Eur. J. Biochem. 228, 981–1008 (1995)

Structural analysis of the sialylated N- and O-linked carbohydrate chains of recombinant human erythropoietin expressed in Chinese hamster ovary cells

Sialylation patterns and branch location of dimeric *N*-acetyllactosamine units

Cornelis H. HOKKE[1], Aldert A. BERGWERFF[1], Gijs W. K. VAN DEDEM[2], Johannis P. KAMERLING[1] and Johannes F. G. VLIEGENTHART[1]

[1] Bijvoet Center, Department of Bio-Organic Chemistry, Utrecht University, The Netherlands
[2] Diosynth BV, Oss, The Netherlands

(Received 7 October 1994) – EJB 94 1525/5

The N-linked carbohydrate chains of recombinant human erythropoietin expressed in CHO cells were quantitatively released with peptide-N^4-(*N*-acetyl-β-glucosaminyl)asparagine amidase F, separated from the remaining O-glycoprotein by gel-permeation chromatography, and subsequently fractionated via FPLC on Mono Q, HPLC on Lichrosorb-NH_2 and high-pH anion-exchange chromatography on CarboPac PA1. The purified sialylated oligosaccharides were analyzed by one-dimensional and two-dimensional 500-MHz ^{1}H-NMR spectroscopy. When necessary, oligosaccharides were treated with endo-β-galactosidase (and *N*-acetyl-β-glucosaminidase) followed by ^{1}H-NMR analysis of the incubation products, to obtain additional structural information. Di-, tri-, tri′- and tetraantennary *N*-acetyllactosamine-type oligosaccharides occur which can be completely (major) or partially (minor) sialylated. Three different types of α2-3-linked sialic acids are present, namely, *N*-acetylneuraminic acid (95 %), *N*-glycolylneuraminic acid (2 %) and *N*-acetyl-9-*O*-acetylneuraminic acid (3 %). In the case of partial sialylation, a non-random distribution of the sialic acids over the branches is observed. One or two extra *N*-acetyllactosamine units, being exclusively located in the branches attached to the α1-6-linked Man residue, can be present in completely or partially sialylated di-, tri′-, and tetraantennary oligosaccharides. Tetraantennary oligosaccharides with *N*-acetyllactosamine repeats could be digested quantitatively with endo-β-galactosidase from *Bacteroides fragilis*, whereas under the same conditions tri′antennary oligosaccharides hardly reacted (< 15 %). Using endo-β-galactosidase from *Escherichia freundii*, these tri′antennary oligosaccharides could be digested more extensively (> 75 %). The O-linked carbohydrate chains were released from the O-glycoprotein by alkaline borohydride treatment, and purified via FPLC on Mono Q and HPLC on Lichrosorb-NH_2. Two O-glycans were found, namely, Neu5Acα2-3Galβ1-3GalNAc-ol and Neu5Acα2-3Galβ1-3(Neu5Acα2-6)GalNAc-ol.

Keywords. Erythropoietin; carbohydrate chains; sialic acids; endo-β-galactosidase.

Human erythropoietin (EPO), a glycoprotein that is synthesized mainly by the kidney, stimulates the proliferation and differentiation of erythroid progenitor cells [1]. Its molecular mass is 34–39 kDa [2], and the polypeptide chain has one O-glycosylation site at Ser126 and three N-glycosylation sites at Asn24, Asn38 and Asn83. The carbohydrate chains account for about 40 % of the molecular mass [3]. EPO has been expressed in various heterologous cell systems and the recombinant human glycoprotein (rEPO) is an important therapeutic agent for the treatment of anemia associated with renal failure [4]. It has been shown that the carbohydrate moiety is of great importance for the biological function of EPO or rEPO. Removal or modification of the carbohydrate chains, or prevention of glycosylation at specific sites by site-directed mutagenesis, results in altered *in vivo* and *in vitro* activity [4–10].

Several studies on the structure of the carbohydrate chains of rEPO, expressed in various cell systems like Chinese hamster ovary (CHO) and baby hamster kidney (BHK) cells, have been carried out [11–18]. Although the results of these studies are to some extent similar, discrepancies are observed on a more detailed level, in particular regarding *N*-acetyllactosamine repeats and sialylation. The latter aspect is especially important since the sialic acid residues influence the activity of rEPO as an *in vivo* therapeutic drug. In order to clarify the uncertainties concerning the structure of the carbohydrate chains of rEPO produced in CHO cells, we undertook a detailed structural analysis, revealing the native sialylation patterns and the branch location of the extra-*N*-acetyllactosamine units.

Correspondence to J. P. Kamerling, Bijvoet Center, Department of Bio-Organic Chemistry, Utrecht University, P. O. Box 80.075, NL-3508 TB Utrecht, The Netherlands
Fax: +31 30 540980.

Abbreviations. BHK, baby hamster kidney; CHO, Chinese hamster ovary; EPO, erythropoietin; HOHAHA, homonuclear Hartmann-Hahn; HPAEC, high-pH anion-exchange chromatography; MLEV, composite pulse devised by M. Levitt; Neu5Ac, *N*-acetylneuraminic acid; Neu5Gc, *N*-glycolylneuraminic acid; Neu5,9Ac_2, *N*-acetyl-9-*O*-acetylneuraminic acid; PAD, pulsed amperometric detection; PNGase-F, peptide-N^4-(*N*-acetyl-β-glucosaminyl)asparagine amidase F; rEPO, recombinant erythropoietin.

Enzymes. Peptide-N^4-(*N*-acetyl-β-glucosaminyl)asparagine amidase F (EC 3.5.1.52); endo-β-galactosidase (EC 3.2.1.103); *N*-acetyl-β-glucosaminidase (EC 3.2.1.30).

Note. This work was presented at the XVIIth International Carbohydrate Symposium (Paris) in July 1992.

MATERIALS AND METHODS

Materials. Recombinant human erythropoietin (rEPO), expressed in CHO cells, was a gift of Organon Teknika BV (Boxtel, The Netherlands). Peptide-N^4-(*N*-acetyl-β-glucosaminyl)-asparagine amidase F (PNGase-F) from *Flavobacterium meningosepticum* and endo-β-galactosidase from *Bacteroides fragilis* were obtained from Boehringer Mannheim. Endo-β-galactosidase from *Escherichia freundii* was purchased from ICN Biochemicals, and *N*-acetyl-β-glucosaminidase from jack beans from Sigma.

Liberation of the N-linked carbohydrate chains. Prior to liberation of the carbohydrate chains, rEPO was separated from mannitol, added as a stabilizer, on a Bio-Gel P-4 column (40×2.5 cm, 200−400 mesh, Bio-Rad). The elution was carried out with 25 mM NH_4HCO_3, adjusted to pH 7.0 with HCl, at a flow rate of 25 ml/h, and monitored using refractive index detection. The void volume peak containing the glycoprotein was collected and lyophilized.

The N-linked carbohydrate chains were enzymically released essentially as described [19, 20], but using incubation conditions optimized for rEPO. Briefly, 50 mg rEPO was dissolved in 5 ml 50 mM Tris/HCl pH 8.4, containing 50 mM EDTA. Subsequently, 1 % (by vol.) 2-mercaptoethanol and 1 % (mass/vol.) SDS were added, and the mixture was kept for 3 min at 100 °C. After cooling to room temperature, the sample was diluted twice with incubation buffer, and PNGase-F was added (0.2 U/mg rEPO). The mixture was incubated for 16 h in an end-over-end mixer at ambient temperature, then the solution was kept for 3 min at 100 °C, and after cooling, another aliquot of PNGase-F (0.2 U/mg rEPO) was added. The incubation was continued for 24 h and the deglycosylation was checked by SDS/PAGE on a 15 % slab gel (2.6 % cross-linking) with Coomassie brilliant blue staining.

The digest was fractionated on a Bio-Gel P-100 column (47×2 cm, 200−400 mesh, Bio-Rad), eluted with 25 mM NH_4HCO_3 pH 7.0, at a flow rate of 22 ml/h. The eluent was monitored at 206 nm (Uvicord SII, LKB) and fractions of 7 ml were collected. The void volume fraction containing the N-deglycosylated glycoprotein was kept apart, and the other carbohydrate-containing fractions (orcinol/H_2SO_4) were pooled, lyophilized and desalted on a Bio-Gel P-2 column (45×1 cm, 200−400 mesh, Bio-Rad), using water as eluent.

Liberation of the O-linked carbohydrate chains. The P-100 void volume fraction was lyophilized and suspended (approx. 10 mg/ml) in 0.1 M NaOH, containing 1 M $NaBH_4$. The solution was kept for 24 h at 40 °C, then cooled on ice and neutralized with 4 M HOAc. Boric acid was removed by repetitive co-evaporation with MeOH, containing 1 % HOAc. Finally, the material was resuspended in water and centrifuged at 12000×*g*, and the supernatant was desalted on a Bio-Gel P-2 column (45×1 cm, 200−400 mesh, Bio-Rad) [21].

FPLC fractionation. Enzymically and chemically released oligosaccharide pools were fractionated on a Mono Q HR 5/5 anion-exchange column using a Pharmacia FPLC system [21]. Fractions were eluted with a gradient of NaCl in H_2O as indicated in the figures, at a flow rate of 2 ml/min. The eluent was monitored at 214 nm and collected fractions were lyophilized, desalted on a Bio-Gel P-2 column (20×1 cm, 200−400 mesh, Bio-Rad) and lyophilized again.

HPLC fractionation. The carbohydrate-containing Mono Q fractions were subfractionated by HPLC on a 10-μm Lichrosorb-NH_2 column (25×0.46 cm, Chrompack) using a Spectroflow 400 HPLC system (ABI Analytical, Kratos Division). Elutions were performed with 30 mM K_2HPO_4/KH_2PO_4, pH 7.0/acetonitrile in a ratio as indicated in the figures, at a flow rate of 2 ml/min. The fractionation was monitored at 205 nm. Fractions were dried under a stream of nitrogen, desalted on a Bio-Gel P-2 column (20×1 cm, 200−400 mesh, Bio-Rad) and lyophilized.

High-pH anion-exchange chromatography. Some of the HPLC fractions were further fractionated by high-pH anion-exchange chromatography (HPAEC) with pulsed amperometric detection (PAD) 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 (25×0.9 cm, Dionex). Elutions were performed with a concentration gradient of NaOAc in 0.1 M NaOH as indicated in the figures, at a flow rate of 4 ml/min. Detection was performed using the following pulse potentials and durations for the PAD detector: $E_1 = 0.05$ V (300 ms); $E_2 = 0.65$ V (60 ms); $E_3 = -0.95$ V (180 ms). Collected fractions were immediately neutralized by addition of 2 M HCl, lyophilized, desalted on a column of Bio-Gel P-2 (45×1 cm, 200−400 mesh, Bio-Rad) eluted with 5 mM NH_4HCO_3, and lyophilized again.

Monosaccharide analysis. Monosaccharides were analyzed by gas chromatography on a capillary CP-Sil 5 WCOT fused silica column (25 m×0.32 mm, Chrompack) using a Varian Aerograph 3700 gas chromatograph. The trimethylsilylated (methyl ester) methyl glycosides were prepared by methanolysis, *N*-(re)acetylation, and trimethylsilylation as reported [22].

Sialic acid analysis. Sialic acids were released by hydrolysis in 2 M acetic acid and converted into fluorescent derivatives with 1,2-diamino-4,5-methylene-dioxybenzene [23]. The sialic acid derivatives were analyzed by reverse-phase HPLC on a Chromspher C_{18} column (25×0.46 cm, Chrompack) as described [24].

^{1}H-NMR spectroscopy. Prior to ^{1}H-NMR spectroscopic analysis, samples were treated twice with 99.9 % ^2H_2O at p^{2}H 7 with intermediate lyophilization. Finally, samples were dissolved in 99.96 % ^2H_2O (MSD Isotopes). ^{1}H-NMR spectra were recorded at 500 MHz on a Bruker AM-500 spectrometer at a probe temperature of 27 °C (N-linked carbohydrate chains) or 22 °C (O-linked carbohydrate chains). Chemical shifts are expressed in ppm by reference to internal acetone (δ 2.225) [25]. Typically, one-dimensional spectra were recorded with a spectral width of 5000 Hz, collecting 64−2500 free induction decays of 8K or 16K complex data points. Suppression of the residual water signal was achieved by applying the water-eliminated Fourier transform pulse sequence as described [26]. The resolution of the one-dimensional spectra was enhanced by Lorentzian-to-Gaussian transformation and the final spectra were baseline-corrected with a polynomal function when necessary.

For the two-dimensional homonuclear Hartmann-Hahn (HOHAHA) measurement [27] a MLEV-17 mixing sequence of 120 ms was used. The 90° pulse width was adjusted to about 26 μs and the spectral width was 3500 Hz in both dimensions. The HO^{2}H signal was presaturated for 1 s during the relaxation delay. In total 360 spectra of 2048 data points with 96 scans/t_1 value were recorded. The two-dimensional NMR data were processed on a VAXstation 3100 using Triton software (Bijvoet Center, Department of NMR spectroscopy, Utrecht University). The time domain data were multiplied with a phase-shifted sine bell. After Fourier transformation, the resulting data set of 1024×2048 points was baseline-corrected in both frequency domains with a fourth-order polynomal fit.

Quantification of oligosaccharides. The molar ratio of the FPLC fractions was calculated from the FPLC peak areas on the basis of the weighted average number of C=O groups (responsible for absorption at 214 nm) being known after structural identification and determination of the relative amounts of each individual component. The molar ratio of the constituent oligosaccharides within each FPLC fraction was determined from the

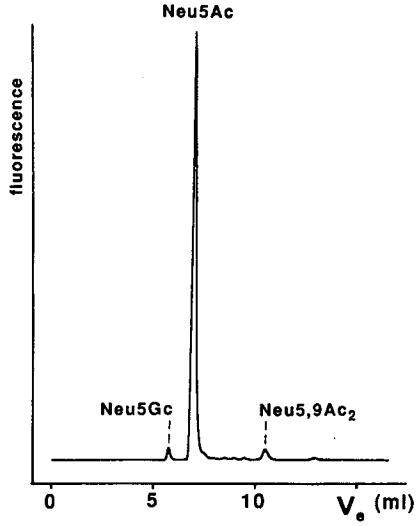


Fig. 1. HPLC profile of the 1,2-diamino-4,5-methylene-dioxybenzene derivatives of the sialic acids released from recombinant human erythropoietin expressed in CHO cells. The Chromspher C_{18} column (25×0.46 cm) was eluted with acetonitrile/methanol/water (9:7:84, by vol.), at a flow rate of 1 ml/min.

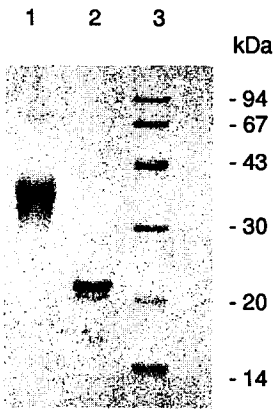


Fig. 2. SDS/PAGE of recombinant human erythropoietin from CHO cells on a 15 % slab gel. The gel was stained with Coomassie brilliant blue. Lane 1, native rEPO; lane 2, N-deglycosylated rEPO; lane 3, molecular mass markers.

HPLC peak areas (corrected for the number of C=O groups) at 205 nm [19]. When HPAEC/PAD was used for further fractionation of HPLC fractions, the PAD response was assumed to be equal for each oligosaccharide present within one HPLC fraction. For mixtures, molar ratios were determined on basis of the ^{1}H-NMR spectra.

Endo-β-galactosidase digestions. Digestions with endo-β-galactosidase from *B. fragilis*, followed by purification of the digestion products on Mono Q and subsequent treatment with *N*-acetyl-β-glucosaminidase, were performed as described [28]. Digestions with endo-β-galactosidase from *E. freundii* were carried out using the same incubation conditions as for endo-β-galactosidase from *B. fragilis*.

RESULTS

Monosaccharide analysis of rEPO revealed the presence of Fuc, Gal, GlcNAc, Man and Neu5Ac in the molar ratio of 1.0:4.3:6.0:3.0:4.1, and a trace amount of GalNAc. Neu5Ac

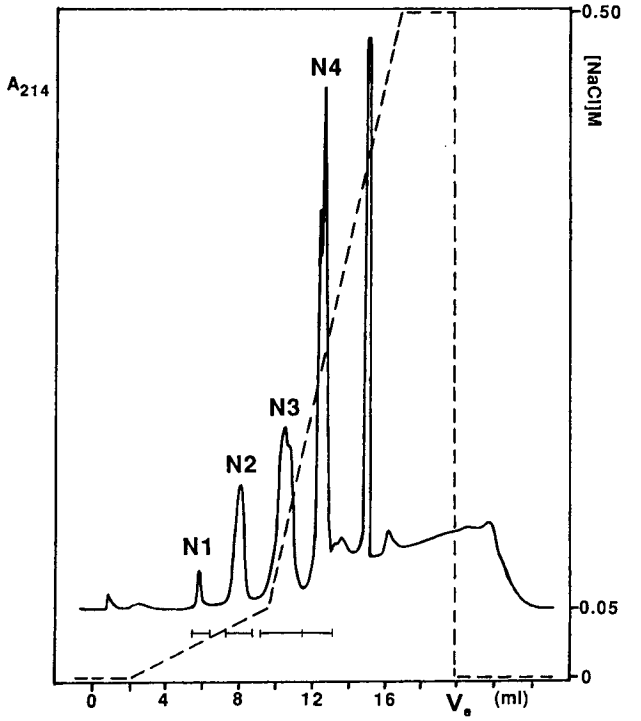


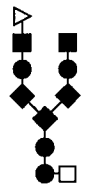
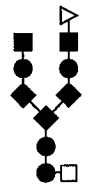
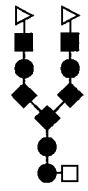
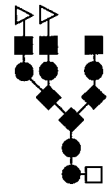
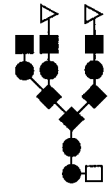
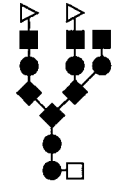
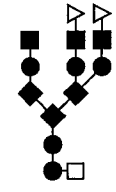
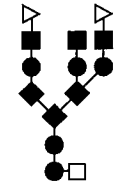
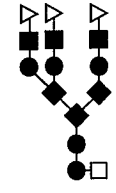
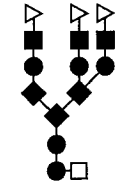
Fig. 3. Fractionation pattern at 214 nm of the PNGase-F-released carbohydrate chains of rEPO on a FPLC HR 5/5 Mono Q column. The column was first eluted with 2 ml H_2O, followed by a linear concentration gradient of 0–50 mM NaCl in 8 ml H_2O, and finally by a steeper gradient of 50–500 mM NaCl in 8 ml H_2O, at a flow rate of 2 ml/min.

represents the total sialic acid content. The value of GlcNAc has been corrected for the relatively stable GlcNAc-Asn linkage. Sialic acid analysis of rEPO demonstrated the occurrence of Neu5Ac, Neu5Gc and $Neu5,9Ac_2$ in the molar ratio of 95:2:3 (Fig. 1). Analysis by SDS/PAGE of a sample of the PNGase-F digestion mixture taken after a 20-h incubation indicated a complete release of the N-linked carbohydrate chains (Fig. 2). The molecular mass of rEPO shifted from 35 kDa (native) to 22 kDa (N-deglycosylated).

