-linked chains of glycoproteins (2, 3, 10, 68), and thus give further evidence for the similarity of different classes of protein- and lipid-linked glycans (4).

Although FAB-mass spectrometry has previously been successfully used to characterize glycopeptides (69), more comprehensive results were obtained for the permethylated glycopeptides with EI-mass spectrometry. Like with large glycolipids studied before with the same technique (70), no peaks containing all the sugars were obtained. However, fragments corresponding to the outer branches confirmed the results obtained by NMR, and in addition clearly indicated the presence of a blood group A trifucosylated terminal. This type of structure has not been described for protein-linked glycans before but is analogous to the trifucosylated glycolipid described by Nudelman *et al.* (71).

The polyfucosylated glycopeptides represent a novel type of blood group active glycans of glycoproteins. The previously characterized glycans include the *O*-glycosidic series of structures of the blood group substances of ovarian cysts and other mucins (72, 73), and the *N*-glycosidic polyglycosyl (poly-*N*-acetyllactosamine) peptides of human erythrocytes (14, 57). The small intestinal glycopeptides differ from the latter by their much smaller size and their higher content of fucose. In the erythrocyte polyglycosylpeptides the blood group determinants only occur in part of the glycans, and these glycans seem to be of the largest size among these glycans (14, 57). In contrast to this, the small intestinal glycopeptides contain a high content of blood group determinant structures bound to a normal small-size backbone structure, which indicates that there are different factors regulating the expression of blood group determinants on the erythrocytes and the small intestinal cells.

TABLE VIII

Density of blood group A, H determinants on the human small intestine epithelial cell surface as compared to erythrocytes

Cell type	Blood group A, H determinants		Sphingolipids[a]
	Total	Ratio protein/lipid	Total
	nmol/ml cells	*mol/mol*	*nmol/ml cells*
Small intestine	1000[b]	15	1000
Erythrocytes[c]	20–60	2	1100[d]

[a] Sphingolipids (glycosphingolipids and sphingomyelin) are localized to the outer leaflet of the plasma membrane and make up ½–⅓ of the lipid molecules (2). The sphingolipid value thus gives a rough measure of the plasma membrane surface area.

[b] The estimate of the A or H determinants was based on the amounts of *N*-acetylgalactosamine and 2-substituted galactose in blood group A and O glycopeptides, respectively, taking into account the amounts of glycopeptides containing these structures and their yields during the isolation procedure.

[c] From Ref. 74.

[d] Calculation based on Refs. 2 and 75.

The present report together with the analyses of the glycolipids of the same source (10) represents the first cell type, besides erythrocytes, where both major types of blood group active glycoconjugates have been characterized in detail in order to get a more complete picture of the cell surface. Therefore, it is of special interest to compare the results with those of erythrocytes. Table VIII shows that the concentration of blood group A,H determinants on the small intestinal cells is 20–50 times that on erythrocytes, as estimated from the amounts of the determinants in the glycopeptides and their approximate yields during the purification. If the carbohydrate is found mainly on the cell surface, as is generally

believed (2), and if the sphingolipids are found in the outer part of the plasma membrane lipid bilayer making up ½–⅓ of the lipid molecules, then the density of blood group determinants on the small intestine epithelial cells is very high indeed (compare first and last columns of Table VIII). The greater contribution of A, H determinants by glycoproteins *versus* glycolipids is similar for the two sources. On the other hand, it is striking that the majority of the determinants of the small intestinal glycopeptides are of type 2, whereas the major blood group glycolipids are of type 1.

An unusual feature of the small intestinal glycopeptides is also the low proportion of concanavalin A-binding glycopeptides. The fraction bound most strongly to this lectin, containing the high-mannose glycopeptides, comprises some 8–25% of the glycopeptide carbohydrate of various tissue and cell types (58). However, only 2–4% of the total carbohydrate is recovered in this fraction from the small intestinal glycopeptides. The low concentration of these chains in the normal small intestinal epithelial cells is in accordance with the proposal that the high-mannose chains could be involved in the pathophysiology of gluten-sensitive enteropathy by serving as binding sites for gluten (76).

