Glycobiology vol. 10 no. 9 pp. 901–917, 2000

Characterization of the carbohydrate chains of the secreted form of the human epidermal growth factor receptor

Corné J.M.Stroop, Wolfgang Weber[2], Gerrit J.Gerwig, Manfred Nimtz[3], Johannis P.Kamerling and Johannes F.G.Vliegenthart[1]

Bijvoet Center, Department of Bio-Organic Chemistry, Utrecht University, P.O. Box 80075, 3508 TB Utrecht, The Netherlands, [2]Institut für Physiologische Chemie, Universitätskrankenhaus Eppendorf, D-20246 Hamburg, Germany, and [3]Gesellschaft für Biotechnologische Forschung mbH, Structure Research, Mascheroder Weg 1, D-38124 Braunschweig, Germany

Received on February 7, 2000; accepted on April 3, 2000

The human epidermal growth factor receptor (EGFR) is a transmembrane glycoprotein having 11 potential N-glycosylation sites in its extracellular domain. N-Glycosylation is needed for proper membrane insertion, EGF binding and receptor functioning. The human epidermoid carcinoma A431 cell line secretes a soluble 105 kDa glycoprotein (sEGFR) that represents the extracellular domain of the membrane-bound form, and its glycosylation pattern has been investigated. After liberation of the oligosaccharides from sEGFR with PNGase F, the glycans were fractionated along different routes, including Concanavalin A affinity chromatography, anion-exchange chromatography, HPLC and high-pH anion-exchange chromatography. The oligosaccharide fractions were characterized by 500- and 600-MHz ^{1}H-NMR spectroscopy and mass spectrometry (FAB, ESI, and MALDI-TOF). The oligomannose-type glycans range from $Man_5GlcNAc_2$ to $Man_8GlcNAc_2$ and account for 17% of the total carbohydrate moiety. Furthermore, di-, tri′- and tetraantennary complex-type structures are present, both neutral and (α2–3)-sialylated (up to tetrasialo), comprising 24 and 59%, respectively, of the total carbohydrate moiety. In this study, 32 new complex-type glycans are characterized containing the Le^x, Le^y, and sialyl-Le^x determinants, the bloodgroup A and H antigens, as well as the ALe^y determinant. This first comprehensive glycosylation study on a human nonrecombinant receptor shows the immense heterogeneity of the glycosylation of sEGFR.

Key words: human epidermal growth factor receptor/human epidermoid carcinoma A431 cell line/glycoprotein/NMR spectroscopy/mass spectrometry

Introduction

The human epidermal growth factor receptor (EGFR) is a 170 kDa glycoprotein, which binds EGF and mediates growth and differentiation of epithelial and fibroblastic cells. The entire cDNA and corresponding amino acid sequence of EGFR have been reported (Ullrich *et al.*, 1984). It has been found that the receptor consists of three functional domains: (1) an EGF-binding domain located on the external cell surface, (2) a single transmembrane domain with high hydrophobicity, and (3) a cytoplasmic domain possessing intrinsic protein tyrosine kinase activity (Ullrich and Schlessinger, 1990). Binding of EGF to the receptor initiates a cascade of intracellular anabolic events, including activation of tyrosine kinase activity leading to self-phosphorylation and phosphorylation of various intracellular substrates. Increased levels of EGFR are associated with different types of cancer, such as brain tumors and squamous carcinomas (Libermann *et al.*, 1984) and different ectodermal-derived malignancies (Dittadi *et al.*, 1991).

Earlier reports showed that the receptor domain is an 70.9 kDa single polypeptide having 10 to 11 potential N-glycosylation sites occupied with complex-type (predominantly tri- and/or tri′- and tetraantennary chains, with terminal Neu5Ac, Fuc, and α-GalNAc) and oligomannose-type (mainly $Man_8GlcNAc_2$) oligosaccharides in a ratio of approximately 2:1. No evidence was found for O-linked oligosaccharides (Mayes and Waterfield, 1984; Cummings *et al.*, 1985). The N-glycosylation and some glycan processing was suggested to be required for proper transport to the membrane and for establishing a functional conformation of the EGFR (Slieker and Lane, 1985; Defize *et al.*, 1988).

The human epidermoid carcinoma A431 cell line, used as a model to study EGFR, has a high density of receptors (2×10^6 receptors/cell) and secretes a soluble 105 kDa glycoprotein, termed secreted EGF receptor (sEGFR), which is related to the cell surface domain of the EGFR (Mayes and Waterfield, 1984; Weber *et al.*, 1984). This protein represents the amino-terminal extracellular domain of the EGFR lacking its transmembrane and cytoplasmic parts, and has an additional potential N-glycosylation site (Ullrich *et al.*, 1984). sEGFR has an intact ligand binding site and dimerizes in the presence of EGF (Lax *et al.*, 1991). The soluble receptor protein can be isolated in relatively large quantities from an A431 variant cell culture supernatant by immunoaffinity chromatography under acidic conditions using monoclonal antibodies against the EGFR. Until now, analysis of the glycosylation of sEGFR has been limited to Con A Sepharose chromatographic data of radioactively labeled sEGFR glycopeptides (Mangelsdorf Soderquist *et al.*, 1988). Crystals of sEGFR could be obtained in the presence of the ligand (Günther *et al.*, 1990; Degenhardt *et al.*, 1998); the 3D structure, however, is still awaited. The protein accounted for less than one-third of the crystal volume

[1]To whom correspondence should be addressed

indicating a high solvent content as might be expected for a highly glycosylated protein.

In this report, we present the fractionation and structure determination by high-resolution ^{1}H-NMR spectroscopy and mass spectrometry of part of the enzymatically released oligosaccharides of sEGFR. For the first time, such a comprehensive glycan analysis has been performed on a human non-recombinant growth factor receptor. It shows one of the most heterogeneous and complex ensembles of N-glycans published so far.

Results

Release and fractionation of the N-linked carbohydrate chains

Monosaccharide analysis of immunopurified sEGFR revealed a carbohydrate content of 32% (by mass) with Man, Fuc, Gal, GlcNAc, GalNAc, and Neu5Ac in the molar ratio of 3.0:1.4:2.1:4.1:0.4:1.0 (Man taken as 3.0). Two batches of sEGFR of 30 mg each were completely de-N-glycosylated by incubation with PNGase F, as was shown by SDS–PAGE. Chromatographic separation on Superdex 75pg yielded a protein, a carbohydrate, and a detergents and salts fraction. De-N-glycosylated sEGFR was subjected to monosaccharide analysis, demonstrating the absence of O-glycans. The pool of liberated oligosaccharides from batch I was applied to anion-exchange chromatography on Resource Q yielding five carbohydrate-positive fractions having elution positions corresponding to those of neutral, mono-, di-, tri/tri′- and tetrasialylated complex-type N-glycans, denoted **QN**–**Q4** (Table I). Gel permeation chromatography on Bio-Gel P-2 of **QN** afforded two neutral fractions, **QN1** and **QN2**, accounting for 24 and 17%, respectively, of the total glycan pool. Fraction **QN2** was subfractionated by HPLC on LiChrospher-NH_2 resulting in subfractions **QN2.1**–**QN2.4** (Figure 1). The oligosaccharides in **Q4** were fractionated by HPAEC on CarboPac PA-1 yielding subfractions **Q4.1**–**Q4.9** (Figure 2).

Because an HPLC chromatogram of fraction **QN1** (batch I) was too complex for fractionation, the pool of N-glycans from incubation batch II was first subjected to Concanavalin A Sepharose affinity chromatography (Con A) prior to separation on Resource Q. The Con A step resulted in three fractions (Figure 3), **QN2***, **Qd**, and **Qt**. Their elution positions correspond to oligomannose-type, diantennary complex-type, and tri- and/or tri′- and tetraantennary complex-type structures,

Table I. Anion-exchange chromatography of glycans liberated from sEGFR

Resource Q fraction[a]	Batch I (%)[b]
QN	41
Q1	23
Q2	18
Q3	12
Q4	6

[a]**QN**, Neutral; **Q1**, monocharged, etc.
[b]The percentage of a glycan fraction was based on the UV absorption at 214 nm and corrected for the average number of carbonyl groups of the components, as determined by HPLC or HPAEC, contributing to the different peaks.

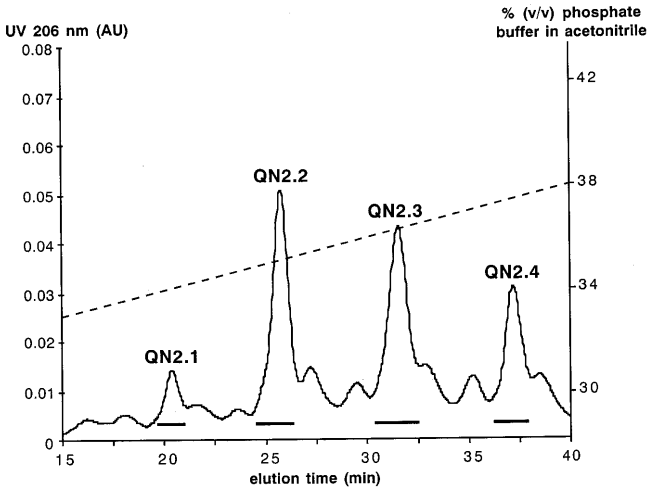


Fig. 1. HPLC elution profile on LiChrospher-NH_2 of the oligomannose-type fraction **QN2**. The eluate was monitored by UV detection at 206 nm. The column was eluted with a linear gradient of 30 to 38% (v/v) 10 mM phosphate buffer, pH 6.5, in acetonitrile for 40 min at a flow rate of 2 ml/min. The unnumbered peaks originated from an oscillation in the baseline.

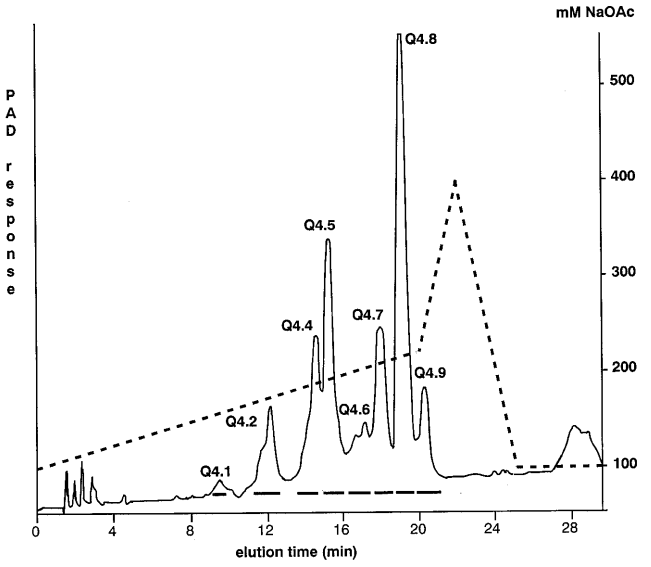


Fig. 2. HPAEC elution profile on CarboPac PA-1 of fraction **Q4**. The column (25 × 0.9 cm) was eluted with a linear gradient of 100 mM to 220 mM NaOAc in 80 ml 0.1 M NaOH, followed by a linear gradient of 220 mM to 400 mM NaOAc in 8 ml 0.1 M NaOH at a flow rate of 4 ml/min. The eluate was monitored by pulsed amperometric detection.

respectively. Fraction **QN2*** was similar to **QN2** from batch I. Separation according to charge on Resource Q of fraction **Qd** yielded three subfractions having elution positions corresponding to those of neutral (**QdN**), monosialylated (**Qd1**), and disialylated (**Qd2**) complex-type glycans (Figure 4). Fractions **QdN** and **Qd1** were subfractionated by HPLC on LiChrosorb-NH_2, whereas fraction **Qd2** was directly analyzed by ^{1}H-NMR spectroscopy. Separation of **QdN** resulted in subfractions **QdN.1**–**QdN.5** (Figure 5), whereas fraction **Qd1** yielded six subfractions denoted **Qd1.3**–**Qd1.8** (Figure 6).