N-linked carbohydrate chains. FPLC on Mono Q of the Bio-Gel P-100 pool of PNGase-F released carbohydrate chains yielded four carbohydrate-containing peaks, denoted **N1**, **N2**, **N3** and **N4**, having elution positions corresponding to mono-, di-, tri- and tetrasialylated *N*-acetyllactosamine-type N-linked oligosaccharides, respectively (Fig. 3).

HPLC on Lichrosorb-NH_2 yielded for fraction **N1** three subfractions **N1.1–N1.3** (Fig. 4A), for **N2** ten subfractions **N2.1–N2.10** (Fig. 4B), for **N3** thirteen subfractions **N3.1–N3.13** (Fig. 4C), and for **N4** eight subfractions **N4.1–N4.8** (Fig. 4D). The HPLC fractions were investigated by 500-MHz ^{1}H-NMR spectroscopy and, if necessary, further fractionated by HPAEC. Some fractions containing compounds with *N*-acetyllactosamine repeats were treated with endo-β-galactosidase to generate specific digestion products, which were subsequently purified by anion-exchange chromatography on Mono Q, and then analyzed by ^{1}H-NMR spectroscopy [28]. ^{1}H-NMR data of di-, tri-, and tri'antennary oligosaccharides, and of tetraantennary oligosaccharides, all without *N*-acetyllactosamine repeats, are compiled in Tables 1 and 2, respectively. The spectral data of oligosaccharides with *N*-acetyllactosamine repeats are presented in Table 3, and those of endo-β-galactosidase digestion products in Table 4.

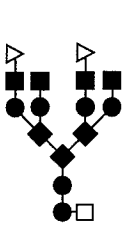
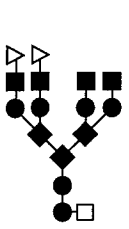
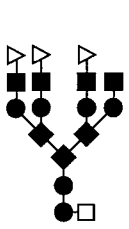
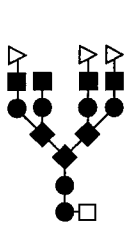
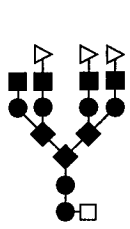
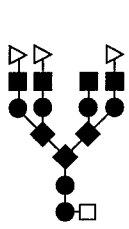
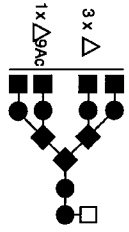
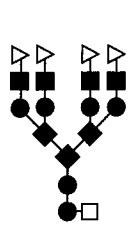
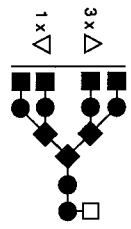
Table 1. ^{1}H-chemical shifts of the structural-reporter-group protons of the constituent monosaccharides of di-, tri- and tri′antennary N-linked oligosaccharides, derived from recombinant human erythropoietin. Chemical shifts are given at 27°C and were measured in ^2H_2O relative to acetone (δ 2.225 [25]). Compounds are represented by short-hand symbolic notation: □, Fuc; ■, Gal; ●, GlcNAc; ◆, Man; △, Neu5Ac. For numbering of the monosaccharide residues, see text. n.d., not determined; α and β stand for the anomeric configuration of GlcNAc-**1**.

Reporter group	Residue	Chemical shift in									
											
		N1.1A	**N1.1B**	**N2.4**	**N2.6B**	**N2.6C**	**N2.7.2A**	**N2.7.2B**	**N2.7.3**	**N3.4**	**N3.5**
		ppm									
H1	GlcNAc-**1**α	5.181	5.181	5.181	5.181	5.181	5.182	5.182	5.182	5.184	5.182
	β	4.693	4.693	4.692	4.691	4.691	4.691	4.691	4.690	4.692	4.690
	GlcNAc-**2**α	4.664	4.664	4.663	4.663	4.663	4.661	4.661	4.662	4.663	4.660
	β	4.664	4.664	4.666	4.663	4.663	4.661	4.661	4.662	4.663	4.664
	Man-**4**	5.120	5.120	5.116	5.117	5.117	5.126	5.126	5.123	5.114	5.123
	Man-**4′**	4.924	4.924	4.921	4.921	4.903	4.867	4.867	4.874	4.904	4.871
	GlcNAc-**5**	4.573	4.582	4.573	4.563	4.563	4.578	4.578	4.577	4.559	4.577
	GlcNAc-**5′**	4.582	4.573	4.573	4.579	4.579	4.582	4.582	4.595	5.572	4.587
	Gal-**6**	4.543	4.467	4.544	4.544	4.544	4.545	4.466	4.545	4.543[a]	4.544
	Gal-**6′**	4.472	4.547	4.549	4.472	4.544	4.545	4.545	4.469	4.548[a]	4.544
	GlcNAc-**7**	–	–	–	4.544	4.544	–	–	–	4.543[a]	–
	GlcNAc-**7′**	–	–	–	–	–	4.545	4.545	4.545	–	4.544
	Gal-**8**	–	–	–	4.544	4.464	–	–	–	4.548[a]	–
	Gal-**8′**	–	–	–	–	–	4.481	4.560	4.561	–	4.560
	Fucα	4.890	4.890	4.892	4.891	4.891	4.899	4.899	4.895	4.892	4.899
	β	4.898	4.898	4.900	4.899	4.899	4.906	4.906	4.903	4.900	4.907
H2	Man-**3**	4.249	4.249	4.248	4.213	4.213	4.249	4.249	4.252	4.210	4.250
	Man-**4**	4.191	4.191	4.191	4.213	4.213	4.198	4.198	4.196	4.210	4.197
	Man-**4′**	4.106	4.106	4.105	4.105	4.105	4.090	4.090	4.090	4.104	4.090

H3	Gal-**6**	4.118	n.d.	4.118	4.114	n.d.	4.117	n.d.	4.117	4.116	4.119
	Gal.**6′**	n.d.	4.118	4.118	n.d.	4.114	4.117	4.117	n.d.	4.116	4.119
	Gal-**8**	–	–	–	4.114	4.114	–	–	–	4.116	–
	Gal-**8′**	–	–	–	–	–	n.d.	4.117	4.117	–	4.119
H3a	Neu5Ac	1.797	1.797	1.798/803	1.800[b]	1.800[b]	1.802[b]	1.802[b]	1.800[b]	1.802[c]	1.803[c]
H3e	Neu5Ac	2.757	2.757	2.758[b]	2.757[b]	2.757[b]	2.757[b]	2.757[b]	2.757[b]	2.756[c]	2.757[c]
H5	Fucα	4.099	4.099	4.096	4.10[d]	4.10	4.10	4.10	4.10	4.097	4.10
	β	4.13	4.13	4.135	4.13	4.13	4.13	4.13	4.13	4.13	4.13
NAc	GlcNAc-**1**	2.039	2.039	2.040	2.038	2.038	2.038	2.038	2.039	2.038	2.039
	GlcNAc-**2**α	2.095	2.095	2.097	2.095	2.095	2.094	2.094	2.094	2.095	2.094
	β	2.095	2.095	2.094	2.095	2.095	2.094	2.094	2.094	2.095	2.091
	GlcNAc-**5**	2.048	2.048	2.048	2.044	2.044	2.052	2.052	2.051	2.043	2.052
	GlcNAc-**5′**	2.039	2.048	2.044	2.044	2.044	2.039	2.039	2.042	2.043	2.039
	GlcNAc-**7**	–	–	–	2.074	2.074	–	–	–	2.073	–
	GlcNAc-**7′**	–	–	–	–	–	2.039	2.039	2.039	–	2.039
	Neu5Ac	2.031	2.031	2.032[e]	2.031[e]	2.031[e]	2.031[e]	2.031[e]	2.031[e]	2.031[f]	2.031[f]
CH_3	Fucα	1.209	1.209	1.212	1.210	1.210	1.211	1.211	1.209	1.211	1.211
	β	1.221	1.221	1.223	1.223	1.223	1.222	1.222	1.221	1.223	1.223

[a] Values may have to be interchanged.
[b] Signal stemming from two protons.
[c] Signal stemming from three protons.
[d] Some values are given with only two decimals because of spectral overlap.
[e] Signal stemming from two NAc groups.
[f] Signal stemming from three NAc groups.

Table 2. ^{1}H-chemical shifts of the structural-reporter-group protons of the constituent monosaccharides of tetraantennary N-linked oligosaccharides derived from recombinant erythropoietin. Chemical shifts are given at 27 °C and were measured in ^2H_2O relative to acetone (δ 2.225 [25]). For the short-hand symbolic notation, see Table 1. (△9Ac, Neu5,9Ac$_2$; ▽, Neu5Gc). For numbering of the monosaccharide residues, see text. n.d., not determined; α and β stand for the anomeric configuration of GlcNAc-**1**.

Reporter group	Residue	Chemical shift in								
										
		N2.10.1	**N2.10.2**	**N3.7.1A**	**N3.7.1B**	**N3.7.1.C**	**N3.7.2A**	**N4.2**	**N4.4**	**N4.5**
		ppm								
H1	GlcNAc-**1**α	5.181	5.182	5.182	5.182	5.182	5.182	5.183	5.182	5.183
	β	4.695	4.691	4.690	4.690	4.690	4.691	4.69[a]	4.688	4.689
	GlcNAc-**2**α	4.66	4.661	4.660	4.660	4.660	4.662	4.661	4.659	4.66
	β	4.66	4.664	4.663	4.663	4.663	4.665	4.661	4.664	4.66
	Man-**4**	5.129	5.126	5.125	5.125	5.125	5.129	5.131	5.131	5.132
	Man-**4′**	4.858	4.867	4.857	4.857	4.857	4.870	4.858	4.857	4.857
	GlcNAc-**5**	4.563	4.565	4.566	4.566	4.566	4.561	4.564	4.563	4.563
	GlcNAc-**5′**	4.59	4.60	4.59	4.59	4.59	4.60	4.59	4.593	4.593
	Gal-**6**	4.470	4.545	4.542[b]	4.472	4.542[b]	4.546	4.544	4.542[b]	4.545
	Gal-**6′**	4.548	4.468	4.546[b]	4.546[b]	4.546[b]	4.470	4.544	4.545[b]	4.545
	GlcNAc-**7**	4.548	4.545	4.542[b]	4.542[b]	4.542[b]	4.546	4.544	4.542[b]	4.545
	GlcNAc-**7′**	4.548	4.545	4.546[b]	4.546[b]	4.546[b]	4.546	4.544	4.545[b]	4.545
	Gal-**8**	4.548	4.545	4.546[b]	4.542[b]	4.465	4.546	4.544	4.542[b]	4.545
	Gal-**8′**	4.481	4.483	4.482	4.560	4.560	4.561	4.560	4.559	4.560
	Fucα	4.900	4.896	4.899	4.899	4.899	4.899	4.902	4.901	4.902
	β	4.908	4.903	4.906	4.906	4.906	4.906	4.910	4.909	4.909
H2	Man-**3**	4.21	4.21	4.21	4.21	4.21	4.20	4.20	4.203	4.207
	Man-**4**	4.22	4.22	4.22	4.22	4.22	4.22	4.22	4.220	4.223
	Man-**4′**	4.092	4.091	4.091	4.091	4.091	4.093	4.091	4.090	4.091
H3	Gal-**6**	n.d.	4.116	4.117	n.d.	4.117	4.117	4.117	4.117	4.115
	Gal-**6′**	4.116	n.d.	4.117	4.117	4.117	n.d.	4.117	4.117	4.115
	Gal-**8**	4.116	4.116	4.117	4.117	n.d.	4.117	4.117	4.117	4.115
	Gal-**8′**	n.d.	n.d.	n.d.	4.117	4.117	4.117	4.117	4.117	4.115
H3a	Neu5Ac	1.800[c]	1.801[c]	1.803[d]	1.803[d]	1.803[d]	1.803[d]	1.804[d]	1.804[e]	1.804[d]
	Neu5Gc	–	–	–	–	–	–	–	–	1.821
	Neu5,9Ac$_2$	–	–	–	–	–	–	1.804	–	–
H3e	Neu5Ac	2.757[c]	2.756[c]	2.757[d]	2.757[d]	2.757[d]	2.757[d]	2.757[d]	2.756[e]	2.757[d]
	Neu5Gc	–	–	–	–	–	–	–	–	2.774
	Neu5,9Ac$_2$	–	–	–	–	–	–	2.757	–	–

H5	Fucα	4.099	4.10	4.10	4.10	4.10	4.10	4.097	4.097	4.098
	β	4.13	4.13	4.13	4.13	4.13	4.13	4.13	4.133	4.13
H7	Neu5,9Ac$_2$	–	–	–	–	–	–	3.65[f]	–	–
H8	Neu5,9Ac$_2$	–	–	–	–	–	–	4.11[f]	–	–
H9	Neu5,9Ac$_2$	–	–	–	–	–	–	4.20[f]	–	–
H9′	Neu5,9Ac$_2$	–	–	–	–	–	–	4.415	–	–
NAc	GlcNAc-**1**	2.038	2.039	2.039	2.039	2.039	2.039	2.039	2.038	2.039
	GlcNAc-**2**α	2.093	2.093	2.094	2.094	2.094	2.094	2.094	2.094	2.094
	β	2.093	2.093	2.090	2.090	2.090	2.094	2.094	2.090	2.094
	GlcNAc-**5**	2.050	2.048	2.049	2.049	2.049	2.049	2.048	2.047	2.048
	GlcNAc-**5′**	2.038	2.039	2.039	2.039	2.039	2.039	2.039	2.038	2.039
	GlcNAc-**7**	2.078	2.075	2.076	2.076	2.076	2.075	2.075	2.075	2.075
	GlcNAc-**7′**	2.038	2.038	2.039	2.039	2.039	2.039	2.038	2.038	2.039
	Nue5Ac	2.031[g]	2.031[g]	2.031[h]	2.031[h]	2.031[h]	2.032[h]	2.031[h]	2.031[i]	2.031[h]
	Neu5,9Ac$_2$	–	–	–	–	–	–	2.038	–	–
NGc	Neu5Gc	–	–	–	–	–	–	–	–	4.121
OAc	Neu5,9Ac$_2$	–	–	–	–	–	–	2.143	–	–
CH_3	Fucα	1.211	1.209	1.212	1.212	1.212	1.210	1.212	1.211	1.212
	β	1.223	1.221	1.223	1.223	1.223	1.222	1.223	1.222	1.223

[a] Some values are given with only two decimals because of spectral overlap.
[b] Values may have to be interchanged.
[c] Signal stemming from two protons.
[d] Signal stemming from three protons.
[e] Signal stemming from four protons.
[f] Value obtained from a 2D-HOHAHA spectrum.
[g] Signal stemming from two NAc groups.
[h] Signal stemming from three NAc groups.
[i] Signal stemming from four NAc groups.

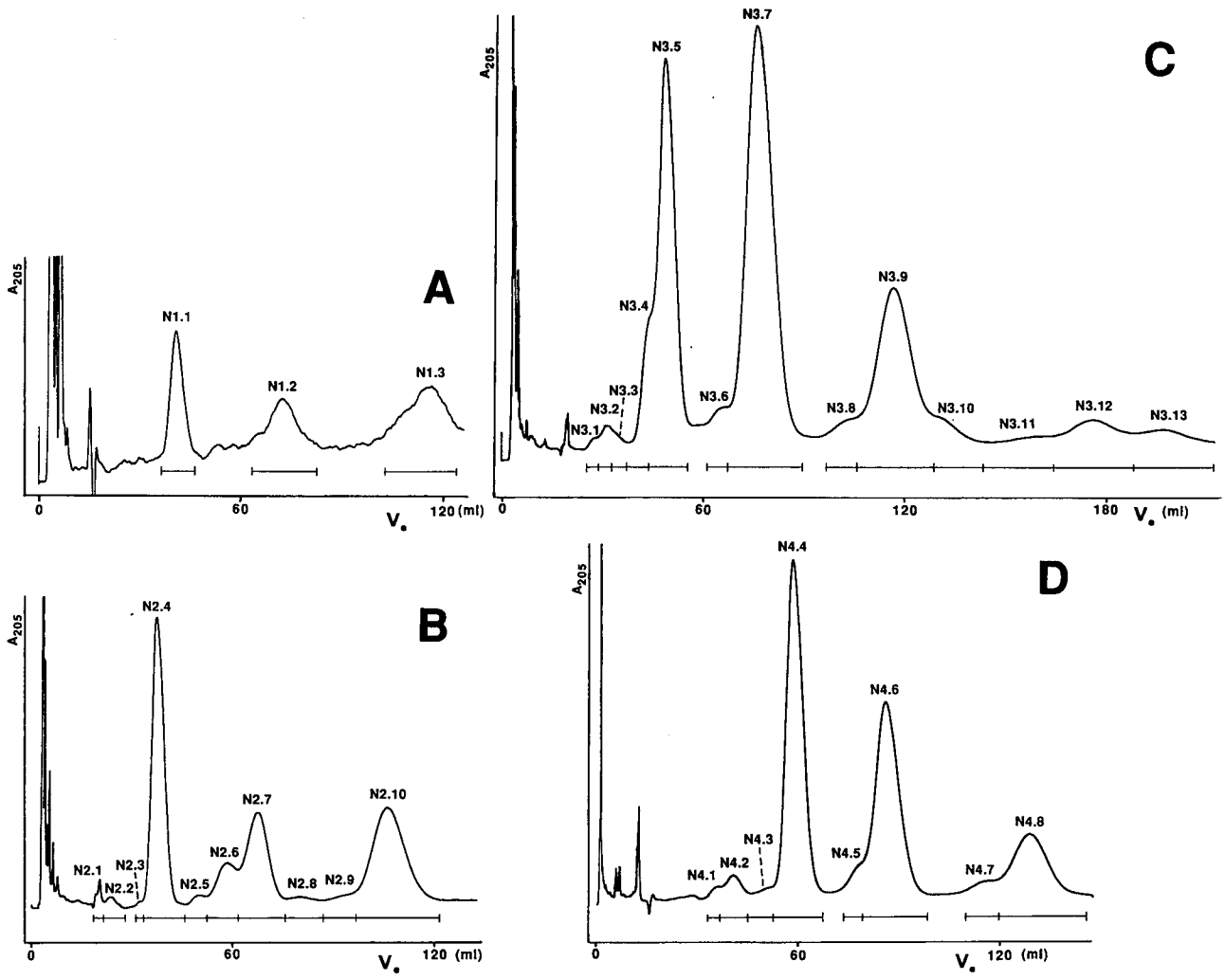


Fig. 4. Fractionation patterns at 205 nm of FPLC fractions N1 (A), N2 (B), N3 (C) and N4 (D), derived from rEPO, on a Lichrosorb-NH_2 column (25×0.46 cm). The column was eluted with 30 mM K_2HPO_4/KH_2PO_4 pH 7.0/acetonitrile (37.5:62.5, by vol.), at a flow rate of 2 ml/min.

