Eur. J. Biochem. *209*, 895–915 (1992)

The Asn-linked carbohydrate chains of human Tamm-Horsfall glycoprotein of one male

Novel sulfated and novel *N*-acetylgalactosamine-containing N-linked carbohydrate chains

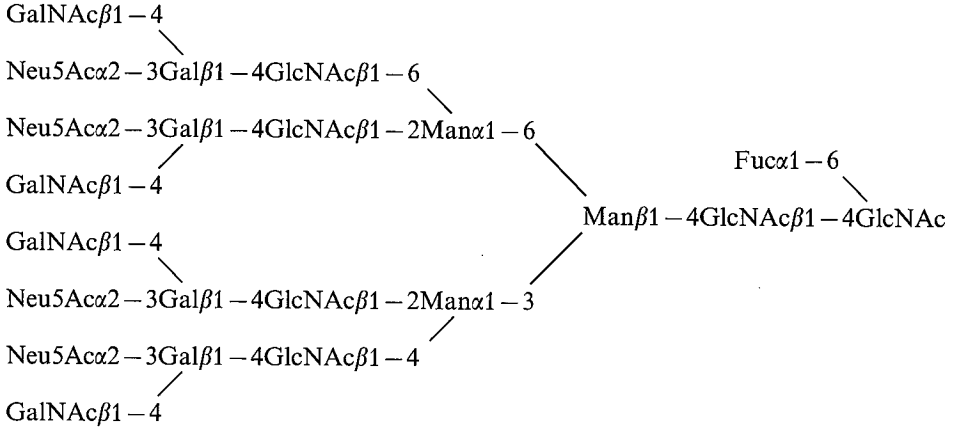
Karl HÅRD, Guus VAN ZADELHOFF, Peter MOONEN, Johannis P. KAMERLING and Johannes F. G. VLIEGENTHART
Bijvoet Center, Department of Bio-Organic Chemistry, Utrecht University, The Netherlands

(Received April 8, 1992) – EJB 92 0491

Human Tamm-Horsfall glycoprotein has been purified from the urine of one male. The Asn-linked carbohydrate chains were enzymically released by peptide-N^4-(*N*-acetyl-β-glucosaminyl)asparagine amidase F, and separated from the remaining protein by gel-permeation chromatography on Bio-Gel P-100. Fractionation of the intact (sulfated) sialylated carbohydrate chains was achieved by a combination of three liquid-chromatographic techniques, namely, anion-exchange FPLC on Q-Sepharose, amine-adsorption HPLC on Lichrospher-NH_2, and high-pH anion-exchange chromatography on CarboPac PA1. In total, more than 150 carbohydrate-containing fractions were obtained, some of which still contained mixtures of oligosaccharides. The primary structure of 30 N-glycans, including 10 novel oligosaccharides, were determined by one- and two-dimensional ^{1}H-NMR spectroscopy at 500 MHz or 600 MHz. The types of compounds identified range from non-fucosylated, monosialylated, diantennary to fucosylated, tetrasialylated, tetraantennary carbohydrate chains, possessing the following terminal structural elements:

Galβ1–4GlcNAcβ1–
Neu5Acα2–3Galβ1–4GlcNAcβ1–
Neu5Acα2–6Galβ1–4GlcNAcβ1–
(SO_4^-)–3Galβ1–4GlcNAcβ1–
(SO_4^-)–4GalNAcβ1–4GlcNAcβ1–
GalNAcβ1–4
\
Neu5Acα2–3Galβ1–4GlcNAcβ1–

The largest GalNAc-containing compound has the following structure:



Correspondence to J. F. G. Vliegenthart, Bijvoet Center, Department of Bio-Organic Chemistry, Utrecht University, P. O. Box 80.075, NL-3508 TB Utrecht, The Netherlands

Abbreviations. Fuc, L-fucose; HOHAHA, homonuclear Hartmann-Hahn; HPAEC, high-pH anion-exchange chromatography; MLEV, composite pulse devised by M. Levitt; NAc, *N*-acetyl; Neu5Ac, *N*-acetylneuraminic acid; NOESY, two-dimensional NOE spectroscopy; PAD, pulsed amperometric detection; PNGase-F, peptide-N^4-(*N*-acetyl-β-glucosaminyl)asparagine amidase F; T-H, Tamm-Horsfall; 1D, one dimensional; 2D, two dimensional.

Enzyme. Peptide-N^4-(*N*-acetyl-β-glucosaminyl)asparagine amidase F (EC 3.5.1.52).

Human Tamm-Horsfall (T-H) glycoprotein is the most abundant protein in normal human urine, being excreted in quantities up to about 100 mg/day (Tamm and Horsfall, 1950, 1952). The glycoprotein is produced by the kidney and has been localized in the thick ascending limb of the loop of Henle and the early, distal, convoluted, tubule segments of the nephron (Sikri et al., 1979). Interestingly, the T-H glycoprotein is a phosphatidylinositol-linked membrane protein (Rindler et al., 1990). There are several examples where the glycoprotein plays a role in pathological conditions. For instance, cell-mediated and humoral responses directed against

T-H glycoprotein have been found in patients with renal tubular acidosis and autoimmune liver disease (Tsantoulas et al., 1974). In patients suffering from hepatitis (Lindberg et al., 1983) or urinary tract infections (Maurier et al., 1978), autoantibodies directed against T-H glycoprotein have been demonstrated. The physiological function of T-H glycoprotein is still controversial, but one possible function is to prevent infection in the urinary bladder and urinary tract, because T-H glycoprotein can inhibit the binding of *Escherichia coli* S fimbriae to epithelial cells (Parkkinen et al., 1988). In recent years, the interest in T-H glycoprotein has also increased due to the discovery that uromodulin, described to be an unique 85-kDa immunosuppressive glycoprotein isolated from the urine of pregnant women, is at least in its protein part, identical to T-H glycoprotein (Muchmore and Decker, 1985; Pennica et al., 1987). The N-linked oligosaccharides of uromodulin were proposed to relate to the immunosuppressive properties by interaction with cytokines like interleukin 1 (Muchmore et al., 1987), although later data could not demonstrate any binding of uromodulin to native cytokines or any difference between uromodulin and T-H glycoprotein in several different assays (Moonen et al., 1988). A remarkable feature of uromodulin/T-H glycoprotein is its strong binding to denatured protein (e.g. tumor necrosis factor) and the absence of binding to native protein (Moonen et al., 1988). The possible physiological importance of the carbohydrate part prompted us to carry out a detailed structural characterization of the carbohydrate chains of human T-H glycoprotein.

The carbohydrate content of T-H glycoprotein is about 25–30%, comprising at least five glycosylated asparagine residues (Afonso et al., 1981). The first studies on the carbohydrate part of this easily aggregating glycoprotein were reported in 1952 (Gottschalk, 1952; Odin, 1952). It has been shown that the Sd^a blood-group determinant, present in over 90% of Caucasian erythrocytes, is also associated with the carbohydrate part of T-H glycoprotein (Soh et al., 1980; Donald et al., 1982), and involves a terminal Neu5Acα2–3(GalNAcβ1–4)Galβ1–4GlcNAcβ1– element (Donald et al., 1983). Furthermore, the structure has been determined of a Sd^a active pentasaccharide cleaved from T-H glycoprotein with endo-β-galactosidase (Donald et al., 1983; Donald and Feeney, 1986). Previous studies on T-H-glycopeptide mixtures have revealed that the Sd^a determinant resides in the non-reducing termini of branched N-glycans (Williams et al., 1984). In addition to sialylated compounds, small amounts of oligomannose-type carbohydrates have been reported to occur in the glycoprotein (Afonso et al., 1981; Dall'Olio et al., 1988). Recently, a pregnancy-associated decrease in the total content of oligomannose chains has been found for both human and bovine T-H glycoprotein (Smagula et al., 1990). However, due to the extreme complexity of the carbohydrates, no data are available about the structure of individual, intact, acidic, carbohydrate chains of T-H glycoprotein.

The present study describes the enzymic release by peptide-N^4-(*N*-acetyl-β-glucosaminyl)asparagine amidase F (PNGase-F), and the liquid-chromatographic fractionation, of more than 150 asparagine-linked carbohydrate chains of human T-H glycoprotein obtained from the urine of one male donor. The primary structures of 30 oligosaccharides, including 10 novel compounds, have been determined.

EXPERIMENTAL PROCEDURES

Materials

Human T-H glycoprotein was purified as described (Moonen et al., 1988) from approximately 100 l urine obtained from one male. The purity of the protein was assessed by several criteria: SDS/PAGE indicated homogenous protein with Coomassie-blue staining; concanavalin-A blotting suggested a purity of more than 99%; Edman degradation showed the absence of amino acids other than from T-H glycoprotein; SDS/PAGE followed by transfer to a nitrocellulose filter and blotting with the lectin *Dolichus biflorus* gave rise to one band coinciding with T-H glycoprotein, thereby proving that all GalNAc-containing oligosaccharides originated solely from T-H glycoprotein. PNGase-F from *Flavobacterium meningosepticum* was obtained from Boehringer Mannheim. ^2H_2O (99.8 and 99.96 atom % ^2H) was purchased from MSD Isotopes. Bio-Gel P-2 and P-100 (200–400 μm mesh) were from Bio-Rad, *n*-octyl α-D-glucopyranoside was from Sigma and acetonitrile (grade HPLC S) from Rathburn.

Liberation of the carbohydrate chains

The N-linked carbohydrate chains were enzymically released from the T-H glycoprotein according to a modified version of a previously described protocol (Hård et al., 1990). Three batches of approximately 100 mg glycoprotein in 10 ml 0.2 M urea were each worked up as follows. After the addition of 1 ml 10% (by vol.) 2-mercaptoethanol and 2 ml 10% (mass/vol.) SDS, the solution was boiled for 3 min and cooled to room temperature. 10 ml 7.1% (mass/vol.) *n*-octyl α-D-glucopyranoside in 50 mM Tris, pH 7.3, containing 50 mM EDTA, was then added, and the solution was incubated with 24 U PNGase-F (120 μl) for 23 h at room temperature. Subsequently, the mixture was boiled for 3 min, another batch of 24 U PNGase-F was added, and the incubation was continued for 6 h at room temperature. After lyophilization and suspension of the residue in water, the mixture was centrifuged and the supernatant collected. Resuspension of the pellet, sonication (2 × 2 min), and centrifugation were repeated three times. The pooled supernatants, containing released carbohydrate chains and soluble T-H glycoprotein, were lyophilized, and applied to a column (3 cm × 62 cm) of Bio-Gel P-100, eluted with 25 mM NH_4HCO_3, pH 8, at a flow rate of 15 ml/h. Aliquots of the 2.5-ml fractions collected were stained for carbohydrate with orcinol/H_2SO_4 and for protein with a protein assay reagent (Pierce). Fractions eluting after the void-volume peak were pooled, lyophilized, and desalted three times on a column (2.5 cm × 45 cm) of Bio-Gel P-2 using water as eluent. Oligosaccharides were detected as a void-volume peak by monitoring the effluent at 206 nm.

FPLC fractionation

Fractionation according to charge of the enzymically released carbohydrate chains was carried out on a Q-Sepharose anion-exchange column (1.8 cm × 10 cm, Pharmacia FPLC system), at a flow rate of 4 ml/min. The column was first eluted with a linear concentration gradient of 0–100 mM NaCl in 150 ml H_2O, followed by linear gradients of 100–150 mM NaCl in 45 ml H_2O, 150–250 mM NaCl in 30 ml H_2O, and finally by 250–500 mM NaCl in 45 ml H_2O. The fractionation was monitored at 214 nm, and carbohydrate-containing fractions (orcinol/H_2SO_4) were lyophilized, desalted on Bio-Gel P-2 and lyophilized again.

HPLC fractionation

The carbohydrate-containing FPLC fractions were subfractionated on a Kratos SF 400 HPLC system (ABI Analytical, Kratos Division) using either a 5-μm Lichrospher-NH_2 column (0.46 cm × 25 cm, Merck) at a flow rate of 1.5 ml/min or a 10-μm Lichrospher-NH_2 column (1.0 cm × 25 cm, Merck) at a flow rate of 4 ml/min. The columns were eluted with linear concentration gradients of either 30 mM potassium phosphate, pH 7.0, and acetonitrile, or 15 mM potassium phosphate, pH 4.7, and acetonitrile (see Fig. 3 for further details). The effluent was monitored at 206 nm, and collected fractions were lyophilized and desalted on Bio-Gel P-2, using water as eluent and lyophilized again.

High-pH anion-exchange chromatography (HPAEC)

Further fractionation of some of the HPLC fractions was carried out by HPAEC using pulsed-amperometric detection (PAD; Townsend and Hardy, 1991) on a Dionex LC system consisting of a Dionex Bio LC quaternary gradient module, a PAD 2 detector, and a CarboPac PA1 pellicular anion-exchange column (0.9 cm × 25 cm, Dionex). Elutions were carried out at a flow rate of 4 ml/min. Detection was performed with a gold electrode and triple-pulse amperometry, comprising the following pulse potentials and durations: E_1 = 0.05 V, 300 ms; E_2 = 0.65 V, 60 ms; E_3 = −0.95 V, 180 ms. For further details concerning solvent systems, see the legend to Fig. 4. Fractions were immediately neutralized by addition of 1 M HCl, and lyophilized. Salts were removed by gel-permeation chromatography on a column (1.5 cm × 45 cm) of Bio-Gel P-2, using 5 mM NH_4HCO_3 as eluent, whereafter the carbohydrate-containing fractions were lyophilized. It should be noted that separation of acetate from acidic carbohydrates is facilitated by using 5 mM NH_4HCO_3 as eluent instead of using water.

^{1}H-NMR spectroscopy

Prior to ^{1}H-NMR spectroscopic analysis the desalted samples were exchanged twice in 99.8% ^2H_2O. Finally, samples were dissolved in 99.96% ^2H_2O and transferred to Wilmad 528 PP or 535 PP NMR tubes. ^{1}H-NMR spectra at 300 MHz were recorded on a Bruker AC-300 spectrometer (Department of Organic Chemistry, Utrecht University), spectra at 500 MHz were recorded on a Bruker AM-500 or on an AMX-500 spectrometer (Bijvoet Center, Department of NMR Spectroscopy, Utrecht University), and spectra at 600 MHz were recorded on a Bruker AM-600 spectrometer (SON hf-NMR facility, Department of Biophysical Chemistry, University of Nijmegen, The Netherlands). The probe temperature was 27°C, unless indicated otherwise. Chemical shifts (ppm) are expressed by reference to internal acetone (δ = 2.225 in ^2H_2O at 27°C) (Vliegenthart et al., 1983). Typically, one-dimensional (1D) spectra at 500 MHz were recorded with a spectral width of 5 kHz collecting between 16 and 2000 free-induction decays of 8−32 k complex data points. The carrier frequency was positioned on the residual water signal. In Fig. 1 the pulse sequence used for recording 1D spectra is depicted. The pulse sequence is based on the previously described WEFT sequence (water eliminated Fourier transform; Patt and Sykes, 1972), but uses a selective pulse for the inversion of the HO^{2}H signal, followed by a homospoil pulse and a composite detection pulse. Resolution enhancement of 1D spectra was achieved by Lorentzian-to-Gaussian transformation. When necessary, the final spectrum was base-line corrected with a polynomal function. For the two-dimensional (2D) homonuclear Hartmann-Hahn (HOHAHA) spectra (Bax and Davis, 1985), a MLEV-17 (MLEV, composite pulse devised by M. Levitt) mixing sequence of 40−100 ms was used. The 90° pulse width was adjusted to about 25 μs, and the spectral width was 4 kHz in both dimensions. The HO^{2}H signal was presaturated for 1 s during the relaxation delay. In total, 368−512 spectra of 2048 data points were recorded, with 16−80 scans/t_1 value. For some samples, a 2D clean-total-correlation-spectroscopy (Griesinger et al., 1988) pulse sequence was used for tracing homonuclear, scalar coupling networks. In these experiments the 90° pulse and the delay, respectively, in the composite pulse used for magnetisation transfer were adjusted to about 15 μs. 2D NOE spectroscopy (NOESY; Jeener et al., 1979) was performed with a 250-ms mixing time. The time-proportional phase increment method was used to obtain phase-sensitive spectra in the t_1 direction (Marion and Wüthrich, 1983). The 2D NMR data were processed on a VAX station 3100 using the TRITON software developed at the Bijvoet Center, Department of NMR Spectroscopy, Utrecht University. The time-domain data were multiplied with a phase-shifted (squared) sine bell. After Fourier transformation, the resulting data set of 1024 × 1024 points was base-line corrected in both frequency domains by a third- or fourth-order polynomal fit.