Another special feature of both classes of the cell-bound glycoconjugates of human small intestine epithelium is their very low content of sialic acid residues. For many cell types the sialic residues are considered the major contributor of cell surface negative charge. Interestingly, the small intestine epithelia of some other species such as the rat contain a high concentration of glycolipid-bound sialic acid (9, 77). Therefore, an interesting question is if the low concentration of sialic acid in the human is connected to some physiological function (*e.g.* absorption of nutrients), and if this function varies between species. On the other hand, even in the rat, there seems to be a postnatal shift from terminal sialylation toward terminal fucosylation of small intestinal brush-border glycoconjugates (78), which suggests some similarity to the human small intestine. The role of cell surface carbohydrates as receptors for the binding of microorganisms and, as such, for the species, tissue, and individual selectivity of host pathogen interaction is currently being discussed (79, 80). Some bacteria (81, 82) and other pathogens (83, 84) are known to use cell surface sialic acid residues as receptors. The low concentration of such residues at the human small intestine epithelial cell surface, together with their presence in the free mucus (5, 56), might serve as a dual mechanism to contribute to the normal bacterium-free conditions of the human small intestine.

Acknowledgments—Thanks are due to Dr. J. H. G. M. Mutsaers for valuable experimental assistance, to Maire Ojala and Hilkka Rönkkö for skillful technical assistance, and to Lisa Maynard for critical review of the manuscript.