The structure determination of the compounds present in the Lichrosorb-NH_2 fractions will be described in an order which is convenient for discussing the ^{1}H-NMR data. The numbering of the monosaccharide residues is exemplified in the structure of compound **N4.8**. The structures of the identified oligosaccharides are summarized in Table 5, together with their relative amounts. Fractions not discussed below did not contain enough material, or were mixtures too heterogeneous for structure determination by ^{1}H-NMR spectroscopy.

All N-linked oligosaccharides identified have the α1-6-fucosylated *N*, *N'*-diacetylchitobiose element at their reducing ends in common. This unit is recognized from the anomeric signals of αGlcNAc-**1**, βGlcNAc-**1**, GlcNAc-**2**, and Fuc, together with the NAc CH_3 signals of GlcNAc-**1** and GlcNAc-**2**, as well as from the Fuc CH_3 signals [26, 29]. Furthermore, the sialic acids occur exclusively in α2-3-linkage to Galβ1-4GlcNAc, as deduced from the H3 and NAc signals of Neu5Ac [25], the H3 and NGc signals of Neu5Gc [24], the H3, NAc and OAc signals of Neu5,9Ac_2 [30], and the H3 signal of the Gal residue to which the sialic acid residue is linked [25]. Throughout this paper, the term 'sialylated' is used for 'substituted with Neu5Ac', unless indicated otherwise.

The ^{1}H-NMR spectrum of fraction **N1.1** shows the presence of a mixture of two monosialylated diantennary oligosaccharides. Compound **N1.1A** carries the Neu5Ac residue on the Manα1-3 branch (cf. compound N1.3 in [26]), whereas in compound **N1.1B**, the Neu5Ac residue is located in the Manα1-6 branch (cf. compound C-Q1-1′ in [31]). The ^{1}H-NMR spectrum of fraction **N2.4**, being the largest **N2** subfraction, indicates the occurrence of a disialylated diantennary oligosaccharide (cf. compound N2.5 in [26]). According to the ^{1}H-NMR spectra, fraction **N3.4** contains a trisialylated triantennary oligosaccharide (cf. compound N6.3.1 in [26]), whereas fraction **N3.5** contains a trisialylated tri'antennary oligosaccharide (cf. compound N3.5 in [20]), which is in fact the main component of fraction **N3**. The ^{1}H-NMR spectrum of fraction **N4.4** demonstrates the presence of a tetrasialylated tetraantennary oligosaccharide, which is the major constituent of all N-linked carbohydrate chains in this study. The spectral data of **N4.4** are in accordance with those of compound N4.4.5 in [26].

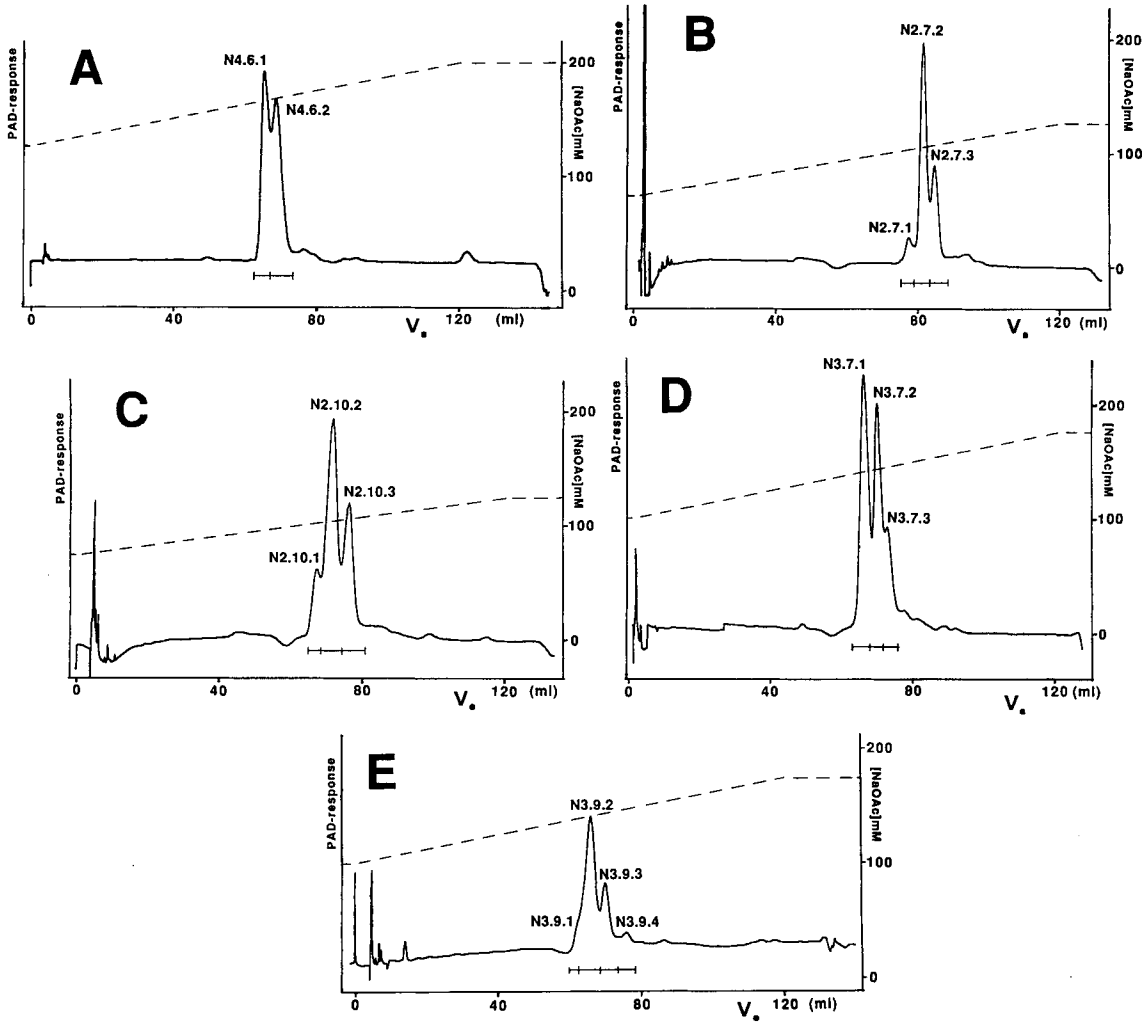


Fig. 5. Fractionation patterns of HPLC fractions N4.6 (A), N2.7 (B), N2.10 (C), N3.7 (D) and N3.9 (E) on a semi-preparative CarboPac PA1 column (25×0.9 cm). The column was eluted with 0.1 M NaOH, at a flow rate of 4 ml/min using a NaOAc gradient as indicated in the figures.

N1.1A

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

N1.1B

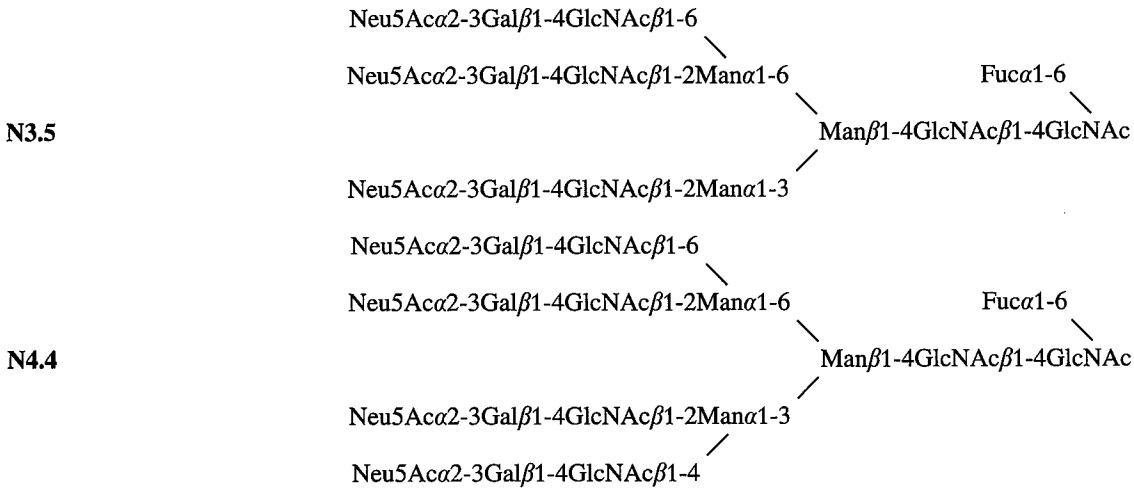
Neu5Acα2-3Galβ1-4GlcNAcβ1-2Manα1-6 Fucα1-6
Manβ1-4GlcNAcβ1-4GlcNAc
Galβ1-4GlcNAcβ1-2Manα1-3

N2.4

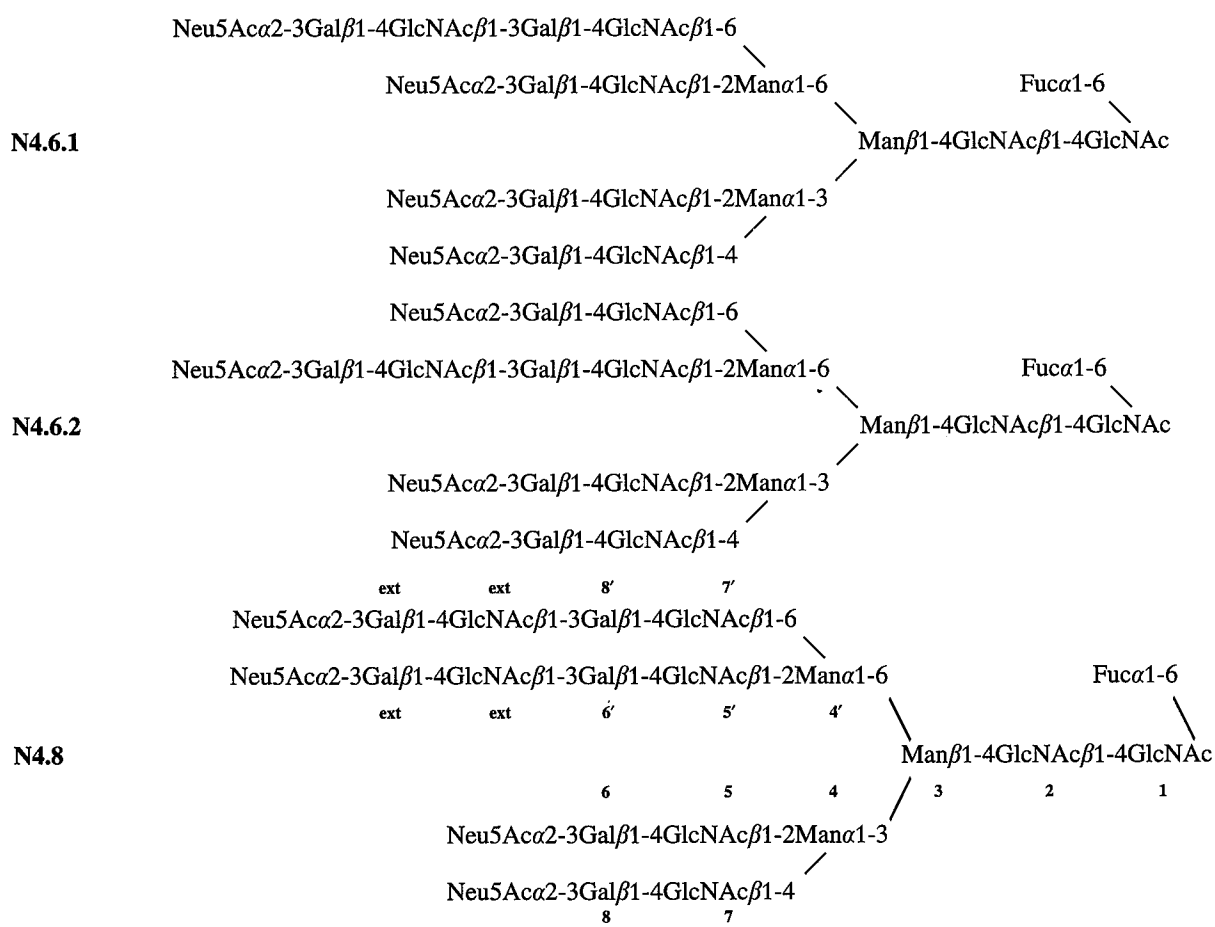
Neu5Acα2-3Galβ1-4GlcNAcβ1-2Manα1-6 Fucα1-6
Manβ1-4GlcNAcβ1-4GlcNAc
Neu5Acα2-3Galβ1-4GlcNAcβ1-2Manα1-3

N3.4

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



From the ^{1}H-NMR spectrum of fraction **N4.6**, the occurrence of a mixture of tetrasialylated tetraantennary oligosaccharides, each containing one *N*-acetyllactosamine repeat, was deduced. The components in fraction **N4.6** were separated using HPAEC, affording the subfractions **N4.6.1** and **N4.6.2** (Fig. 5A). As has been proven earlier, the ^{1}H-NMR spectrum of fraction **N4.6.1** demonstrates the presence of a tetrasialylated tetraantennary oligosaccharide having an additional *N*-acetyllactosamine unit linked to Gal-**8′**, whereas the NMR data of fraction **N4.6.2** show that it contains an isomer of **N4.6.1** having the extra unit linked to Gal-**6′** (cf. compounds B and C in [28]). The branch locations of the repeats are reflected by the typical H1 signals of Gal-**8′** at δ 4.467 in **N4.6.1** and of Gal-**6′** at δ 4.455 in **N4.6.2**. Previously, the ^{1}H-NMR spectrum of fraction **N4.8** has been proven to reflect the presence of a tetrasialylated tetraantennary oligosaccharide with two elongated branches linked to Man-**4′** (cf. compound D in [28]). The branch location of the *N*-acetyllactosamine repeats is shown by the Gal-**6′** and Gal-**8′** H1 signals at δ 4.454 and δ 4.468, respectively.



In fact, the determination of the branch locations of the additional *N*-acetyllactosamine units in **N4.6.1**, **N4.6.2** and **N4.8** was based on the ^{1}H-NMR analysis of the digestion products obtained after subsequent endo-β-galactosidase (*B. fragilis*) and *N*-acetyl-β-glucosaminidase (jack bean) treatments [28]. This made the assignment of the specific structural reporters of the intact

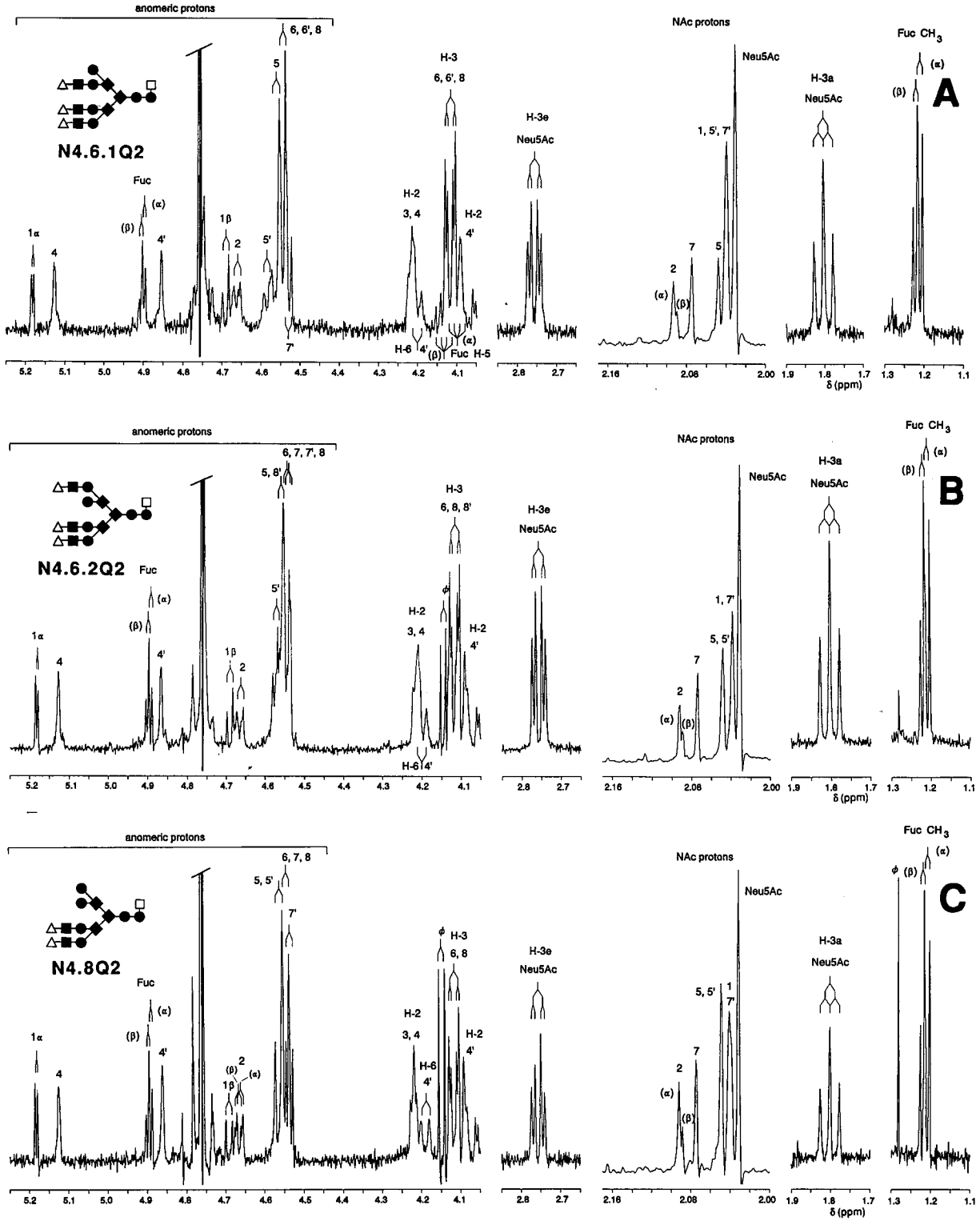
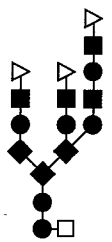
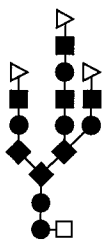
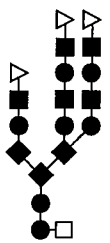
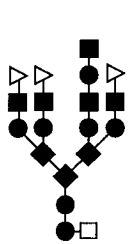
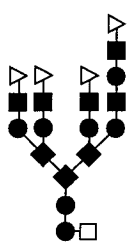
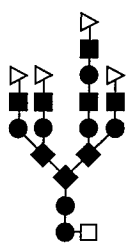
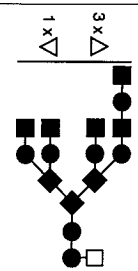
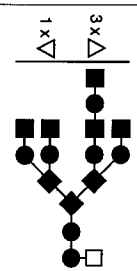
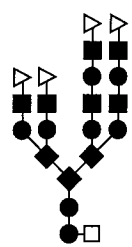


Fig. 6. Structural-reporter-group regions of the 500-MHz ^{1}H-NMR spectra of fractions N4.6.1Q2 (A), N4.6.2Q2 (B) and N4.8Q2 (C), recorded in ^2H_2O at 27°C. The relative intensity of the NAc proton region differs from that of the rest of the spectrum. ϕ denotes impurity. H6 of **4'** is not given in the tables since it is not considered to be a structural-reporter-group signal.