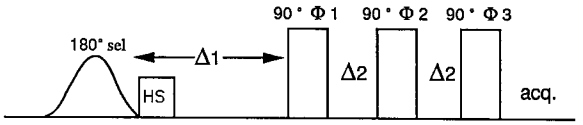


Fig. 1. Pulse sequence used for water suppression and hump elimination in 1D ^{1}H-NMR experiments in ^2H_2O. The 180° pulse is applied selectively on the HO^{2}H resonance. The best results have been obtained with a hyperbolic secant-shaped pulse. HS, 50-ms homospoil pulse used to eliminate residual transverse magnetization; Δ_1, delay adjusted so that the longitudinal z-magnetization of the HO^{2}H signal is zero prior to the detection pulse train; Δ_2, 10-μs delay for phase switching. The phase cycling used is as follows: Φ_1 = 2(*x*), 4(-*x*), 2(*x*), 2(-*y*), 4(*y*), 2(-*y*); Φ_2 = -*x*, 2(*x*), -*x*, *x*, 2(-*x*), *x*, *y*, 2(-*y*), *y*, -*y*, 2(*y*), -*y*; Φ_3 = *x*, -*x*, *x*, 2(-*x*), *x*, -*x*, *x*, -*y*, *y*, -*y*, 2(*y*), -*y*, *y*, -*y*; acquisition (acq.) = 4(*x*), 4(-*x*), 4(-*y*), 4(*y*).

Monosaccharide analysis

Monosaccharide analysis was carried out by gas/liquid chromatography on a capillary SE-30 fused silica column (0.32 mm × 25 m, Pierce) using a Varian Aerograph 3700 gas chromatograph. The trimethylsilylated (methyl ester) methyl glycosides were prepared by methanolysis, *N*-(re)acetylation and trimethylsilylation (Kamerling and Vliegenthart, 1989).

RESULTS

Monosaccharide analysis of human T-H glycoprotein, with a total carbohydrate content of 25% (by mass), revealed the presence of L-fucose (Fuc)/Gal/GalNAc/GlcNAc/Man/*N*-acetylneuraminic acid (Neu5Ac) in the molar ratio 0.6:3.0:0.7:5.7:3.0:3.2. The N-linked carbohydrate chains were split off enzymically with PNGase-F, and the released oligosaccharides were separated from the protein by gel-permeation chromatography on Bio-Gel P-100 (not shown).

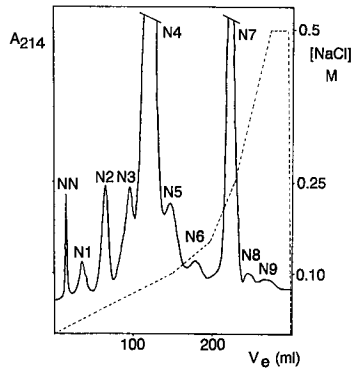


Fig. 2. Fractionation pattern at 214 nm on an FPLC Q-Sepharose anion-exchange column of the carbohydrate-containing Bio-Gel P-100 fraction, derived from PNGase-F-treated human T-H glycoprotein. The column was first eluted with a linear concentration gradient of 0–100 mM NaCl in 150 ml H_2O, followed by linear gradients of 100–150 mM NaCl in 45 ml H_2O, 150–250 mM NaCl in 30 ml H_2O and, finally, by 250–500 mM NaCl in 45 ml H_2O, at a flow rate of 4 ml/min.

Monosaccharide analysis of the Bio-Gel P-100 void-volume peak showed the presence of about 4% (by mass) carbohydrate on the remaining protein, indicating that 88% (by mass) of the carbohydrates had been cleaved by PNGase-F. The identity of the carbohydrate chains not released by PNGase-F is still under investigation. The desalted carbohydrate-containing material, eluting after the void-volume peak in the Bio-Gel P-100 separation, was fractionated by FPLC on a Q-Sepharose anion-exchange column. This separation yielded ten carbohydrate-positive fractions, denoted **NN** and **N1–N9** (Fig. 2). It should be noted that the high intensity of **N7** mainly results from non-carbohydrate material. ^{1}H-NMR spectroscopy of FPLC fraction **NN** showed that it contained a small amount of a heterogeneous mixture of sialylated oligosaccharides, as well as traces of neutral carbohydrates, including oligomannose-type carbohydrate chains. Due to the low amount of material, Q-Sepharose fractions **NN**, **N8** and **N9** were not investigated further. Fractions **N1–N5** were subfractionated by HPLC on a Lichrospher-NH_2 column (Fig. 3), and the major carbohydrate structures identified in these fractions will be discussed. Structure elucidation has not yet been carried out for the highly charged FPLC fractions **N6** and **N7**, due to the high complexity of the structures of the carbohydrate chains in these multicomponent mixtures. Relevant ^{1}H chemical shifts of the major carbohydrates of T-H glycoprotein are compiled in Tables 1–4, and the structures together with their relative amounts are shown in Tables 5–7.

In general, the carbohydrate chains have either an α1–6-fucosylated *N*,*N*′-diacetylchitobiose or a non-fucosylated *N*,*N*′-diacetylchitobiose element at their reducing ends. The fucosylated unit is recognized from the anomeric signals of αGlcNAc-**1** at δ = 5.18, GlcNAc-**2** at δ = 4.66–4.67 and Fuc at δ = 4.89–4.91, together with the *N*-acetyl (NAc) methyl signals of GlcNAc-**1** at δ = 2.038–2.039 and of GlcNAc-**2** at δ = 2.09–2.10, as well as from the Fuc CH_3 signals at δ = 1.21 (αGlcNAc-**1**) and δ = 1.22 (βGlcNAc-**1**), and the Fuc H-5 multiplets at δ = 4.10 (αGlcNAc-**1**) and δ = 4.13 (βGlcNAc-**1**; de Waard et al., 1991). The non-fucosylated *N*,*N*′-diacetylchitobiose unit is recognized from the anomeric

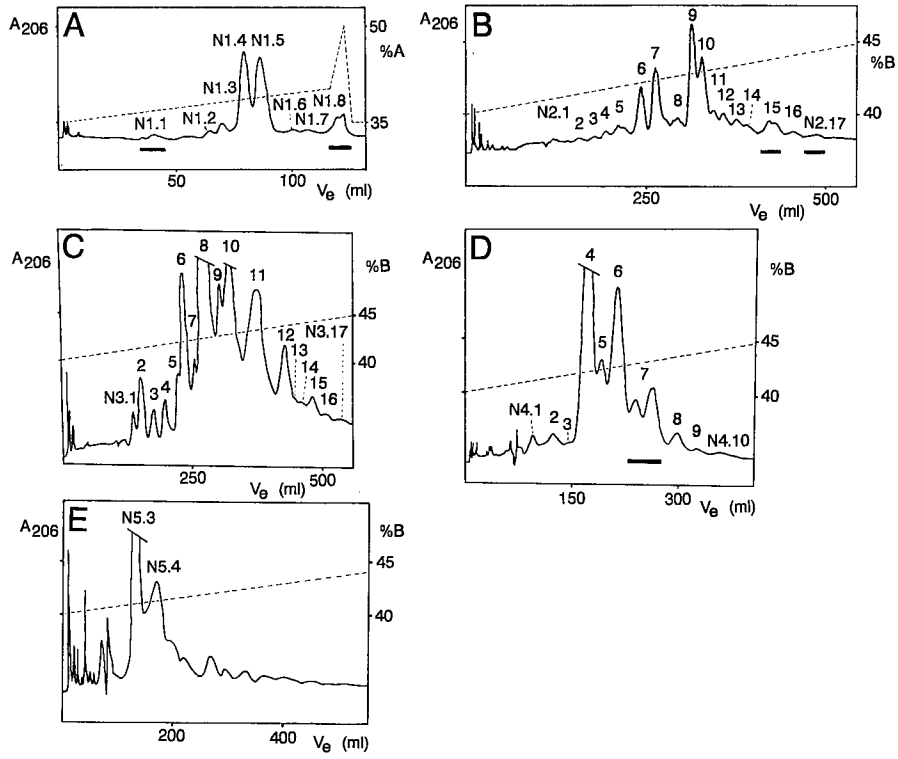
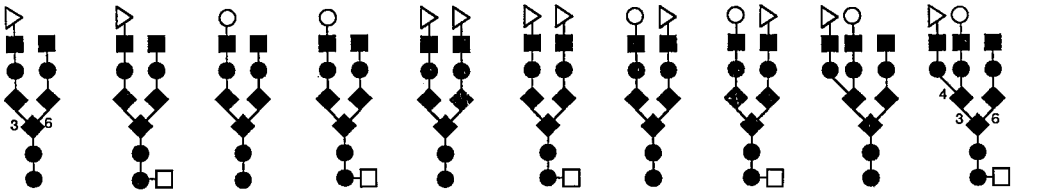


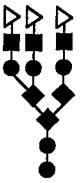
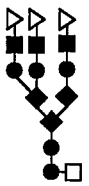
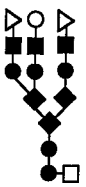
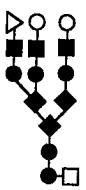
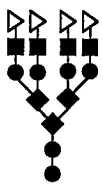
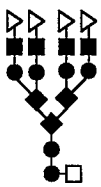
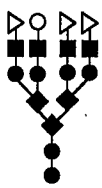
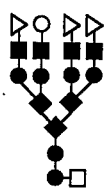
Fig. 3. Fractionation patterns at 206 nm on an HPLC Lichrospher-NH_2 column of FPLC oligosaccharide fractions N1–N5 obtained from human T-H glycoprotein. (A) 5-μm Lichrospher-NH_2 (0.46 cm × 25 cm); (B–E) 10-μm Lichrospher-NH_2 (1.0 cm × 25 cm); solvent A, 15 mM potassium phosphate, pH 4.7; solvent B, 30 mM potassium phosphate, pH 7.0; Elutions were performed with mixtures of acetonitrile and either solvent A or B, as indicated; A, FPLC fraction **N1**; B, fraction **N2**; C, fraction **N3**; D, fraction **N4**; E, fraction **N5**.

Table 1. ^{1}H chemical shifts of structural-reporter-group protons of the constituent monosaccharides of monosialylated and disialylated, non-sulfated, N-linked oligosaccharides, lacking GalNAc, derived from human T-H glycoprotein. Chemical shifts are given at 27 °C and were measured in ^2H_2O relative to internal acetone (δ = 2.225; Vliegenthart et al., 1983). Compounds are represented by short-hand symbolic notation: □, Fuc; ■, Gal; ●, GlcNAc; ◆, Man; △, Neu5Acα2−3; ○, Neu5Acα2−6. For numbering of the monosaccharide residues, see Table 5. n.d., not determined. α and β stand for the anomeric configuration of GlcNAc-**1**.

Reporter group	Residue	Chemical shift in									
		N1.2	**N1.3**	**N1.4**	**N1.5**	**N2.4A**	**N2.5**	**N2.6**	**N2.7**	**N2.9**	**N2.10**
											
		ppm									
H-1	GlcNAc-**1**α	5.188	5.180	5.190	5.181	5.19	5.182	5.190	5.182	5.191	5.182
	β	n.d.	4.694	4.697	4.694	n.d.	4.693	4.698	4.693	4.695	4.692
	GlcNAc-**2**α	4.61[a]	4.67	4.615	4.665	4.613	4.67	4.614	4.664	4.613	4.665
	β	4.61	4.67	4.606	4.668	4.605	4.67	4.605	4.666	4.605	4.669
	Man-**3**	n.d.	n.d.	4.771	4.770	n.d.	n.d.	n.d.	4.769	n.d.	n.d.
	Man-**4**	5.119	5.119	5.137	5.136	5.118	5.117	5.134	5.134	5.130	5.130
	Man-**4′**	4.927	4.928	4.931	4.928	4.925	4.923	4.926	4.922	4.925	4.923
	GlcNAc-**5**	4.57	4.57	4.606	4.606	4.574	4.573	4.605	4.605	4.594	4.594
	GlcNAc-**5′**	4.582	4.582	4.583	4.583	4.574	4.573	4.575	4.574	4.581	4.581
	Gal-**6**	4.544	4.544	4.446	4.446	4.544	4.544	4.445	4.444	4.446	4.446
	Gal-**6′**	4.473	4.473	4.474	4.473	4.550	4.549	4.550	4.549	4.473	4.473
	GlcNAc-**7**	—	—	—	—	—	—	—	—	4.549[c]	4.549[c]
	Gal-**8**	—	—	—	—	—	—	—	—	4.545[c]	4.545[c]
	Fucα	—	4.890	—	4.890	—	4.893	—	4.893	—	4.890
	β	—	4.897	—	4.897	—	4.900	—	4.900	—	4.897
H-2	Man-**3**	4.248	4.250	4.253	4.254	4.248	4.250	4.251	4.253	4.215	4.215
	Man-**4**	4.191	4.193	4.197	4.196	4.194	4.190	4.196	4.196	4.215	4.215
	Man-**4′**	4.11	4.11	4.111	4.109	4.11	4.11	4.12	4.11	4.11	4.11
H-3	Man-**4**	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	4.045	4.044
	Gal-**6**	4.11	4.11	n.d.	n.d.	4.11	4.113	n.d.	n.d.	n.d.	n.d.
	Gal-**6′**	n.d.	n.d.	n.d.	n.d.	4.12	4.118	4.117	4.116	n.d.	n.d.
	Gal-**8**	—	—	—	—	—	—	—	—	4.115	4.115
H-3a	Neu5Ac	1.796	1.797	1.719	1.719	1.798	1.798	1.716	1.717	1.718	1.718
	Neu5Ac′	—	—	—	—	1.798	1.802	1.800	1.802	—	—
	Neu5Ac*	—	—	—	—	—	—	—	—	1.800	1.800
H-3e	Neu5Ac	2.757	2.757	2.668	2.668	2.757	2.759	2.668	2.668	2.669	2.669
	Neu5Ac′	—	—	—	—	2.757	2.759	2.758	2.758	—	—
	Neu5Ac*	—	—	—	—	—	—	—	—	2.757	2.757
H-5	Fucα	—	4.094	—	4.097	—	4.098	—	4.097	—	4.097
	β	—	4.135	—	4.134	—	4.134	—	4.133	—	4.133
NAc	GlcNAc-**1**	2.038	2.039	2.039	2.039	2.040	2.039	2.039	2.039	2.038	2.038
	GlcNAc-**2**α	2.081	2.096	2.082	2.096	2.081	2.096	2.081	2.096	2.081	2.095
	β	2.081	2.093	2.082	2.093	2.081	2.093	2.081	2.093	2.081	2.092
	GlcNAc-**5**	2.047	2.048	2.069	2.069	2.048	2.048	2.068	2.068	2.066	2.066
	GlcNAc-**5′**	2.047	2.048	2.048	2.048	2.042	2.044	2.043	2.044	2.047	2.047
	GlcNAc-**7**	—	—	—	—	—	—	—	—	2.073	2.074
	Neu5Ac	2.030	2.031	2.030	2.030	2.031[b]	2.032[b]	2.031[b]	2.031[b]	2.031[b]	2.031[b]
CH_3	Fucα	—	1.210	—	1.209	—	1.211	—	1.211	—	1.210
	β	—	1.220	—	1.219	—	1.221	—	1.221	—	1.220

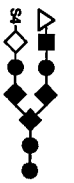
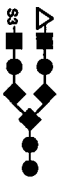
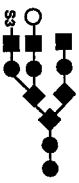
[a] Some values are given with only two decimals because of spectral overlap or virtual couplings.
[b] Signal belonging to two Neu5Ac residues.
[c] Assignments may have to be interchanged.