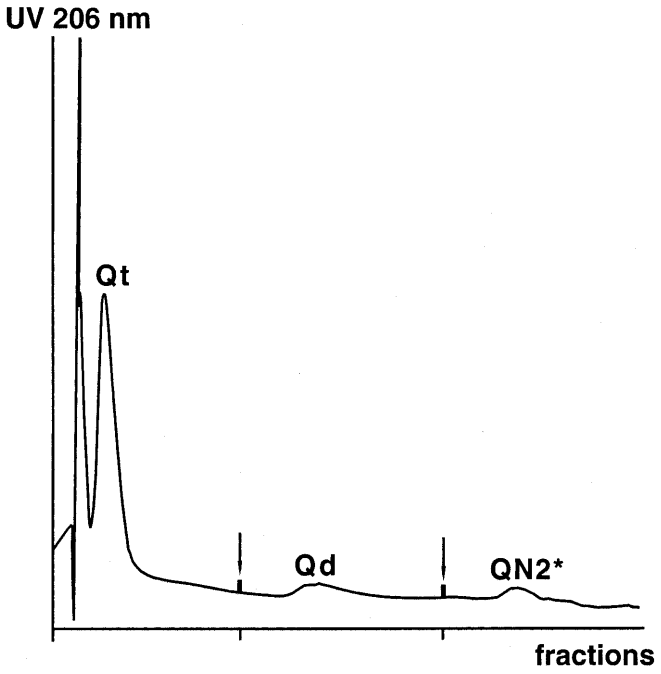


Fig. 3. Separation pattern at 206 nm on Concanavalin A Sepharose of the total glycan pool liberated from sEGFR. The elution was performed with 20 ml TBS buffer to collect fraction **Qt**. At the first arrow, 20 ml 10 mM methyl α-D-glucopyranoside in TBS was added to collect fraction **Qd**. At the second arrow, 20 ml 100 mM methyl α-D-mannopyranoside in TBS was added to collect fraction **QN2***.

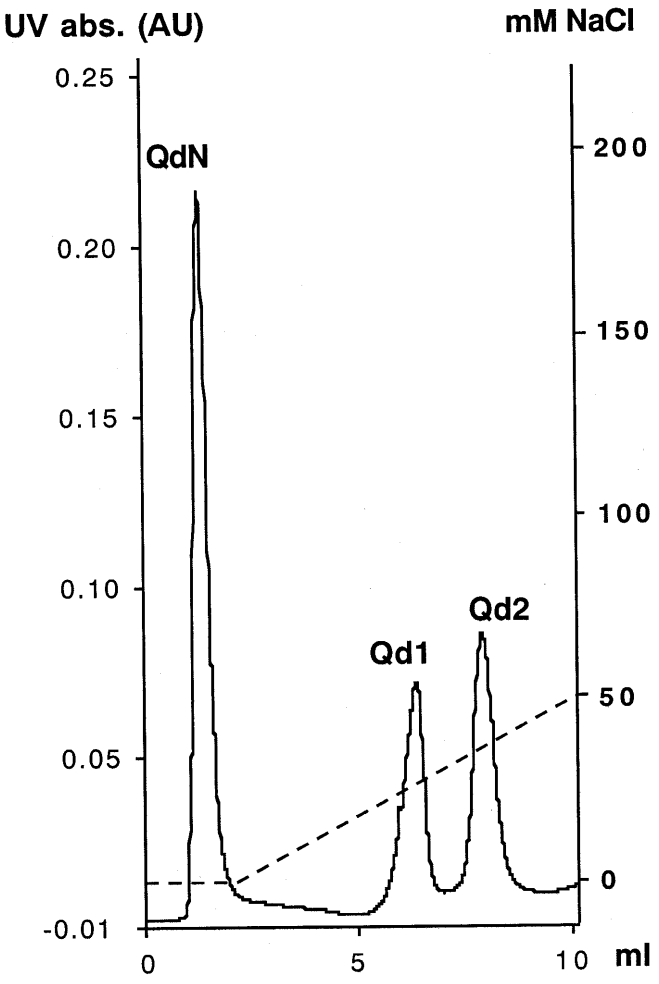


Fig. 4. Elution profile on Resource Q of the diantennary complex-type glycan fraction **Qd**. The column (3 × 0.64 cm) was eluted with 2 ml H_2O, followed by a linear gradient of 0 to 50 mM NaCl in 8 ml H_2O and a linear gradient of 50 to 250 mM NaCl in 8 ml H_2O at a flow rate of 4 ml/min. The eluate was monitored by conductivity measurement and UV (214 nm).

Resource Q chromatography was used for further fractionation of fraction **Qt** resulting in five fractions eluting at positions corresponding to 0–4 sialic acid residues, **QtN** and **Qt1**–**Qt4** (Figure 7), respectively. Then, fraction **QtN** was subfractionated on HPLC (LiChrosorb-NH_2) yielding subfractions **QtN.2**–**QtN.8** (Figure 8). Fraction **Qt4** was similar to **Q4** from batch I. The HPLC fractions **QtN.4**–**QtN.7** were too heterogeneous to be characterized by NMR spectroscopy and fractions **QtN.8**, **Q4.1**, and **Q4.6** did not contain enough carbohydrate material for a reliable structural analysis. Fractions **Q1**, **Q2**, and **Q3** from batch I and **Qt1**, **Qt2**, and **Qt3** from batch II are not discussed. These fractions are highly complex and are expected to contain the structural elements that are present in the fractions that are described in this report.

Neutral oligosaccharides

The relevant ^{1}H-NMR parameters of fractions **QN2.1**–**QN2.4** (Figure 1) in terms of structural-reporter-group signals are listed in Table II, which also includes the numbering of the residues. The ^{1}H-NMR spectra of all four fractions show chemical shifts typical for oligomannose-type structures. The NMR data of **QN2.1** are in accordance with those of glycan *U8* in Hård *et al.* (1991), indicating the $Man_5GlcNAc_2$ structure with two Man residues linked to Man-**4′**. Fraction **QN2.2** contains a $Man_6GlcNAc_2$ oligosaccharide, identified as an extension of structure **QN2.1** with Man-**C** (α1–2)-linked to Man-**4** (cf. compound *Man6*; Fu *et al.*, 1994). Fraction **QN2.3** contains a mixture of two structural isomers of a $Man_7GlcNAc_2$ glycan in about equimolar amounts. The NMR data of the two isomers, **QN2.3A** (**QN2.2** + Man-$\mathbf{D_3}$) and **QN2.3B** (**QN2.2** + Man-$\mathbf{D_1}$), match those of *U9.3* and *U9.1*, respectively (Hård *et al.*, 1991). Finally, the NMR data of **QN2.4** fit those of glycan *Man8(1)* (Fu *et al.*, 1994), $Man_8GlcNAc_2$ containing both Man-$\mathbf{D_1}$ and Man-$\mathbf{D_3}$.

The ^{1}H-NMR data of the neutral diantennary fractions **QdN.1**, **QdN.3**, **QdN.4** and **QdN.5** (Figure 5) are presented in Table III, and the full structures are depicted in Table IV. The fraction **QdN.2** is contaminated with **Qd1** as is evident from the typical NMR signals of a Neu5Ac residue. The NMR data of fraction **QdN.1** match those of a neutral complex-type (α1–6)-fucosylated diantennary structure with terminal β-Gal at both antennae (cf. compound *CN4*; Bergwerff *et al.*, 1995). Fraction **QdN.3** consists of an equimolar mixture of two diantennary isomers, **QdN.3A** and **QdN.3B**, differing in the antennary location of the blood group A type 2 determinant, GalNAc(α1–3)[Fuc(α1–2)]Gal(β1–4)GlcNAc. The occurrence of such an element is characterized by the structural-reporter-group signals of GalNAc (H-1, δ 5.180; H-2, δ 4.236; NAc, δ 2.039), Fuc (H-1, δ 5.350; H-5, δ 4.314; CH_3, δ 1.254), and Gal (H-1, δ 4.598; H-4, δ 4.212) (Dua *et al.*, 1986; Kamerling and Vliegenthart, 1992). The antennary position of the

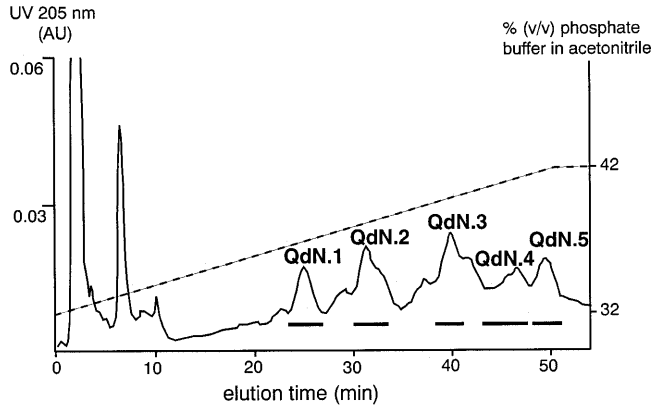


Fig. 5. HPLC elution profile on LiChrosorb-NH_2 of the neutral diantennary complex-type fraction **QdN**. The eluate was monitored by UV detection at 206 nm. The column was eluted with a linear gradient of 32 to 42% (v/v) 10 mM phosphate buffer, pH 6.5, in acetonitrile for 50 min at a flow rate of 2 ml/min.

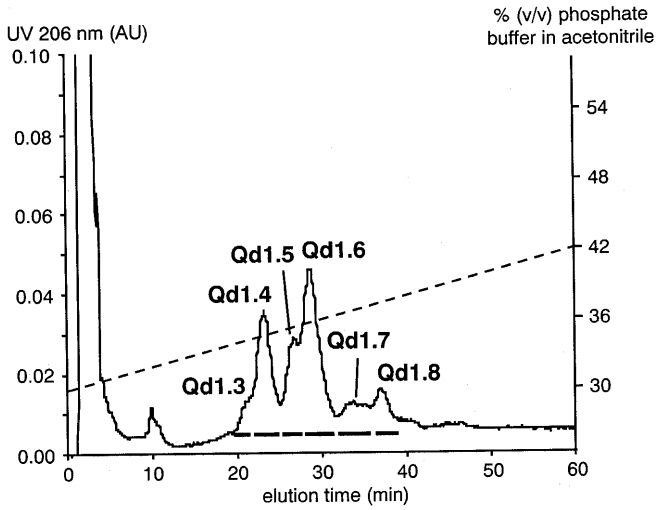


Fig. 6. HPLC elution profile on LiChrosorb-NH_2 of the monosialylated diantennary complex-type fraction **Qd1**. The eluate was monitored by UV detection at 206 nm. The column was eluted with a linear gradient of 30 to 44% (v/v) 10 mM phosphate buffer, pH 6.5, in acetonitrile for 70 min at a flow rate of 2 ml/min.

A type 2 determinant can be derived from the δ values of Man-**4′** H-1 and H-2, and of terminal Gal-**6/6′** H-1. In **QdN.3A**, the determinant is linked to Man-**4** (Man-**4′** H-1, δ 4.926 and H-2, δ 4.104; Gal-**6′** H-1, δ 4.474) and, consequently, in **QdN.3B** to Man-**4′** (Man-**4′** H-1, δ 4.909 and H-2, δ 4.091; Gal-**6** H-1, δ 4.469). Fraction **QdN.4** mainly consists of a diantennary oligosaccharide with the Le^x determinant in each of the antennae (cf. compound *GP22*; Michalski *et al.*, 1991). The NMR spectrum of fraction **QdN.5** shows a new combination of anomeric signals that fit the set of structural reporters of the ALe^y antigen, earlier established for this epitope in mucin-derived O-glycans (Strecker *et al.*, 1992): (α1–3)-linked GalNAc H-1 (δ 5.198), (α1–2)-linked Fuc H-1 (δ 5.295) and CH_3 (δ 1.284), and (α1–3)-linked Fuc H-1

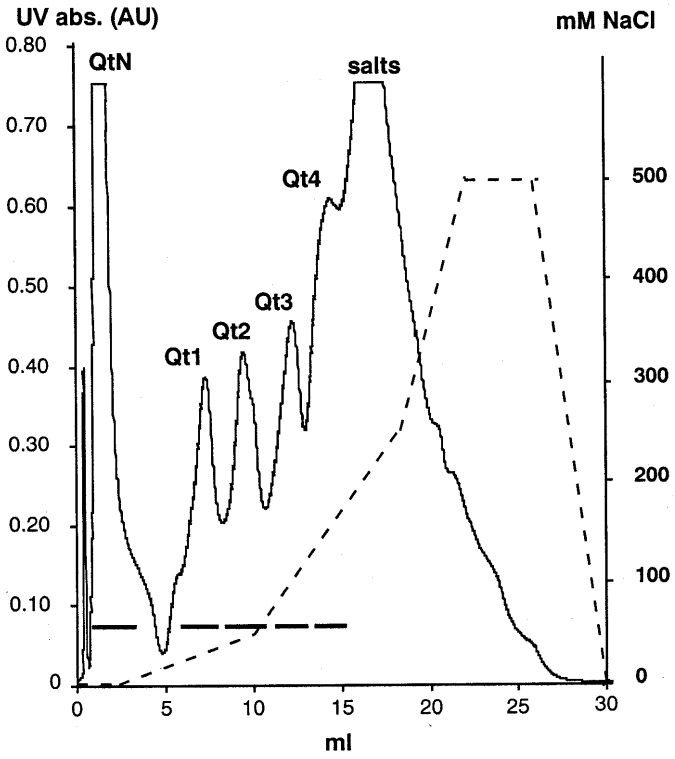


Fig. 7. Elution profile on Resource Q of the tri/tri′- and tetraantennary complex-type glycan fraction **Qt**. The column (3 × 0.64 cm) was eluted with 2 ml H_2O, followed by a linear gradient of 0 to 50 mM NaCl in 8 ml H_2O and a linear gradient of 50 to 250 mM NaCl in 8 ml H_2O and, finally, a linear gradient of 250 to 500 mM NaCl in 4 ml H_2O, at a flow rate of 4 ml/min. The eluate was monitored by conductivity measurement and UV (214 nm).