REFERENCES

1. McKibbin, J. M. (1978) *J. Lipid. Res.* **19,** 131–147
2. Karlson, K.-A. (1982) in *Biological Membranes* (Chapman, D., ed) Vol. 4, pp. 1–74, Academic Press, London
3. Björk, S., Breimer, M. E., Hansson, G. C., Karlsson, K.-A., Larson, G., and Leffler, H. (1983) in *Red Cell Membrane Glycoconjugates and Related Genetic Markers* (Cartron, J.-P., Rouger, P., and Salmon, C., eds), pp. 125–137, Librairie Arnette, Paris
4. Rauvala, H., and Finne, J. (1979) *FEBS Lett.* **97,** 1–8
5. Mantle, M., Forstner, G. G., and Forstner, J. F. (1984) *Biochem. J.* **224,** 345–354
6. Triadou, N., Audran, E., Rousset, M., Zweibaum, A., and Oriol, R. (1983) *Biochim. Biophys. Acta* **761,** 231–236
7. Breimer, M. E., Hansson, G. C., Karlsson, K.-A., and Leffler, H. (1981) *Exp. Cell Res.* **135,** 1–13
8. Karlsson, K.-A. (1987) *Methods Enzymol.* **138,** 212–220
9. Breimer, M. E., Hansson, G. C., Karlsson, K.-A., and Leffler, H. (1982) *J. Biol. Chem.* **257,** 906–912
10. Björk, S., Breimer, M. E., Hansson, G. C., Karlsson, K.-A., and Leffler, H. (1987) *J. Biol. Chem.* **262,** 6758–6765
11. Svennerholm, L., and Fredman, P. (1980) *Biochim. Biophys. Acta* **617,** 97–109
12. Finne, J., and Krusius, T. (1982) *Methods Enzymol.* **83,** 269–277
13. Hansson, G. C., Karlsson, K.-A., Larson, G., McKibbin, J. M., Blaszczyk, M., Herlyn, M., Steplewski, Z., and Koprowski, H. (1983) *J. Biol. Chem.* **258,** 4091–4097
14. Finne, J., Krusius, T., Rauvala, H., Kekomäki, R., and Myllylä, G. (1978) *FEBS Lett.* **89,** 111–115
15. Bhatti, T., Chambers, R. E., and Clamp, J. R. (1970) *Biochim. Biophys. Acta* **222,** 339–347
16. Nikkilä, E. A., and Pesola, R. (1960) *Scand. J. Clin. Lab. Invest.* **12,** 209–220
17. Finne, J., Krusius, T., and Rauvala, H. (1980) *Carbohydr. Res.* **80,** 336–339
18. Krusius, T., and Finne, J. (1977) *Eur. J. Biochem.* **78,** 369–379
19. Finne, J., Tao, T.-W., and Burger, M. M. (1980) *Cancer Res.* **40,** 2580–2587
20. Haasnoot, C. A. G. (1983) *J. Magn. Reson.* **52,** 153–158
21. Vliegenthart, J. F. G., Dorland, L., and Van Halbeek, H. (1983) *Adv. Carbohydr. Chem. Biochem.* **41,** 209–374
22. Larson, G., Karlsson, H., Hansson, G. C., and Pimlott, W. (1987) *Carbohydr. Res.* **161,** 281–290
23. Karlsson, K.-A. (1974) *Biochemistry* **13,** 3643–3647
24. Stanley, P., Vivona, G., and Atkinson, P. H. (1984) *Arch. Biochem. Biophys.* **230,** 363–374
25. Pierce-Cretel, A., Debray, H., Montreuil, J., Spik, G., Van Halbeek, H., Mutsaers, J. H. G. M., and Vliegenthart, J. F. G. (1984) *Eur. J. Biochem.* **139,** 337–349