Table 2. ^1H chemical shifts of structural-reporter-group protons of the constituent monosaccharides of trisialylated, and tetrasialylated non-sulfated, N-linked oligosaccharides, lacking GalNAc, derived from human T-H glycoprotein. Chemical shifts are given at 27 °C and were measured in ^2H_2O relative to internal acetone (δ = 2.225; Vliegenthart et al., 1983). For the short-hand symbolic notation, see Table 1. For numbering of the monosaccharide residues, see Table 5. n.d., not determined. α and β stand for the anomeric configuration of GlcNAc-1.

Reporter group	Residue	Chemical shift in							
		N3.5	**N3.6.1**	**N3.8**	**N3.10.3**	**N4.4.6**	**N4.4.5**	**N4.6.5**	**N4.6.3**
									
		ppm							
H-1	GlcNAc-1α	5.190	5.182	5.183	5.181	5.191	5.181	5.192	5.183
	β	4.695	4.690	4.691	n.d.	4.692	4.689	4.691	4.688
	GlcNAc-2α	4.610	4.663	4.664	4.67[a]	4.61	4.658	4.60	4.659
	β	n.d.	4.666	4.668	4.67	4.61	4.662	4.60	4.663
	Man-3	n.d.	n.d.	4.763	n.d.	n.d.	4.752	n.d.	n.d.
	Man-4	5.116	5.114	5.130	5.128	5.130	5.130	5.140	5.138
	Man-4′	4.910	4.904	4.907	4.933	4.859	4.857	4.859	4.858
	GlcNAc-5	4.563	4.559	4.593	4.591	4.563	4.564	4.594	4.594
	GlcNAc-5′	4.572	4.570	4.572	4.600	4.593	4.593	4.594	4.594
	Gal-6	4.541[b]	4.541[b]	4.444	4.443	4.542[b]	4.542[b]	4.439	4.439
	Gal-6′	4.546[b]	4.546[b]	4.546	4.446	4.547[b]	4.545[b]	4.547	4.546
	GlcNAc-7	4.550[b]	4.548[b]	4.546	4.546	4.542[b]	4.542[b]	4.547	4.546
	GlcNAc-7′	—	—	—	—	4.547[b]	4.545[b]	4.547	4.546
	Gal-8	4.541[b]	4.541[b]	4.546	4.546	4.542[b]	4.542[b]	4.547	4.546
	Gal-8′	—	—	—	—	4.558	4.559	4.558	4.558
	Fucα	—	4.892	4.893	4.886	—	4.900	—	4.901
	β	—	4.900	4.901	4.902	—	4.908	—	4.909
H-2	Man-3	4.211	4.211	4.215	4.22	4.204	4.204	4.21	4.21
	Man-4	4.211	4.211	4.215	4.22	4.220	4.220	4.22	4.22
	Man-4′	4.11	4.11	4.11	4.11	4.095	4.090	4.09	4.091
H-3	Man-4	4.042	4.040	4.041	4.042	4.048	4.044	4.049	4.048
	Gal-6	4.116	4.116	n.d.	n.d.	4.118	4.117	n.d.	n.d.
	Gal-6′	4.116	4.116	4.117	n.d.	4.118	4.117	4.119	4.118
	Gal-8	4.116	4.116	4.117	4.117	4.118	4.117	4.119	4.118
	Gal-8′	—	—	—	—	4.118	4.117	4.119	4.118
H-3a	Neu5Ac	1.801	1.802	1.719	1.719	1.804	1.803	1.720	1.719
	Neu5Ac′	1.801	1.802	1.802	1.719	1.804	1.803	1.804	1.804
	Neu5Ac*	1.801	1.802	1.802	1.802	1.804	1.803	1.804	1.804
	Neu5Ac^	—	—	—	—	1.804	1.803	1.804	1.804
H-3e	Neu5Ac	2.756	2.756	2.668	2.669	2.756	2.758	2.669	2.669
	Neu5Ac′	2.756	2.756	2.757	2.672	2.756	2.758	2.756	2.756
	Neu5Ac*	2.756	2.756	2.757	2.756	2.756	2.758	2.756	2.756
	Neu5Ac^	—	—	—	—	2.756	2.758	2.756	2.756
H-5	Fucα	—	4.095	4.096	4.097	—	4.10	—	4.10
	β	—	4.134	4.134	4.136	—	4.137	—	4.138
NAc	GlcNAc-1	2.039	2.039	2.039	2.038	2.038	2.038	2.039	2.039
	GlcNAc-2α	2.081	2.096	2.096	2.098[b]	2.080	2.094	2.080[b]	2.094
	β	2.081	2.093	2.093	2.094[b]	2.080	2.090	2.078[b]	2.090
	GlcNAc-5	2.044	2.044	2.066	2.066	2.048	2.047	2.069	2.069
	GlcNAc-5′	2.044	2.044	2.044	2.066	2.038	2.038	2.039	2.039
	GlcNAc-7	2.072	2.073	2.074	2.075	2.074	2.074	2.075	2.076
	GlcNAc-7′	—	—	—	—	2.038	2.038	2.039	2.039
	Neu5Ac	2.031[c]	2.031[c]	2.031[c]	2.030[c]	2.031[d]	2.031[d]	2.031[d]	2.031[d]
CH_3	Fucα	—	1.211	1.212	1.210	—	1.211	—	1.212
	β	—	1.222	1.222	1.222	—	1.221	—	1.222

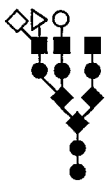
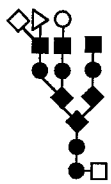
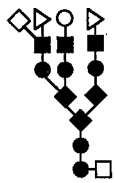
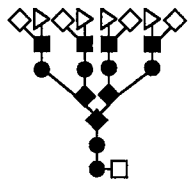
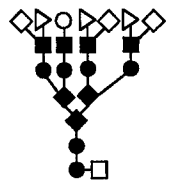
[a] Some values are given with only two decimals because of spectral overlap or virtual couplings.
[b] Assignments may have to be interchanged.
[c] Signal belonging to three Neu5Ac residues.
[d] Signal belonging to four Neu5Ac residues.

Table 3. ^{1}H chemical shifts of structural-reporter-group protons of the constituent monosaccharides of sulfated, sialylated, N-linked oligosaccharides derived from human T-H glycoprotein. Chemical shifts are given at 27 °C and were measured in ^2H_2O relative to internal acetone (δ = 2.225; Vliegenthart et al., 1983). For the short-hand symbolic notation, see Table 1 (◇, GalNAc). For numbering of the monosaccharide residues, see Table 5. n.d., not determined. α and β stand for the anomeric configuration of GlcNAc-**1**.

Reporter group	Residue	Chemical shift in						
		N2.4B	**N3.2.1**	**N3.2.2A**	**N3.2.2B**	**N3.2.3**	**N3.2.4**	**N3.6.5**
								
		ppm						
H-1	GlcNAc-1α	5.18	5.181	5.180	5.180	5.190	5.190	5.190
	β	n.d.	4.693	n.d.	n.d.	n.d.	n.d.	4.694
	GlcNAc-2α	4.67[a]	4.663	4.664	4.664	4.612	4.613	4.613
	β	4.67	4.667	4.667	4.667	n.d.	4.604	4.604
	Man-**4**	5.134	5.109	5.118	5.118	5.110	5.120	5.130
	Man-**4′**	4.925	4.923	4.922	4.922	4.927	4.924	4.925
	GlcNAc-**5**	4.605	4.557	4.583	4.572	4.557	4.582	4.596
	GlcNAc-**5′**	4.579	4.572	4.572	4.579	4.572	4.572	4.581
	Gal-**6**	4.444	–	4.583	4.544	–	4.582	4.445
	Gal-**6′**	4.589	4.549	4.549	4.589	4.549	4.549	4.473
	GlcNAc-**7**	–	–	–	–	–	–	4.554
	Gal-**8**	–	–	–	–	–	–	4.584
	GalNAc	–	4.585	–	–	4.586	–	–
	Fucα	4.891	4.893	4.892	4.892	–	–	–
	β	4.898	4.900	4.900	4.900	–	–	–
H-2	Man-**3**	4.25	4.245	4.248	4.248	4.243	4.246	4.22
	Man-**4**	4.20	4.180	4.191	4.191	4.178	4.192	4.22
	Man-**4′**	4.11	4.11	4.11	4.11	4.11	4.110	4.110
H-3	Man-**4**	n.d.	n.d.	n.d.	n.d.	n.d.	n.d.	4.047
	Gal-**6**	n.d.	–	4.335	4.11	–	4.335	n.d.
	Gal-**6′**	4.341	4.117	4.116	4.441	4.117	n.d.	n.d.
	Gal-**8**	–	–	–	–	–	–	4.335
H-3a	Neu5Ac	1.717	–	–	1.797	–	–	1.719
	Neu5Ac′	–	1.803	1.801	–	1.800	1.800	–
H-3e	Neu5Ac	2.667	–	–	2.758	–	–	2.668
	Neu5Ac′	–	2.759	2.758	–	2.759	2.757	–
H-4	Gal-**6**	n.d.	–	4.291	n.d.	–	4.291	n.d.
	Gal-**6′**	4.296	n.d.	n.d.	4.296	n.d.	n.d.	–
	Gal-**8**	–	–	–	–	–	–	4.293
	GalNAc	–	4.692	–	–	4.692	–	–
H-5	Fucα	4.095	4.095	4.094	4.094	–	–	–
	β	n.d.	4.135	n.d.	n.d.	–	–	–
NAc	GlcNAc-**1**	2.040	2.040	2.039	2.039	2.039	2.039	2.039
	GlcNAc-2α	2.095	2.096	2.096	2.096	2.081	2.081	2.081
	β	2.095	2.093	2.093	2.093	2.081	2.081	2.081
	GlcNAc-**5**	2.069	2.044	2.051	2.047	2.044	2.051	2.067
	GlcNAc-**5′**	2.048	2.044	2.043	2.047	2.044	2.042	2.047
	GlcNAc-**7**	–	–	–	–	–	–	2.077
	GalNAc	–	2.069	–	–	2.069	–	–
	Neu5Ac	2.031	2.032	2.032	2.032	2.032	2.031	2.030
CH_3	Fucα	1.210	1.211	1.211	1.211	–	–	–
	β	1.220	1.222	1.221	1.221	–	–	–

[a] Some values are given with only two decimals because of spectral overlap or virtual couplings.

Table 4. ^{1}H chemical shifts of structural-reporter-group protons of the constituent monosaccharides of GalNAc-containing, sialylated, N-linked oligosaccharides, derived from human T-H glycoprotein. Chemical shifts are given at 27°C and were measured in ^2H_2O relative to internal acetone (δ = 2.225; Vliegenthart et al., 1983). For the short-hand symbolic notation, see Table 1 (◇, GalNAc). For numbering of the monosaccharide residues, see Table 5. α and β stand for the anomeric configuration of GlcNAc-**1**. n.d., not determined.

Reporter group	Residue	Chemical shift in				
		N2.11.2 	**N2.12** 	**N3.9.2** 	**N4.7.2** 	**N4.9.2** 
		ppm				
H-1	GlcNAc-**1**α	5.190	5.183	5.183	5.182	5.181
	β	4.694	4.692	n.d.	4.685	4.686
	GlcNAc-**2**α	4.613	4.666	4.664	4.668	4.668
	β	4.604	4.669	4.668	4.671	4.672
	Man-**3**	4.751[a]	4.755[a]	4.758[a]	4.75[b]	4.75[a]
	Man-**4**	5.130	5.130	5.130	5.130	5.140
	Man-**4′**	4.924	4.923	4.905	4.861	4.860
	GlcNAc-**5**	4.596	4.595	4.593	4.56[c]	4.59
	GlcNAc-**5′**	4.580	4.581	4.572	4.60[c]	4.60
	Gal-**6**	4.445	4.446	4.443	4.537[d]	4.438
	Gal-**6′**	4.473	4.473	4.549[d]	4.543[d]	4.54
	GlcNAc-**7**	4.544[d]	4.545[d]	4.542[d]	4.537[d]	4.54
	GlcNAc-**7′**	—	—	—	4.543[d]	4.54
	Gal-**8**	4.541[d]	4.541[d]	4.542[d]	4.537[d]	4.54
	Gal-**8′**	—	—	—	4.556	4.557
	GalNAc	4.758[a]	4.754[a]	4.756[a]	4.737[g]/4.743[g]	4.750/4.754/4.758
	Fucα	—	4.891	4.893	4.904	4.903
	β	—	4.898	4.901	4.911	4.912
H-2	Man-**3**	4.216	4.217	4.216	4.20[b]	4.21
	Man-**4**	4.216	4.217	4.216	4.22[c]	4.22
	Man-**4′**	4.11	4.11	4.11	4.08[c]	4.08
	Gal-**8**	3.357	3.357	3.357	3.356[e]	3.357[f]
H-3	Man-**4**	4.043	4.043	4.040	4.044	4.047
	Gal-**6**	n.d.	n.d.	n.d.	4.15[c]	n.d.
	Gal-**6′**	n.d.	n.d.	4.118	4.15[c]	4.15
	Gal-**8**	4.151	4.151	4.153	4.15[c]	4.15
	Gal-**8′**	—	—	—	4.15[c]	4.15
H-3a	Neu5Ac	1.720	1.719	1.720	1.930	1.720
	Neu5Ac′	—	—	1.803	1.930	1.93
	Neu5Ac*	1.927	1.926	1.931	1.930	1.93
	Neu5Ac^	—	—	—	1.930	1.93
H-3e	Neu5Ac	2.668	2.669	2.669	2.66	2.669
	Neu5Ac′	—	—	2.758	2.66	2.66
	Neu5Ac*	2.662	2.663	2.664	2.66	2.66
	Neu5Ac^	—	—	—	2.66	2.66
H-4	Gal-**8**	4.11	4.11	4.11	4.12[e]	4.12[i]
H-5	Fucα	—	4.099	4.096	4.095	n.d.
	β	—	4.134	n.d.	4.13[b]	n.d.
NAc	GlcNAc-**1**	2.038	2.039	2.039	2.04	2.038
	GlcNAc-**2**α	2.081	2.096	2.097	2.096	2.096
	β	2.081	2.093	2.094	2.093	2.093
	GlcNAc-**5**	2.066	2.066	2.066	2.049	2.069
	GlcNAc-**5′**	2.047	2.047	2.045	2.04	2.04
	GlcNAc-**7**	2.073	2.074	2.075	2.076	2.076
	GlcNAc-**7′**	—	—	—	2.04	2.04
	GalNAc	2.015	2.015	2.015	2.015[f]/2.018	2.015[g]/2.019
	Neu5Ac	2.031[g]	2.031[g]	2.032[f]	2.032[h]	2.032[h]
CH_3	Fucα	—	1.210	1.212	1.210	1.210
	β	—	1.221	1.222	1.221	1.222

[a] Measured at 42°C.
[b] Value obtained from 2D HOHAHA experiment at 37°C.
[c] Value obtained from 2D NOE experiment.
[d] Assignments may have to be interchanged.
[e] Also stemming from Gal-**6**, -**6′**, and -**8′**.
[f] Signal stemming from three residues.
[g] Signal stemming from two residues.
[h] Signal stemming from four residues.
[i] Also stemming from Gal-**6′** and -**8′**.