(δ 5.126) and CH_3 (δ 1.274). When compared with the A determinant (see **QdN.3**), shifts for the GalNAc H-1 signal of +0.018 p.p.m. and for the (α1–2)-linked Fuc H-1 signal of –0.055 p.p.m. are due to the proximity of the (α1–3)-linked Fuc residue. The ALe^y determinant occurs exclusively in the lower branch as indicated by the shifts of Man-**4** H-1 (δ 5.095) and H-2 (δ 4.157), when compared with the NMR data of **QdN.1** (Table III).

The ^{1}H-NMR data of the neutral tri′antennary fraction **QtN.2** (Figure 8) are presented in Table III, and the full structures are depicted in Table IV. Characteristic H-1 signals for Man-**4** and Man-**4′** (δ 5.127 and δ 4.870, respectively) combined with the Man-**3**, -**4**, and -**4′** H-2 signals show that fraction **QtN.2** consists of structures having three antennae composed of *N*-acetyllactosamine (LacNAc) units [Gal(β1–4)GlcNAc], including one antenna (β1–6)-linked to Man-**4′** (cf. tri′antennary structure *11*; Vliegenthart *et al.*, 1983). The major oligosaccharides in fraction **QtN.2** have a tri′antennary structure with (**QtN.2A**; GlcNAc-**2** H-1, δ 4.666 and NAc, δ 2.095) and without (**QtN.2B**; GlcNAc-**2** H-1, δ 4.615 and NAc, δ 2.079) core fucosylation. Minor structures are present in fraction **QtN.2** containing an (α1–2)-linked Fuc residue in a blood group H determinant (Michalski *et al.*, 1991; Kamerling and Vliegenthart, 1992), which is reflected by the set of Fuc H-1 (δ 5.303), H-5 (δ 4.219), and CH_3 (δ 1.235) structural-reporter-group signals (2D-TOCSY NMR analysis). Another set of Fuc H-1 (δ 5.351), H-5 (δ 4.312), and CH_3 (δ 1.250) signals

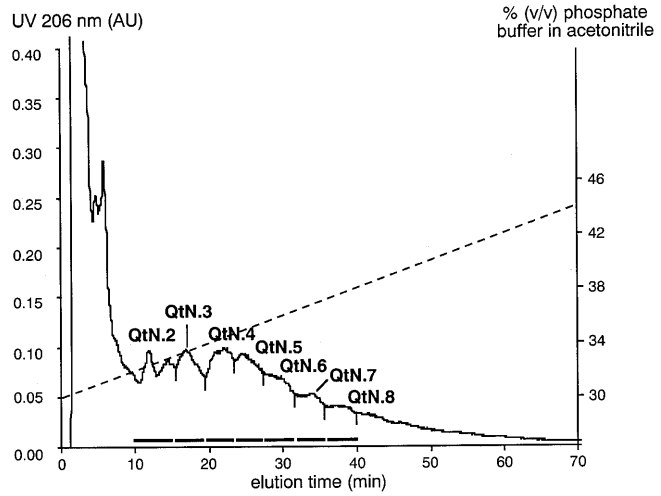


Fig. 8. HPLC elution profile on LiChrosorb-NH_2 of the neutral tri/tri′- and tetraantennary complex-type fraction **QtN**. The eluate was monitored by UV detection at 206 nm. The column was eluted with a linear gradient of 30 to 44% (v/v) 10 mM phosphate buffer, pH 6.5, in acetonitrile for 70 min at a flow rate of 2 ml/min.

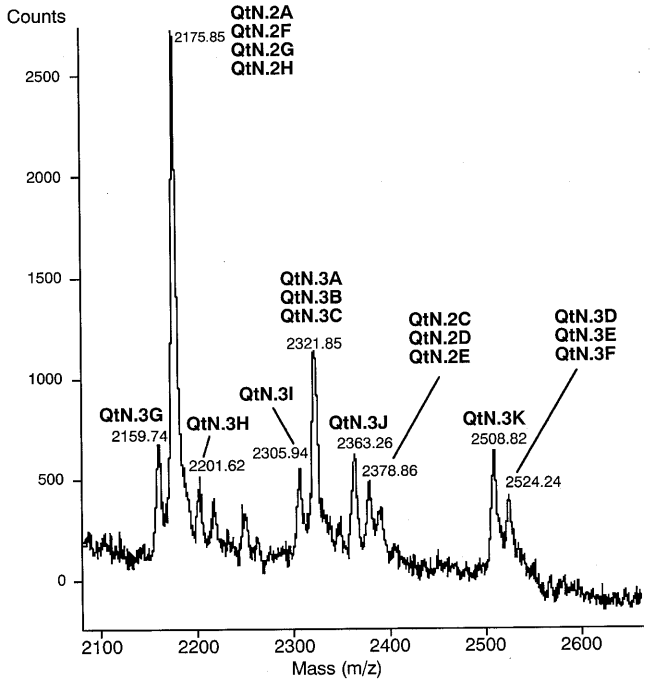


Fig. 9. MALDI-TOF-MS spectrum of neutral complex-type glycan fraction **QtN.3**. The codes refer to the structures summarized in Tables IV and V.

together with the GalNAc set (H-1, δ 5.182; H-2, δ 4.236; NAc, δ 2.039) indicates the presence of a blood group A determinant (see **QdN.3**). The attachment of both GalNAc and Fuc to terminal Gal results in a shift of the NAc signal of the adjacent GlcNAc residue to δ 2.061 for GlcNAc-**5** (in **QtN.2C**) and to δ 2.056 for GlcNAc-**5′** (in **QtN.2D**) and/or GlcNAc-**7′** (in **QtN.2E**). The change in the chemical shift of Gal H-1 upon addition of an (α1–2)-linked Fuc residue (in an H epitope) is 0.063 p.p.m. downfield (Vliegenthart *et al.*, 1983) resulting in signals at δ 4.532 for both Gal-**6** (in **QtN.2F**) and Gal-**6′** (in **QtN.2G**), and at δ 4.546 for Gal-**8′** (in **QtN.2H**). The presence of an (α1–2)-linked Fuc to a Gal residue, in both an A and H determinant, influences the H-1 and H-2 signals of the Man residue in the corresponding antenna, as is shown by a 2D-TOCSY experiment: Man-**4′** H-1 (δ 4.85) and H-2 (δ 4.07) for structures **QtN.2D** and **QtN.2G**, and Man-**4** H-1 (δ 5.12) and H-2 (δ 4.18) for structure **QtN.2C** and **QtN.2F**. Based on these findings, the presence of structures **QtN.2C**–**QtN.2H** is indicated (Table IV). For a discussion of the core-fucosylated analogues, that cannot be excluded as minor components, see the description of fraction **QtN.3**.

Fraction **QtN.3** was analyzed by MALDI-TOF-MS showing overlap with fraction **QtN.2**. The mass spectrum (Figure 9) shows two major sodiated molecular ions at *m/z* 2175 [$Hex_6HexNAc_5dHex$ + Na] and *m/z* 2321 [$Hex_6HexNAc_5$ $dHex_2$ + Na] corresponding to glycan **QtN.2A** and glycans **QtN.3A**–**QtN.3C**, respectively. Structures **QtN.3A**–**QtN.3C** only differ in the position of the (α1–2)-linked Fuc residue, but are all core-fucosylated. The peak at *m/z* 2175 can also originate from structures **QtN.2F** – **QtN.2H**, although the NMR data showed that they cannot be major components. Minor sodiated molecular ions at *m/z* 2378 [$Hex_6HexNAc_6dHex$ + Na] and *m/z* 2524 [$Hex_6HexNAc_6dHex_2$ + Na] are in agreement with the structures **QtN.2C**–**QtN.2E** and **QtN.3D**–**QtN.3F**, respectively. Several other minor sodiated molecular ions (*m/z* 2159, 2201, 2305, 2363, and 2508) indicate the presence of glycans with antennae that lack one or two terminal Gal residues (structures **QtN.3G** through **QtN.3K** in Figure 9, proposed structures in Table V). ^{1}H-NMR analysis (1D and 2D-TOCSY) revealed that fraction **QtN.3** contains the same structural-reporter-group signals (Table III), including those of A and H determinants, as fraction **QtN.2**. However, the relatively low intensities of the GlcNAc-**2** NAc (δ 2.080) and H-1 (δ 4.615) signals, specific for a fucose-free core, are indicative of a higher amount of core fucosylation in **QtN.3**, and the MALDI-TOF data have been interpreted accordingly. Glycans **QtN.3A**–**QtN.3C** (Table IV) are the core-fucosylated forms of glycans **QtN.2F**–**QtN.2H**, whereas oligosaccharides **QtN.3D**–**QtN.3F** (Table IV) are the core-fucosylated forms of oligosaccharides **QtN.2C**–**QtN.2E**. The positions of the H and A determinants are indicated by the H-1 and H-2 signals of the Man residue in the corresponding antenna, as is shown by a 2D-TOCSY experiment: Man-**4′** H-1 (δ 4.85) for structures **QtN.3B** and **QtN.3E**, and Man-**4** H-1 (δ 5.116) and H-2 (δ 4.17) for structure **QtN.3A** and **QtN.3D**. It may be evident that on the basis of the present information structural variants of **QtN.3A**–**QtN.3F** with an extra H determinant instead of core fucosylation cannot be excluded.

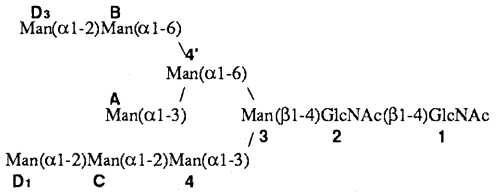
Monosialylated diantennary oligosaccharides

The ^{1}H-NMR data of fractions **Qd1.3**–**Qd1.6** (Figure 6) are presented in Table VI, and a full structure including the numbering of the residues is depicted below (structure **Qd1.5D**). Fraction **Qd1.3** contains a non-fucosylated (α2–3)-monosialylated diantennary structure (cf. glycan *N1.2*; Hård *et al.*, 1990). Fraction **Qd1.4** contains the (α1–6)-fucosylated

Table II. ^{1}H-NMR chemical shifts of structural-reporter-group protons of the constituent monosaccharides of oligomannose-type N-glycans liberated from sEGFR. Chemical shifts are given relative to internal acetone (δ 2.225) in ^2H_2O at 300 K and p^{2}H 7. Compounds are represented by symbolic shorthand notation: solid diamonds, Man; solid circles, GlcNAc.

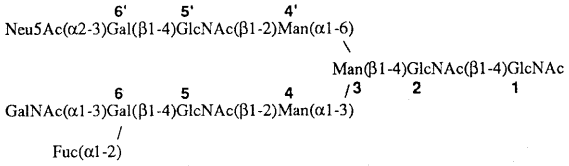
		QN2.1	QN2.2	QN2.3A	QN2.3B	QN2.4[b]
Relative amount[a]:		7%	35%	18%	18%	21%
Reporter group	Residue	Chemical shift (ppm)				
H-1	GlcNAc-1α	5.189	5.184	5.187	5.187	5.183
	Man-**4**	5.095	5.347	5.347	5.339	5.338
	Man-**4'**	4.870	4.872	4.869	4.869	4.867
	Man-**A**	5.095	5.093	5.090	5.090	5.085
	Man-**B**	4.906	4.908	5.145	4.909	5.145
	Man-**C**	–	5.052	5.054	5.306	5.304
	Man-**D_1**	–	–	–	5.042	5.041
	Man-**D_3**	–	–	5.042	–	5.041
H-2	Man-**3**	4.253	4.233	4.234	4.234	4.231
	Man-**4**	4.076	4.115	4.110	4.110	4.104
	Man-**4'**	4.146	4.147	4.148	4.148	4.149
	Man-**A**	4.068	4.064	4.068	4.068	4.067
	Man-**B**	3.981	3.982	4.020	3.990	4.025
	Man-**C**	–	4.064	4.068	4.110	4.104
	Man-**D_1**	–	–	–	4.068	4.067
	Man-**D_3**	–	–	4.068	–	4.067
NAc	GlcNAc-**1**	2.037	2.038	2.038	2.038	2.036
	GlcNAc-**2**	2.063	2.063	2.063	2.063	2.061

[a] The amounts are given as molar percentages of the total **QN2** fraction. Results are based on peak areas of HPLC peaks on LiChrospher-NH_2.