26. Paz Parente, J., Wieruszeski, J. M., Strecker, G., Montreuil, J., Fournet, B., Van Halbeek, H., Dorland, L., and Vliegenthart, J. F. G. (1982) *J. Biol. Chem.* **257,** 13173–13176
27. Paz Parente, J., Strecker, G., Leroy, Y., Montreuil, J., Fournet, B., Van Halbeek, H., Dorland, L., and Vliegenthart, J. F. G. (1983) *FEBS Lett.* **152,** 145–152
28. Van Halbeek, H., Vliegenthart, J. F. G., Iwase, H., Li, S.-C., and Li, Y.-T. (1985) *Glycoconjugate J.* **2,** 235–253
29. Bernard, N., Engler, R., Strecker, G. Montreuil, J., Van Halbeek, H., and Vliegenthart, J. F. G. (1984) *Glycoconjugate J.* **1,** 123–140
30. Van Halbeek, H., Rijkse, I., Van Beek, W. P., and Vliegenthart, J. F. G. (1985) in *Proceedings of the VIIIth International Symposium on Glycoconjugates* (Davidson, E. A., Williams, J. C., and DiFerrante, N. M., eds) Vol. 1, pp. 120–121, Praeger, New York
31. Townsend, R. R., Hardy, M. R., Wong, T. C., and Lee, Y. C. (1986) *Biochemistry* **25,** 5716–5725
32. Dua, V. K., and Bush, C. A. (1983) *Anal. Biochem.* **133,** 1–8
33. Mutsaers, J. H. G. M., Van Halbeek, H., Vliegenthart, J. F. G., Wu, A. M., and Kabat, E. A. (1986) *Eur. J. Biochem.* **157,** 139–146
34. Lamblin, G., Boersma, A., Lhermitte, M., Roussel, P., Mutsaers, J. H. G. M., Van Halbeek, H., and Vliegenthart, J. F. G. (1984) *Eur. J. Biochem.* **143,** 227–236
35. Derappe, C., Bauvy, C., Lemonnier, M., Platzer, N., Lhermitte, M., Egge, H., Peter-Katalinic, J., Van Halbeek, H., Kamerling, J. P., and Vliegenthart, J. F. G. (1986) *Carbohydr. Res.* **150,** 273–284
36. Lemieux, R. U., Rock, K., Delbaere, L. T. J., Koto, S., and Rao, V. S. (1980) *Can. J. Chem.* **58,** 631–653
37. Hindsgaul, O., Norberg, T., Lependu, J., and Lemieux, R. U. (1982) *Carbohydr. Res.* **109,** 109–142
38. Tanaka, M., Dube, V. E., and Anderson, B. (1984) *Biochim. Biophys. Acta* **798,** 283–290
39. Van Halbeek, H., Dorland, L., Vliegenthart, J. F. G., Hull, W. E., Lamblin, G., Lhermitte, M., Boersma, A., and Roussel, P. (1982) *Eur. J. Biochem.* **127,** 7–20
40. Lamblin, G., Boersma, A., Klein, A., Roussel, P., Van Halbeek, H., and Vliegenthart, J. F. G. (1984) *J. Biol. Chem.* **259,** 9051–9058
41. Van Halbeek, H., Dorland, L., Vliegenthart, J. F. G., Kochetkov, N. K., Arbatsky, N. P., and Derevitskaya, V. A. (1982) *Eur. J. Biochem.* **127,** 21–29
42. Van Halbeek, H., Gerwig, G. J., Vliegenthart, J. F. G., Tsuda, R.,