Table 5. Sialylated, non-sulfated carbohydrate chains, lacking GalNAc, obtained from human T-H glycoprotein. For a given pair of compounds the bold part indicates an invariant core structure. For example, **N1.2** is an α2−3-monosialylated, diantennary compound, whereas **N1.4** is the α2−6-monosialylated analogue. The numbering of the monosaccharide residues is illustrated for **N4.4.5** and **N4.6.3**.

Code	Amount	Structure
	mol/100 mol	
N1.2	<1	Neu5Acα2−3 | **Galβ1−4GlcNAcβ1−2Manα1−6** \ **Manβ1−4GlcNAcβ1−4GlcNAc** / Neu5Acα2−6 | **Galβ1−4GlcNAcβ1−2Manα1−3**
N1.4	<1	
N1.3	<1	Neu5Acα2−3 | **Galβ1−4GlcNAcβ1−2Manα1−6** **Fucα1−6** \ \ **Manβ1−4GlcNAcβ1−4GlcNAc** / Neu5Acα2−6 | **Galβ1−4GlcNAcβ1−2Manα1−3**
N1.5	<1	
N2.4A	<1	**Neu5Acα2−3Galβ1−4GlcNAcβ1−2Manα1−6** \ **Manβ1−4GlcNAcβ1−4GlcNAc** / Neu5Acα2−3 | Neu5Acα2−6 | **Galβ1−4GlcNAcβ1−2Manα1−3**
N2.6	<1	
N2.5	<1	**Neu5Acα2−3Galβ1−4GlcNAcβ1−2Manα1−6** **Fucα1−6** \ \ **Manβ1−4GlcNAcβ1−4GlcNAc** / Neu5Acα2−3 | Neu5Acα2−6 | **Galβ1−4GlcNAcβ1−2Manα1−3**
N2.7	<1	
N2.9	<1	**Galβ1−4GlcNAcβ1−2Manα1−6** \ **Manβ1−4GlcNAcβ1−4GlcNAc** / **Neu5Acα2−6Galβ1−4GlcNAcβ1−2Manα1−3** / **Neu5Acα2−3Galβ1−4GlcNAcβ1−4**
N2.10	<1	**Galβ1−4GlcNAcβ1−2Manα1−6** **Fucα1−6** \ \ **Manβ1−4GlcNAcβ1−4GlcNAc** / **Neu5Acα2−6Galβ1−4GlcNAcβ1−2Manα1−3** / **Neu5Acα2−3Galβ1−4GlcNAcβ1−4**
N3.5	<1	**Neu5Acα2−3Galβ1−4GlcNAcβ1−2Manα1−6** \ **Manβ1−4GlcNAcβ1−4GlcNAc** / **Neu5Acα2−3Galβ1−4GlcNAcβ1−2Manα1−3** / **Neu5Acα2−3Galβ1−4GlcNAcβ1−4**

Table 5. (continued)

Code	Amount	Structure
	mol/100 mol	
N3.6.1	<1	**Neu5Acα2–3Galβ1–4GlcNAcβ1–2Manα1–6** **Fucα1–6** **Manβ1–4GlcNAcβ1–4GlcNAc** **Neu5Acα2–3Galβ1–4GlcNAcβ1–2Manα1–3** **Neu5Acα2–3Galβ1–4GlcNAcβ1–4**
N3.8 **N3.10.3**	2 1	Neu5Acα2–3 Neu5Acα2–6 **Galβ1–4GlcNAcβ1–2Manα1–6** **Fucα1–6** **Manβ1–4GlcNAcβ1–4GlcNAc** **Neu5Acα2–6Galβ1–4GlcNAcβ1–2Manα1–3** **Neu5Acα2–3Galβ1–4GlcNAcβ1–4**
N4.4.6 **N4.6.5**	5 2	**Neu5Acα2–3Galβ1–4GlcNAcβ1–6** **Neu5Acα2–3Galβ1–4GlcNAcβ1–2Manα1–6** **Manβ1–4GlcNAcβ1–4GlcNAc** Neu5Acα2–3 Neu5Acα2–6 **Galβ1–4GlcNAcβ1–2Manα1–3** **Neu5Acα2–3Galβ1–4GlcNAcβ1–4**
N4.4.5 **N4.6.3**	20 5	**Neu5Acα2–3Galβ1–4GlcNAcβ1–6** (8′, 7′) **Neu5Acα2–3Galβ1–4GlcNAcβ1–2Manα1–6** (6′, 5′, 4′) **Fucα1–6** **Manβ1–4GlcNAcβ1–4GlcNAc** (3, 2, 1) Neu5Acα2–3 Neu5Acα2–6 **Galβ1–4GlcNAcβ1–2Manα1–3** (6, 5, 4; 8, 7) **Neu5Ac*α2–3Galβ1–4GlcNAcβ1–4**

signals of αGlcNAc-**1** at $\delta = 5.19$ and GlcNAc-**2** at $\delta = 4.60-4.61$, together with the NAc methyl signals of GlcNAc-**1** at $\delta = 2.038-2.039$ and GlcNAc-**2** at $\delta = 2.081-2.082$ (Hård et al., 1990).

HPLC on Lichrospher-NH_2 of Q-Sepharose fraction **N1** yielded eight subfractions, denoted **N1.1–N1.8** (Fig. 3A). Due to the low amount of material, structure elucidation could not be carried out for fractions **N1.1** and **N1.6–N1.8**. The ^{1}H-NMR spectra of the two minor fractions **N1.2** and **N1.3**, and of the two major fractions **N1.4** and **N1.5**, indicate the presence of monosialylated diantennary compounds with a terminal β-galactosyl residue on the Manα1–6 branch. The structures of the carbohydrates present in these fractions are depicted in Table 5, and relevant ^{1}H chemical shifts are compiled in Table 1. From the ^{1}H-NMR spectrum of fraction **N1.2**, the presence of a non-fucosylated, α2–3-monosialylated, diantennary oligosaccharide is deduced (cf. compound **N1.2** in Hård et al., 1990). In addition to this compound, the fraction contains minor amounts of carbohydrate chains with α2–6-linked Neu5Ac and α1–6-linked Fuc. In fraction **N1.3**, the main compound is the α1–6-fucosylated analogue of **N1.2** (cf. compound **C-Q1-1** in Spellman et al., 1989). There are also indications of α2–6-linked Neu5Ac and non-fucosylated carbohydrate chains in fraction **N1.3**. Fraction **N1.4** contains a non-fucosylated, α2–6-monosialylated, diantennary carbohydrate chain (cf. compound **N1.3** in Damm et al., 1989a), whereas the fucosylated analogue of **N1.4** is present in fraction **N1.5** (cf. compound **1-3** in de Waard et al., 1991).

HPLC of Q-Sepharose fraction **N2** yielded 17 subfractions, denoted **N2.1–N2.17** (Fig. 3B). The structures could be established of the carbohydrates in fractions **N2.4–N2.7** and **N2.9–N2.12**, and are shown in Tables 5–7 (for ^{1}H-NMR data, see Tables 1, 3 and 4). ^{1}H-NMR spectroscopy of fraction

Table 6. Sulfated, sialylated carbohydrate chains obtained from human T-H glycoprotein.

Code	Amount	Structure
	mol/100 mol	
N2.4B	<1	**(SO_4^-)–3Galβ1–4GlcNAcβ1–2Manα1–6** **Fucα1–6** **Manβ1–4GlcNAcβ1–4GlcNAc** **Neu5Acα2–6Galβ1–4GlcNAcβ1–2Manα1–3**
N3.2.3	<1	**Neu5Acα2–3Galβ1–4GlcNAcβ1–2Manα1–6** **Manβ1–4GlcNAcβ1–4GlcNAc** (SO_4^-)–4GalNAcβ1–4 | **GlcNAcβ1–2Manα1–3**
N3.2.4	<1	**Neu5Acα2–3Galβ1–4GlcNAcβ1–2Manα1–6** **Manβ1–4GlcNAcβ1–4GlcNAc** (SO_4^-)–3Galβ1–4 | **GlcNAcβ1–2Manα1–3**
N3.2.1	<1	**Neu5Acα2–3Galβ1–4GlcNAcβ1–2Manα1–6** **Fucα1–6** **Manβ1–4GlcNAcβ1–4GlcNAc** (SO_4^-)–4GalNAcβ1–4 | **GlcNAcβ1–2Manα1–3**
N3.2.2A	<1	**Neu5Acα2–3Galβ1–4GlcNAcβ1–2Manα1–6** **Fucα1–6** **Manβ1–4GlcNAcβ1–4GlcNAc** (SO_4^-)–3Galβ1–4 | **GlcNAcβ1–2Manα1–3**
N3.2.2B	<1	**(SO_4^-)–3Galβ1–4GlcNAcβ1–2Manα1–6** **Fucα1–6** **Manβ1–4GlcNAcβ1–4GlcNAc** **Neu5Acα2–3Galβ1–4GlcNAcβ1–2Manα1–3**
N3.6.5	<1	**Galβ1–4GlcNAcβ1–2Manα1–6** **Manβ1–4GlcNAcβ1–4GlcNAc** **Neu5Acα2–6Galβ1–4GlcNAcβ1–2Manα1–3** **(SO_4^-)–3Galβ1–4GlcNAcβ1–4**

N2.4 indicates the presence of two oligosaccharides, namely an α2–3-disialylated, diantennary oligosaccharide (**N2.4A** in Table 5; cf. compound **N2.1** in Hård et al., 1990) and a sulfated, fucosylated, α2–6-monosialylated, diantennary carbohydrate chain (**N2.4B** in Table 6; cf. compound **S_1-1** in de Waard et al., 1991) in a molar ratio of 5:1. Analytical HPAEC on CarboPac PA1 shows that **N2.4A** coelutes with its corresponding reference compound, and the same holds for **N2.4B** (Fig. 4A). In fraction **N2.5**, the fucosylated analogue of **N2.4A** is identified as the major component (Table 5; cf. compound **N2.2** in Hård et al., 1990). The major compound in fraction **N2.6** is a non-fucosylated, disialylated, diantennary carbohydrate chain, bearing α2–3- and α2–6-linked sialic acid in the Manα1–6 and Manα1–3 branches, respectively (Table 5; cf. compound **N2.1.1** in Damm et al., 1989b). The fucosylated analogue of **N2.6** is present in fraction **N2.7** (Table 5; cf. compound **2-1** in de Waard et al., 1991). The ^{1}H-NMR data of fraction **N2.9** indicate the occurrence of a non-fucosylated, disialylated, triantennary oligosaccharide (**N2.9** in Table 5). The triantennary nature of **N2.9** is deduced from the typical Man H-1/H-2 pattern, the H-3 signal of **Man-4**, and the NAc methyl singlet of GlcNAc-**7** (Table 1). To establish the branch location of the two sialic acids, the ^{1}H-NMR data of **N2.9** are compared with those of a fucosylated, monosialylated, triantennary reference compound having an α2–6-sialylated Gal-**6** (cf. compound **1-6** in de Waard et al., 1991), and with a trisialylated, triantennary oligosaccharide having an α2–6-sialylated Gal-**6** and α2–3-sialylated Gal-**6′** and Gal-**8** residues (cf. compound **Tri-S 2A** in Bendiak et al., 1989). By inspection of the chemical shifts of the Neu5Ac H-3e and H-3a signals, in combination with the elution position on Q-Sepharose, it is evident that **N2.9** contains two sialic acids: one α2–6 linked and one α2–3 linked. Guided by the Man-**4** and Man-**4′** H-1 resonances at δ = 5.130 and δ = 4.925, respectively, it is concluded that the α2–6-linked Neu5Ac resides on Gal-**6** and the α2–3-linked one on Gal-**8**. In case