[b]


analogue of compound **Qd1.3** (cf. glycan *N1.3*; Hård *et al.*, 1992).

^{1}H-NMR spectroscopy, including 2D-TOCSY, indicates that fraction **Qd1.6** consists of the (α1–6)-fucosylated analogues of the glycans in **Qd1.5** (Table VI). Each of the two fractions includes four structures, in which a Neu5Ac residue is (α2–3)-linked to one of the antennae. The other antenna carries the

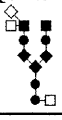
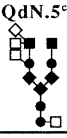
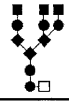
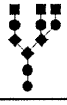
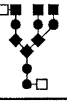


Qd1.5D

blood group A (Fuc H-1, δ 5.343, H-5, δ 4.309, and CH_3, δ 1.249; GalNAc H-1, δ 5.179, H-2, δ 4.236, and NAc, δ 2.040; Gal H-1, δ 4.600 and H-4, δ 4.222) or Le^x (Fuc H-1, δ 5.125, H-5, δ 4.827, and CH_3, δ 1.176) determinant.

In oligosaccharide **Qd1.6A**, the presence of the Le^x determinant in the (α1–6) branch is indicated by the structural-reporter-group signals of Gal-**6'** (H-1, δ 4.450) and Man-**4'** (H-1, δ 4.909) (compare with **QdN.4** and **Qd1.4**). Not only the presence of a Fuc residue on GlcNAc-**5'** influences the chemical shift of Man-**4'** H-1, but also Fuc (α1–2)-linked (in an A epitope) to Gal-**6'** has such an effect (see **QdN.3B**), resulting in δ 4.909 for Man-**4'** H-1 in structure **Qd1.6B**. Neu5Ac in structures **Qd1.6C** and **Qd1.6D** is (α2–3)-linked to Gal-**6'**, as is shown by the chemical shift of Gal-**6'** H-1 at δ 4.549 (cf. glycan *N2.5*; Hård *et al.*, 1992). The Gal-**6** H-1 chemical shifts,

Table III. ^{1}H-NMR chemical shifts of structural-reporter-group protons of the constituent monosaccharides of neutral complex-type N-glycans liberated from sEGFR. Chemical shifts are given relative to internal acetone (δ 2.225) in ^2H_2O at 300 K and p^{2}H 7. Compounds are represented by symbolic shorthand notation: solid circles, GlcNAc; solid diamonds, Man; open squares, Fuc; solid squares, Gal; open diamonds, GalNAc. α and β stand for the anomeric configuration of GlcNAc-**1**. For numbering of the monosaccharide residues, see Table IV.

Reporter group	Residue	Chemical shift (ppm)								
		QdN.1	QdN.3A	QdN.3B	QdN.4[c]	QdN.5[c]	QtN.2A	QtN.2B	QtN.3A	QtN.3D
										
H-1	GlcNAc-1α	5.178	5.180	5.180	5.179	5.178	5.182	5.182	5.180	5.180
	GlcNAc-1β	4.696	4.693	4.693	nd	nd	4.695	4.695	4.692	4.692
	GlcNAc-2[α]	4.667	4.664	4.664	4.667	4.664	4.663	4.615	4.663	4.663
	GlcNAc-2[β]	4.672	4.669	4.669		4.668	4.669		4.668	4.668
	Man-**4**	5.121	5.118	5.118	5.107	5.095	5.127	5.127	5.116	5.116
	Man-**4'**	4.923	4.926	4.909	4.910	4.926	4.870	4.870	4.868	4.868
	GlcNAc-**5**	4.582	4.583	4.583	4.586	4.586	4.581	4.581	nd	nd
	GlcNAc-**5'**	4.582	4.583	4.583	4.586	4.584	4.595[e]	4.595[e]	4.596	4.596
	Gal-**6**	4.468	4.598	4.469	4.441	4.550	4.469	4.469	4.536	4.599
	Gal-**6'**	4.471	4.474	4.598	4.446	4.474	4.469	4.469	4.470	4.470
	GlcNAc-**7'**	–	–	–	–	–	4.546	4.546	4.552	4.552
	Gal-**8'**	–	–	–	–	–	4.483	4.483	4.482	4.482
	(α1-6)Fuc[α]	4.886	4.890	4.890	4.888	4.890	4.896	–	4.894	4.894
	(α1-6)Fuc[β]	4.895	4.897	4.897	4.894	4.896		–	4.900	4.900
	(α1-3)Fuc	–	–	–	5.125[d]	5.126	–	–	–	–
	(α1-2)Fuc	–	5.350	5.350	–	5.295	–	–	5.303	5.350
	GalNAc	–	5.180	5.180	–	5.198	–	–	–	5.180
H-2	Man-**3**	4.247	4.252	4.252	4.247	4.247	4.254	4.254	4.252	4.252
	Man-**4**	4.189	4.188	4.188	4.187	4.157	4.196	4.196	4.17[f]	4.17[f]
	Man-**4'**	4.106	4.104	4.091	4.095	4.106	4.091	4.091	4.093	4.093
	GalNAc	–	4.236	4.236	–	4.235	–	–	–	4.233[f]
H-3	Gal-**6**	nd[a]	nd	nd	nd	nd	3.66[f]	3.66[f]	nd	3.98[f]
	Gal-**6'**	nd	nd	nd	nd	nd	3.66[f]	3.66[f]	nd	3.67[f]
H-4	Gal-**6**	nd	4.212	nd	nd	4.196	3.92[f]	3.92[f]	nd	4.223[f]
	Gal-**6'**	nd	nd	4.212	nd	nd	3.92[f]	3.92[f]	nd	3.93[f]
	Gal-**8'**	–	–	–	–	–	3.94[f]	3.94[f]	nd	3.93[f]
H-5	(α1-6)Fuc[α]	4.10[b]	4.096	4.096	4.101	4.100	4.095[f]	–	–	4.094[f]
	(α1-6)Fuc[β]	4.13[b]	4.123	4.123	4.13[b]	4.125	4.133	–	–	4.130[f]
	(α1-3)Fuc	–	–	–	4.83[b]	4.86[b]	–	–	–	–
	(α1-2)Fuc	–	4.314	4.314	–	4.324	–	–	4.221[f]	4.314[f]
CH_3	(α1-6)Fuc[α]	1.208	1.209	1.209	1.207	1.208	1.209	–	–	1.208
	(α1-6)Fuc[β]	1.222	1.221	1.221	1.22[b]	1.219	1.221	–	–	1.220
	(α1-3)Fuc (at **5**)	–	–	–	1.171	1.274	–	–	–	–
	(at **5'**)	–	–	–	1.179	–	–	–	–	–
	(α1-2)Fuc	–	1.254	1.254	–	1.284	–	–	1.237[f]	1.253[f]
NAc	GlcNAc-**1**	2.038	2.039	2.039	2.040	2.039	2.039	2.039	2.040	2.040
	GlcNAc-**2**	2.094	2.096	2.096	2.095	2.094	2.095	2.079	2.095	2.095
	GlcNAc-**5**	2.047	2.056	2.048	2.040	2.047	2.051	2.051	2.055	2.055
	GlcNAc-**5'**	2.047	2.048	2.056	2.040	2.047	2.039	2.039	2.040	2.040
	GlcNAc-**7'**	–	–	–	–	–	2.039	2.039	2.040	2.040
	GalNAc	–	2.039	2.039	–	2.034	–	–	–	2.040

[a]nd, not determined
[b]Value given in two decimals due to spectral overlap
[c]Spectrum recorded at 600 MHz
[d]Signal stemming from two protons
[e]Value obtained from a 2D-NOESY experiment
[f]Value obtained from a 600-MHz 2D-TOCSY experiment

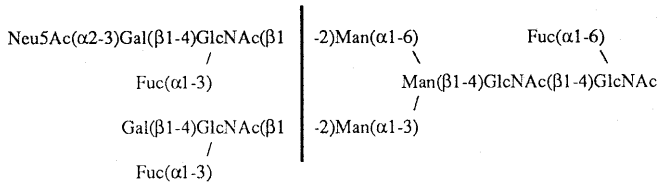
δ 4.445 for **Qd1.6C** and δ 4.600 for **Qd1.6D**, indicate the Le^x and the blood group A epitope, respectively.

The reduced and permethylated fraction **Qd1.7** shows a major sodiated molecular ion in MALDI-TOF-MS at *m/z* 2970 [Neu5AcHex$_5$HexNAc$_3$dHex$_3$HexNAc-ol + Na] and a minor sodiated molecular ion at *m/z* 2796 [Neu5AcHex$_5$ HexNAc$_3$dHex$_2$HexNAc-ol + Na] differing only by the mass of a dHex residue (Fuc, see monosaccharide analysis). The composition of the antennae in reduced and permethylated fraction **Qd1.7** was determined in more detail by methylation analysis and ESI-MS-MS. Methylation analysis revealed the presence of terminal Fuc, 3,4-disubstituted GlcNAc, terminal Gal, and 3-substituted Gal residues, supporting structure **Qd1.7A** (with the sialyl-Le^x and the Le^x determinant). In ESI-MS, the major disodiated molecular ion at *m/z* 1496 [Neu5AcHex$_5$HexNAc$_3$dHex$_3$HexNAc-ol + 2 Na] was selected for further fragmentation. The presence of a fragment at *m/z* 490 [dHexHexNAc-ol + Na] shows that the major molecular-ion peak corresponds to core-fucosylated glycans. Fragments at *m/z* 1021 [Neu5AcHexHexNAcdHex + Na] (and *m/z* 646: loss of

Table IV. Neutral complex-type glycans linked to sEGFR

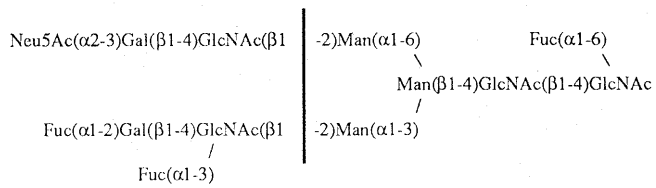
Code	Glycan structure
QdN.1	Gal(β1-4)GlcNAc(β1-2)Man(α1-6) Fuc(α1-6) \ \ Man(β1-4)GlcNAc(β1-4)GlcNAc / Gal(β1-4)GlcNAc(β1-2)Man(α1-3)
QdN.3A	Gal(β1-4)GlcNAc(β1-2)Man(α1-6) Fuc(α1-6) \ \ Man(β1-4)GlcNAc(β1-4)GlcNAc / GalNAc(α1-3)Gal(β1-4)GlcNAc(β1-2)Man(α1-3) / Fuc(α1-2)
QdN.3B	GalNAc(α1-3)Gal(β1-4)GlcNAc(β1-2)Man(α1-6) Fuc(α1-6) / \ Fuc(α1-2) Man(β1-4)GlcNAc(β1-4)GlcNAc / Gal(β1-4)GlcNAc(β1-2)Man(α1-3)
QdN.4	Gal(β1-4)GlcNAc(β1-2)Man(α1-6) Fuc(α1-6) / \ \ Fuc(α1-3) Man(β1-4)GlcNAc(β1-4)GlcNAc / Gal(β1-4)GlcNAc(β1-2)Man(α1-3) / Fuc(α1-3)
QdN.5	Gal(β1-4)GlcNAc(β1-2)Man(α1-6) Fuc(α1-6) \ \ Man(β1-4)GlcNAc(β1-4)GlcNAc / GalNAc(α1-3)Gal(β1-4)GlcNAc(β1-2)Man(α1-3) / / Fuc(α1-2) Fuc(α1-3)
QtN.2A	Gal(β1-4)GlcNAc(β1-6) \ Gal(β1-4)GlcNAc(β1-2)Man(α1-6) Fuc(α1-6) \ \ Man(β1-4)GlcNAc(β1-4)GlcNAc / Gal(β1-4)GlcNAc(β1-2)Man(α1-3)
QtN.2B	Gal(β1-4)GlcNAc(β1-6) \ Gal(β1-4)GlcNAc(β1-2)Man(α1-6) \ Man(β1-4)GlcNAc(β1-4)GlcNAc / Gal(β1-4)GlcNAc(β1-2)Man(α1-3)
QtN.2C	Gal(β1-4)GlcNAc(β1-6) \ Gal(β1-4)GlcNAc(β1-2)Man(α1-6) \ Man(β1-4)GlcNAc(β1-4)GlcNAc / GalNAc(α1-3)Gal(β1-4)GlcNAc(β1-2)Man(α1-3) / Fuc(α1-2)

Neu5Ac) and *m/z* 660 [HexHexNAcdHex + Na] indicate the presence of the sialyl-Le^x determinant and a fucosylated Gal-GlcNAc unit in different antennae. A vertical bar is used to indicate that the position at the antennae is not specified.