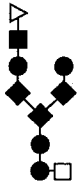
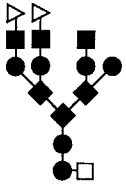
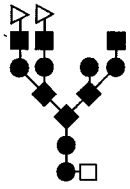
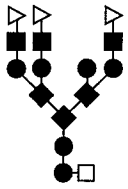
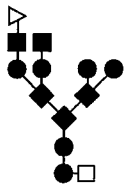
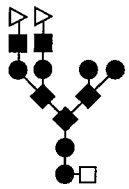
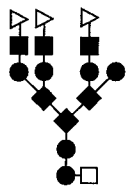
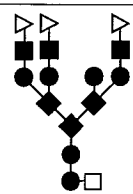
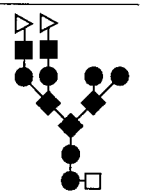
Table 3. ^{1}H-chemical shifts of the structural-reporter-group protons of the constituent monosaccharides of tri'- and tetraantennary N-linked oligosaccharides, containing *N*-acetyllactosamine repeats, derived from recombinant human erythropoietin. Chemical shifts are given at 27°C and were measured in ^2H_2O relative to acetone (δ 2.225 [25]). For the short-hand symbolic notation, see Tables 1 and 2. For numbering of the monosaccharide residues, see text. n.d., not determined; α and β stand for the anomeric configuration of GlcNAc-**1**.

Reporter group	Residue	Chemical shift in								
										
		N3.7.2B	**N3.7.3**	**N3.9.3**	**N3.10**	**N4.6.1**	**N4.6.2**	**N4.7A**	**N4.7B**	**N4.8**
		ppm								
H1	GlcNAc-**1**α	5.183	5.182	5.182	5.182	5.182	5.182	5.183	5.183	5.182
	β	4.69[a]	4.69	4.69	4.69	4.690	4.689	4.689	4.689	4.689
	GlcNAc-**2**α	4.66	4.66	4.660	4.66	4.659	4.659	4.661	4.661	4.659
	β	4.66	4.66	4.660	4.66	4.663	4.664	4.664	4.664	4.664
	Man-**4**	5.125	5.123	5.126	5.131	5.130	5.130	5.130	5.130	5.129
	Man-**4'**	4.867	4.870	4.866	4.866	4.854	4.858	4.858	4.858	4.855
	GlcNAc-**5**	4.58	4.577	4.578	4.560	4.564	4.563	4.564	4.564	4.561
	GlcNAc-**5'**	4.59	4.59	4.59	4.60	4.589	4.602	4.59	4.60	4.60
	Gal-**6**	4.544	4.544	4.543	4.544	4.542[b]	4.542[b]	4.543	4.543	4.541[b]
	Gal-**6'**	4.544	4.455	4.454	4.456	4.547[b]	4.455	4.543	4.456	4.454
	GlcNAc-**7**	–	–	–	4.544	4.542[b]	4.542[b]	4.543	4.543	4.541[b]
	GlcNAc-**7'**	4.544	4.544	4.543	4.544	4.547[b]	4.545[b]	4.543	4.543	4.546[b]
	Gal-**8**	–	–	–	4.544	4.542[b]	4.542[b]	4.543	4.543	4.541[b]
	Gal-**8'**	4.469	4.560	4.468	4.560	4.467	4.557	4.469	4.557	4.468
	$GlcNAc_{ext}$	4.700	4.700	4.699[c]	4.70	4.701	4.699	4.699	4.699	4.700[c]
	Gal_{ext}	4.558	4.560	4.556[c]	4.479	4.556	4.557	4.557	4.557	4.556[c]
	Fucα	4.898	4.895	4.895	4.899	4.900	4.898	4.900	4.900	4.897
	β	4.906	4.903	4.903	4.907	4.908	4.905	4.908	4.908	4.905
H2	Man-**3**	4.247	4.251	4.252	4.21	4.206	4.205	4.21	4.21	4.21
	Man-**4**	4.199	4.199	4.199	4.22	4.219	4.219	4.22	4.22	4.22
	Man-**4'**	4.091	4.08	4.080	4.08	4.091	4.081	4.09	4.08	4.080
H3	Gal-**6**	4.115	4.117	4.115	4.117	4.117	4.116	4.117	4.117	4.116
	Gal-**6'**	4.115	n.d.	n.d.	n.d.	4.117	n.d.	4.177	n.d.	n.d.
	Gal-**8**	–	–	–	4.117	4.117	4.116	4.117	4.117	4.116
	Gal-**8'**	n.d.	4.117	n.d.	4.117	n.d.	4.116	n.d.	4.117	n.d.
	Gal_{ext}	4.115	4.117	4.115[c]	n.d.	4.117	4.116	4.117	4.117	4.116[c]

H3a	Neu5Ac	1.800[d]	1.800[d]	1.799[d]	1.802[d]	1.802[e]	1.803[e]	1.803[d]	1.803[d]	1.801[e]
	Neu5Gc	–	–	–	–	–	–	1.820	1.820	–
H3e	Neu5Ac	2.757[d]	2.757[d]	2.758[d]	2.757[d]	2.757[e]	2.757[e]	2.757[d]	2.757[d]	2.757[e]
	Neu5Gc	–	–	–	–	–	–	2.774	2.774	–
H4	Gal-**6′**	n.d.	4.158	4.159	4.156	n.d.	4.161	n.d.	4.160	4.161
	Gal-**8′**	4.155	n.d.	4.159	n.d.	4.159	n.d.	4.160	n.d.	4.161
H5	Fuc*α*	4.10	4.10	4.095	4.098	4.10	4.10	4.10	4.10	4.10
	β	4.13	4.13	4.13	4.13	4.13	4.13	4.13	4.13	4.13
NAc	GlcNAc-**1**	2.038	2.038	2.037	2.039	2.037	2.038	2.038	2.038	2.037
	GlcNAc-**2***α*	2.093	2.092	2.091	2.092	2.094	2.092	2.093	2.093	2.091
	β	2.093	2.092	2.091	2.092	2.090	2.088	2.093	2.093	2.088
	GlcNAc-**5**	2.051	2.051	2.051	2.047	2.047	2.047	2.048	2.048	2.047
	GlcNAc-**5′**	2.038	2.038	2.037	2.039	2.037	2.038	2.038	2.038	2.037
	GlcNAc-**7**	–	–	–	2.075	2.075	2.074	2.075	2.075	2.075
	GlcNAc-**7′**	2.038	2.038	2.037	2.039	2.037	2.038	2.038	2.038	2.037
	$GlcNAc_{ext}$	2.038	2.038	2.037[f]	2.039	2.037	2.038	2.038	2.038	2.037[f]
	Neu5Ac	2.031[g]	2.031[g]	2.031[g]	2.031[g]	2.031[h]	2.031[h]	2.031[g]	2.031[g]	2.031[h]
NGc	Neu5Gc	–	–	–	–	–	–	4.122	4.122	–
CH_3	Fuc*α*	1.210	1.209	1.209	1.210	1.211	1.210	1.211	1.211	1.210
	β	1.222	1.221	1.221	1.221	1.223	1.222	1.222	1.222	1.222

[a] Some values are given with only two decimals because of spectral overlap.
[b] Values may have to be interchanged.
[c] Signal stemming from two protons.
[d] Signal stemming from three protons.
[e] Signal stemming from four protons.
[f] Signal stemming from two NAc groups.
[g] Signal stemming from three NAc groups.
[h] Signal stemming from four NAc groups.

Table 4. ^{1}H-chemical shifts of the structural-reporter-group protons of the constituent monosaccharides of endo-β-galactosidase digestion products of N-linked oligosaccharides, derived from recombinant human erythropoietin. Chemical shifts are given at 27 °C and were measured in ^2H_2O relative to acetone (δ 2.225 [25]). For the short-hand symbolic notation, see Table 1. For numbering of the monosaccharide residues, see text. n.d., not determined; α and β stand for the anomeric configuration of GlcNAc-**1**.

Reporter group	Residue	Chemical shift in								
										
		N2.6Q1	**N3.9.2Q2A**	**N3.9.2Q2B**	**N3.10Q2**	**N3.12Q1**[a]	**N3.12Q3**	**N4.6.1Q2**	**N4.6.2Q2**	**N4.8Q2**
		ppm								
H1	GlcNAc-**1**α	5.181	5.18[b]	5.18	5.182	5.181	5.182	5.182	5.182	5.182
	β	n.d.	n.d.	n.d.	4.693	4.692	4.691	4.690	4.690	4.691
	GlcNAc-**2**α	4.664	4.66	4.66	4.664	4.663	4.663	4.661	4.664	4.663
	β	4.664	4.66	4.66	4.664	4.663	4.663	4.664	4.664	4.667
	Man-**4**	5.118	5.128	5.128	5.127	5.123	5.124	5.128	5.128	5.124
	Man-**4′**	4.917	4.864	4.858	4.866	4.861	4.860	4.855	4.866	4.861
	GlcNAc-**5**	4.575	4.562	4.562	4.562	4.562	4.562	4.563	4.561	4.563
	GlcNAc-**5′**	4.554	4.59	4.562	4.58	4.562	4.562	4.585	4.574	4.563
	Gal-**6**	4.545	4.545	4.545	4.546	4.468	4.545	4.546	4.546	4.545
	Gal-**6′**	–	4.468	–	–	–	–	4.546	–	–
	GlcNAc-**7**	–	4.545	4.545	4.546	4.545	4.545	4.546	4.546	4.545
	GlcNAc-**7′**	–	4.530	4.545	4.546	4.534	4.534	4.530	4.546	4.536
	Gal-**8**	–	4.545	4.545	4.546	4.545	4.545	4.546	4.546	4.545
	Gal-**8′**	–	–	4.482	4.562	–	–	–	4.561	–
	Fucα	4.886	4.90	4.90	4.893	4.891	4.891	4.899	4.893	4.891
	β	4.894	4.90	4.90	4.901	4.899	4.898	4.907	4.900	4.899

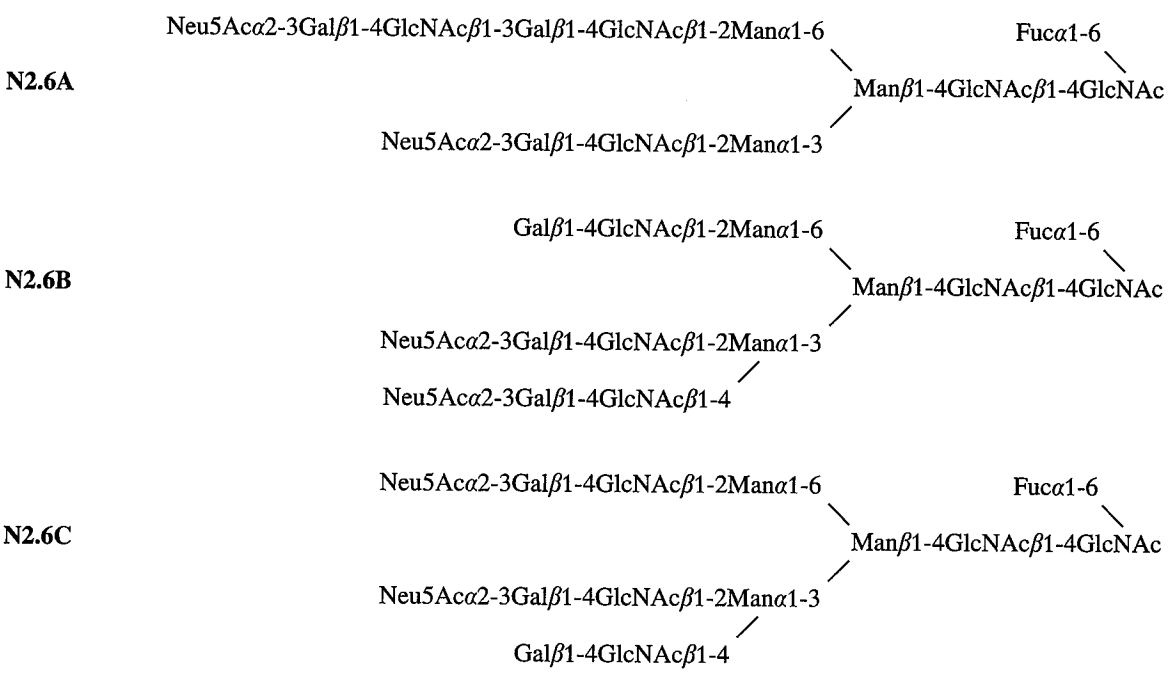
H2	Man-**3**	4.250	4.21	4.21	4.215	4.216	4.215	4.214	4.21	4.216
	Man-**4**	4.186	4.22	4.22	4.215	4.216	4.215	4.214	4.21	4.216
	Man-**4′**	4.107	4.090	4.090	4.091	4.091	4.091	4.092	4.091	4.092
H3	Gal-**6**	4.114	4.114	4.114	4.117	n.d.	4.116	4.117	4.117	4.116
	Gal-**6′**	–	n.d.	–	–	–	–	4.117	–	–
	Gal-**8**	–	4.114	4.114	4.117	4.113	4.116	4.117	4.117	4.116
	Gal-**8′**	–	–	n.d.	4.117	–	–	–	4.117	–
H3a	Neu5Ac	1.797	1.802[c]	1.802[c]	1.803[d]	1.798	1.800[c]	1.802[d]	1.803[d]	1.800[c]
H3e	Neu5Ac	2.758	2.757[c]	2.757[c]	2.756[d]	2.756	2.757[c]	2.756[d]	2.756[d]	2.756[c]
H5	Fucα	4.10	4.10	4.10	4.10	4.10	4.10	4.098	4.10	4.10
	β	4.13	4.13	4.13	4.13	4.13	4.13	4.135	4.13	4.13
NAc	GlcNAc-**1**	2.039	2.039	2.039	2.038	2.039	2.039	2.040	2.038	2.040
	GlcNAc-**2**α	2.094	2.093	2.093	2.092	2.091	2.091	2.093	2.093	2.091
	β	2.094	2.093	2.093	2.092	2.091	2.091	2.090	2.089	2.088
	GlcNAc-**5**	2.042	2.048	2.048	2.048	2.049	2.048	2.048	2.048	2.048
	GlcNAc-**5′**	2.052	2.042	2.042	2.048	2.049	2.048	2.040	2.048	2.048
	GlcNAc-**7**	–	2.075	2.075	2.074	2.075	2.074	2.075	2.074	2.074
	GlcNAc-**7′**	–	2.039	2.039	2.038	2.039	2.039	2.040	2.038	2.040
	Neu5Ac	2.031	2.032[e]	2.032[e]	2.031[f]	2.031	2.031[e]	2.031[f]	2.031[f]	2.031[e]
CH_3	Fucα	1.209	1.210	1.210	1.209	1.209	1.209	1.211	1.209	1.209
	β	1.220	1.221	1.221	1.221	1.220	1.220	1.223	1.211	1.220

[a] It was chosen arbitrarily that the Neu5Ac residue is linked to Gal-**8** (see text).
[b] Some values are given with only two decimals because of spectral overlap.
[c] Signal stemming from two protons.
[d] Signal stemming from three protons.
[e] Signal stemming from two NAc groups.
[f] Signal stemming from three NAc groups.

compounds possible in a retrospective way. In the context of this report, the ^{1}H-NMR data of the intermediate products **N4.6.1Q2** (trisialylated), **N4.6.2Q2** (trisialylated) and **N4.8Q2** (disialylated) (see Fig. 6), formed after endo-β-galactosidase digestion of **N4.6.1**, **N4.6.2** and **N4.8**, respectively, and isolated via Mono Q, will be discussed. These data (Table 4) will be used for the identification of the endo-β-galactosidase digestion products obtained from fractions **N3.9.2**, **N3.10** and **N3.12** (see below). The tetraantennary nature of **N4.6.1Q2**, **N4.6.2Q2** and **N4.8Q2** is in each case deduced from the chemical shifts of the set of H1 signals of Man-**4** and Man-**4′**, and H2 signals of Man-**3**, Man-**4** and Man-**4′** [25]. The terminal GlcNAc-**7′** residue in **N4.6.1Q2** gives rise to a typical H1 signal at δ 4.530 (cf. compound 13 {26,24} in [32]), whereas the terminal GlcNAc-**5′** residue in **N4.6.2Q2** is reflected by its H1 signal at δ 4.574. GlcNAc-**7′** H1 in **N4.6.1Q2** and GlcNAc-**5′** H1 in **N4.6.2Q2** show similar upfield shifts ($\Delta\delta$−0.019 and $\Delta\delta$−0.015, respectively), as compared to the corresponding signals in the tetrasialylated tetraantennary compound **N4.4**. In the ^{1}H-NMR spectrum of **N4.8Q2**, in which both GlcNAc-**5′** and GlcNAc-**7′** are in terminal position, the typical H1 signals are observed at δ 4.563 (GlcNAc-**5′**) and δ 4.536 (GlcNAc-**7′**).