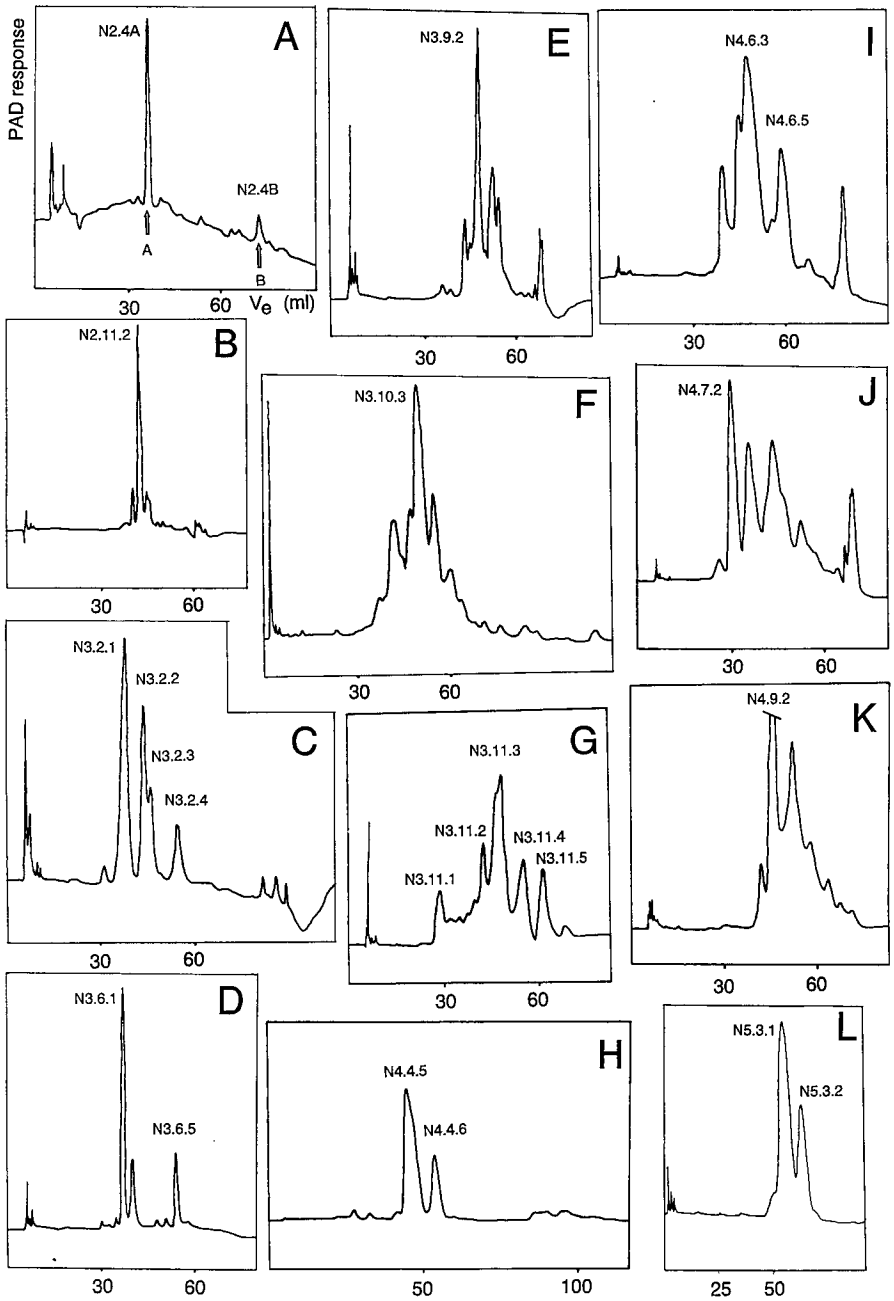


Fig. 4. Separation of HPLC fractions by high-pH anion-exchange chromatography on CarboPac PA1 with pulsed-amperometric detection. Elutions were performed in 100 mM NaOH, applying linear concentration gradients of 1 M sodium acetate in 100 mM NaOH (solvent A). For further details, see Experimental Procedures. A, fraction **N2.4** [t = 0, 10% (by vol.) solvent A; t = 20 min, 30% (by vol.) solvent A]. The arrow marked **A** indicates the elution position of reference compound Neu5Acα2 – 3Galβ1 – 4GlcNAcβ1 – 2Manα1 – 3(Neu5Acα2 – 3Galβ1 – 4GlcNAcβ1 – 2Manα1 – 6)Manβ1 – 4GlcNAcβ1 – 4GlcNAc (**N2.1** in Hård et al., 1990), and the arrow marked **B** indicates the elution position of reference compound Neu5Acα2 – 6Galβ1 – 4GlcNAcβ1 – 2Manα1 – 3[(SO_4^-)-3Galβ1 – 4GlcNAcβ1 – 2Manα1 – 6]Manβ1 – 4GlcNAcβ1 – 4(Fucα1 – 6)GlcNAc (**S_1-1** in de Waard et al., 1991); B, fraction **N2.11** [t = 0, 6% (by vol.) solvent A; t = 20 min, 30% (by vol.) solvent A]; C, the pool of fractions **N3.1** and **N3.2** [t = 0, 20% (by vol.) solvent A; t = 20 min, 30% (by vol.) solvent A]; D, fraction **N3.6** [t = 0, 10% (by vol.) solvent A; t = 20 min, 40% (by vol.) solvent A]; E, fraction **N3.9** [t = 0, 10% (by vol.) solvent A; t = 20 min, 25% (by vol.) solvent A]; F, fraction **N3.10** [t = 0, 12% (by vol.) solvent A; t = 30 min, 22% (by vol.) solvent A]; G, fraction **N3.11** [t = 0, 12% (by vol.) solvent A; t = 30 min, 22% (by vol.) solvent A]; H, fraction **N4.4** [t = 0, 15% (by vol.) solvent A; t = 30 min, 30% (by vol.) solvent A]; I, fraction **N4.6** [t = 0, 15% (by vol.) solvent A; t = 30 min, 25% (by vol.) solvent A]; J, fraction **N4.7** [t = 0, 15% (by vol.) solvent A; t = 30 min, 25% (by vol.) solvent A]; K, fraction **N4.9** [t = 0, 12% (by vol.) solvent A; t = 30 min, 25% (by vol.) solvent A]; L, fraction **N5.3** [t = 0, 20% (by vol.) solvent A; t = 30 min, 45% (by vol.) solvent A].

of the element Neu5Acα2 – 3Galβ1 – 4GlcNAcβ1 – 2Manα1 – 6Manβ1 –, the βMan (i.e. Man-**4′**) H-1 resonates at δ = 4.91 (cf. compound **Tri-S 2A** in Bendiak et al., 1989). Green et al. have published ^{1}H-NMR data for an oligosaccharide alditol having the same structure as **N2.9** [see compound **N-2/T(G)** in Green et al., 1988], occurring in a 1:1 mixture with an isomeric carbohydrate. A significant discrepancy is noted in their assignment of GlcNAc-**5′** H-1 (δ =

Table 7. Sialylated, GalNAc-containing carbohydrate chains obtained from human T-H glycoprotein.

Code	Amount	Structure
	mol/100 mol	
N2.11.2	<1	**Galβ1–4GlcNAcβ1–2Manα1–6** **Manβ1–4GlcNAcβ1–4GlcNAc** **Neu5Acα2–6Galβ1–4GlcNAcβ1–2Manα1–3** **Neu5Acα2–3Galβ1–4GlcNAcβ1–4** **GalNAcβ1–4**
N2.12	<1	H **Galβ1–4GlcNAcβ1–2Manα1–6** **Fucα1–6**
N3.9.2	<1	Neu5Acα2–3 **Manβ1–4GlcNAcβ1–4GlcNAc** **Neu5Acα2–6Galβ1–4GlcNAcβ1–2Manα1–3** **Neu5Acα2–3Galβ1–4GlcNAcβ1–4** **GalNAcβ1–4**
N4.7.2	1	**GalNAcβ1–4** **Neu5Acα2–3Galβ1–4GlcNAcβ1–6** **Neu5Acα2–3Galβ1–4GlcNAcβ1–2Manα1–6** **Fucα1–6** **GalNAcβ1–4** **Manβ1–4GlcNAcβ1–4GlcNAc** GalNAcβ1–4 Neu5Acα2–3Galβ1–4 | **GlcNAcβ1–2Manα1–3**
N4.9.2	<1	Neu5Acα2–6Galβ1–4 | **Neu5Acα2–3Galβ1–4GlcNAcβ1–4** **GalNAcβ1–4**

4.570). The value given in the present report ($\delta = 4.581$) is in agreement with Cumming et al. (1989), Hård et al. (1990) and de Waard et al. (1991). The ^{1}H-NMR spectrum of fraction **N2.10** demonstrates the occurrence of the novel α1–6-fucosylated analogue of **N2.9** (Table 5). Compound **N2.10** is characterized by the same structural-reporter-group signals as **N2.9**, except for the effects of the fucosylation (Table 1).

N2.10

Galβ1–4GlcNAcβ1–2Manα1–6 Fucα1–6
Manβ1–4GlcNAcβ1–4GlcNAc
Neu5Acα2–6Galβ1–4GlcNAcβ1–2Manα1–3
Neu5Acα2–3Galβ1–4GlcNAcβ1–4

Fraction **N2.11** was subfractionated by HPAEC, yielding one major fraction denoted **N2.11.2** (Fig. 4B). The ^{1}H-NMR spectrum of **N2.11.2** shows the presence of a novel GalNAc-containing, non-fucosylated, disialylated, triantennary oligosaccharide (Fig. 5A; Table 7).

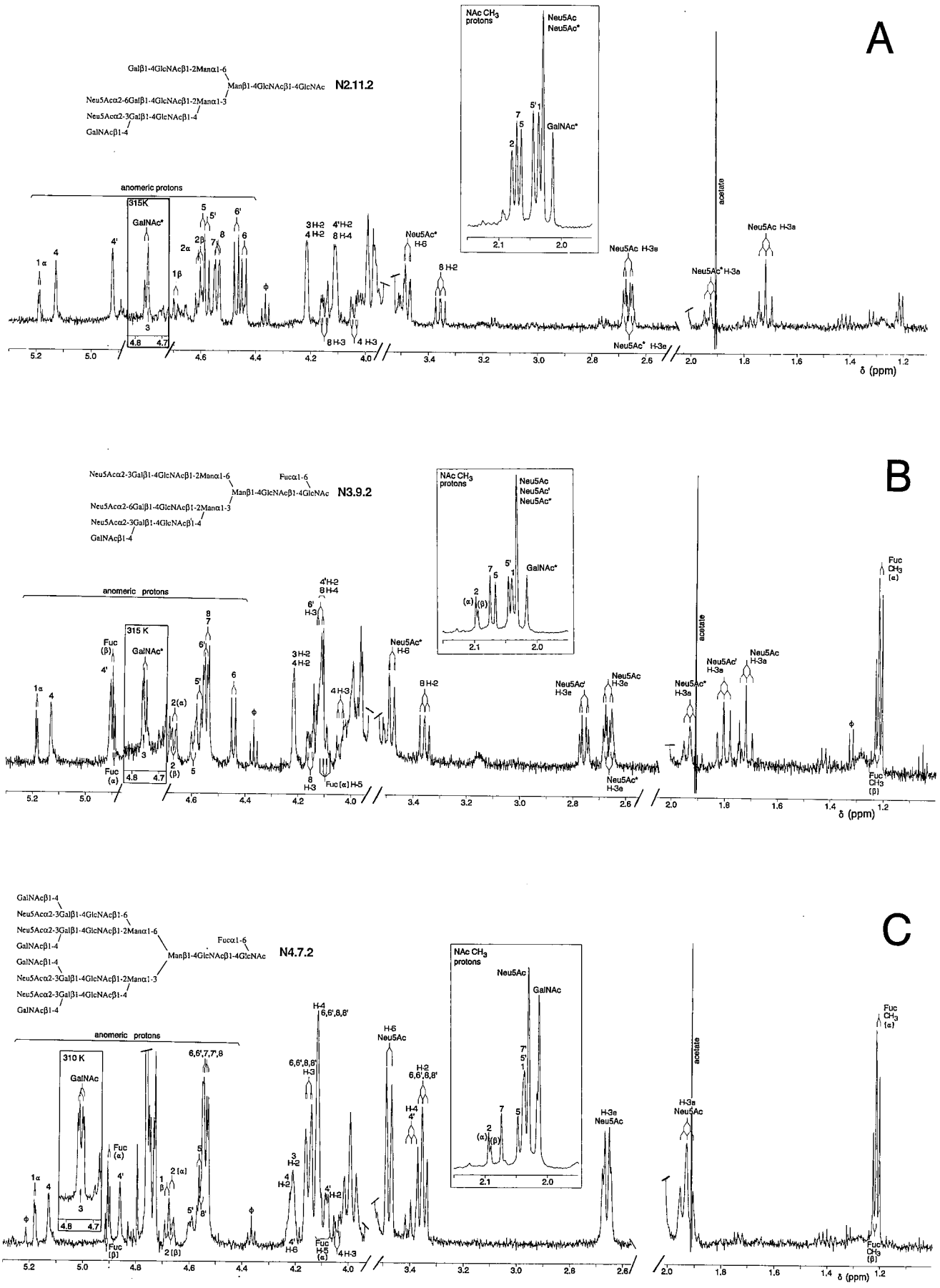


Fig. 5. Structural-reporter-group regions of the resolution-enhanced ^{1}H-NMR spectra of novel GalNAc-containing oligosaccharides obtained from human T-H glycoprotein. A, fraction **N2.11.2**; B, fraction **N3.9.2**; C, fraction **N4.7.2**; D, fraction **N4.9.2**.

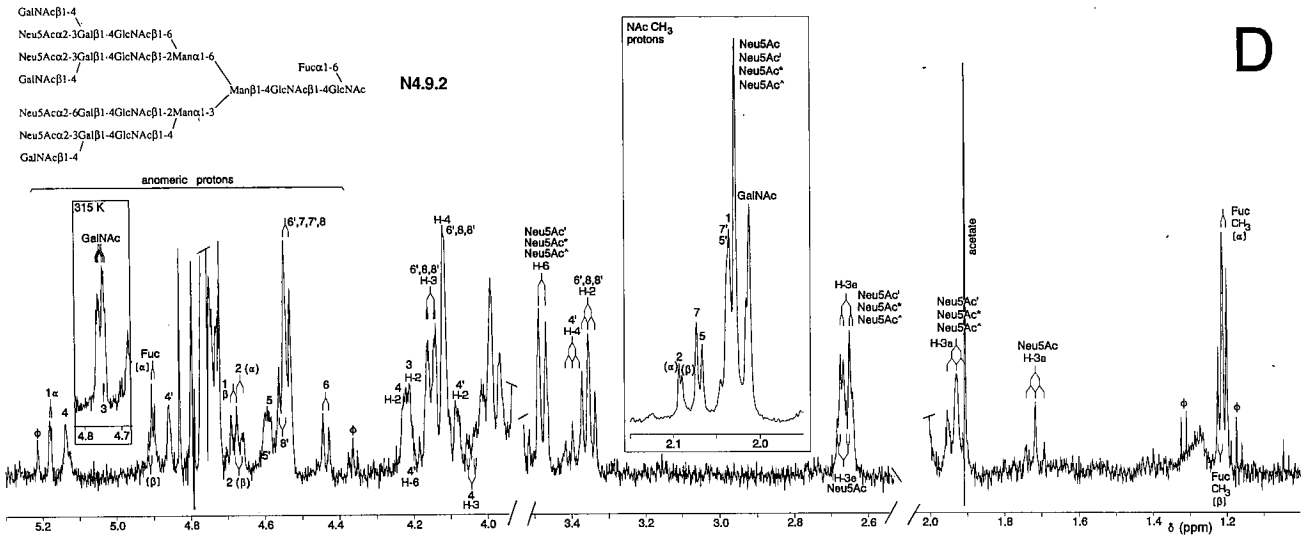
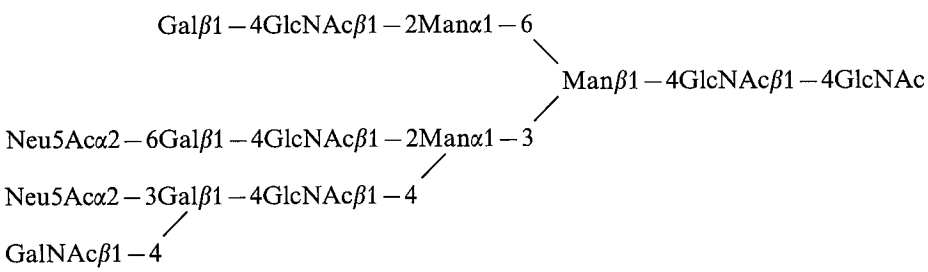


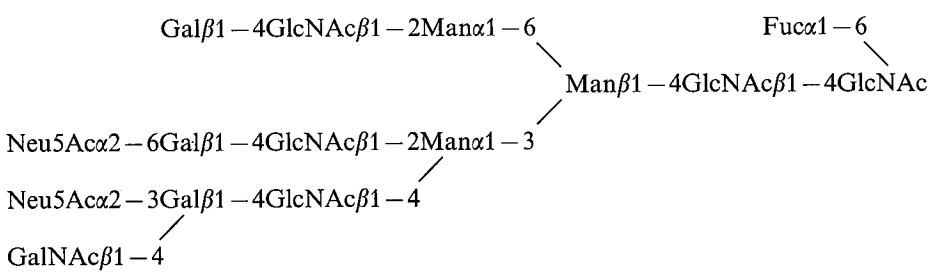
Fig. 5. (continued)

N2.11.2



The triantennary branching pattern of **N2.11.2** is deduced from the characteristic set of chemical shifts of the Man H-1 and H-2 signals (cf. **N2.9**; Table 4). The structural element Neu5Acα2-3(GalNAcβ1-4)Galβ1-4GlcNAcβ1- is recognized from the typical signals for Neu5Ac H-3a at $\delta = 1.927$ and H-3e at $\delta = 2.662$, in conjunction with those for GalNAc H-1 at $\delta = 4.758$ (42°C), GalNAc NAc methyl at $\delta = 2.015$, as well as Gal H-3 at $\delta = 4.151$ (cf. Blanchard et al., 1983; Williams et al., 1984; Donald and Feeney, 1986; Nasir-Ud-Din et al., 1986; Soh et al., 1989). Since structural-reporter-group signals are present for both a Neu5Acα2-6Galβ1-4GlcNAcβ1-2Manα1-3Manβ1- and a Galβ1-4GlcNAcβ1-2Manα1-6Manβ1- element, the GalNAc must be located on Gal-**8** (Table 4). HPLC fraction **N2.12** contains as major component the novel α1-6-fucosylated analogue of **N2.11.2** (Table 7).