Qd1.7A

Other fragments in the ESI-MS-MS spectrum of *m/z* 1496 of fraction **Qd1.7** at *m/z* 847 [Neu5AcHexHexNAc + Na] and *m/z* 472 (loss of Neu5Ac from *m/z* 847) originate from a sialylated, non-fucosylated antenna, whereas the peaks at *m/z* 834 [HexHexNAcdHex$_2$ + Na] and *m/z* 433 [HexdHex + Na] represent a Gal-GlcNAc unit bearing two Fuc residues. In the methylation analysis, apart from the before-mentioned residues, also a 2-substituted Gal residue is present. This structural information gives rise to structure **Qd1.7B**, containing the Le^y antigen.



Qd1.7B

Methylation analysis shows that 3-substituted Gal (the attachment site for Neu5Ac) is present in a ratio of 2:1 to terminal Gal, implying that structure **Qd1.7A**, with a terminal Gal, accounts for 50% of the structures in fraction **Qd1.7**. The fragment ions at *m/z* 834 (typical for **Qd1.7B**) and *m/z* 660 (typical for **Qd1.7A**) are equally intense, therefore, a structural isomer of glycan

Table IV. *Continued*

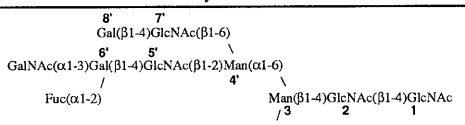
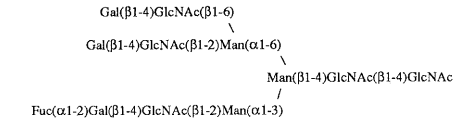
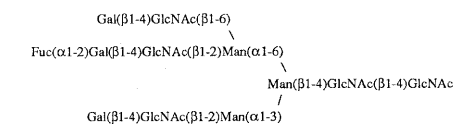
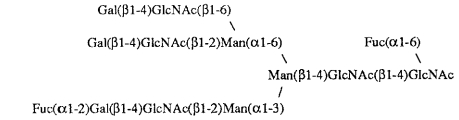
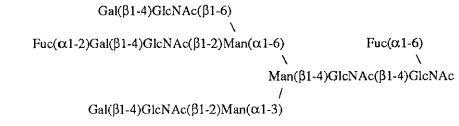
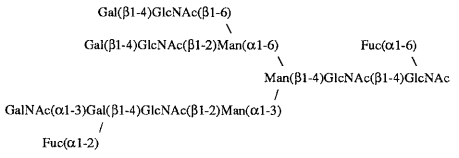
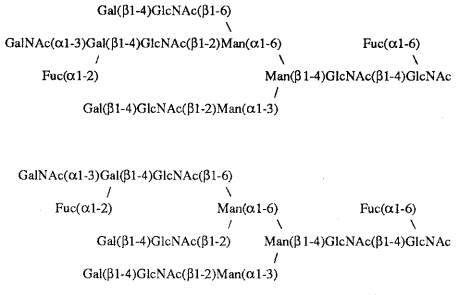
Code	Glycan structure
QtN.2D and/or **QtN.2E**	
QtN.2F	
QtN.2G and/or **QtN.2H**	
QtN.3A	
QtN.3B and/or **QtN.3C**	
QtN.3D	
QtN.3E and/or **QtN.3F**	

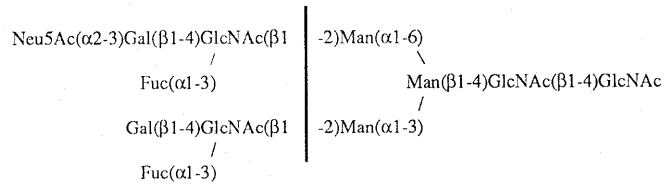
Table V. Proposed structures for tri′antennary neutral complex-type N-linked carbohydrate chains based on MALDI-TOF-MS. A vertical bar is used to indicated that the antennae is not specified.

Code	*m/z*	Composition	Proposed structure
QtN.3G	2159	$Hex_5HexNAc_5dHex_2$ + Na	Fuc(α1-2)Gal(β1-4) Gal(β1-4) | GlcNAc(β1-6) GlcNAc(β1-2)Man(α1-6) Fuc(α1-6) Man(β1-4)GlcNAc(β1-4)GlcNAc GlcNAc(β1-2)Man(α1-3)
QtN.3H	2201	$Hex_4HexNAc_6dHex_2$ + Na	GalNAc(α1-3)Gal(β1-4) Fuc(α1-2) | GlcNAc(β1-6) GlcNAc(β1-2)Man(α1-6) Fuc(α1-6) Man(β1-4)GlcNAc(β1-4)GlcNAc GlcNAc(β1-2)Man(α1-3)
QtN.3I	2305	$Hex_5HexNAc_5dHex_3$ + Na	Fuc(α1-2)Gal(β1-4) Fuc(α1-2)Gal(β1-4) | GlcNAc(β1-6) GlcNAc(β1-2)Man(α1-6) Fuc(α1-6) Man(β1-4)GlcNAc(β1-4)GlcNAc GlcNAc(β1-2)Man(α1-3)
QtN.3J	2363	$Hex_5HexNAc_6dHex_2$ + Na	GalNAc(α1-3)Gal(β1-4) Fuc(α1-2) Gal(β1-4) | GlcNAc(β1-6) GlcNAc(β1-2)Man(α1-6) Fuc(α1-6) Man(β1-4)GlcNAc(β1-4)GlcNAc GlcNAc(β1-2)Man(α1-3)
QtN.3K	2508	$Hex_5HexNAc_6dHex_3$ + Na	GalNAc(α1-3)Gal(β1-4) Fuc(α1-2) Fuc(α1-2)Gal(β1-4) | GlcNAc(β1-6) GlcNAc(β1-2)Man(α1-6) Fuc(α1-6) Man(β1-4)GlcNAc(β1-4)GlcNAc GlcNAc(β1-2)Man(α1-3)

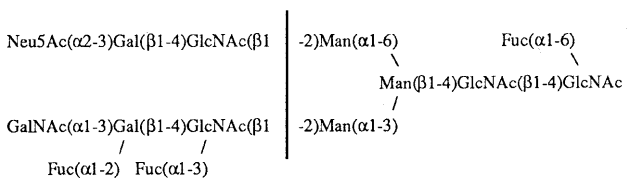
Qd1.7A with the blood group H determinant, instead of the Le^x epitope, is excluded as a major component of fraction **Qd1.7**.

The minor disodiated molecular ion in the ESI-MS spectrum at *m/z* 1409 [Neu5AcHex$_5$HexNAc$_3$dHex$_2$HexNAc-ol + 2 Na] in **Qd1.7** was also further fragmented. The presence of 4-substituted GlcNAc-ol in the methylation analysis indicates the absence of core fucosylation. ESI-MS-MS peaks at *m/z* 1021, 646, 660, and 316 [HexNAc-ol + Na] confirm the presence of structure **Qd1.7C** as an analogue of **Qd1.7A** without core fucosylation.

The MALDI-TOF mass spectrum of fraction **Qd1.8** (after reduction and permethylation) displays a sodiated molecular ion at *m/z* 3216 [Neu5AcHex$_5$HexNAc$_4$dHex$_3$HexNAc-ol + Na]. Methylation analysis yields 2,3- and 3-substituted Gal, terminal Fuc, 3,4-di- and 4-substituted GlcNAc, and terminal GalNAc residues, confirming structure **Qd1.8**. Fragmentation of the disodiated ESI-MS molecular ion at *m/z* 1620 [Neu5AcHex$_5$HexNAc$_4$dHex$_3$HexNAc-ol + 2 Na] results in fragments at *m/z* 847 and 472 representing an antenna composed of Neu5AcHexHexNAc. Fragments at *m/z* 1079 [HexHexNAc$_2$dHex$_2$ + Na] and 678 [HexHexNAcdHex + Na] correspond to an antenna ending in the ALe^y determinant. The presence of a fragment at *m/z* 490 and the absence of a fragment at *m/z* 316 show that glycan **Qd1.8** is core-fucosylated.



Qd1.7C



Qd1.8

Table VI. 600-MHz ^{1}H-NMR chemical shifts of structural-reporter-group protons of the constituent monosaccharides of mono- and disialylated diantennary complex-type N-glycans liberated from sEGFR. Chemical shifts are given relative to internal acetone (δ 2.225) in ^2H_2O at 300 K and p^{2}H 7. Compounds are represented by symbolic shorthand notation: solid circles, GlcNAc; solid diamonds, Man; open squares, Fuc; solid squares, Gal; open diamonds, GalNAc; open triangles, Neu5Ac(α2–3). α and β stand for the anomeric configuration of GlcNAc-**1**.