The ^{1}H-NMR spectrum of fraction **N2.6** provides indications for the presence of di- and triantennary types of branching, and repeating *N*-acetyllactosamine units. Fraction **N2.6** was treated with endo-β-galactosidase from *B. fragilis*, and the digest was applied to Mono Q (not shown), yielding two fractions (**N2.6Q1** and **N2.6Q2**) eluting in the monosialo region and one fraction (**N2.6Q3**) eluting in the disialo region. The ^{1}H-NMR data of fraction **N2.6Q2** are in accordance with Neu5Acα2-3Galβ1-4GlcNAcβ1-3Gal [28], whereas those of fraction **N2.6Q1** demonstrate the occurrence of a monosialylated diantennary oligosaccharide having a terminal GlcNAc-**5′** residue. The diantennary character of **N2.6Q1** is reflected by the chemical shifts of the set of Man-**4** and Man-**4′** H1, and Man-**3**, Man-**4** and Man-**4′** H2 signals [25]. As compared with compound **N1.1A** (Table 1), upfield shifts are observed for Man-**4′** H1 ($\Delta\delta$−0.007) and GlcNAc- **5′** H1 ($\Delta\delta$−0.028). These data indicate that GlcNAc-**5′** is in terminal position (cf. the α2-6 sialylated isomer 1-2 in [29]), which is confirmed by the absence of an H1 signal of Gal-**6′**. Taking together the findings for **N2.6Q1** and **N2.6Q2**, it is concluded that in the intact diantennary component of fraction **N2.6**, denoted **N2.6A**, an elongated branch was located at GlcNAc-**5′**. The ^{1}H-NMR spectrum of fraction **N2.6Q3** shows the presence of a mixture of two disialylated triantennary oligosaccharides of which the branching pattern is reflected by their sets of Man-**4** and Man-**4′** H1, and Man-**3**, Man-**4** and Man-**4′** H2 signals. It should be noted that two different Man-**4′** H1 resonances are observed at δ 4.903 and δ 4.921, respectively. The compound denoted **N2.6B** has a terminal Gal-**6′** residue as is concluded from the Gal-**6′** H1 signal at δ 4.472, in combination with the Man-**4′** H1 signal at δ 4.921 which shifted downfield ($\Delta\delta$ +0.017) as compared to the trisialylated triantennary compound **N3.4** (cf. compound N2.3 in [20]). Compound **N2.6C** contains a terminal Gal-**8** residue which is characterized by the H1 signal at δ 4.463 (compare Gal-**8** H1 in the asialo triantennary compound 10 in [25]) in combination with Man-**4′** H1 at δ 4.903 which remains unchanged as compared to compound **N3.4**.

^{1}H-NMR analysis of fraction **N2.7** showed the presence of a complex mixture of oligosaccharides. Further fractionation of **N2.7** by HPAEC on CarboPac PA1, yielded the subfractions **N2.7.1**–**N2.7.3** (Fig. 5B). The ^{1}H-NMR spectrum of fraction **N2.7.2** (Fig. 7A) reveals the presence of a mixture of disialylated tri′antennary oligosaccharides, the branching pattern being characterized by the chemical shifts of the set of Man-**4** and Man-**4′** H1, and Man-**3**, Man-**4** and Man-**4′** H2 signals [25]. The presence of two different isomers is deduced from the Gal H1 signals in the δ 4.48−4.46 region of which the signal at δ 4.481 stems from a terminal Gal-**8′** residue (**N2.7.2A**; cf. compound N2.4A in [30]) and that at δ 4.466 from terminal Gal-**6** (**N2.7.2B**; compare the Gal H1 values of the asialo tri′antennary compound 11 in [25] and the ^{1}H-NMR data of **N2.7.3**). As shown by its ^{1}H-NMR spectrum (Fig. 7B), fraction **N2.7.3** contains the third possible isomer of the disialylated tri′antennary oligosaccharides. The Man-**4′** H1 signal at δ 4.874 has shifted downfield ($\Delta\delta$+0.003) as compared to compound **N3.5** (Table 1), which is, together with the Gal H1 signal at δ 4.469 (cf. compound 11 in [25]), indicative of a Gal-**6′** residue in terminal position. This reasoning is based on the observation that desialy-



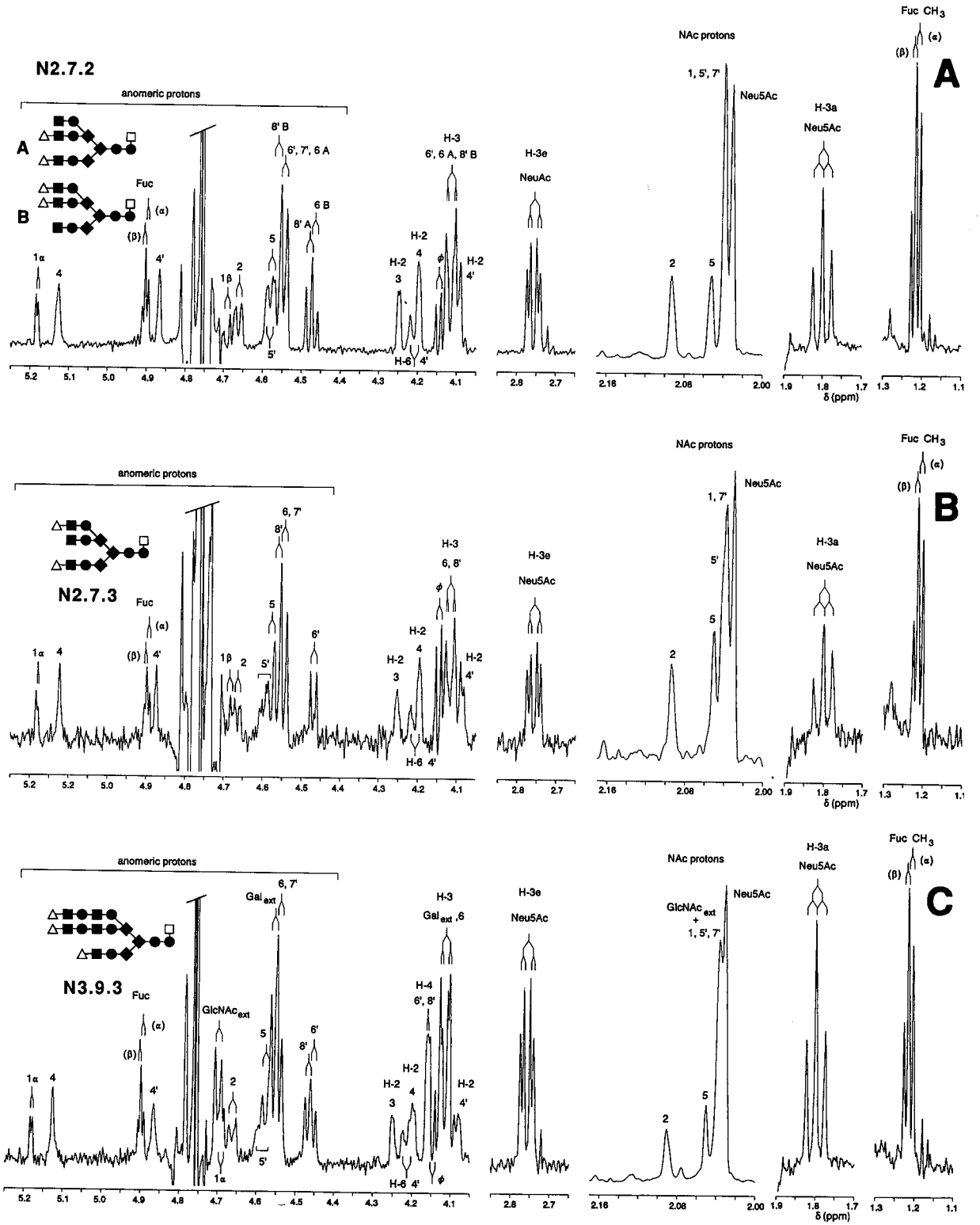


Fig. 7. Structural-reporter-group regions of the 500-MHz ^{1}H-NMR spectra of fractions N2.7.2 (A), N2.7.3 (B), N3.9.3 (C), N4.2 (D) and N4.5 (E), recorded in ^2H_2O at 27°C. The relative intensity of the NAc proton region differs from that of the rest of the spectrum. ϕ denotes impurity. H6 of **4′** is not given in the tables since it is not considered to be a structural-reporter-group signal.

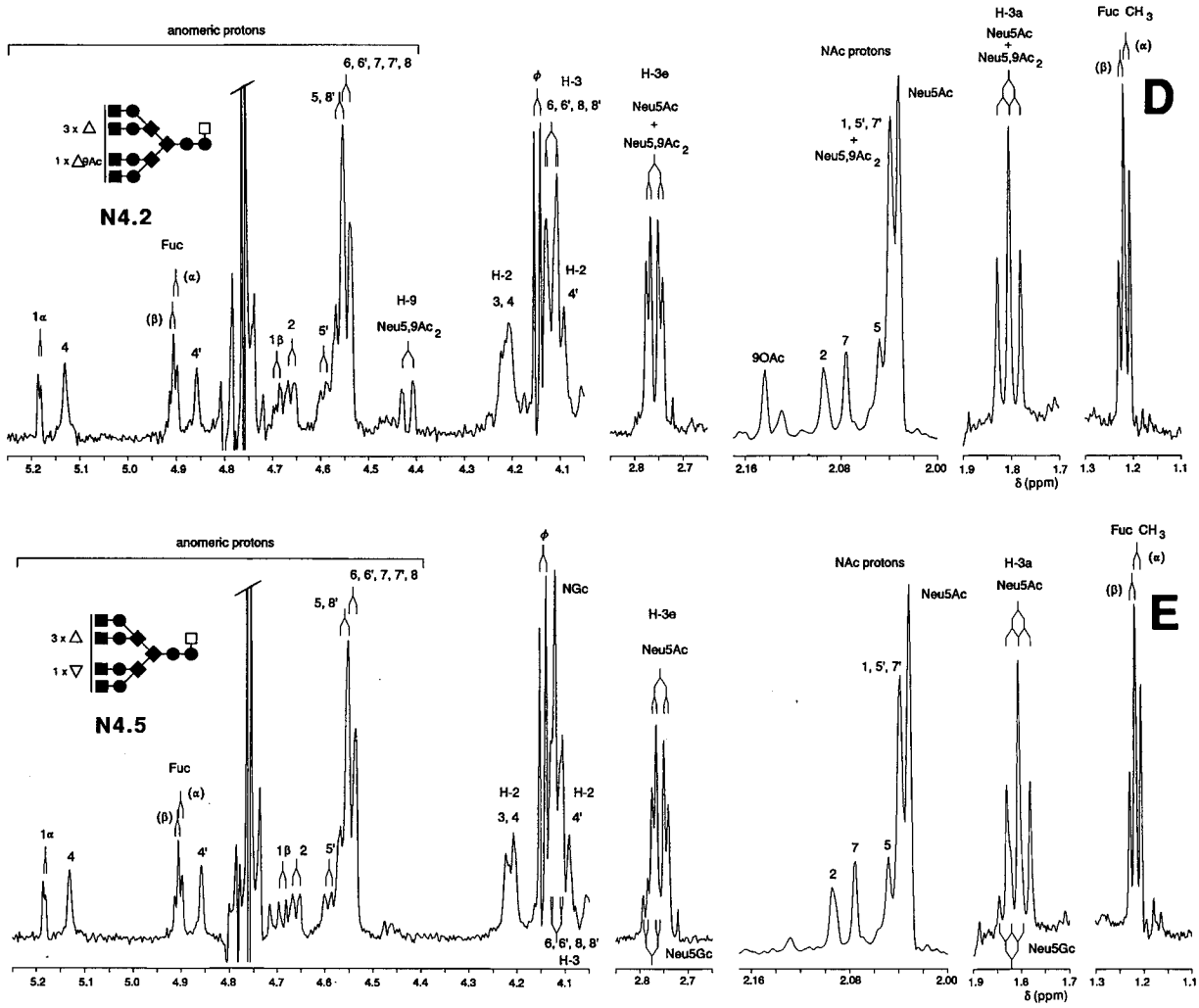


Fig. 7 D, E.

lation at Gal-**6′** in a trisialylated triantennary compound is accompanied by a downfield shift for Man-**4′** H1 (compare compounds **N2.6B** and **N3.4**, see [20]).

The ^{1}H-NMR data of fraction **N2.10** showed indications for the presence of tri′- and tetraantennary types of branching, and repeating *N*-acetyllactosamine sequences. Further fractionation

N2.7.2A

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

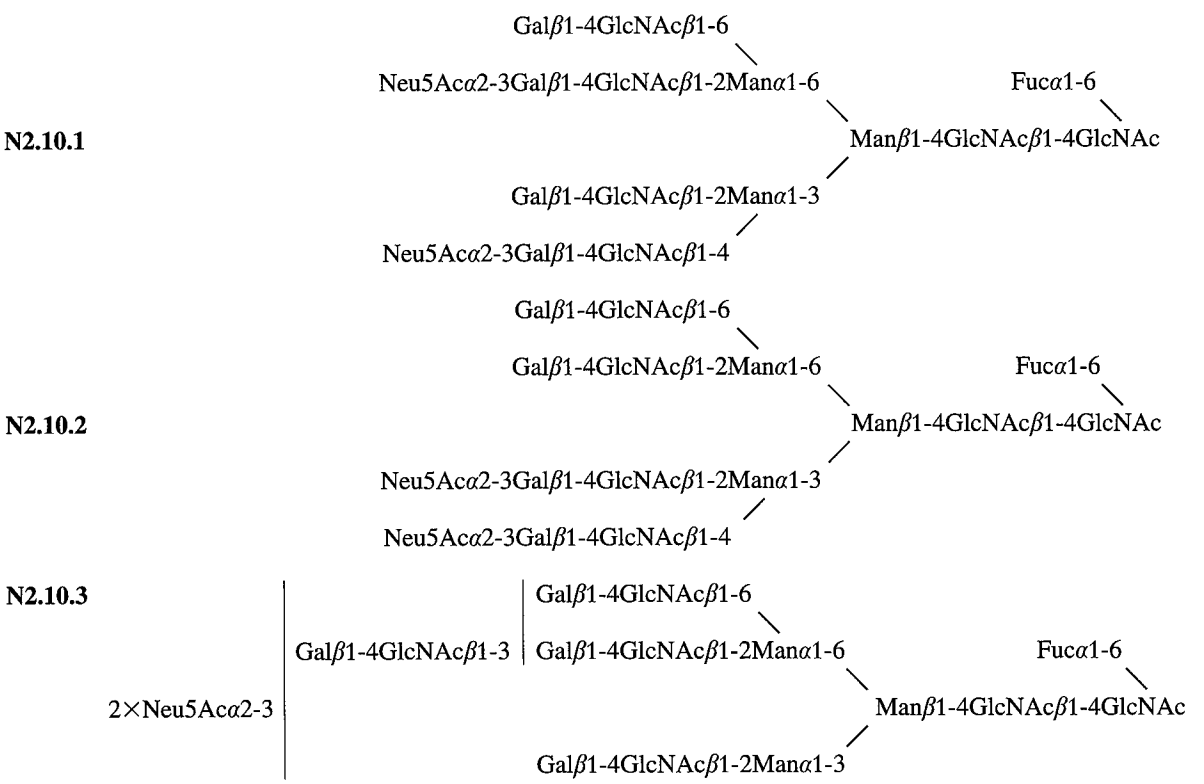
N2.7.2B

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

N2.7.3

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

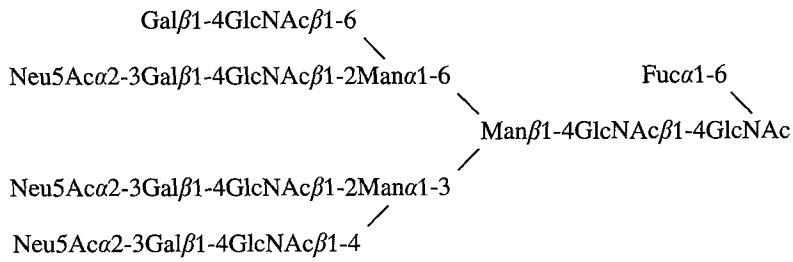
of fraction **N2.10** by HPAEC yielded three subfractions, denoted **N2.10.1–N2.10.3** (Fig. 5C). The ^{1}H-NMR spectrum of fraction **N2.10.1** demonstrates the presence of a disialylated tetraantennary oligosaccharide, of which the branching pattern is reflected by the chemical shifts of the set of Man-**4** and Man-**4′** H1 signals, and Man-**3**, Man-**4** and Man-**4′** H2 signals [25]. The terminal Gal-**6** and Gal-**8′** residues are recognized from their H1 signals at δ 4.470 (Gal-**6**, taking into account the ^{1}H-NMR data of **N2.10.2**) and δ 4.481 (Gal-**8′**; cf. compound 12 in [25]). The ^{1}H-NMR spectrum of **N2.10.2** shows the presence of an isomer of compound **N2.10.1**, containing terminal Gal-**6′** and Gal-**8′** residues. In a similar way as discussed for **N2.7.3** and **N2.7.2A/B**, the downfield shift of Man-**4′** H1 ($\Delta\delta$ +0.010) as compared to the Man-**4′** H1 signal in the fully sialylated compound **N4.4** (Table 2), in combination with the position of the Gal H1 signal at δ 4.468 (cf. compound 12 in [25]) proves that Gal-**6′** is in terminal position. The H1 signal at δ 4.483 stems from the terminal Gal-**8′** residue (cf. compound 12 in [25]). The ^{1}H-NMR data of fraction **N2.10.3** suggest the presence of a mixture of tri′antennary oligosaccharides. Because of the complexity of the mixture, in combination with the low amount of material, no definite structures can be presented. The chromatographic data indicate that the components in fraction **N2.10.3** are disialylated (elution position of **N2** on Mono Q) and contain one additional *N*-acetyllactosamine element as compared to the compounds in fraction **N2.7** (elution position of **N2.10** on Lichrosorb-NH_2, $R_{N2.10} = 1.58 \times R_{N2.7}$; cf. $R_{N3.7} = 1.58 \times R_{N3.5}$). Based on the results obtained for other fractions, it is suggested that in each compound the *N*-acetyllactosamine repeat is located in one of the branches linked to Man-**4′**.