N2.12



The structure of compound **N2.12** is deduced from the same structural-reporter-group signals as were used for **N2.11.2**, taking into account the effects associated with the α1-6 fucosylation of GlcNAc-**1** (Table 4).

HPLC of Q-Sepharose fraction **N3** yielded 17 subfractions, denoted **N3.1** to **N3.17** (Fig. 3C). The structure has been determined for the carbohydrate chains present in the major fractions **N3.6**, and **N3.8**-**N3.10**, as well as for the minor fractions **N3.1**, **N3.2** and **N3.5** (Tables 5-7). The ^{1}H-NMR spectra of fractions **N3.1** and **N3.2** indicated in each case the presence of a mixture of two monosulfated, monosialylated compounds, the α1-6-fucosylated forms being present in **N3.2** and the non-fucosylated compounds in **N3.1**. In order to achieve unambiguous NMR assignments for the carbohydrates in these fractions, a 2D HOHAHA spectrum was recorded for **N3.2**, and a 2D NOE spectrum was recorded for the pool of **N3.1** and **N3.2** (not shown). Then, the pool was fractionated by HPAEC, giving rise to four major fractions, denoted **N3.2.1**-**N3.2.4** (Fig. 4C). From the ^{1}H-NMR data of fraction **N3.2.1**, the presence of a (SO_4^-)-4GalNAc-containing, fucosylated, α2-3-monosialylated, diantennary oligosaccharide can be deduced (Tables 3 and 6; cf. compound **S-^{3}N** in Weisshaar et al., 1991).

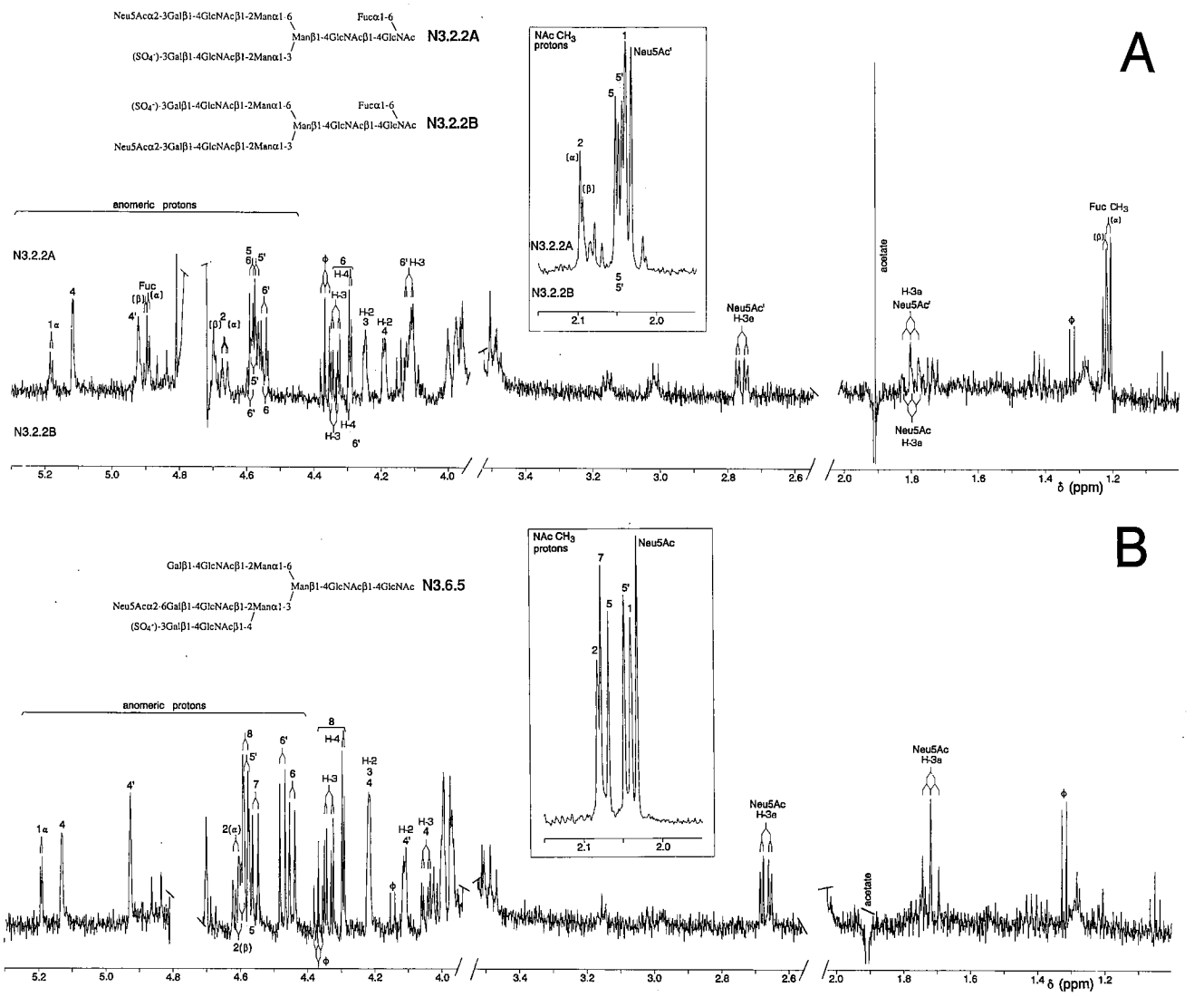
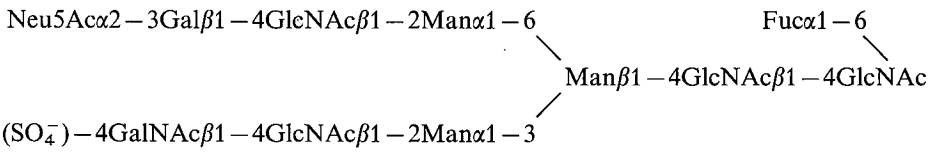


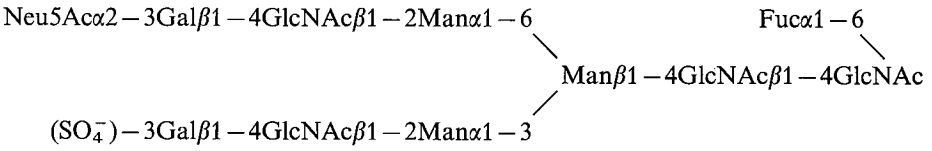
Fig. 6. Structural-reporter-group regions of the resolution-enhanced ^{1}H-NMR spectra of novel sulfated oligosaccharides obtained from human T-H glycoprotein. A, fraction **N3.2.2**; B, fraction **N3.6.5**.

N3.2.1

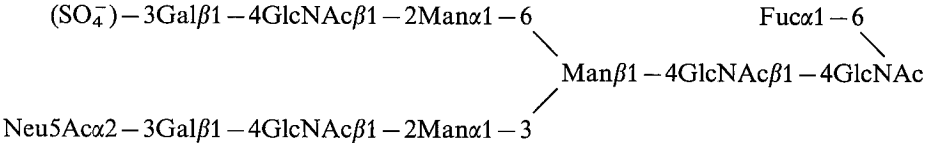


The spectrum of fraction **N3.2.2** (Fig. 6A) indicates the occurrence of two novel isomeric $(SO_4^-)-3$Gal-containing, fucosylated, α2 – 3-monosialylated, diantennary oligosaccharides, denoted **N3.2.2A** and **N3.2.2B**, in a molar ratio of approximately 5:2 (Table 6).

N3.2.2A

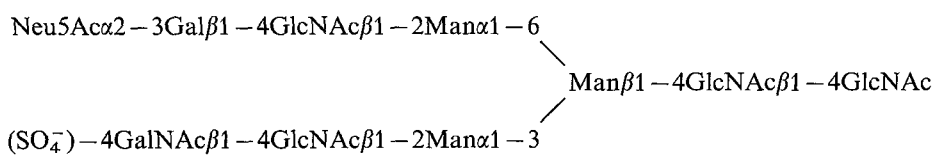


N3.2.2B

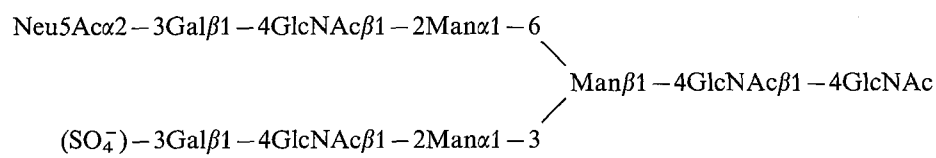


Relevant ^{1}H chemical shifts of **N3.2.2A** and **N3.2.2B** are compiled in Table 3. The spectral data of **N3.2.2A** resemble those of **N2.5**. The Neu5Ac H-3a at δ = 1.801, Gal-**6′** H-1 at δ = 4.549, and GlcNAc-**5′** NAc methyl at δ = 2.044, are typical of the structural element Neu5Acα2−3Galβ1−4GlcNAcβ1−, when present in the Manα1−6 branch (cf. compound **N2.5**; see Tables 1 and 5). Additional signals are observed at δ = 4.335 and δ = 4.291, which can be assigned to Gal H-3 and H-4, respectively, in a 3-sulfated Gal, and by 2D HOHAHA spectroscopy they are interconnected to Gal-**6** H-1 at δ = 4.583. The GlcNAc-**5** H-1 resonance at δ = 4.583, which is found from a NOE to Man-**4** H-1 at δ = 5.118 and the GlcNAc-**5** NAc methyl singlet at δ = 2.051, indicate that the 3-sulfated Gal is attached to this residue (cf. compound S_1-**1** in de Waard et al., 1991; see also Ruiz Contreras et al., 1988; Lamblin et al. 1991). The set of chemical shifts for the (SO_4^-)−3Galβ1−4GlcNAcβ1−2Manα1− element in the Manα1−3 branch is significantly different from, but resemble closely, that observed for the same structural element present in the Manα1−6 branch (cf. compound S_1-**1** in de Waard et al., 1991). The presence of compound **N3.2.2B** is demonstrated from the separate set of (SO_4^-)−3Gal H-1 at δ = 4.589, H-3 at δ = 4.441 and H-4 at δ = 4.296, together with the GlcNAc-**5′** H-1 at δ = 4.579 and NAc methyl at δ = 2.047, characteristic of a (SO_4^-)−3Galβ1−4GlcNAcβ1−2Manα1− element in the Manα1−6 branch (cf. compound S_1-**1** in de Waard et al., 1991). Furthermore, the Gal-**6** H-1 at δ = 4.544 is indicative of the presence of a terminal α2−3-linked Neu5Ac in the Manα1−3 branch. HPAEC fraction **N3.2.3** contained the non-fucosylated analogue of **N3.2.1**, whereas the novel non-fucosylated analogue of **N3.2.2A** is present in fraction **N3.2.4** (Table 6).

N3.2.3



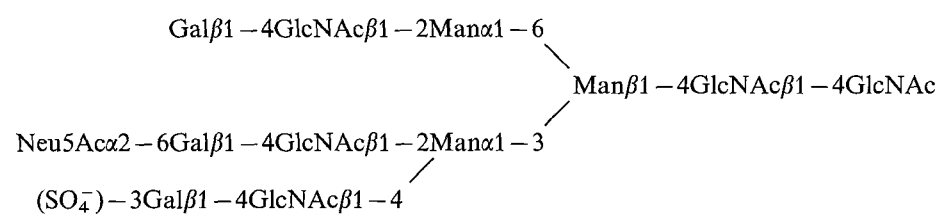
N3.2.4



The structure of **N3.2.3** is deduced from the same structural-reporter-group signals used for **N3.2.1**. The structure of **N3.2.4** is inferred from a comparison with the ^{1}H-NMR data of **N3.2.2A**, taking into account the effects associated with the α1−6 fucosylation of GlcNAc-**1** (Table 3). In the spectrum of fraction **N3.2.4**, there are also indications for a compound with a 3-sulfated Gal-**6′**, evidenced by a Gal H-1 at δ = 4.589, by a Gal H-4-like signal at δ = 4.295 and by a GlcNAc-**5′** NAc methyl resonance at δ = 2.047 (cf. **N.3.2.2B**). However, due to the small amount of material, no conclusive structure could be established for this component.

The ^{1}H-NMR spectrum of fraction **N3.5** indicates the presence of a non-fucosylated, α2−3-trisialylated, triantennary oligosaccharide (Tables 2 and 5; cf. compound **N3.1** in Hård et al., 1990). Fraction **N3.6** contained a mixture of carbohydrates and was therefore subfractionated by HPAEC (Fig. 4D). The α1−6-fucosylated analogue of **N3.5** occurs in the major fraction **N3.6.1** (Tables 2 and 5; cf. compound **N3.2A** in Hård et al., 1990). The minor fraction **N3.6.5** contains a novel sulfated, α2−6-monosialylated, triantennary carbohydrate chain (Table 6).