		Qd1.3	**Qd1.4**	**Qd1.5A**	**Qd1.5B**	**Qd1.5C**	**Qd1.5D**[b]	**Qd1.6A**	**Qd1.6B**	**Qd1.6C**	**Qd1.6D**	**Qd2A**	**Qd2B**	**Qd2C**	**Qd2D**
Relative amount[a]:		4%	27%	5%	4%	3%	3%	14%	11%	9%	8%	60%	25%	15%	
Reporter group	Residue	Chemical shift (ppm)													
H-1	GlcNAc-1α	5.184	5.180	5.179	5.179	5.179	5.179	5.179	5.179	5.179	5.179	5.182	5.182	5.182	5.182
	GlcNAc-1β	nd[c]	4.700	nd	nd	nd	nd	4.701	4.701	4.701	4.701	nd	nd	nd	nd
	GlcNAc-2[α]	4.61[d]	4.665	4.615	4.615	4.615	4.615	4.664	4.664	4.664	4.664	4.664	4.612	4.664	4.664
	GlcNAc-2[β]							4.670	4.670	4.670	4.670				
	Man-**4**	5.116	5.119	5.117	5.117	5.105	5.117	5.117	5.117	5.107	5.117	5.115	5.115	5.115	5.115
	Man-**4'**	4.926	4.925	4.909	4.909	4.924	4.924	4.909	4.909	4.924	4.924	4.922	4.922	4.922	4.91[d]
	GlcNAc-**5**	4.576	4.579	4.576	4.576	4.586	4.586	4.578	4.578	4.582	4.582	4.573	4.573	4.573	4.573
	GlcNAc-**5'**	4.581	4.583	4.586	4.586	4.576	4.576	4.582	4.582	4.578	4.578	4.573	4.573	4.573	4.573
	Gal-**6**	4.543	4.543	4.543	4.543	4.441	4.598	4.543	4.543	4.445	4.600	4.546	4.546	4.510	4.546
	Gal-**6'**	4.473	4.475	4.449	4.598	4.551	4.551	4.450	4.600	4.549	4.549	4.549	4.549	4.549	4.519
	(α1-6)Fuc[α]	–	4.889	–	–	–	–	4.888	4.888	4.888	4.888	4.893	–	4.893	4.893
	(α1-6)Fuc[β]	–	4.896	–	–	–	–	4.894	4.894	4.894	4.894	4.901	–	4.901	4.901
	(α1-3)Fuc	–	–	5.129	–	5.129	–	5.125	–	5.125	–	–	–	5.108	5.108
	(α1-2)Fuc	–	–	–	5.346	–	5.346	–	5.343	–	5.343	–	–	–	–
	GalNAc	–	–	–	5.179	–	5.179	–	5.179	–	5.179	–	–	–	–
H-2	Man-**3**	4.247	4.246	4.247	4.247	4.247	4.247	4.246	4.246	4.246	4.246	4.248	4.248	4.248	4.248
	Man-**4**	4.189	4.191	4.188	4.188	4.188	4.188	4.190	4.190	4.190	4.190	4.188	4.188	4.188	4.188
	Man-**4'**	4.106	4.107	4.083	4.083	4.106	4.106	4.081[e]	4.081[e]	4.109[e]	4.109[e]	4.110[e]	4.110[e]	4.110[e]	4.09[d]
	GalNAc	–	–	–	4.235	–	4.235	–	4.236[e]	–	4.236[e]	–	–	–	–
H-3	Gal-**6**	4.115	4.114	4.11[d]	4.11[d]	nd	nd	4.11[d]	4.11[d]	3.660[e]	3.981[e]	4.117	4.117	4.089[e]	4.117
	Gal-**6'**	nd	nd	nd	nd	4.11[d]	4.11[d]	3.660[e]	3.981	4.11[d]	4.11[d]	4.117	4.117	4.117	4.089[e]
H-3a	Neu5Ac	1.795	1.795	1.795	1.795	1.795	1.795	1.795	1.795	1.795	1.795	1.798[f]	1.798[f]	1.798[f]	1.798[f]
H-3e	Neu5Ac	2.756	2.755	2.756	2.756	2.756	2.756	2.755	2.755	2.755	2.755	2.758[f]	2.758[f]	2.758[f]	2.758[f]
H-4	Gal-**6**	nd	nd	nd	nd	nd	4.216	3.961[e]	3.961[e]	3.906[e]	4.222[e]	nd	nd	nd	nd
	Gal-**6'**	nd	nd	nd	4.216	nd	nd	3.906[e]	4.222[e]	3.961[e]	3.961[e]	nd	nd	nd	nd
H-5	(α1-6)Fuc[α]	–	4.102	–	–	–	–	4.098[e]	4.098[e]	4.098[e]	4.098[e]	4.098	–	4.098	4.098
	(α1-6)Fuc[β]	–	4.127	–	–	–	–	4.126[e]	4.126[e]	4.126[e]	4.126[e]	4.132	–	4.132	4.132
	(α1-3)Fuc	–	–	4.816	–	4.816	–	4.827[e]	–	4.827[e]	–	–	–	4.816[e]	4.816[e]
	(α1-2)Fuc	–	–	–	4.311	–	4.311	–	4.309	–	4.309	–	–	–	–
CH_3	(α1-6)Fuc[α]	–	1.209	–	–	–	–	1.212	1.212	1.212	1.212	1.211	–	1.211	1.211
	(α1-6)Fuc[β]	–	1.220	–	–	–	–	1.222	1.222	1.222	1.222	1.223	–	1.223	1.223
	(α1-3)Fuc	–	–	1.180	–	1.180	–	1.176	–	1.176	–	–	–	1.166	1.166
	(α1-2)Fuc	–	–	–	1.254	–	1.254	–	1.249[e]	–	1.249[e]	–	–	–	–
NAc	GlcNAc-**1**	2.037	2.040	2.039	2.039	2.039	2.039	2.040	2.040	2.040	2.040	2.040	2.040	2.040	2.040
	GlcNAc-**2**	2.079	2.094	2.081	2.081	2.081	2.081	2.095	2.095	2.095	2.095	2.096	2.082	2.096	2.096
	GlcNAc-**5**	2.046	2.048	2.048	2.048	2.039	2.055	2.049	2.049	2.040	2.056	2.047	2.047	2.040	2.047
	GlcNAc-**5'**	2.046	2.048[#]	2.039	2.055	2.048	2.048	2.040	2.056	2.044	2.044	2.042	2.042	2.042	2.034
	Neu5Ac	2.029	2.031	2.030	2.030	2.030	2.030	2.032	2.032	2.032	2.032	2.031[g]	2.031[g]	2.031[g]	2.031[g]
	GalNAc	–	–	–	2.039	–	2.039	–	2.040	–	2.040	–	–	–	–

[a]The amount is given as molar percentages of the total **Qd1** or **Qd2** fraction. The percentages are based on peak areas of HPLC peaks on LiChrosorb-NH_2 and, for mixtures, on NMR intensities.
[b]See structure **Qd1.5D** in the text; bloodgroup A: [see structure below]; sialyl-Le^x: [see structure below]
[c]nd, not determined
[d]Value given in two decimals due to spectral overlap
[e]Value obtained from 2D-TOCSY experiment
[f]Signal stemming from two protons
[g]Signal stemming from two NAc groups
[#]In Hokke *et al.* (1995), the GlcNAc-**5′** NAc signals for *N1.1A* and *N1.1B* are interchanged

bloodgroup A:

GalNAc(α1-3)Gal(β1-
/
Fuc(α1-2)

sialyl-Le^x:

Neu5Ac(α2-3)Gal(β1-4)GlcNAc
/
Fuc(α1-3)

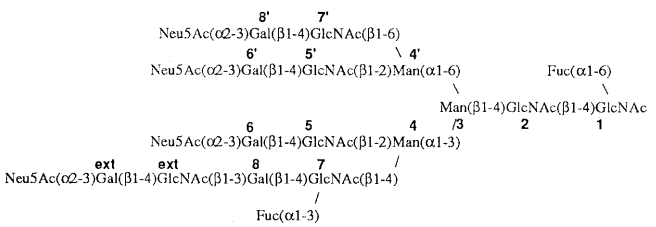
It should be noted that for fractions **Qd1.7** and **Qd1.8** the MS data indicating a fucosylated Gal-GlcNAc sequence have been interpreted as Gal(β1–4)[Fuc(α1–3)]GlcNAc and not as Gal(β1–3)[Fuc(α1–4)]GlcNAc. This choice is based on the fact that all structures analyzed by NMR only contain Gal(β1–4)GlcNAc sequences.

Disialylated oligosaccharides

The ^{1}H-NMR data of fraction **Qd2** (Figure 4) are presented in Table VI, and the numbering of the residues is depicted in structure **Qd1.5D** above. The structural-reporter-group signals of the major component of fraction **Qd2**, glycan **Qd2A** (60%), indicate a core-fucosylated (α2–3)-disialylated diantennary structure (cf. glycan *N2.4*; Hokke *et al.*, 1995), and about 25% of fraction **Qd2** consists of **Qd2B**, the non-fucosylated analogue of glycan **Qd2A** (cf. glycan *N2B*; Damm *et al.*, 1987). The remaining 15% of the structures have the sialyl-Le^x determinant (Fuc H-1, δ 5.108, H-5, δ 4.816, and CH_3, δ 1.166) linked to one of the antennae. The Gal-**6** H-1 and Gal-**6′** H-1 signals at 4.510 and 4.519 p.p.m., respectively, show that the sialyl-Le^x epitope is linked to the (α1–3) branch in glycan **Qd2C** and to the (α1–6) branch in glycan **Qd2D**. Because the intensity of the Gal-**6** H-1 signal is higher than that of Gal-**6′** H-1, a structure with sialyl-Le^x epitopes linked to both branches is excluded as a prominent minor fraction. In view of the fact that mixtures were analyzed, the occurrence of relatively small amounts of non-core-fucosylated analogues of **Qd2C** and **Qd2D** cannot be excluded.

Tetrasialylated oligosaccharides

^{1}H-NMR data of the HPAEC fractions **Q4.4**, **Q4.5**, **Q4.7**, and **Q4.8** (Figure 2) are presented in Table VII, and a full structure and its residue numbering is depicted in structure **Q4.4A** below.



Q4.4A

NMR analysis shows that all fractions contain tetraantennary oligosaccharides, being evident from the typical set of values of the Man-**4** and Man-**4′** H-1 signals and the Man-**3**, Man-**4**, and Man-**4′** H-2 signals (Vliegenthart *et al.*, 1983; Hokke *et al.*, 1995). All glycans in the **Q4** fractions have an (α1–6)-linked Fuc residue at GlcNAc-**1**, as follows from the typical set of Fuc H-1, H-5, and CH_3 δ values, and the GlcNAc-**2** NAc signal (Hokke *et al.*, 1995). Neu5Ac residues are (α2–3)-linked to Gal residues at all nonreducing termini, as indicated by the combined set of δ values of the Neu5Ac H-3a/3e, Neu5Ac NAc, and Gal H-3 signals (Vliegenthart *et al.*, 1983).

Oligosaccharide **Q4.8**, the major constituent in **Q4**, is a tetrasialylated tetraantennary glycan with all four antennae composed of Neu5Ac(α2–3)Gal(β1–4)GlcNAc(β1- (cf. glycan *N4.4*; Hokke *et al.*, 1995). The NMR spectra of fractions **Q4.8** and **Q4.9** are alike indicating that the two fractions are derived from the same glycan. Considering the alkaline conditions (pH 13) during the HPAEC separation, epimerization at C-2 of the reducing GlcNAc into ManNAc is obvious (Toida *et al.*, 1996). This structural change is confirmed by monosaccharide analysis, and by incubation of comparable compounds in 0.1 M NaOH followed by reinjection on an HPAEC column (data not shown).

Fraction **Q4.7** contains a mixture of isomeric structures with an extra LacNAc unit in one of the antennae of structure **Q4.8**. First of all, this is shown by the reporter signals of the Gal residue extended with the LacNAc unit: Gal-**8′** H-1 (δ 4.468) and H-4 (δ 4.161) in **Q4.7A**. The Gal-**8′** H-1 signal in case of the absence of an extra LacNAc unit (δ 4.557) differs from those of Gal-**6**, -**6′**, and –**8**, and can therefore be used for the assignment of the location of the extra LacNAc element in structure **Q4.7A** (Hokke *et al.*, 1991, 1995). The other isomers, with a LacNAc unit linked to Gal-**6**, –**6′**, or –**8**, are summarized in structure **Q4.7B** (Table VII). The ^{1}H-NMR data fit those of structure *C* in Hokke *et al.* (1991), in which the extra LacNAc element is linked to Gal-**6′** (H-1, δ 4.542), but since no data are available on the other two isomers, their absence or presence cannot be confirmed. As has been reported before (Hermentin *et al.*, 1992), the extra LacNAc unit in a sialylated oligosaccharide leads to a decrease of its retention time on HPAEC, as is the case for fraction **Q4.7** that elutes 1 min earlier than **Q4.8** (Figure 2).

The second major fraction, **Q4.5**, differs from **Q4.8** only by the presence of an extra (α1–3)-linked Fuc residue in one of the antennae (H-1, δ 5.103; CH_3, δ 1.169) constituting the sialyl-Le^x determinant. Based on a comparison of the δ values of the NAc signals of antennary (α1–3)-fucosylated GlcNAc and nonfucosylated GlcNAc residues (e.g. **QdN.4** versus **QdN.1** and **Qd1.5A** versus **Qd1.4**), a shift of about –0.008 p.p.m. is expected. Comparing the NAc regions of **Q4.5** and **Q4.8** shows the presence of an additional signal for **Q4.5** at δ 2.065, indicating an (α1–3)-fucosylated GlcNAc-**7** (Δδ –0.009). The intensities of the GlcNAc-**7** NAc signals at δ 2.074 and δ 2.065 show that fraction **Q4.5** consists of an isomeric mixture of about 65% **Q4.5A** (extra Fuc residue linked to GlcNAc-**7**) and 35% **Q4.5B**, having the sialyl-Le^x determinant at one of the other antennae. Shifts in the GlcNAc-**5**, -**5′**, and -**7′** NAc signals due to an extra Fuc residue cannot be distinguished because of overlap with other NAc signals. Further evidence for glycan **Q4.5A** is shown by the Gal-**8** H-1 signal at δ 4.517 (Maaheimo *et al.*, 1995). Because of the signal of Gal-**8′** H-1 at δ 4.530, **Q4.5B** contains at least the isomer with the Fuc (α1–3)-linked to GlcNAc-**7′**. The addition of a Fuc residue to an oligosaccharide also results in the decrease of HPAEC retention time, as is seen for **Q4.5**, having an elution time of 3.8 min lower than **Q4.8**, in Figure 2 (Hardy and Townsend, 1988; Basa and Spellman, 1990).