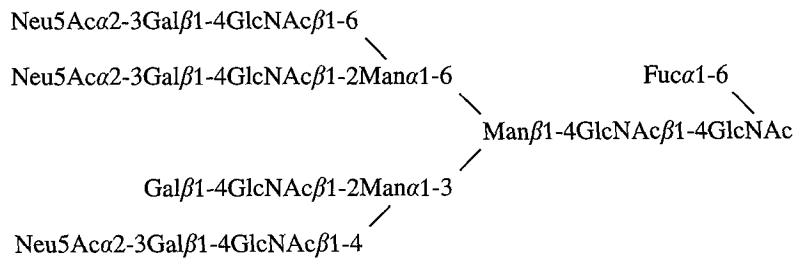
Fraction **N3.7** contained a complex mixture of trisialylated oligosaccharides (^{1}H-NMR analysis) and was subfractionated by HPAEC, yielding three subfractions, denoted **N3.7.1–N3.7.3** (Fig. 5D). The ^{1}H-NMR spectrum of fraction **N3.7.1** demonstrates the occurrence of three isomeric trisialylated tetraantennary oligosaccharides of which the branching patterns are inferred by their common set of Man H1 and H2 signals [25]. The difference in structure between the oligosaccharides is reflected by the chemical shift values of the anomeric protons of the terminal Gal residues (cf. compound 12 in [25]), indicating that compound **N3.7.1A** contains terminal Gal-**8′** (H1, δ 4.482), compound **N3.7.1B** has Gal-**6** in terminal position (H1, δ 4.472) and compound **N3.7.1C** contains terminal Gal-**8** (H1, δ 4.465). As is evident from the ^{1}H-NMR data of compound 12 in [25], the position of the H1 resonance of terminal Gal-**6** is nearly identical to that of a terminal Gal-**6′**. However, the presence of the isomer with a terminal Gal-**6′** residue would have affected the position of Man-**4′** H1 by $\Delta\delta$ +0.01 (cf. compounds **N2.10.2** and **N3.7.2A**; Table 2; see below). Inspection of the ^{1}H-NMR spectrum of fraction **N3.7.2** indicated the presence of tri′- and tetraantennary types of branching, and repeating *N*-acetyllactosamine units. Incubation of fraction **N3.7.2** with endo-β-galactosidase from *B. fragilis* and subsequent fractionation of the digestion mixture on Mono Q (not shown) afforded two fractions, denoted **N3.7.2Q1** and **N3.7.2Q2**, eluting very close to each other in the trisialo region. ^{1}H-NMR spectroscopy of fraction **N3.7.2Q2** revealed that it contains the fourth possible isomeric structure of the trisialylated tetraantennary oligosaccharides (see **N3.7.1A–C**), denoted **N3.7.2A**. The presence of the terminal Gal-**6′** residue is reflected by its H1 signal at δ 4.470, in combination with the H1 signal of Man-**4′** at δ 4.870, which shifted downfield ($\Delta\delta$+0.013) as compared to compound **N4.4** (see Discussion for compound **N2.10.2**). The ^{1}H-NMR spectrum of fraction **N3.7.2Q1** demonstrates the occurrence of a trisialylated oligosaccharide having an additional *N*-acetyllactosamine unit attached to Gal-**8′**. The location of the repeat is inferred from the Gal-**8′** H1 signal at δ 4.469 which is assigned by comparing the NMR data obtained for tetraantennary oligosaccharides having additional *N*-acetyllactosamine units in the branches linked to Man-**4′** (see compounds **N4.6.1**, **N4.6.2**, and **N4.8**; Table 3). Furthermore, typical signals of Gal-**8′** H4 (δ 4.155) and GlcNAc$_{ext}$ H1 (δ 4.700) are present. Evidently, compound **N3.7.2B**

was not sensitive to treatment with endo-β-galactosidase from *B. fragilis*, and Mono Q fractionation could have been carried out directly on fraction **N3.7.2.** ^{1}H-NMR spectroscopy of fraction **N3.7.3** demonstrates that it contains an isomer of **N3.7.2B** having an elongated GlcNAc-**5′** branch. The location of the additional *N*-acetyllactosamine unit is reflected by the typical position of Gal-**6′** H1 at δ 4.455 (cf. compound **N4.6.2**) together with that of the Man-**4′** H2 signal at δ 4.08 which has shifted upfield ($\Delta\delta$ −0.01) compared to compound **N3.7.2B** (cf. compounds **N4.6.2** and **N4.6.1**). As observed for **N3.7.2B**, treatment of **N3.7.3** with endo-β-galactosidase from *B. fragilis* did not result in the release of Neu5Acα2-3Galβ1-4GlcNAcβ1-3Gal.

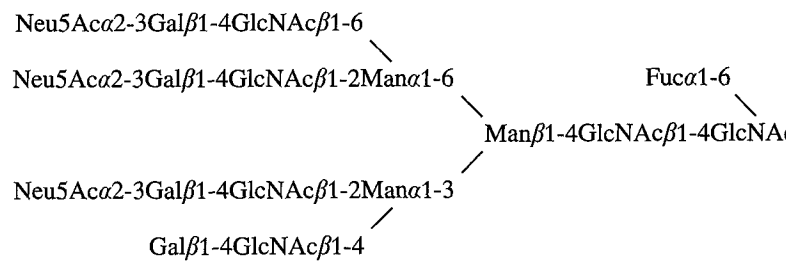
N3.7.1A



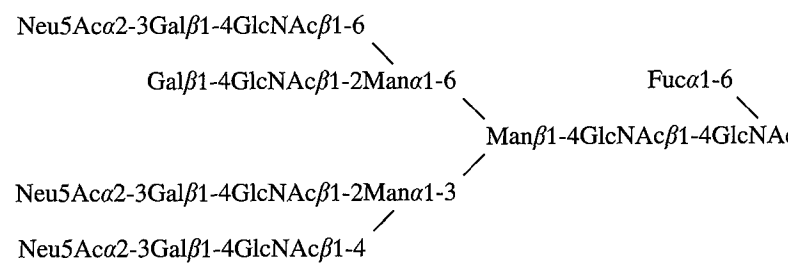
N3.7.1B



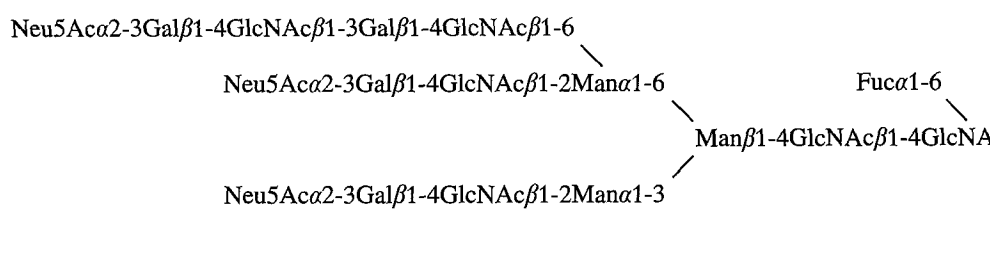
N3.7.1C



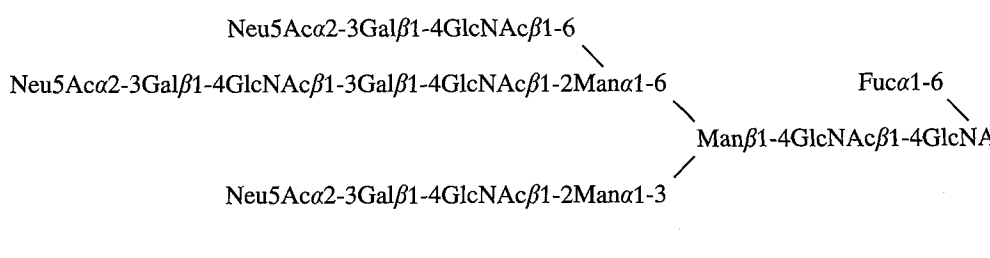
N3.7.2A



N3.7.2B



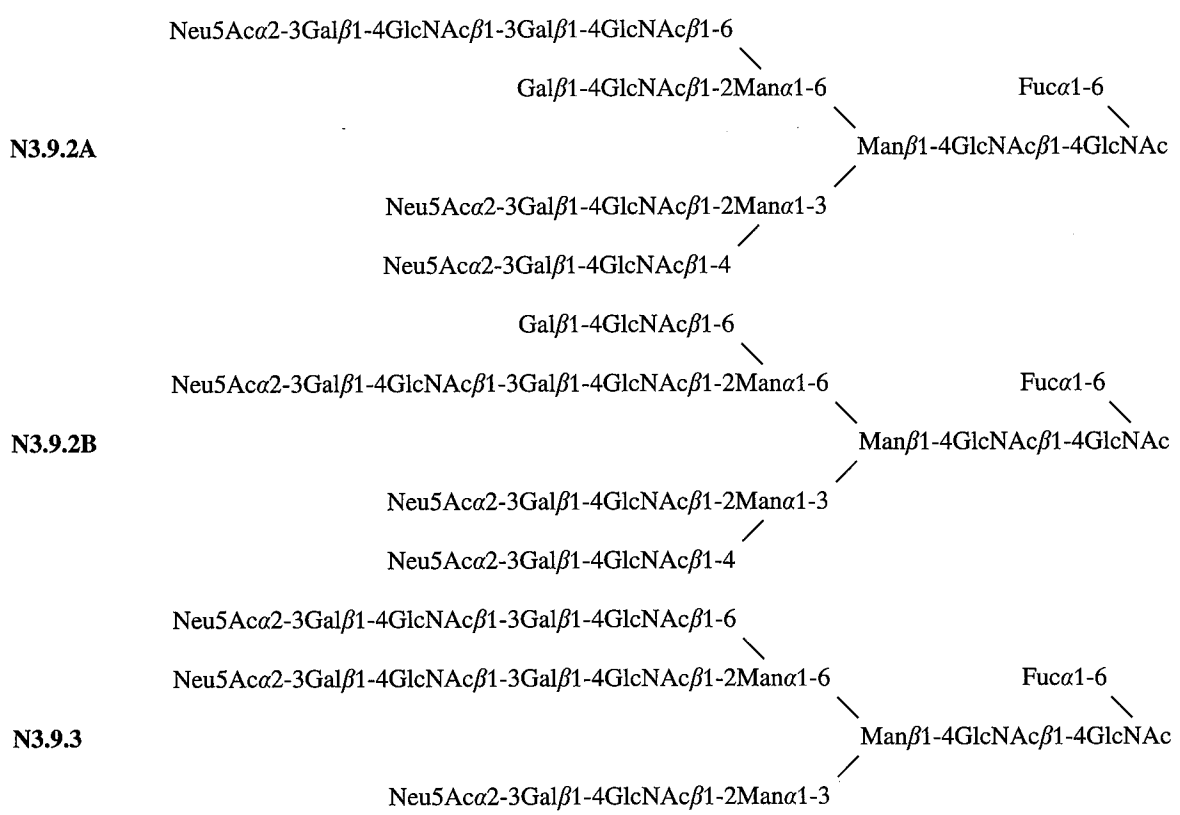
N3.7.3

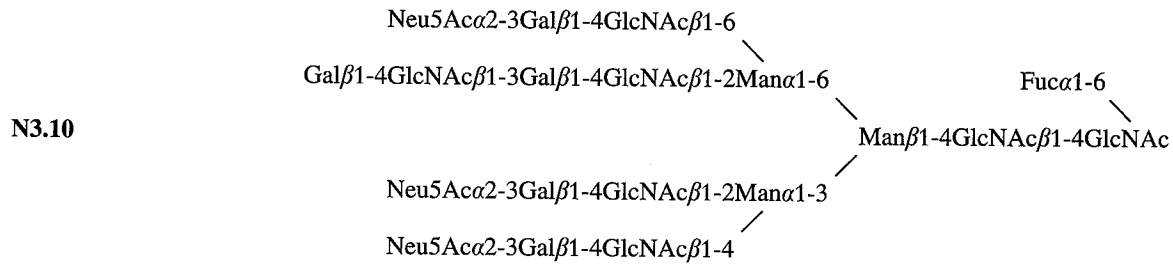


The ^{1}H-NMR spectrum of fraction **N3.9** showed the presence of tri′- and tetraantennary types of branching, and of *N*-acetyllactosamine repeating units. Further fractionation of fraction **N3.9** by HPAEC yielded four subfractions, denoted **N3.9.1–N3.9.4** (Fig. 5E). ^{1}H-NMR spectroscopy of fraction **N3.9.2** demonstrates the occurrence of a mixture of trisialylated tetraantennary oligosaccharides having one extra *N*-acetyllactosamine unit, but no assignment could be made at this stage concerning the location of this unit and the Neu5Ac residues in the different components. Therefore, fraction **N3.9.2** was incubated with endo-β-galactosidase from *B. fragilis*, and then subjected to FPLC on Mono Q (not shown), affording two subfractions, denoted **N3.9.2Q1** and **N3.9.2Q2**, eluting in the mono- and disialo region, respectively. ^{1}H-NMR spectroscopy shows that fraction **N3.9.2Q1** contains Neu5Acα2-3Galβ1-4GlcNAcβ1-3Gal [28], whereas fraction **N3.9.2Q2** contains a mixture of two disialylated tetraantennary oligosaccharides. The first compound, denoted **N3.9.2Q2A**, contains a terminal GlcNAc-**7′** residue (H1, δ 4.530; cf. compound **N4.6.1Q2**; Table 4) and a terminal Gal-**6′** residue as shown by its anomeric signal at δ 4.468 in combination with Man-**4′** H1 at δ 4.864. The latter signal shows a similar downfield shift compared to **N4.6.1Q2** (Δδ +0.009) as **N3.7.2A** compared to **N4.4** (Δδ +0.013). The second compound, denoted **N3.9.2Q2B**, has terminal GlcNAc-**5′** (H1, δ 4.562; cf. compound **N4.6.2Q2**; Table 4) and Gal-**8′** (H1, δ 4.482; cf. compound **N3.7.1A**; Table 2) residues. From these data it is concluded that, prior to digestion, fraction **N3.9.2** contained a mixture of two trisialylated tetraantennary oligosaccharides, one with a terminal Gal-**6′** residue and an elongated GlcNAc-**7′** branch (**N3.9.2A**) and one with terminal Gal-**8′** and an elongated GlcNAc-**5′** branch (**N3.9.2B**). The ^{1}H-NMR spectrum of fraction **N3.9.3** (Fig. 7C) reveals the presence of a trisialylated tri′antennary oligosaccharide containing two elongated branches. The occurrence and location of these branches is reflected by the $GlcNAc_{ext}$ H1 signals at δ 4.699, the Gal H1 signals at δ 4.454 (Gal-**6′**) and δ 4.468 (Gal-**8′**), and the Gal H4 resonances at δ 4.159 (cf. compounds **N3.7.2B**, **N3.7.3** and **N4.8**; Table 3). Compound **N3.9.3** was not susceptible to endo-β-galactosidase from *B. fragilis*.

The ^{1}H-NMR spectrum of fraction **N3.10** shows similar characteristics as the spectrum of fraction **N3.9.2.** Incubation of **N3.10** with endo-β-galactosidase from *B. fragilis* and subsequent purification of the incubation mixture on Mono Q yielded, in addition to a neutral fraction, one fraction (**N3.10Q2**) eluting in the trisialo region. The ^{1}H-NMR data of fraction **N3.10Q2** match those of the terminal-GlcNAc-**5′**-containing trisialylated tetraantennary oligosaccharide **N4.6.2Q2** (Table 4), indicating that in the undigested oligosaccharide, the additional *N*-acetyllactosamine unit was located in the GlcNAc-**5′** branch, and that the Gal_{ext} residue was in terminal position. The structure of **N3.10Q2** was confirmed by further digestion with *N*-acetyl-β-glucosaminidase which splits off the terminal GlcNAc-**5′** residue, yielding an oligosaccharide identical to compound C.Q4* in [28], having branches linked to C2 and C4 of Man-**4**, and to C6 of Man-**4′**. Using these data, the two Gal H1 signals in the δ 4.48−4.45 region of the ^{1}H-NMR spectrum of the intact compound **N3.10** can be assigned to Gal-**6′** (δ 4.456; reflecting the location of the repeat) and Gal_{ext} (δ 4.479; reflecting the terminal position of Gal_{ext}).

The ^{1}H-NMR spectrum of fraction **N3.12** provides indications for the presence of a mixture of trisialylated tetraantennary oligosaccharides with two additional *N*-acetyllactosamine units, the latter feature being judged from the intensities of the characteristic Gal H1, Gal H4, and $GlcNAc_{ext}$ H1 signals. Because no definite structures could be deduced at this stage, fraction **N3.12** was treated with endo-β-galactosidase from *B. fragilis*. The digestion products were separated by anion-exchange chromatography on Mono Q (not shown), affording two fractions eluting in the monosialo region (**N3.12Q1** and **N3.12Q2**), and one fraction eluting in the disialo region (**N3.12Q3**). Fraction **N3.12Q2** contains Neu5Acα2-3Galβ1-4GlcNAcβ1-3Gal [28]. The ^{1}H-NMR data of **N3.12Q1**, together with the elution position on Mono Q, indicate the occurrence of a monosialylated tetraantennary oligosaccharide with two terminal GlcNAc residues. The

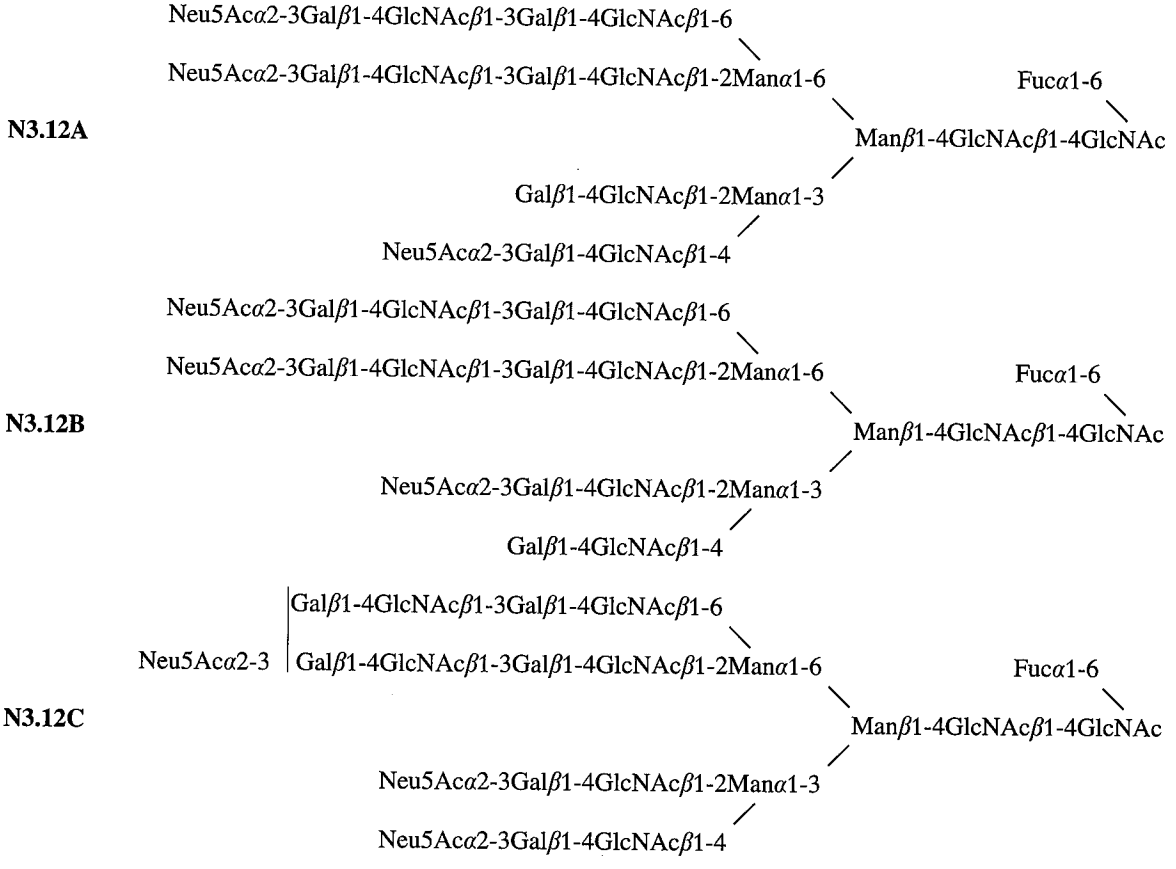




N3.10

terminal GlcNAc-**5′** and GlcNAc-**7′** residues are reflected by their H1 signals at δ 4.562 and δ 4.534, respectively (cf. compound **N4.8Q2**; Table 4). The position of the single Neu5Ac residue cannot be directly deduced since the observed chemical shift value of the H1 resonance of the terminal Gal residue (δ 4.468), appearing as a relatively broad signal, lies in between the values of the Gal H1 signals expected in case of a terminal Gal-**6** (δ 4.472, cf. compound **N3.7.1C**; Table 2) or a terminal Gal-**8** (δ 4.465, cf. compound **N3.7.1B**; Table 2) residue. Digestion of **N3.12Q1** with *N*-acetyl-β-glucosaminidase yielded a mixture of two diantennary compounds with a terminal Man-**4′** residue and a Neu5Ac residue linked to Gal-**6** or Gal-**8**. The NMR data of the digest are similar to those of compound D.Q3* in [28], but two different H1 signals characteristic for terminal Gal residues are observed, indicating that two isomers are present, one having Gal-**6** (H1, δ 4.468) in terminal position and one having Gal-**8** (H1, δ 4.464) in terminal position. The ^{1}H-NMR spectrum of fraction **N3.12Q3** demonstrates the presence of a disialylated tetraantennary oligosaccharide having terminal GlcNAc-**5′** and GlcNAc-**7′** residues, and is in accordance with that of compound **N4.8Q2** (Table 4). Digestion of **N3.12Q3** with *N*-acetyl-β-glucosaminidase yielded a disialylated diantennary oligosaccharide with a terminal Man-**4′** residue (cf. compound D.Q3* in [28]). From the above-mentioned data, it is concluded that before endo-β-galactosidase treatment of fraction **N3.12** in each constituent oligosaccharide two extra *N*-acetyllactosamine units were present, one in the GlcNAc-**5′** and one in the GlcNAc-**7′** branch. Compound **N3.12A** contains terminal Gal-**6**, compound **N3.12B** has Gal-**8** in terminal position, and in compound **N3.12C** one of the Gal_{ext} residues is not sialylated. From the available data it can not be determined whether a specific branch in compound **N3.12C** carries the terminal Gal_{ext} residue.