N3.6.5

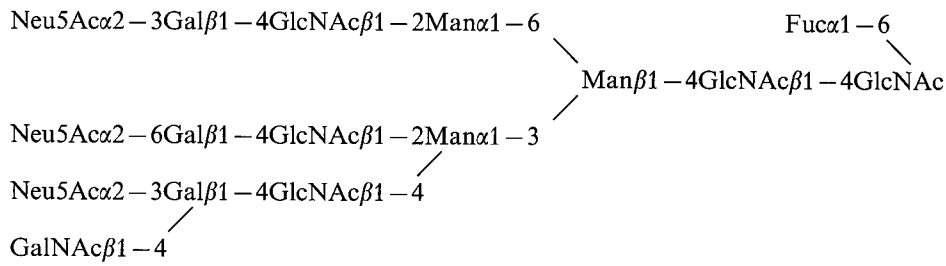


The structural-reporter-group regions of the ^{1}H-NMR spectrum of **N3.6.5** are shown in Fig. 6B, and relevant ^{1}H chemical shifts are given in Table 3. The triantennary character of **N3.6.5** is inferred from the typical combination of Man H-1 and H-2 signals. The chemical-shift values and intensities of the Gal H-4 at δ = 4.293 and H-3 at δ = 4.335 point to the presence of one terminal (SO_4^-)−3Gal residue. By inspection of the Neu5Ac H-3a, H-3e and NAc methyl signals, it can be deduced that one α2−6-linked Neu5Ac occurs. The location of the α2−6-linked Neu5Ac on Gal-**6** is evident from the set of Gal-**6** H-1 at δ = 4.445, GlcNAc-**5** H-1 at δ = 4.596, GlcNAc-**5** NAc methyl at δ = 2.067 and Man-**4** H-1 at δ = 5.130 (cf. compound **N2.9** in Table 1). The combination of Gal-**6′** H-1 at δ = 4.473, GlcNAc-**5′** H-1 at δ = 4.581, GlcNAc-**5′** NAc methyl at δ = 2.047 and Man-**4′** H-1 at δ = 4.925 are characteristic of the non-sialylated element Galβ1−4GlcNAcβ1−2Manα1−6 (cf. compound **N2.9** in Table 1; see also compound **1-6** in de Waard et al., 1991). Thus, it is concluded that the sulfate group must be located on Gal-**8**. In case of a terminal Galβ1−4GlcNAcβ1−4Manα1−3 element, the Gal-**8** H-1 resonates at δ = 4.468, and the GlcNAc-**7** H-1 at δ = 4.545 (cf. compound **1-6** in de Waard et al.,

1991), indicating that 3-*O*-sulfation of Gal-**8** gives rise to downfield shifts of $\Delta\delta = +0.116$ for Gal-**8** H-1 and $\Delta\delta = +0.009$ for GlcNAc-**7**. In case of sulfation of Gal-**6** and Gal-**6′**, similar downfield shifts have been observed for Gal-**6** H-1 and Gal-**6′** H-1, respectively (Table 3).

In fraction **N3.8**, a fucosylated, trisialylated, triantennary oligosaccharide occurs as a major component (Tables 2 and 5; cf. compound **Tri-S 2A** in Bendiak et al., 1989). A non-fucosylated carbohydrate chain is present as minor component in this fraction. Subfractionation of **N3.9** by HPAEC yielded five major fractions, denoted **N3.9.1 – N3.9.5** (Fig. 4E). The ^{1}H-NMR spectrum of the major fraction **N3.9.2** (Fig. 5B) indicates the occurrence of the following novel GalNAc-containing, fucosylated, trisialylated, triantennary carbohydrate chain (Table 7).

N3.9.2



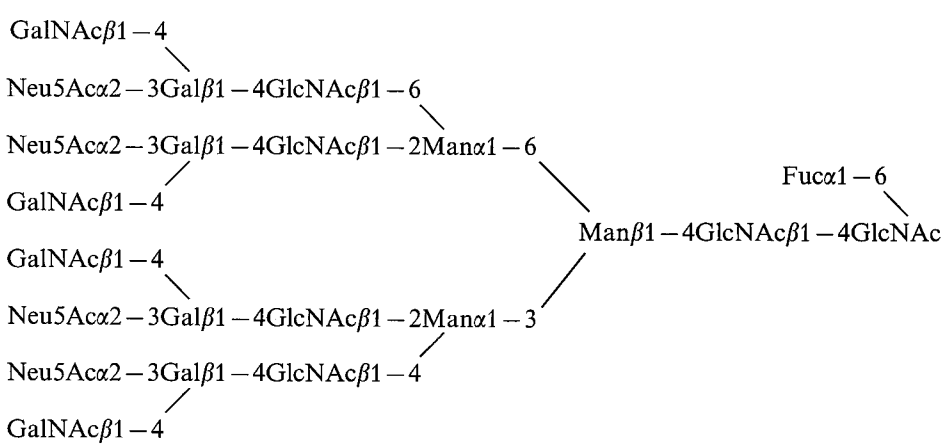
Relevant ^{1}H chemical shifts of **N3.9.2** are given in Table 4. In the spectrum of **N3.9.2**, additional signals, compared to the spectrum of **N2.12**, are observed at $\delta = 1.803$ and $\delta = 2.758$, typical of the H-3a and H-3e resonances, respectively, of α2 – 3-linked Neu5Ac. The location of Neu5Ac at Gal-**6′** is inferred from the Man-**4′** H-1 at $\delta = 4.905$, GlcNAc-**5′** H-1 at $\delta = 4.572$ and Gal-**6′** H-1 at $\delta = 4.549$ (cf. compound **N3.8**, Table 2). The other relevant structural-reporter-group signals of **N3.9.2** fit those of **N2.12**.

The ^{1}H-NMR spectrum of HPLC fraction **N3.10** indicated a complex mixture of sialylated oligosaccharides. Therefore, this fraction was subfractionated by HPAEC, giving rise to five major peaks, denoted **N3.10.1 – N3.10.5** (Fig. 4F). The most abundant fraction **N3.10.3** contains as major component a fucosylated, trisialylated, triantennary carbohydrate chain. Relevant ^{1}H chemical shifts are presented in Table 2, and the structure of **N3.10.3** is depicted in Table 5. The structure of **N3.10.3** is evident by comparing its structural-reporter-group signals with those of a non-fucosylated analogue (cf. **Tri-S 4B** in Bendiak et al., 1989). In addition to signals attributable to compound **N3.10.3**, the ^{1}H-NMR spectrum of this fraction contains indications for a trisialylated, tetraantennary oligosaccharide. This is evidenced by low-intensity signals of Man-**4′** H-1 at $\delta = 4.858$, of terminal Gal-**6′** H-1 at $\delta = 4.473$, and of GlcNAc-**5** NAc methyl at $\delta = 2.048$ (Gal-**6** α2 – 3-sialylated). HPAEC of fraction **N3.11** yields five main peaks, denoted **N3.11.1 – N3.11.5** (Fig. 4G), but despite the combination of FPLC on Q-Sepharose, HPLC on Lichrospher-NH_2 and HPAEC on CarboPac PA1, these fractions still contain mixtures of carbohydrates. Therefore, no structure elucidation could be carried out.

HPLC of Q-Sepharose fraction **N4** yields ten subfractions, denoted **N4.1 – N4.10** (Fig. 3D). The structures could be elucidated for the major carbohydrate chains present in fractions **N4.4**, **N4.6**, **N4.7** and **N4.9**. HPAEC of **N4.4** afforded two major peaks, denoted **N4.4.5** and **N4.4.6** (Fig. 4H). **N4.4.5** contains a fucosylated, α2 – 3-tetrasialylated, tetraantennary oligosaccharide (Tables 2 and 5; cf. compound **A** in Hokke et al., 1991), and fraction **N4.4.6** contains the corresponding non-fucosylated analogue. Subfractionation of **N4.6** by HPAEC yields seven fractions, denoted **N4.6.1 – N4.6.7** (Fig. 4I). The ^{1}H-NMR data of the major fraction **N4.6.3** reveals the presence of a fucosylated, tetrasialylated, tetraantennary oligosaccharide (Table 5). From the intensity ratio of the Neu5Ac H-3 structural-reporter-group signals, it is evident that one α2 – 6-linked and three α2 – 3-linked Neu5Ac residues are present. The resonances of Man-**4** H-1 at $\delta = 5.138$, Gal-**6** H-1 at $\delta = 4.439$ and GlcNAc-**5** NAc methyl at $\delta = 2.069$ indicate that the α2 – 6-linked Neu5Ac resides at Gal-**6** (cf. compound **242** in van Pelt et al., 1988). Fraction **N4.6.5** contains the non-fucosylated analogue of **N4.6.3** (Tables 2 and 5).

Subfractionation by HPAEC of **N4.7** yields five major fractions, denoted **N4.7.1 – N4.7.5** (Fig. 4J). The ^{1}H-NMR spectrum of the most abundant fraction, **N4.7.2** (Fig. 5C), indicates the occurrence of the following novel GalNAc-containing, fucosylated, tetrasialylated, tetraantennary oligosaccharide (Table 7).

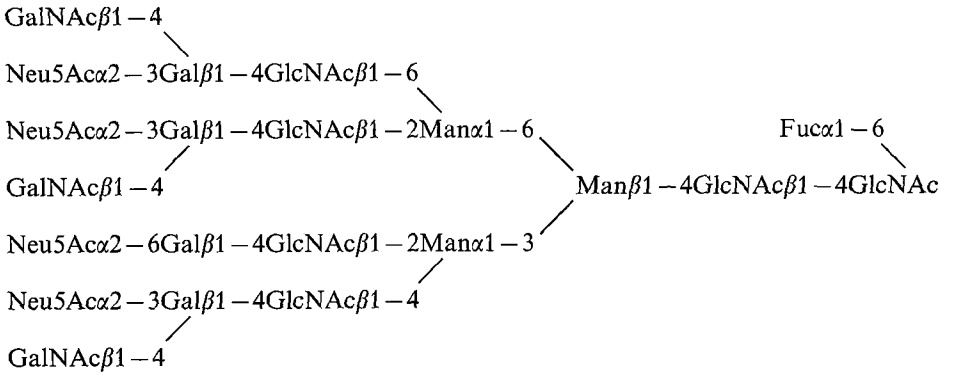
N4.7.2



The combination of Man H-1 and H-2 signals is typical for a tetraantennary oligosaccharide (Table 4). From the sialic acid H-3 structural-reporter-group signals (H-3a, $\delta = 1.930$; H-3e, $\delta = 2.66$), it is evident that all sialic acids are present in a Neu5Acα2−3(GalNAcβ1−4)Galβ1−4GlcNAcβ1− element. This is corroborated by the GalNAc H-1 peaks at $\delta = 4.737/4.743$ and GalNAc NAc methyl signals at $\delta = 2.015/2.018$, as well as the Gal H-3 resonances at $\delta = 4.15$ and Gal H-4 at $\delta = 4.12$. In the 2D HOHAHA spectrum of **N4.7.2** (not shown), all Man H-2 atoms can be found from cross-peaks to their corresponding H-1 atoms, and the Gal H-2 signals at $\delta = 3.356$ can be interconnected to the H-4, H-3 and H-1 resonances of the corresponding Gal residues. The H-6 signal of Man-**4′** ($\delta = 4.21$) is identified from a HOHAHA connectivity to the Man-**4′** H-2, in combination with its multiplet pattern. The chemical-shift positions of, for example, GlcNAc-**5** and GlcNAc-**5′** H-1, are obtained by locating the characteristic NOESY cross-peaks to Man-**4** and Man-**4′** H-1, respectively (not shown).

HPAEC of fraction **N4.9** yields seven fractions, denoted **N4.9.1**−**N4.9.7** (Fig. 4K). The ^{1}H-NMR spectrum of the major fraction **N4.9.2** (Fig. 5D) indicates the presence of the following novel, GalNAc-containing, fucosylated, tetrasialylated, tetraantennary carbohydrate chain.

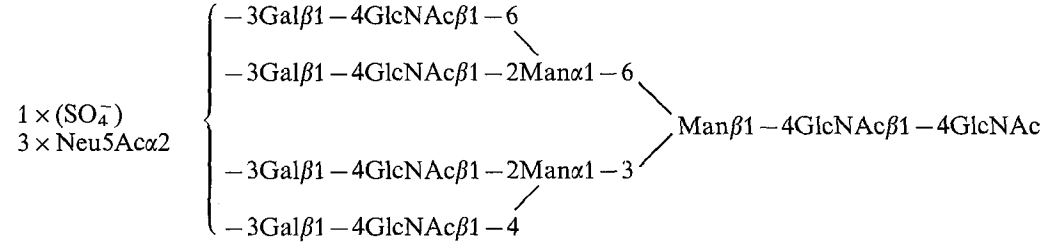
N4.9.2



To deduce the structure of **N4.9.2**, its structural-reporter-group signals are compared with those of **N4.6.3** (Table 2) and **N4.7.2** (Table 4). The intensities of the Neu5Ac H-3a signals at $\delta = 1.720$ and $\delta = 1.93$ indicate that the ratio of Neu5Ac in the element Neu5Acα2−6Galβ1− and in the element Neu5Acα2−3(GalNAcβ1−4)Galβ1− amounts to 1:3. The branch location of the α2−6-linked sialic acid is inferred from the Man-**4** H-1 resonance at $\delta = 5.140$.

HPLC of Q-Sepharose fraction **N5** yields two major peaks, denoted **N5.3** and **N5.4** (Fig. 3E). HPAEC of the major fraction **N5.3** yields two subfractions, **N5.3.1** and **N5.3.2** (Fig. 4L). The ^{1}H-NMR spectrum of the major fraction **N5.3.1** (not shown) suggests the occurrence of a mixture of novel isomeric monosulfated, trisialylated, tetraantennary carbohydrate chains.

N5.3.1



The tetraantennary branching of **N5.3.1** is deduced from the characteristic combination of Man-**4** H-1 at $\delta = 5.130$, Man-**4′** H-1 at $\delta = 4.858$, Man-**3** H-2 at $\delta = 4.205$, Man-**4** H-2 at $\delta = 4.220$ and Man-**4′** H-2 at $\delta = 4.090$. The presence of one 3-sulfated Gal residue is inferred from the relative intensity of the typical H-3 and H-4 signals at $\delta = 4.338$ and $\delta = 4.294$, respectively. From the Neu5Ac H-3a at $\delta = 1.805$ and H-3e at $\delta = 2.756$, it is evident that three α2−3-linked sialic acids occur. In the 2D NOE spectrum of **N5.3.1** the Man-**4** H-1/H-2, the Man-**4′** H-1/H-2 and the (SO_4^-)−3Gal H-3/H-1 cross-peaks are all heterogeneous (not shown). This is most likely due to various locations of the 3-sulfated Gal residue. Due to the heterogeneity of **N5.3.1**, no conclusive structure has yet been determined.

DISCUSSION

Despite numerous studies on the carbohydrate moieties of human T-H glycoprotein, no structures have been established for its intact acidic Asn-linked carbohydrate chains (Gottschalk, 1952; Odin, 1952; Fletcher et al., 1970; Afonso and Marshall, 1979; Soh et al., 1980; Afonso et al., 1981; Donald et al., 1982, 1983; Williams et al., 1984; Serafini-Cessi et al., 1984a, b; Donald and Feeney, 1986; Dall'Olio et al., 1988; Smagula et al., 1990). In the present study the N-linked carbohydrates were enzymically released from T-H glycoprotein of one male donor by PNGase-F, and fractionated into more than 150 fractions by a combination of three liquid-chromatographic techniques. Detailed structural characterization has been carried out for 30 oligosaccharides, including 10 novel compounds (Tables 5−7). Sialylated, tetraantennary carbohydrate chains constitute the major part of the carbohydrates, whereas smaller amounts of triantennary and diantennary oligosaccharides occur. No tri′-antennary carbohydrates, i.e. compounds containing a 2,6-branched Man-**4′**, could be detected. Previously, T-H glycoprotein has been shown to contain 1−20% oligomannose-type carbohydrate chains (Serafini-Cessi et al., 1984a; Dall'Olio et al., 1988;

Smagula et al., 1990). In the present investigation, no significant amounts of oligomannose-type carbohydrates were found. This observation may reflect the unique source of the material being one male donor.