After permethylation, fractions **Q4.2** and **Q4.4** have been analyzed by FAB-MS. The spectra do not contain molecular ions, but typical fragments of the antennae can be identified. In the spectrum of fraction **Q4.4** peaks are present at *m/z* 825 [Neu5AcHexHexNAc + H] and *m/z* 1448 [$Neu5AcHex_2$ $HexNAc_2dHex$ + H] showing that one of the antennae consists of a sialylated dimeric LacNAc element with one Fuc residue attached. The absence of a fragment at *m/z* 999 [Neu5AcHexHexNAcdHex + H] indicates that the Fuc residue is linked to the inner of the two LacNAc units. ^{1}H-NMR analysis of fraction **Q4.4** shows the presence of an (α1–3)-linked Fuc residue and an extra LacNAc unit (Table VII). FAB-MS shows that the two elements are located on one antenna, and NMR data provide the positions of the Fuc residue and the LacNAc unit. In glycan **Q4.4A**, the position of

Table VII. ^{1}H-NMR chemical shifts of structural-reporter-group protons of the constituent monosaccharides of tetrasialylated tetraantennary complex-type N-glycans liberated from sEGFR. Chemical shifts are given relative to internal acetone (δ 2.225) in ^2H_2O at 300 K and p^{2}H 7. Compounds are represented by symbolic shorthand notation: solid circles, GlcNAc; solid diamonds, Man; open squares, Fuc; solid squares, Gal; open triangles, Neu5Ac(α2–3). α and β stand for the anomeric configuration of GlcNAc-**1**.

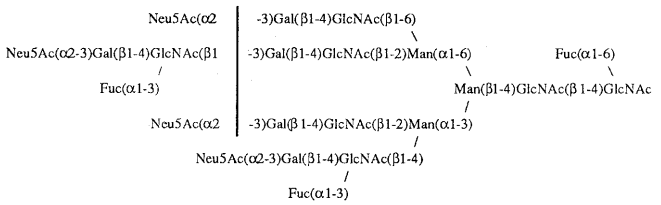
		Q4.4A[b]	**Q4.4B**	**Q4.5A**	**Q4.5B**	**Q4.7A**	**Q4.7B**	**Q4.8**
Relative amount[a]:		7%	<4%	12%	8%	4%	9%	39%
Reporter group	Residue	Chemical shift (ppm)						
H-1	GlcNAc-1α	5.182	5.182	5.182	5.182	5.182	5.182	5.183
	GlcNAc-1β	nd[c]	nd	nd	nd	nd	nd	4.689
	GlcNAc-2$^{\alpha}$	4.66[e]	4.66[e]	4.663	4.663	4.659	4.659	4.66[e]
	GlcNAc-2$^{\beta}$					4.665	4.665	
	Man-**4**	5.125	5.125	5.130	5.130	5.129	5.129	5.131
	Man-**4'**	4.856	4.856	4.857	4.857	4.856	4.856	4.857
	GlcNAc-**5**	4.558	4.558	4.559	4.559	4.56[e]	4.56[e]	4.56[e]
	GlcNAc-**5'**	4.590	4.590	4.593	4.593	4.59[e]	4.59/4.60[e]	4.595
	Gal-**6**	4.543	4.543	4.544	4.544/4.517	4.542	4.542/4.452	4.543
	Gal-**6'**	4.543	4.543	4.544	4.544/4.517	4.542	4.542/4.452	4.543
	GlcNAc-**7**	4.543	4.543	4.544	4.544	4.542	4.542	4.543
	GlcNAc-**7'**	4.543	4.543	4.544	4.544	4.542	4.542	4.543
	Gal-**8**	4.438	4.543	4.517	4.544	4.542	4.542/4.452	4.543
	Gal-**8'**	4.558	4.450	4.559	4.559/4.530	4.468	4.557	4.559
	GlcNAc-**ext**	nd	nd	–	–	4.703	4.703	–
	Gal-**ext**	4.558	4.558	–	–	4.557	4.557	–
	(α1-6)Fuc$^{\alpha}$	4.901	4.901	4.901	4.901	4.903	4.903	4.902
	(α1-6)Fuc$^{\beta}$	4.908	4.908	4.908	4.908			4.909
	(α1-3)Fuc	5.104	5.104	5.103	5.103	–	–	–
H-2	Man-**3**	4.203	4.203	4.203	4.203	4.206	4.206	4.205
	Man-**4**	4.213	4.213	4.217	4.217	4.221	4.221	4.22[e]
	Man-**4'**	4.089	4.089	4.089	4.089	4.091	nd	4.091
H-3	Man-**4**	4.050	4.050	4.053	4.053	nd	nd	4.055
	Gal-**6**	4.116	4.116	4.117	4.117/4.089	4.117	4.117/nd	4.118
	Gal-**6'**	4.116	4.116	4.117	4.117/4.089	4.117	4.117/nd	4.118
	Gal-**8**	nd	4.116	4.089	4.117	4.117	4.117/nd	4.118
	Gal-**8'**	4.116	nd	4.117	4.117/4.089	nd	4.117	4.118
	Gal-**ext**	4.116	4.116	–	–	4.117	4.117	–
H-3a	Neu5Ac	1.802[d]	1.802[d]	1.803[d]	1.803[d]	1.802[d]	1.802[d]	1.804[d]
H-3e	Neu5Ac	2.758[d]	2.758[d]	2.757[d]	2.757[d]	2.756[d]	2.756[d]	2.756[d]
H-4	Gal-**6/6'/8**	4.159	nd	nd	nd	nd	4.161	nd
	Gal-**8'**	nd	4.159	nd	nd	4.161	nd	nd
H-5	(α1-6)Fuc$^{\alpha}$	4.103	4.103	4.104	4.104	4.105	4.105	4.10[e]
	(α1-6)Fuc$^{\beta}$	4.128	4.128	4.129	4.129	4.129	4.129	4.13[e]
	(α1-3)Fuc	nd	nd	nd	nd	–	–	–
CH_3	(α1-6)Fuc$^{\alpha}$	1.211	1.211	1.211	1.211	1.210	1.210	1.211
	(α1-6)Fuc$^{\beta}$	1.222	1.222	1.222	1.222	1.222	1.222	1.223
	(α1-3)Fuc	nd	nd	1.169	1.169	–	–	–
NAc	GlcNAc-**1**	2.037	2.037	2.038	2.038	2.037	2.037	2.038
	GlcNAc-**2**	2.092	2.092	2.093	2.093	2.092	2.092	2.094
	GlcNAc-**5**	2.048	2.048	2.048	2.048/2.038	2.047	2.047	2.048
	GlcNAc-**5'**	2.037	2.037	2.038	2.038/2.031	2.037	2.037	2.038
	GlcNAc-**7**	2.066	2.075	2.065	2.074	2.075	2.075	2.075
	GlcNAc-**7'**	2.037	2.031	2.038	2.038/2.031	2.037	2.037	2.038
	GlcNAc-**ext**	2.037	2.037	–	–	2.037	2.037	–
	Neu5Ac	2.031[f]	2.031[f]	2.031[f]	2.031[f]	2.031[f]	2.031[f]	2.031[f]

[a]The amounts are given as molar percentages of the total **Q4** fraction. The results are based on peak areas of HPAEC peaks on CarboPac PA-1.
[b]See structure **Q4.4A** in the text
[c]nd, not determined
[d]Signal stemming from four protons
[e]Values given in two decimals due to spectral overlap
[f]Signal stemming from four NAc groups

Fuc at GlcNAc-**7** is determined by the characteristic shift of the GlcNAc-**7** NAc signal (δ 2.066, see **Q4.5A**), and by the Gal-**8** H-1 signal due to extension with a LacNAc element (δ 4.438). On the other hand, in glycan **Q4.4B** the position of the extra LacNAc unit at Gal-**8′** is characterized by the Gal-**8′** H-1 signal (δ 4.450), which differs from the value for Gal-**8′** in **Q4.7A** (δ 4.468) because of the Fuc residue linked to GlcNAc-**7′**.

The FAB-MS spectrum of fraction **Q4.2** contains fragment ions at *m/z* 825, 999, and 1448, which demonstrates that one antenna contains a dimeric LacNAc element and a Fuc residue

linked at one of the two LacNAc units. The elution time on HPAEC (Figure 2), 6.9 min lower than that of fraction **Q4.8**, implies that fraction **Q4.2** includes a glycan with two (α1–3)-linked Fuc residues and one extra LacNAc unit (Basa and Spellman, 1990; Hermentin *et al.*, 1992). The ^{1}H-NMR data show that the ratio between the CH_3 signals of the (α1–6)-linked Fuc at GlcNAc-**1** and the (α1–3)-linked Fuc is 1.0:1.5, indicating that, next to components with one (α1–3)-linked Fuc, structures with two (α1–3)-Fuc residues are present. Furthermore, 80% of the GlcNAc-**7** NAc signal in fraction **Q4.2** resonates at δ 2.064, which indicates that in 80% of the glycans a Fuc residue is (α1–3)-linked to GlcNAc-**7**. Combining these data with those from FAB-MS and the HPAEC profile leads to a major component (40–50% of fraction **Q4.2**) with an (α1–3)-fucosylated GlcNAc-**7** and an (α1–3)-fucosylated dimeric LacNAc element (**Q4.2A**). Fucosylation at the inner of the two LacNAc units cannot be excluded. The position of the extended antenna is not specified, as is indicated by the vertical bar.



Q4.2A

Discussion

In the present study, the oligosaccharides linked to sEGFR were characterized by a combination of ^{1}H-NMR spectroscopy and mass spectrometry. For the first time, such a comprehensive glycan analysis has been performed on a human non-recombinant growth factor receptor. The investigation resulted in the identification of 55 structures, many of them bearing blood group determinants (Le^x, Le^y, ALe^y, sialyl-Le^x, and blood group H and A antigens). The combination of that many determinants on sEGFR has only been rarely reported for other glycoproteins, e.g., the carcinoembryonic antigen (Kobata *et al.*, 1989). The immense heterogeneity of glycosylation, comparable to that of human Tamm-Horsfall glycoprotein (Hård *et al.*, 1992; Van Rooijen *et al.*, 1998), is presented by highlighting the different features of the oligosaccharides to cover the full range of structures. The diversity of glycans was already suggested by isoelectric focussing gel electrophoresis of sEGFR, produced in A431 cells, giving rise to 24 distinct bands ranging in pI from 3.7 to 7.3 (Hurwitz *et al.*, 1991). The heterogeneity of the glycosylation is obviously a major contributor to the difficulties in growing sEGFR crystals (Weber *et al.*, 1994).

In this study, oligomannose-type structures (5) account for about 17% of the glycans and range from $Man_5GlcNAc_2$ to $Man_8GlcNAc_2$, but being mainly $Man_6GlcNAc_2$ and $Man_7GlcNAc_2$. This is different from the 33% oligomannose-type glycans on membrane-bound EGFR (Cummings *et al.*, 1985) and the 31% on sEGFR (Mangelsdorf Soderquist *et al.*, 1988). The complex-type glycans on membrane-bound EGFR were described by Cummings *et al.* (1985) as predominantly asialo tri- and/or tri′- and tetraantennary. If sialic acid was present, its occurrence was limited to one or two residues per glycan chain. The predominance of tri- and/or tri′- and tetraantennary glycans is confirmed for sEGFR, but 59% of the total glycan pool (**Q1**–**Q4** in Table I) is sialylated up to four sialic acids per oligosaccharide chain. Man-6-phosphate (Todderud and Carpenter, 1988) and LacNAc repeats (Cummings *et al.*, 1985) are other structural elements of the membrane-bound form. Only the LacNAc repeats are found in sEGFR and account for about a third of the **Q4** fraction representing 2% of the total carbohydrate moiety. The absence of Man-6-phosphate in sEGFR shows that phosphorylation of oligomannose-type glycans is specific for membrane-bound EGFR (Todderud and Carpenter, 1988).

The sialyl-Le^x epitope is present in all sialylated glycan pools of sEGFR (**Qd1**, **Qd2** and **Q4**), but has not been reported before. On the other hand, Le^a and Le^b determinants reported by Childs *et al.* (1984), for the membrane-bound form, have not been found. The diversity in expressed carbohydrate antigens finds its origin in the A431 cells. sEGFR expressed in insect cell line Sf9 does not show this variety in epitopes and only shows two bands on isoelectric focusing gel electrophoresis (Hurwitz *et al.*, 1991). The high structural diversity of the N-glycans linked to sEGFR reflects the high number of expressed and active glycosyltransferases in A431 cells. Since this cell line is derived from carcinoma cells, it is interesting to compare the glycosylation of sEGFR (and EGFR) with that of EGFR from normal cells. Especially, since EGFR lacking the blood group A antigen has a different receptor functioning, including increase in receptor turnover and the number of high-affinity receptors (Defize *et al.*, 1988; Engelmann and Schumacher, 1993). However, detailed information on the glycosylation of EGFR in normal cells is not yet available.

It was concluded earlier that the glycosylations of EGFR and sEGFR are very similar (Mangelsdorf Soderquist *et al.*, 1988). Although the glycan structures clearly show many common features, there are some marked differences. The transport routes from the ER to the cell exterior and the kinetics of the transport are similar for both forms (Mangelsdorf Soderquist *et al.*, 1988). Therefore, EGFR and sEGFR are exposed to the same glycosylation machinery. Hence, differences in glycosylation can be traced back to the differences between secreted and membrane-bound forms.