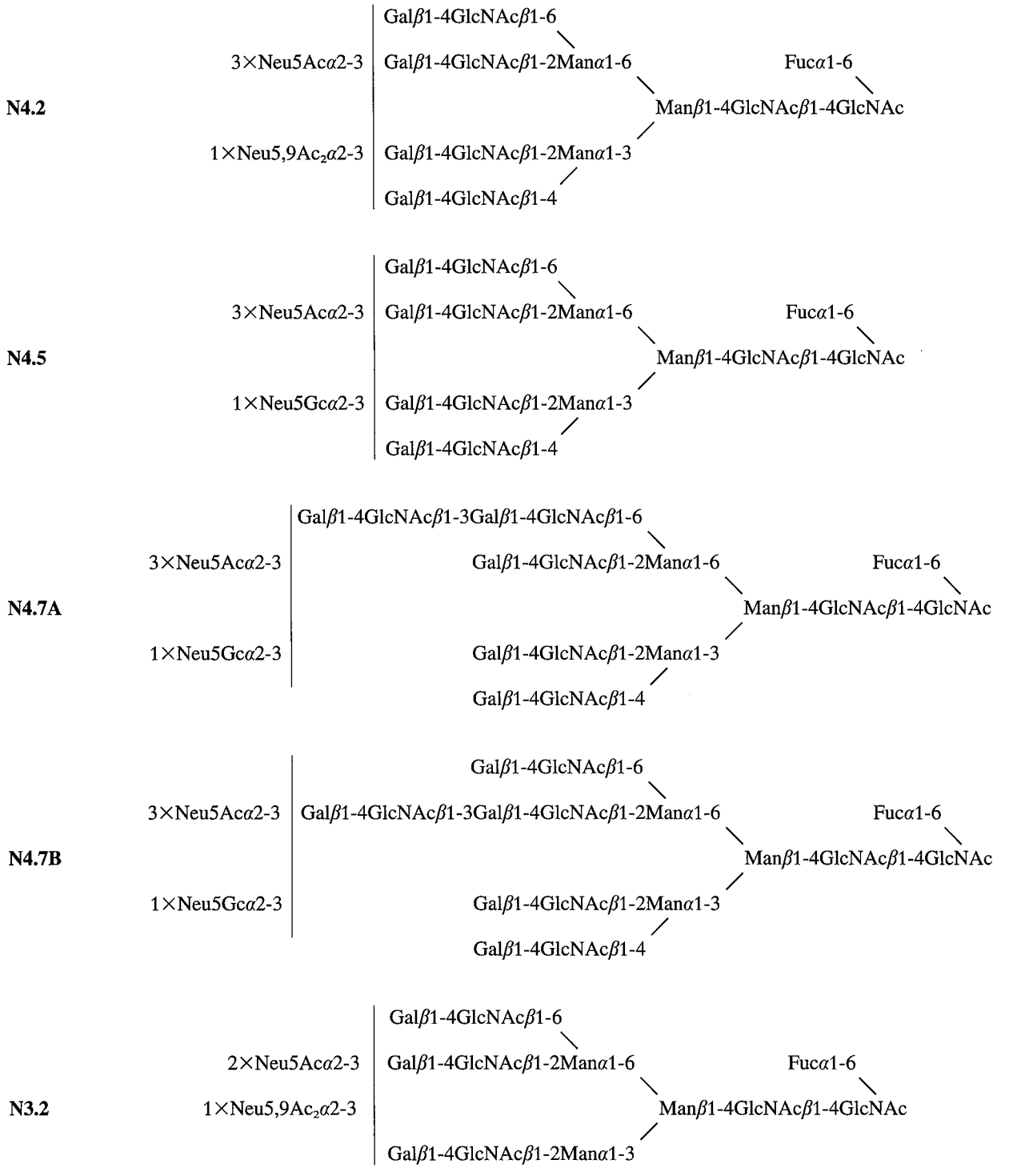
The ^{1}H-NMR spectrum of fraction **N4.2** (Fig. 7D) demonstrates the presence of a tetrasialylated tetraantennary oligosaccharide carrying three Neu5Ac residues and one $Neu5,9Ac_2$ residue. The spectral data of **N4.2** are in close agreement with those of compound **N4.4,** except for the presence of an additional signal at δ 2.143 stemming from the OAc group of the $Neu5,9Ac_2$ residue, whereas the H9 signal of $Neu5,9Ac_2$ shifted out of the bulk region to δ 4.415 (cf. compound N4.1 in [30]). The presence of the $Neu5,9Ac_2$ residue is confirmed by the two-dimensional HOHAHA spectrum of **N4.2** (not shown), in which the characteristic spin system of its acetylated glycerol side chain is observed (H9, δ 4.42; H9′, δ 4.20; H8, δ 4.11; H7, δ 3.65 [30]). From the available data, it can not be deduced if the $Neu5,9Ac_2$ residue has a specific location or is randomly distributed over the branches. ^{1}H-NMR spectroscopy of fraction **N4.5** (Fig. 7E) indicates the occurrence of a tetrasialylated tetraantennary oligosaccharide containing three Neu5Ac residues and one Neu5Gc residue. The structure is deduced by comparison with the



N3.12A

N3.12B

N3.12C



spectral data of **N4.4** (Table 2), taking into account the presence of the typical NGc singlet at δ 4.121, and the altered position of one of the four sets of sialic acid H3 signals (H3a, δ 1.821; H3e, δ 2.774; cf. compound N4.3 in [30]). The location of the Neu5Gc residue cannot be deduced from the available data, but in view of the biosynthetic pathway in which hydroxylation of Neu5Ac takes place before it is transferred to the oligosaccharide [33, 34], a random distribution is most likely.

The ^{1}H-NMR spectrum of fraction **N4.7** is similar to that of fraction **N4.6**, except for the presence of a NGc singlet at δ 4.122, and Neu5Gc H3a and H3e signals at δ 1.820 and δ 2.774, respectively (cf. compound **N4.5**; Table 2). The specific H3 signals of Neu5Gc and Neu5Ac occur in the intensity ratio of 1:3. Therefore, it is concluded that fraction **N4.7** contains a mixture of two tetrasialylated tetraantennary oligosaccharides having an additional *N*-acetyllactosamine unit linked to Gal-**8′** (**N4.7A**; cf. compound **N4.6.1**; Table 3) or Gal-**6′** (**N4.7B**; cf. compound **N4.6.2**; Table 3), and carrying three Neu5Ac residues and one Neu5Gc residue.

Fraction **N3.2** did not contain sufficient material to obtain a good quality ^{1}H-NMR spectrum, but the NAc and OAc singlets in the δ 2.03−2.15 region are clearly observed. In addition to a NAc pattern indicative for a Neu5Ac sialylated tri′antennary oligosaccharide (GlcNAc-**1**, δ 2.039; GlcNAc-**2**, δ 2.095/91; GlcNAc-**5**, δ 2.052; GlcNAc-**5′**, δ 2.039; GlcNAc-**7′**, δ 2.039; Neu5Ac, δ 2.031), the OAc resonance of a Neu5,9Ac$_2$ residue is present at δ 2.142 (cf. compound N3.1A in [30]). Together with the chromatographic behaviour of fraction **N3.2** as compared to

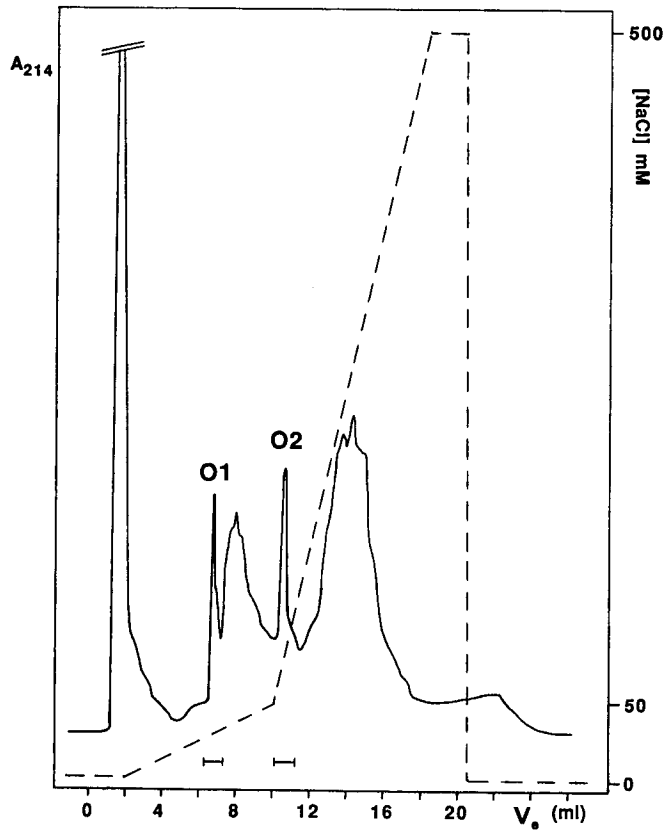


Fig. 8. Fractionation pattern at 214 nm of the β-elimination products derived from N-deglycosylated rEPO on a FPLC Mono Q HR 5/5 anion-exchange column. For details, see Fig. 3.

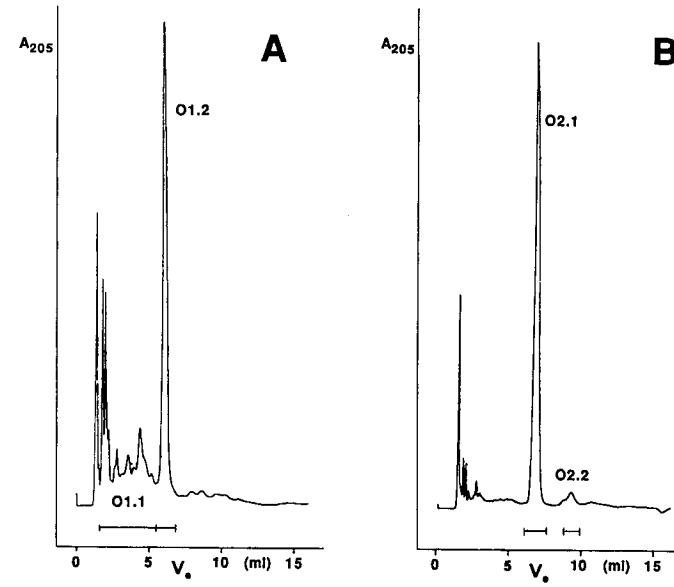


Fig. 9. Fractionation patterns at 205 nm of FPLC fractions O1 (A) and O2 (B) on a Lichrosorb-NH_2 column (25×0.46 cm). The column was eluted isocratically with 30 mM K_2HPO_4/KH_2PO_4 pH 7.0/acetonitrile (70:30, by vol.; A) or (67:33, by vol.; B), at a flow rate of 2 ml/min.

N3.5 (compare compounds **N4.2** and **N4.4**), these data suggest that compound **N3.2** is a trisialylated tri'antennary oligosaccharide carrying two Neu5Ac residues and one Neu5,9Ac_2 residue.

O-linked carbohydrate chains. FPLC on Mono Q of the mixture of O-linked oligosaccharide alditols obtained from N-deglycosylated rEPO by alkaline borohydride treatment gave rise to two carbohydrate-containing fractions, denoted **O1** and **O2** (Fig. 8) which were further fractionated by HPLC on Lichrosorb-NH_2, yielding two subfractions for **O1** (denoted **O1.1** and **O1.2**; Fig. 9A), and two subfractions for **O2** (denoted **O2.1** and **O2.2**; Fig. 9B). The structure of fractions **O1.2** and **O2.1** was determined by ^{1}H-NMR spectroscopy. Fraction **O1.1** did not contain carbohydrate material and for fraction **O2.2** the amount of material was too low for structure determination. The ^{1}H-NMR spectrum of fraction **O1.2** indicated the presence of the monosialylated trisaccharide alditol Neu5Acα2-3Galβ1-3GalNAc-ol (cf. compound 78 in [35]). ^{1}H-NMR spectroscopy of fraction **O2.1** revealed the presence of the disialylated tetrasaccharide alditol Neu5Acα2-3Galβ1-3(Neu5Acα2-6)GalNAc-ol (cf. compound 85 in [35]). The molar ratio of **O1.2/O2.1** is approximately 4:6.

DISCUSSION

The structures of the sialylated N- and O-linked carbohydrate chains of recombinant human erythropoietin produced in CHO cells have been determined. Over 35 different N-linked oligosaccharides were identified, representing at least 95 mol/100 mol carbohydrate chains. These compounds are presented in Table 5, together with their relative amounts. In agreement with previous studies on the carbohydrate moiety of rEPO from both CHO and BHK cells [11–15, 17, 18], the major portion of the identified carbohydrate chains comprise sialylated tetraantennary oligosaccharides with 0–2 *N*-acetyllactosamine repeats. From a biological point of view, a high content of tetraantennary carbohydrate chains is important since there is a positive correlation between the *in vivo* activity of rEPO and the ratio of tetra- to diantennary oligosaccharides [36].

Also, a high degree of sialylation is of great importance since terminal Gal residues are recognized by the hepatic Gal (asialoglycoprotein) receptor resulting in rapid clearance of poorly sialylated EPO from the circulation, having a negative effect on the *in vivo* activity [4]. Concerning the degree of sialylation, contrasting results have appeared in the literature. The first structural study on carbohydrate chains of rEPO from CHO cells showed that almost all oligosaccharides contained at least one non-sialylated terminal Gal residue [11], but more recently it was reported that over 90 % of the oligosaccharides are fully sialylated [15, 18]. The results of the present study show that of the tetraantennary oligosaccharides, which comprise about 62 % of the N-linked carbohydrate chains, 68 % are tetrasialylated, 24 % are trisialylated, and 6 % are disialylated. Of the tri'antennary oligosaccharides, which account for 23 % of the N-linked carbohydrate chains, 70 % are trisialylated and 26 % are disialylated. The above-mentioned inconsistencies may have been caused partly by the use of different methods for the release and the purification of the carbohydrate chains. In particular the use of hydrazinolysis to release the N-linked oligosaccharides [11] may have caused some desialylation. However, in the present study applying PNGase-F for release of the N-glycans, substantial amounts of incompletely sialylated oligosaccharides were also found, showing that these are in fact present in the original ensemble of carbohydrate chains on the isolated glycoprotein. The contrasting results concerning sialylation can therefore more likely be explained by the fact that all studies were performed on different rEPO samples produced in different laboratories. The use of different purification methods to obtain rEPO may

Table 5. Sialylated N-linked carbohydrate chains obtained from recombinant human erythropoietin expressed in CHO cells, together with their relative amounts. Compounds are represented by short-hand symbolic notation (see Tables 1 and 2). The structure of the remaining approximately 5 % was not determined.