Hara, M., Akiyama, K., and Schmidt, K. (1985) *Biochem. Biophys. Res. Commun.* **131,** 507–514
43. Van Halbeek, H. (1984) *Biochem. Soc. Trans.* **12,** 601–605
44. Nasir-Ud-Din, Jeanloz, R. W., Lamblin, G., Roussel, P., Van Halbeek, H., Mutsaers, J. H. G. M., and Vliegenthart, J. F. G. (1986) *J. Biol. Chem.* **261,** 1992–1997
45. Dua, V. K., Dube, V. E., Li, Y.-T., and Bush, C. A. (1985) *Glycoconjugate J.* **2,** 17–30
46. Yates, A. D., Feeney, J., Donald, A. S. R., and Watkins, W. M. (1984) *Carbohydr. Res.* **130,** 251–260
47. Wieruszeski, J. M., Michalski, J.-C., Montreuil, J., Strecker, G., Peter-Katalinic, J., Egge, H., Van Halbeek, H., Mutsaers, J. H. G. M., and Vliegenthart, J. F. G. (1987) *J. Biol. Chem.* **262,** 6650–6657
48. Breg, J., Van Halbeek, H., Vliegenthart, J. F. G., Lamblin, G., Houvenaghel, M.-C., and Roussel, P. (1987) *Eur. J. Biochem.* **168,** 57–68
49. Klein, A., Lamblin, G., Lhermitte, M., Roussel, P., Breg, J., Van Halbeek, H., and Vliegenthart, J. F. G. (1988) *Eur. J. Biochem.* **171,** 631–642
50. Breg, J., Van Halbeek, H., Vliegenthart, J. F. G., Klein, A., Lamblin, G., and Roussel, P. (1988) *Eur. J. Biochem.* **171,** 643–654
51. Clausen, H., Levery, S. B., McKibbin, J. M., and Hakomori, S. (1985) *Biochemistry* **24,** 3578–3586
52. Yan, Z. Y., Rao, B. N. N., and Bush, C. A. (1987) *J. Am. Chem. Soc.* **109,** 7663–7669
53. Karlsson, K.-A. (1976) in *Glycolipid Methodology* (Witting, L. A., eds), pp. 97–122, American Oil Chemists' Society, Champaign, IL
54. Karlsson, K.-A., Pascher, I., Samuelsson, B. E., Finne, J., Krusius, T., and Rauvala, H. (1978) *FEBS Lett.* **94,** 413–417
55. Lloyd, K. O., Larson, G., Strömberg, N., Thurin, J., and Karlsson, K.-A. (1983) *Immunogenetics* **17,** 537–541
56. Wesley, A., Mantle, M., Man, D., Qureshi, R., Forstner, G., and Forstner, J. (1985) *J. Biol. Chem.* **260,** 7955–7959
57. Finne, J., Krusius, T., Rauvala, H., and Järnefelt, J. (1980) *Blood Transfus. Immunohaematol.* **23,** 545–552
58. Finne, J., Leinonen, M., and Mäkelä, P. H. (1983) *Lancet* **ii,** 355–357
59. Finne, J., and Krusius, T. (1979) *Eur. J. Biochem.* **102,** 583–588
60. Krusius, T., and Finne, J. (1977) *Eur. J. Biochem.* **78,** 369–379
61. Van Beek, W. P., Smets, L. A., and Emmelot, P. (1973) *Cancer Res.* **33,** 2913–2922
62. Warren, L., Buck, C. A., and Tuszynski, G. P. (1978) *Biochim. Biophys. Acta* **516,** 97–127
63. Spooncer, E., Fukuda, M., Klock, J. C., Oates, J. E., and Dell, A. (1984) *J. Biol. Chem.* **259,** 4792–4801
64. Fukuda, M., Spooncer, E., Oates, J. E., Dell, A., and Klock, J. C. (1984) *J. Biol. Chem.* **259,** 10925–10935
65. Kobata, A. (1979) *Anal. Biochem.* **100,** 1–14
66. Fournet, B., Montreuil, J., Strecker, G., Dorland, L., Haverkamp, J., Vliegenthart, J. F. G., Binette, J. P., and Schmid, K. (1978) *Biochemistry* **17,** 5206–5214
67. Yamashita, K., Ueda, I., Kuroki, M., Matsuoka, Y., and Kobata, A. (1985) in *Proceedings of the VIIIth International Symposium on Glycoconjugates* (Davidson, E. A., Williams, J. C., and DiFerrante, N. M., eds) Vol. 2, pp. 441–442, Praeger, New York
68. Feizi, T. (1985) *Nature* **314,** 53–57
69. Fukuda, M., Dell, A., and Fukuda, M. N. (1984) *J. Biol. Chem.* **259,** 4782–4791
70. Breimer, M. E., Hansson, G. C., Karlsson, K.-A., Larson, G., Leffler, H., Pimlott, W., Samuelsson, B. E., Strömberg, N., Teneberg, S., and Thurin, J. (1983) *Int. J. Mass. Spectr. Ion Physics* **48,** 133–166
71. Nudelman, E., Levery, S. B., Kaizu, T., and Hakomori, S. (1986) *J. Biol. Chem.* **261,** 11247–11253
72. Carlson, D. M. (1968) *J. Biol. Chem.* **243,** 616–626
73. Lloyd, K. O., and Kabat, E. A. (1968) *Proc. Natl. Acad. Sci. U. S. A.* **61,** 1470–1477
74. Finne, J. (1980) *Eur. J. Biochem.* **104,** 181–189
75. Dodge, J. T., and Phillips, G. B. (1967) *J. Lipid Res.* **8,** 667–675
76. Köttgen, E., Volk, F., and Gerok, W. (1982) *Biochem. Biophys. Res. Commun.* **109,** 168–173
77. Bouhours, J. F., and Glickman, R. M. (1976) *Biochim. Biophys. Acta* **441,** 123–133
78. Torres-Pinedo, R., and Mahmood, A. (1984) *Biochem. Biophys. Res. Commun.* **125,** 546–553
79. Mirelman, D., ed (1986) *Microbial Lectins and Agglutinins*, John Wiley & Sons, New York
80. Leffler, H., Svanborg-Edén, C., Schoolnik, G., and Wadström, T. (1984) in *Attachment of Organisms to the Gut Mucosa* (Boedeker, E. C., ed) Vol. 2, pp. 177–187, CRC Press, Boca Raton
81. Murray, P. A., Levine, M. J., Tabak, L. A., and Reddy, M. S. (1982) *Biochem. Biophys. Res. Commun.* **106,** 390–396
82. Parkkinen, J., Rogers, G. N., Korhonen, T., Dahr, W., and Finne, J. (1986) *Infect. Immun.* **54,** 37–42
83. Loomes, L. M., Uemura, K., Childs, R. A., Paulson, J. C., Rogers, G. N., Scudder, P. R., Michalski, J.-C., Hounsell, E. F., Taylor-Robinson, D., and Feizi, T. (1984) *Nature* **307,** 560–563
84. Paulson, J. C. (1985) in *The Receptors* (Conn, P. M., ed) Vol. 2, pp. 131–219, Academic Press, Orlando