A characteristic feature of the T-H glycoprotein is the presence of GalNAc in the N-glycans. This monosaccharide is part of the determinant for the blood group Sd^a (Soh et al., 1980, 1989). Previously, a Sd^a-active pentasaccharide, released from T-H glycoprotein by endo-β-galactosidase, has been characterized (Donald et al., 1983; Donald and Feeney, 1986). Furthermore, studies on glycopeptide mixtures indicated that the Sd^a determinant is present in triantennary and tetraantennary compounds (Williams et al., 1984). Here, novel structures are presented of three triantennary and of two tetraantennary, intact, N-linked, GalNAc-containing carbohydrate chains (Table 7). Interestingly, none of these compounds contain the Sd^a-active pentasaccharide structural element. It is unlikely that this pentasaccharide has been lost by the action of a urinary endo-β-galactosidase (DeGasperi et al., 1986), because this would generate structures bearing a terminal GlcNAc, and there are no indications for the occurrence of such compounds. If the element does occur in the investigated T-H glycoprotein preparation, it is present in very minor amounts. In view of the recent data by Soh et al. (1989), who have found that the synthetic oligosaccharide Neu5Acα2$-$3(GalNAcβ1$-$4)Galβ1$-$4GlcNAc has strong Sd^a activity, it has to be expected that also the GalNAc-containing oligosaccharides isolated in the present study have Sd^a activity. As in previous reports on Cad-active mucins (Blanchard et al., 1983; Herkt et al., 1985; Nasir-Ud-Din et al., 1986) and Sd^a-active T-H glycoprotein (Donald et al., 1983; Donald and Feeney, 1986; Williams et al., 1984), no terminal GalNAcβ1$-$4Galβ1$-$ or Neu5Acα2$-$6(GalNAcβ1$-$4)Galβ1$-$ element was found. This is in agreement with biosynthetic studies with β-*N*-acetylgalactosaminyltransferases from guinea-pig kidney (Serafini-Cessi and Dall'Olio, 1983), human kidney (Piller et al., 1986), human blood plasma (Takeya et al., 1987) and human urine (Serafini-Cessi et al., 1988), which have demonstrated that the transferases require α2$-$3-linked sialic acid in the acceptor.

T-H glycoprotein is the major human urinary inhibitor of the binding of *E. coli* S fimbriae to epithelial cells of the urinary tract. It has been shown that the interaction with S fimbriae is probably mediated by its sialyloligosaccharide chains (Parkkinen et al., 1988), as demonstrated by the partial inhibition of the binding in the presence of 20 mM Neu5Acα2$-$3Galβ1$-$4Glc. It should be noted that many pulmonary pathogenic bacteria bind specifically to the carbohydrate sequence GalNAcβ1$-$4Gal found in some glycolipids (Krivan et al., 1988). Therefore, T-H glycoprotein, bearing the same structural element, might play a protective role in preventing adhesion of GalNAc-binding bacteria to the urinary tract.

T-H glycoprotein is known to contain sulfate (Odin, 1952), but it has not been localized, neither in the protein nor in the carbohydrate part. In Table 6, the structures of seven sulfated carbohydrates obtained from T-H glycoprotein, including four novel compounds, are given. Two types of sulfated residues were found, namely, terminal (SO_4^-)$-$4GalNAc and terminal (SO_4^-)$-$3Gal. Until now, N-linked carbohydrates with a terminal 4-*O*-sulfated GalNAc have been regarded as unique for some pituitary glycoprotein hormones (Baenziger and Green, 1988; Weisshaar et al., 1991). However, the presence of **N3.2.1** and **N3.2.3** show that also T-H glycoprotein carries 4-*O*-sulfated GalNAc. The presence of glycoprotein N-glycans containing 3-*O*-sulfated Gal has only been observed in thyroglobulin of calf, human and porcine origin (Spiro and Bhoyroo, 1988; Kamerling et al., 1988; de Waard et al., 1991), and only one intact (SO_4^-)$-$3Gal-containing oligosaccharide structure has been characterized, namely a diantennary compound in which Gal-**6′** is 3-*O*-sulfated and Gal-**6** is 6-*O*-sialylated (S_1-**1** from porcine thyroglobulin; Kamerling et al., 1988; de Waard et al., 1991). In T-H glycoprotein, diantennary carbohydrates have been found in which one branch is 3-*O*-sulfated (Gal-**6′** or Gal-**6**) and the other 3-*O*-sialylated (Gal-**6** or Gal-**6′**; see **N3.2.2A**, **N3.2.4** and **N3.2.2B** in Table 6). Furthermore, a novel sulfated, sialylated, triantennary compound occurs, in which Gal-**8** is 3-*O*-sulfated and Gal-**6** 6-*O*-sialylated (see **N3.6.5** in Table 6). Interestingly, immunohistochemical studies using antibodies against 3-*O*-sulfated galactosylceramide and T-H glycoprotein, respectively, have shown a tightly linked topological association of 3-*O*-sulfated galactosylceramide and T-H glycoprotein in the kidney as well as in the brain (Zalc et al., 1984). In view of the presence of 3-*O*-sulfated galactose in T-H glycoprotein, attention has to be paid to the possibility that antibodies recognizing 3-*O*-sulfated galactosyl residues may also cross-react with T-H glycoprotein.

The results presented here indicate an impressive repertoire of sialylated, sulfated, and GalNAc-containing N-linked carbohydrates in T-H glycoprotein of one male donor, thereby excluding the possibility of heterogeneity stemming from the use of a pool of T-H glycoprotein from several persons. By consequence, T-H glycoprotein exhibits the largest heterogeneity thus far described for N-glycoproteins.

The authors thank Ms Annsofi Nihlén for assistance with the purification of human T-H glycoprotein and Dr Jan Damm for performing preliminary experiments. This investigation was supported by the Netherlands Program for Innovation Oriented Carbohydrate Research (IOP-k) with financial aid from the Ministry of Economic Affairs and the Ministry of Agriculture, Nature Management and Fisheries, and by the Netherlands Foundation for Chemical Research (SON) with financial aid from the Netherlands Organization for Scientific Research (NWO). K. H. was supported in part by the Finnish Cultural Foundation, the Magnus Ehrnrooth Foundation and the Oskar Öflund Foundation. Present address of Dr. P. Moonen: European Patent Office, Erhardtstrasse 27, W-8000, München, Federal Republic of Germany.

REFERENCES

Afonso, A.-M. M. & Marshall, R. D. (1979) *Biochem. Soc. Trans. 7*, 170–173.

Afonso, A.-M. M., Charlwood, P. A. & Marshall, R. D. (1981) *Carbohydr. Res. 89*, 309–319.

Baenziger, J. U. & Green, E. D. (1988) *Biochim. Biophys. Acta 974*, 287–306.

Bax, A. & Davis, D. G. (1985) *J. Magn. Reson. 65*, 355–360.

Bendiak, B., Harris-Brandts, M., Michnick, S. W., Carver, J. P. & Cumming, D. A. (1989) *Biochemistry 28*, 6491–6499.

Blanchard, D., Cartron, J.-P., Fournet, B., Montreuil, J., van Halbeek, H. & Vliegenthart, J. F. G. (1983) *J. Biol. Chem. 258*, 7691–7695.

Cumming, D. A., Hellerqvist, C. G., Harris-Brandts, M., Michnick, S. W., Carver, J. P. & Bendiak, B. (1989) *Biochemistry 28*, 6500–6512.

Dall'Olio, F., de Kanter, F. J. J., van den Eijnden, D. H. & Serafini-Cessi, F. (1988) *Carbohydr. Res. 178*, 327–332.

Damm, J. B. L., Voshol, H., Hård, K., Kamerling, J. P. & Vliegenthart, J. F. G. (1989a) *Eur. J. Biochem. 180*, 101–110.

Damm, J. B. L., Bergwerff, A. A., Hård, K., Kamerling, J. P. & Vliegenthart, J. F. G. (1989b) *Recl. Trav. Chim. Pays-Bas 108*, 351–359.

DeGasperi, R., Li, Y.-T. & Li, S.-C. (1986) *J. Biol. Chem. 261*, 5696–5698.
De Waard, P., Koorevaar, A., Kamerling, J. P. & Vliegenthart, J. F. G. (1991) *J. Biol. Chem. 266*, 4237–4243.
Donald, A. S. R., Soh, C. P. C., Watkins, W. M. & Morgan, W. T. J. (1982) *Biochem. Biophys. Res. Commun. 104*, 58–65.
Donald, A. S. R., Yates, A. D., Soh, C. P. C., Morgan, W. T. J. & Watkins, W. M. (1983) *Biochem. Biophys. Res. Commun. 115*, 625–631.
Donald, A. S. R. & Feeney, J. (1986) *Biochem. J. 236*, 821–828.
Fletcher, A. P., Neuberger, A. & Ratcliffe, W. A. (1970) *Biochem. J. 120*, 417–424.
Gottschalk, A. (1952) *Nature 170*, 662–663.
Green, E. D., Adelt, G., Baenziger, J. U., Wilson, S. & van Halbeek, H. (1988) *J. Biol. Chem. 263*, 18253–18268.
Griesinger, C., Otting, G., Wüthrich, K. & Ernst, R. R. (1988) *J. Am. Chem. Soc. 110*, 7870–7872.
Hård, K., Mekking, A., Damm, J. B. L., Kamerling, J. P., de Boer, W., Wijnands, R. A. & Vliegenthart, J. F. G. (1990) *Eur. J. Biochem. 193*, 263–271.
Herkt, R., Parente, J. P., Leroy, Y., Fournet, B., Blanchard, D., Cartron, J. P., van Halbeek, H. & Vliegenthart, J. F. G. (1985) *Eur. J. Biochem. 146*, 125–129.
Hokke, C. H., Kamerling, J. P., van Dedem, G. W. K. & Vliegenthart, J. F. G. (1991) *FEBS Lett. 286*, 18–24.
Jeener, J., Meier, B. H., Bachmann, P. & Ernst, R. R. (1979) *J. Chem. Phys. 71*, 4546–4553.
Kamerling, J. P. & Vliegenthart, J. F. G. (1989) in *Clinical biochemistry; principles, methods, applications* (Lawson, A. M., ed.) vol. 1, pp. 175–263, Walter de Gruyter, Berlin.
Krivan, H. C., Roberts, D. D. & Ginsburg, V. (1988) *Proc. Natl Acad. Sci. USA 85*, 6157–6161.
Lamblin, G., Rahmoune, H., Wieruszeski, J.-M., Lhermitte, M., Strecker, G. & Roussel, P. (1991) *Biochem. J. 275*, 199–206.
Lindberg, J., Fasth, A., Norkrans, G. & Nilsson, L.-Å. (1983) *Int. Arch. Allergy Appl. Immunol. 70*, 146–150.
Marion, D. & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967–974.
Marier, R., Fong, E., Jansen, M., Hodson, C. J., Richards, F. & Andriole, V. T. (1978) *J. Infect. Dis. 138*, 781–790.
Moonen, P., Gaffner, R. & Wingfield, P. (1988) *FEBS Lett. 226*, 314–318.
Muchmore, A. V. & Decker, J. M. (1985) *Science 229*, 479–481.
Muchmore, A. V., Shifrin, S. & Decker, J. M. (1987) *J. Immunol. 138*, 2547–2553.
Nasir-Ud-Din, Jeanloz, R., Lamblin, G., Roussel, P., van Halbeek, H., Mutsaers, J. H. G. M. & Vliegenthart, J. F. G. (1986) *J. Biol. Chem. 261*, 1992–1997.
Odin, L. (1952) *Nature 170*, 663–664.
Parkkinen, J., Virkola, R. & Korhonen, T. K. (1988) *Infect. Immun. 56*, 2623–2630.
Patt, S. L. & Sykes, B. D. (1972) *J. Chem. Phys. 56*, 3182–3184.
Pennica, D., Kohr, W. J., Kuang, W.-J., Glaister, D., Aggarwal, B. B., Chen, E. Y. & Goeddel, D. V. (1987) *Science 236*, 83–88.
Piller, F., Blanchard, D., Huet, M. & Cartron, J.-P. (1986) *Carbohydr. Res. 149*, 171–184.
Rindler, M. J., Naik, S. S., Li, N., Hoops, T. C. & Peraldi, M.-N. (1990) *J. Biol. Chem. 265*, 20784–20789.
Ruiz Contreras, R., Kamerling, J. P., Breg, J. & Vliegenthart, J. F. G. (1988) *Carbohydr. Res. 179*, 411–418.
Serafini-Cessi, F. & Dall'Olio, F. (1983) *Biochem. J. 215*, 483–489.
Serafini-Cessi, F., Dall'Olio, F. & Malagolini, N. (1984a) *Biosci. Rep. 4*, 269–274.
Serafini-Cessi, F., Malagolini, N. & Dall'Olio, F. (1984b) *Biosci. Rep. 4*, 973–978.
Serafini-Cessi, F., Malagolini, N. & Dall'Olio, F. (1988) *Arch. Biochem. Biophys. 266*, 573–582.
Sikri, K. L., Foster, C. L., Bloomfield, F. J. & Marshall, R. D. (1979) *Biochem. J. 181*, 525–532.
Smagula, R. M., van Halbeek, H., Decker, J. M., Muchmore, A. V., Moody, C. E. & Sherblom, A. P. (1990) *Glycoconjugate J. 7*, 609–624.
Soh, C. P. C., Morgan, W. T. J., Watkins, W. M. & Donald, A. S. R. (1980) *Biochem. Biophys. Res. Commun. 93*, 1132–1139.
Soh, C. P. C., Donald, A. S. R., Feeney, J., Morgan, W. T. J. & Watkins, W. M. (1989) *Glycoconjugate J. 6*, 319–332.
Spellman, M. W., Basa, L. J., Leonard, C. K., Chakel, J. A., O'Connor, J. V., Wilson, S. & van Halbeek, H. (1989) *J. Biol. Chem. 264*, 14100–14111.
Spiro, R. G. & Bhoyroo, V. D. (1988) *J. Biol. Chem. 263*, 14351–14358.
Takeya, A., Hosomi, O. & Kogure, T. (1987) *J. Biochem. (Tokyo) 101*, 251–259.
Tamm, I. & Horsfall, F. L. (1950) *Proc. Soc. Exp. Biol. Med. 74*, 108–114.
Tamm, I. & Horsfall, F. L. (1952) *J. Exp. Med. 95*, 71–97.
Townsend, R. R. & Hardy, M. R. (1991) *Glycobiology 1*, 139–147.
Tsantoulas, D. C., McFarlane, I. G., Portmann, B., Eddleston, A. L. W. F. & Williams, R. (1974) *Br. Med. J. 4*, 491–494.
Van Pelt, J., van Kuik, J. A., Kamerling, J. P., Vliegenthart, J. F. G., Reuser, A. J. J. & Galjaard, H. (1988) *Eur. J. Biochem. 177*, 327–338.
Vliegenthart, J. F. G., Dorland, L. & van Halbeek, H. (1983) *Adv. Carbohydr. Chem. Biochem. 41*, 209–374.
Weisshaar, G., Hiyama, J., Renwick, A. G. C. & Nimtz, M. (1991) *Eur. J. Biochem. 195*, 257–268.
Williams, J., Marshall, R. D., van Halbeek, H. & Vliegenthart, J. F. G. (1984) *Carbohydr. Res. 134*, 141–155.
Zalc, B., Collet, A., Monge, M., Ollier-Hartmann, M. P., Jacque, C., Hartmann, L. & Baumann, N. A. (1984) *Brain Res. 291*, 182–187.