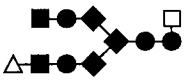
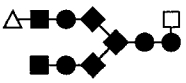
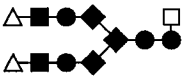
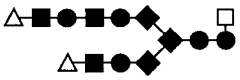
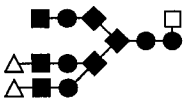
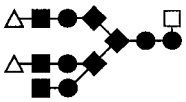
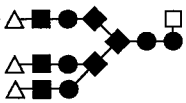
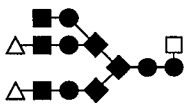
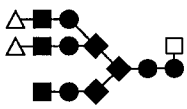
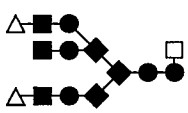
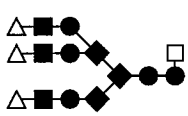
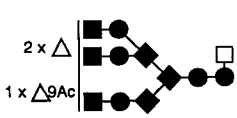
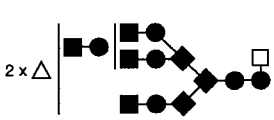
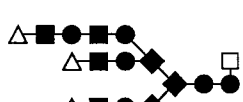
Structure	Code	Amount
		mol/100 mol
	**N1.1A**	<1
	**N1.1B**	<1
	**N2.4**	7
	**N2.6A**	<1
	**N2.6B**	<1
	**N2.6C**	<1
	**N3.4**	2
	**N2.7.2A**	2
	**N2.7.2B**	1
	**N2.7.3**	1
	**N3.5**	9
	**N3.2**	<1
	**N2.10.3**	1
	**N3.7.2B**	3

Table 5. (Continued)

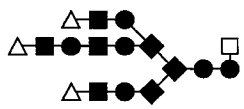
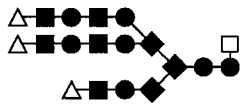
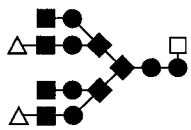
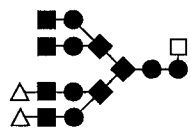
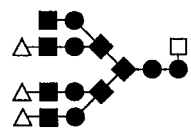
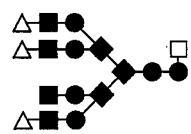
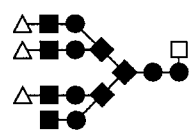
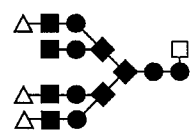
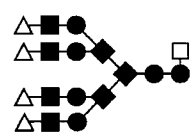
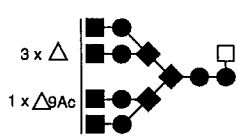
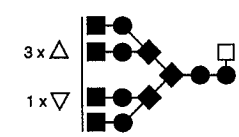
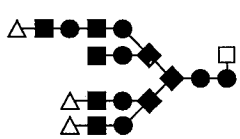
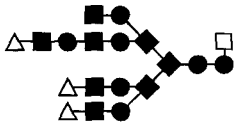
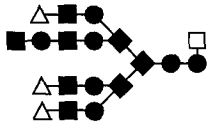
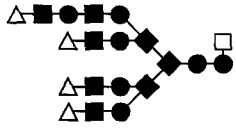
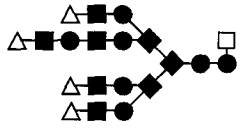
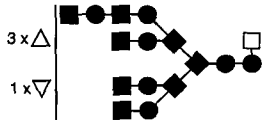
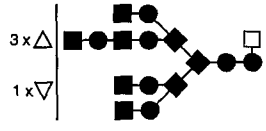
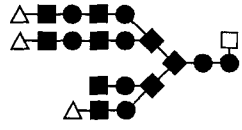
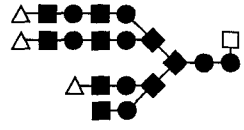
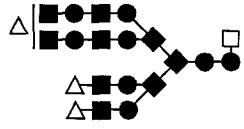
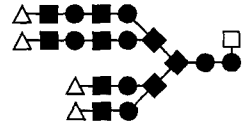
Structure	Code	Amount
		mol/100 mol
	**N3.7.3**	2
	**N3.9.3**	2
	**N2.10.1**	<1
	**N2.10.2**	3
	**N3.7.1A**	5
	**N3.7.1B**	1
	**N3.7.1C**	1
	**N3.7.2A**	3
	**N4.4**	19
	**N4.2**	2
	**N4.5**	2
	**N3.9.2A**	2

Table 5. (Continued)

Structure	Code	Amount
		mol/100 mol
	**N3.9.2B**	1
	**N3.10**	1
	**N4.6.1**	5
	**N4.6.2**	5
	**N4.7A**	< 1
	**N4.7B**	< 1
	**N3.12A**	< 1
	**N3.12B**	< 1
	**N3.12C**	< 1
	**N4.8**	7

influence the distribution of glycoforms which make up the final product. Furthermore, variations found in the degree of sialylation may also have been caused by the use of various CHO cell clones, different culturing conditions which can lead to differences in the activity of the βGal:α2-3-sialyltransferase, and the recently reported possibility for the presence of CHO sialidase activity in the culture fluid [37]. Differences between other glycosyltransferase activities are indicated by differences in the relative amounts of triantennary (2,4-branched) and tri'antennary (2,6-branched) oligosaccharides in various rEPO samples. Whereas previous studies on rEPO from CHO cells showed ratios for tri/tri'antennary oligosaccharides between 1:1 and 2:1 [11, 13, 18], the present study shows a ratio of approximately 1:6. These different ratios cannot be artefacts caused by analytical problems as in the case of partial desialylation. They may be explained by differences in the activities of the involved glycosyltransferases for the various CHO cell lines and culture conditions.

For the incompletely sialylated N-linked carbohydrate chains, the following distribution of sialic acids over the different branches was observed (see Table 5). In tetraantennary oligosaccharides, Gal-**6** and Gal-**8** are sialylated to the highest degree, followed by Gal-**6'**, and then by Gal-**8'**. In tri'antennary oligosaccharides, Gal-**6** and Gal-**6'** are about twice as often sialylated as Gal-**8'**. For triantennary carbohydrate chains, the distribution could not be determined accurately because of the low amounts of material, and in diantennary oligosaccharides, the sialic acid residues are equally divided over the two branches. The occurrence of this non-random distribution of sialic acids can be explained by the branch-specific action of the involved sialyltransferases, although the possibility that some sialic acid residues have been released by sialidase cannot be excluded. However, to our knowledge, a branch-specific action of sialidase has not been reported. On the other hand, it is interesting that the observed distribution in tetra- and diantennary oligosaccharides is in good agreement with the branch specificity of human placental α2-3-sialyltransferase [38]. For tetraantennary oligosaccharides it was determined that the preferred acceptor sites are Gal-**6** and Gal-**8**, followed by Gal-**6'** and then by Gal-**8'**, whereas in diantennary oligosaccharides this enzyme has no preference for a specific branch. In tri'antennary oligosaccharides, the preferred acceptor site is Gal-**6**, followed by Gal-**8'**, and then by Gal-**6'**. The sialic acid distributions found in the present study contrast strongly with previous results obtained for rEPO from CHO cells, in which Gal-**6** was found to be terminal for most oligosaccharides [11].

Three different types of exclusively α2-3-linked sialic acids were found, namely, Neu5Ac (95%), Neu5Gc (2%) and $Neu5,9Ac_2$ (3%). It has been shown that Neu5Gc generally occurs in CHO-cell-derived recombinant glycoproteins [24]. $Neu5,9Ac_2$ has recently been found in a chimeric tissue plasminogen activator produced in CHO cells [30]. It should be noted that the present results indicate that the small peak eluting at about 11 ml in the sialic acid analysis elution profile of recombinant follicle-stimulating hormone from CHO cells (Fig. 1C in [24]) represents the 1,2-diamino-4,5-methylene-dioxybenzene derivative of $Neu5,9Ac_2$. Therefore, it is suggested that $Neu5,9Ac_2$ is also generally present in recombinant glycoproteins expressed in CHO cells. The implications of the presence of the different sialic acids for clinical use of these glycoproteins have been discussed [24, 30].

Repeating *N*-acetyllactosamine units were found to be present exclusively in the branches arising from the α1-6-linked Man residue in di-, tri'-, and tetraantennary oligosaccharides. The amount of oligosaccharides containing elongated branches is similar to that in previous studies. The carbohydrate chains of urinary EPO have in general a lower content of *N*-acetyllactosamine repeats [11, 13, 14]. In the present study, oligosaccharides containing more than two additional *N*-acetyllactosamine units were not detected, in contrast to previous investigations describing the finding of small amounts of carbohydrate chains with three repeats [11, 14, 18, 39]. The present results concerning the location of the repeats are in agreement with those reported in [13] but a somewhat different distribution of the re-

peats over the branches was reported in [11]. The location of the *N*-acetyllactosamine repeats in the branches attached to the α1-6-linked Man residue is in accordance with the reported specificity of GlcNAc-VII transferase from Novikoff tumor cell ascites fluid which synthesizes the GlcNAcβ1-3Gal linkage in repeating *N*-acetyllactosamine sequences [40].

No indications have been found for the presence of sulfated oligosaccharides in rEPO from CHO cells, which contrasts with the suggestion made in [41]. All Asn-linked GlcNAc residues were substituted with an α1-6-linked Fuc residue. Quantitative α1-6-fucosylation of recombinant glycoprotein glycans has also been reported for the N-linked carbohydrate chains of CHO-cell-derived chimeric tissue plasminogen activator [30] and rEPO from BHK cells [17]. Previous analyses of rEPO from CHO cells have shown a fucosylation of at least 85 % [11, 13, 18]. It is not known whether α1-6-fucosylation plays any role in the biological activity of rEPO.

Two O-linked carbohydrate chains, namely, Neu5Acα2-3Galβ1-3GalNAc-ol and Neu5Acα2-3Galβ1-3(Neu5Acα2-6)GalNAc-ol were identified. These oligosaccharides were also found in previous investigations on both CHO- and BHK-cell-derived rEPO [11, 17, 18]. In recent studies on the O-glycans of rEPO from CHO cells also Galβ1-3(Neu5Acα2-6)GalNAc-ol [16] and Galβ1-3GalNAc-GP, GalNAc-GP and Neu5Acα2-6GalNAc-GP [39] (GP = glycopeptide) were detected in minor amounts. No indications for the occurrence of these carbohydrate structures were found in the present study.

Comparison of the carbohydrate chains of rEPO with those of urinary EPO shows that in general the same type of N-linked oligosaccharides in terms of branching patterns are present, but the relative amounts of the individual structures are different. For urinary EPO a somewhat lower content of tetraantennary oligosaccharides has been found, although large differences can occur between different individuals [13, 14]. The earlier statement that the O-glycans of rEPO seem to be completely different from those of urinary EPO containing only GalNAc and Neu5Acα2-6GalNAc [16] contrasts with the findings in [39]. It should be noted that for urinary EPO contrasting results have been found concerning α2-3/α2-6-sialylation [11, 13].

By using endo-β-galactosidase to specifically degrade oligosaccharides containing *N*-acetyllactosamine repeats, typical ^{1}H-NMR data were obtained, useful for the identification of oligosaccharides that were not susceptible to endo-β-galactosidase. It is interesting that di- and tetraantennary oligosaccharides containing *N*-acetyllactosamine repeats could be degraded quantitatively by endo-β-galactosidase from *B. fragilis*, whereas tri'antennary oligosaccharides were digested for less than 15 %. In an experiment performed on an analytical scale (results not shown), it was observed that these tri'antennary oligosaccharides could be degraded by endo-β-galactosidase from *E. freundii* to a much higher extent (> 75 %), suggesting a specific difference in substrate specificity between the two enzymes.

Specific knowledge was collected about the chromatographic behaviour of closely related structures on Lichrosorb-NH_2 and CarboPac PA1. Whereas Lichrosorb-NH_2 is very well suited for separation based on the type and linkage of the sialic acid residues [19, 42, 43], chromatography on CarboPac PA1 is useful in many cases for separating isomeric carbohydrate chains [44] (and results in this study).

In conclusion, this study has shown that by using PNGase-F to release N-linked oligosaccharides combined with extensive purification and high-resolution ^{1}H-NMR spectroscopy for analysis, the native sialylation patterns of glycoprotein glycans can be determined, even when mixtures are present. Combination of ^{1}H-NMR spectroscopy with specific enzymic degradation is a powerful tool for the structure determination of carbohydrate chains which contain *N*-acetyllactosamine repeats. Using the above-mentioned methodology, the present study provides the most detailed analysis so far of the carbohydrate chains of rEPO from CHO cells, paying attention to the many aspects of carbohydrate chain microheterogeneity in glycoproteins.

This investigation was supported by the Netherlands Foundation for Chemical Research (SON) with financial aid from the Netherlands Organization for Scientific Research (NWO), 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 Organon International BV (Oss, The Netherlands).

REFERENCES

1. Krantz, S. B. (1991) *Blood 77*, 419–434.
2. Miyake, T., Kung, C. K.-H. & Goldwasser, E. (1977) *J. Biol. Chem. 252*, 5558–5564.
3. Lai, P.-H., Everett, R., Wang, F.-F., Arakawa, T. & Goldwasser, E. (1986) *J. Biol. Chem. 261*, 3116–3121.
4. Takeuchi, M. & Kobata, A. (1991) *Glycobiology 1*, 337–346.
5. Imai, N., Higuchi, M., Kawamura, A., Tomonoh, K., Oh-eda, M., Fujiwara, M., Shimonaka, Y. & Ochi, N. (1990) *Eur. J. Biochem. 194*, 457–462.
6. Owers Narhi, L., Arakawa, T., Aoki, K. H., Elmore, R., Rohde, M. F., Boone, T. & Strickland, T. W. (1991) *J. Biol. Chem. 266*, 23022–23036.
7. Sytkowski, A. J., Feldman, L. & Zurbuch, D. J. (1991) *Biochem. Biophys. Res. Commun. 176*, 698–704.
8. Yamaguchi, K., Akai, K., Kawanishi, G., Ueda, M., Masuda, S. & Sasaki, R. (1991) *J. Biol. Chem. 266*, 20434–20439.
9. Delorme, E., Lorenzini, T., Giffin, J., Martin, F., Jacobsen, F., Boone, T. & Elliott, S. (1992) *Biochemistry 31*, 9871–9876.
10. Higuchi, M., Oh-eda, M., Kuboniwa, H., Tomonoh, K., Shimonaka, Y. & Ochi, N. (1992) *J. Biol. Chem. 267*, 7703–7709.
11. Sasaki, H., Bothner, B., Dell, A. & Fukuda, M. (1987) *J. Biol. Chem. 262*, 12059–12076.
12. Sasaki, H., Ochi, N., Dell, A. & Fukuda, M. (1988) *Biochemistry 27*, 8618–8626.
13. Takeuchi, M., Takasaki, S., Miyazaki, H., Kato, T., Hoshi, S., Kochibe, N. & Kobata, A. (1988) *J. Biol. Chem. 263*, 3657–3663.
14. Tsuda, E., Goto, M., Murakami, A., Akai, K., Ueda, M., Kawanishi, G., Takahashi, N., Sasaki, H., Chiba, H., Ishihara, H., Mori, M., Tejima, S., Endo, S. & Arata, Y. (1988) *Biochemistry 27*, 5646–5654.
15. Rice, K. G., Takahashi, N., Namiki, Y., Tran, A. D., Lisi, P. J. & Lee, Y. C. (1992) *Anal. Biochem. 206*, 278–287.
16. Inoue, N., Takeuchi, M., Asano, K., Shimizu, R., Takasaki, S. & Kobata, A. (1993) *Arch. Biochem. Biophys. 301*, 375–378.
17. Nimtz, M., Martin, W., Wray, V., Klöppel, K.-D., Augustin, J. & Conradt, H. S. (1993) *Eur. J. Biochem. 213*, 39–56.
18. Watson, E., Bhide, A. & Van Halbeek, H. (1994) *Glycobiology 4*, 227–237.
19. Damm, J. B. L., Voshol, H., Hård, K., Kamerling, J. P. & Vliegenthart, J. F. G. (1989) *Eur. J. Biochem. 180*, 101–110.
20. 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.
21. Damm, J. B. L., Kamerling, J. P., Van Dedem, G. W. K. & Vliegenthart, J. F. G. (1987) *Glycoconjugate J. 4*, 129–144.
22. Kamerling, J. P. & Vliegenhart, J. F. G. (1989) in *Clinical biochemistry; principles, methods, applications.* vol. 1, *Mass Spectrometry* (Lawson, A. M., ed.) pp. 175–263, Walter de Gruyter, Berlin.
23. Hara, S., Yamaguchi, M., Takemori, Y., Furuhata, K., Ogura, H. & Nakamura, M. (1989) *Anal. Biochem. 179*, 162–166.
24. Hokke, C. H., Bergwerff, A. A., Van Dedem, G. W. K., Van Oostrum, J., Kamerling, J. P. & Vliegenthart, J. F. G. (1990) *FEBS Lett. 275*, 9–14.

25. Vliegenthart, J. F. G., Dorland, L. & Van Halbeek, H. (1983) *Adv. Carbohydr. Chem. Biochem. 41*, 209–374.
26. Hård, K., Van Zadelhoff, G., Moonen, P., Kamerling, J. P. & Vliegenthart, J. F. G. (1992) *Eur. J. Biochem. 209*, 895–915.
27. Bax, A. & Davis, D. G. (1985) *J. Magn. Reson. 65*, 355–360.
28. Hokke, C. H., Kamerling, J. P., Van Dedem, G. W. K. & Vliegenthart, J. F. G. (1991) *FEBS Lett. 286*, 18–24.
29. De Waard, P., Koorevaar, A., Kamerling, J. P. & Vliegenthart, J. F. G. (1991) *J. Biol. Chem. 266*, 4237–4243.
30. Bergwerff, A. A., Van Oostrum, J., Asselbergs, F. A. M., Bürgi, R., Hokke, C. H., Kamerling, J. P. & Vliegenthart, J. F. G. (1993) *Eur. J. Biochem. 212*, 639–656.
31. 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.
32. Brockhausen, I., Grey, A. A., Pang, H., Schachter, H. & Carver, J. P. (1988) *Glycoconjugate J. 5*, 419–448.
33. Shaw, L. & Schauer, R. (1988) *Biol. Chem. Hoppe-Seyler 369*, 477–486.
34. Shaw, L. & Schauer, R. (1989) *Biochem. J. 263*, 355–363.
35. Kamerling, J. P. & Vliegenthart, J. F. G. (1992) *Biol. Magn. Reson. 10*, 1–194.
36. Takeuchi, M., Inoue, N., Strickland, T. W., Kubota, M., Wada, M., Shimizu, R., Hoshi, S., Kozutsumi, H., Takasaki, S. & Kobata, A. (1989) *Proc. Natl Acad. Sci. USA 86*, 7819–7822.
37. Warner, T. G., Chang, J., Ferrari, J., Harris, R., McNerney, T., Bennet, G., Burnier, J. & Sliwkowski, M. B. (1994) *Glycobiology 3*, 455–463.
38. Nemansky, M. (1993) PhD Thesis, Vrije Universiteit van Amsterdam, ch. 4.
39. Linsley, K. B., Chan, S.-Y., Chan, S., Reinhold, B. B., Lisi, P. J. & Reinhold, V. N. (1994) *Anal. Biochem. 219*, 207–217.
40. Van den Eijnden, D. H., Koenderman, A. H. L. & Schiphorst, W. E. C. M. (1988) *J. Biol. Chem. 263*, 12461–12471.
41. Strickland, T. W., Adler, B., Aoki, K., Asher, S., Derby, P., Goldwasser, E. & Rogers, G. (1992) *J. Cell. Biochem. Suppl. 16D*, 167.
42. Van Pelt, J., Van Kuik, J. A., Kamerling, J. P., Vliegenthart, J. F. G., Van Diggelen, O. P. & Galjaard, H. (1988) *Eur. J. Biochem. 177*, 327–338.
43. Damm, J. B. L., Bergwerff, A. A., Hård, K., Kamerling, J. P. & Vliegenthart, J. F. G. (1988) *Recl. Trav. Chim. Pays-Bas 108*, 351–359.
44. Townsend, R. R. & Hardy, M. R. (1991) *Glycobiology 1*, 139–147.