Since more than 15 years, attempts are being made to elucidate the 3D structure of the extracellular domain of the receptor and its complex with EGF. Electron microscopic images (Lax *et al.*, 1991) and small angle X-ray scattering (Lemmon *et al.*, 1997) support a 3D model of sEGFR as a four-lobed, doughnut-like overall shape. However, X-ray crystallography is a necessity to determine in detail the 3D structure, with and without the ligand. Progress has been made on the crystallization (Günther *et al.*, 1990; Degenhardt *et al.*, 1998), but the structure remains to be resolved. The immense heterogeneity of the glycosylation of sEGFR, as presented here, clearly points towards the application of glycan processing inhibitors for obtaining less heterogeneous sEGFR molecules.

Materials and methods

Purification of sEGFR

The purification of sEGFR was performed as described previously (Günther *et al.*, 1990). Confluent cultures of A431 variant cells (Institut für Physiologische Chemie, UKE Hamburg) were grown in serum-free Dulbecco's modified Eagle's/F-12 medium supplemented with 0.5 mg/l human transferrin, 50 nM hydrocortisone, 0.025 mg/l sodium selenite, and 0.1 g/l bovine serum albumin. Prior to immunoaffinity chromatography, sEGFR was freed from cellular debris, stabilized with 1 mM EDTA containing 100 U/ml Trasylol (Bayer), and concentrated by ultrafiltration. sEGFR was eluted from the immune absorbent with 0.1 M acetic acid containing 0.05 M NaCl, and immediately neutralized with Na_3PO_4. $(NH_4)_2SO_4$ precipitation served to concentrate the protein to 8–12 mg/ml.

In order to remove buffer salts, glycoprotein samples were passed over a Bio-Gel P-2 column (1.5 × 45 cm, Bio-Rad) eluted with 5 mM NH_4HCO_3. The void volume was collected and lyophilized.

Monosaccharide analysis

Desalted sEGFR samples (400 µg) were subjected to methanolysis (1.0 M methanolic HCl, 24 h, 85°C) followed by re-*N*-acetylation and trimethylsilylation. The trimethylsilylated (methyl ester) methyl glycosides were analyzed by GLC on a capillary DB-1 fused silica column (30 m × 0.32 mm, J & W Scientific). The value of GlcNAc has been corrected for the relatively stable GlcNAc-Asn linkage (Kamerling and Vliegenthart, 1989).

Release of the carbohydrate chains

The N-linked carbohydrate chains were enzymatically released from the protein by recombinant peptide-N^4-(*N*-acetyl-β-glucosaminyl)asparagine amidase F (PNGase F, EC 3.5.1.52; Boehringer Mannheim), according to a modified version of a previously described protocol (Damm *et al.*, 1989b). sEGFR, in two batches of 30 mg, was dissolved in 10 ml 50 mM Tris, adjusted to pH 7.4 with HCl, containing 50 mM EDTA and 1% (v/v) β-mercaptoethanol. After adding SDS (34 mg), the mixture was heated for 1 h at 40°C followed by cooling to room temperature. Then, the detergent decanoyl-*N*-methylglucamide (68 mg; MEGA-10, Boehringer Mannheim) was added and the solution was incubated with 6 U PNGase F for 24 h at room temperature in an end-over-end mixer. With a second portion of 6 U PNGase F the incubation was continued for another 24 h. The incubation was monitored with SDS–PAGE, on a 10% slab gel, and Coomassie brilliant blue staining.

The incubation mixture was fractionated on a Superdex 75pg column (60 × 2.6 cm, Pharmacia) using 150 mM NH_4HCO_3 (pH 7 with HCl) at a flow rate of 4 ml/min (Pharmacia FPLC system). The elution was monitored by conductivity and UV detection (214 nm). Fractions were collected and aliquots stained for carbohydrate with orcinol/H_2SO_4. Carbohydrate-positive fractions were pooled and lyophilized twice, then desalted on three connected HiTrap Desalting columns (Pharmacia) at a flow rate of 3 ml/min, using 5 mM NH_4HCO_3 as eluent and subsequent lyophilization. Oligosaccharides were detected as a void-volume peak by monitoring the effluent at 206 nm.

FPLC fractionation

The Superdex carbohydrate-positive fraction derived from the first sEGFR batch was fractionated according to charge on a 1 ml Resource Q anion-exchange column (Pharmacia) at a flow rate of 4.0 ml/min (Pharmacia FPLC system). 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, a gradient of 50–250 mM NaCl in 8 ml H_2O, and, finally, by a steeper gradient of 250–500 mM NaCl in 4 ml H_2O. The eluent was monitored at 214 nm. After lyophilization, the five fractions obtained were desalted on a Bio-Gel P-2 column (45 × 1.6 cm), and lyophilized again. This gel permeation step was also used to separate the neutral fraction roughly into two fractions (complex- and oligomannose-type glycans).

Concanavalin A fractionation

The Superdex carbohydrate-positive fraction derived from the second sEGFR batch was subjected to Concanavalin A Sepharose (Pharmacia) lectin chromatography. For this purpose, a 3 ml column was used, equilibrated with TBS binding buffer (pH 8.0) containing 0.15 M NaCl, 10 mM Tris, 1 mM $CaCl_2$, and 1 mM $MgCl_2$. The column was first eluted with 20 ml binding buffer at a flow rate of 14 ml/h, yielding tri/tri′- and tetraantennary complex-type glycans. Elution with 20 ml binding buffer containing 10 mM methyl α-D-glucopyranoside was used, at the same flow rate, to elute the diantennary complex-type glycans. Finally, 20 ml binding buffer containing 100 mM methyl α-D-mannopyranoside, at 60°C, eluted the oligomannose-type glycans. The elution was monitored at 206 nm. The three resulting fractions were separated from the salts and monosaccharides using four connected HiTrap Desalting columns (5 mM NH_4HCO_3, 3 ml/min) and subsequent lyophilization, followed by fractionation on Resource Q, as described above for the first sEGFR batch.

HPLC fractionation

The oligomannose-type glycan fraction was subfractionated using a Kratos SF 400 HPLC system (ABI Analytical, Kratos Division) equipped with a 100-µm LiChrospher-NH_2 column (25 × 0.46 cm, Merck). The column was eluted with a linear concentration gradient of 30 to 38% (v/v) 10 mM phosphate buffer (K_2HPO_4/KH_2PO_4, pH 6.5) in acetonitrile for 40 min at a flow rate of 2 ml/min.

The complex-type glycan fractions (neutral and charged) were subfractionated on the same HPLC system equipped with a 10 µm LiChrosorb-NH_2 column (25 × 0.46 cm, Merck). The column was eluted with an acetonitrile/phosphate buffer (10 mM potassium phosphate, pH 6.5) gradient program at a flow rate of 2 ml/min. For details concerning the solvent systems, see the legends to Figures 5, 6, and 8.

The eluents were monitored at 206 nm. The HPLC fractions were desalted on three connected HiTrap Desalting columns (flow rate 3 ml/min) using 5 mM NH_4HCO_3 as eluent and subsequent lyophilization.

HPAEC fractionation

Fractionation of highly charged Resource Q fractions was carried out by high-pH anion-exchange chromatography with pulsed amperometric detection using a Dionex LC system consisting of a Dionex Bio LC quaternary gradient module, a

PAD 2 detector, and a CarboPac PA-1 column (25 × 0.9 cm, Dionex) (Damm *et al.*, 1989a). Elutions were carried out at a flow rate of 4 ml/min. Detection was performed with a gold electrode and triple-pulse amperometry, comprising the following pulse potentials and durations at 300 mA: E_1 = 0.05 V, 480 ms; E_2 = 0.60 V, 120 ms; E_3 = –0.60 V, 60 ms. For further details concerning the gradients of the 0.1 M NaOH/0.5 M NaOAc solvent system, see caption to Figure 2. Fractions were immediately neutralized by addition of 1 M HCl and lyophilized. Desalting was carried out on three connected HiTrap Desalting columns (flow rate 3 ml/min) using 5 mM NH_4HCO_3 as eluent and subsequent lyophilization.

Permethylation and methylation analysis

Desalted and dried HPAEC samples were dissolved in 500 µl dry dimethyl sulfoxide, and freshly powdered NaOH (~10 mg) was added. The solution was kept for 10 min at room temperature under an inert atmosphere, then 100 µl CH_3I was added (Jay, 1996). Extra portions of 100 µl CH_3I were added after 10 and 20 min. Then, the reaction was quenched by the addition of 1 ml freshly prepared 4 mM sodium thiosulfate. The permethylated oligosaccharides or oligosaccharide-alditols were isolated by extraction with chloroform (3 × 1 ml). The organic phase was washed with water (3 × 0.5 ml) and concentrated.

For methylation analysis, oligosaccharides were permethylated, purified on Sepharose LH-20 (run with ethyl acetate), hydrolyzed, reduced, and peracetylated as described previously (Geyer *et al.*, 1983). Separation and identification of partially methylated alditol acetates was performed on a Finnigan gas chromatograph (Finnigan MAT Corp.), equipped with a 30 m DB-5 capillary column, connected to a Finnigan GCQ ion-trap mass spectrometer (Structure Research, GBF Braunschweig).

^{1}H-NMR spectroscopy

The oligosaccharide samples were exchanged twice in ^2H_2O (99.9 atom% ^{2}H, Cambridge Isotopes Ltd.) with intermediate lyophilization, then dissolved in ^2H_2O (99.96 atom% ^{2}H, Isotec Inc.). 1D ^{1}H-NMR spectra were recorded at 500 MHz on a Bruker AMX500 instrument, and at 600 MHz on Bruker AMX600 (Bijvoet Center, Department of NMR Spectroscopy, Utrecht University) or Bruker Avance600 (NSR Center, SON NMR facility, Nijmegen University, the Netherlands) spectrometers at probe temperatures of 300 K. Chemical shifts (δ) are expressed in p.p.m. by reference to internal acetone (δ 2.225) in ^2H_2O (Vliegenthart *et al.*, 1983).

2D-TOCSY spectra at 500 or 600 MHz were recorded using Bruker software with MLEV-17 mixing sequence cycles of 100 ms. Data matrices of 512 × 2048 or 256 × 1024 points were collected representing a spectral width of 4800 Hz in each dimension. The ^{2}HO^{1}H signal was suppressed by presaturation for 1 s during the relaxation delay. Phase-sensitive handling of the data was performed by the time-proportional phase increment method implemented in the Bruker software. The time domain data were zero-filled to data matrices of 1024 × 2048 or 512 × 1024 points, respectively, prior to multiplication with a squared-bell function phase shifted by π/3 (Hård *et al.*, 1992).

Mass spectrometry

For analysis by positive-ion MALDI-TOF-MS (matrix-assisted laser desorption ionization time of flight mass spectrometry), 2,5-dihydroxybenzoic acid (10 mg/ml 10% v/v ethanol in water) was used as UV-absorbing matrix. One µl of a sample was spotted onto the target, mixed with the matrix solution in a 1:2 ratio and dried at room temperature. The concentrations of the analyte mixtures were approximately 150 pmol/ml. Measurements were performed on a PerSeptive Biosystems Voyager-DE MALDI-TOF mass spectrometer with implemented delayed extraction technique using a N_2 laser (337 nm) with 3 ns pulse width. Spectra were recorded in the linear mode at an accelerating voltage of 24 kV using an extraction delay of 150 ns for enhanced resolution.

Positive-ion FAB (fast atom bombardment) mass spectra of permethylated oligosaccharides were obtained using MS1 of a JEOL JMS-SX/SX102A tandem mass spectrometer (Bijvoet Center, Department of Mass Spectrometry, Utrecht University) operated at an accelerating voltage set around 6 kV. The matrix used was thioglycerol and the bombarding gas was Xe.

For ESI-MS-MS (electrospray ionization tandem mass spectrometry), a Finnigan MAT TSQ 700 triple quadrupole mass spectrometer (Structure Research, GBF Braunschweig) equipped with a nanospray ion source (Protana) was used. The reduced and permethylated samples were dissolved in methanol saturated with NaCl (about 10 pmol/ml) and ~3 µl of solution was filled into gold-coated nanospray glas capillaries (Protana). The tip of the capillary was placed directly in front of the entrance hole of the heated transfer line of the mass spectrometer and a voltage of 800 V applied leading to flow rates of ~50 nl/min. For collision-induced dissociation experiments, parent ions were selectively transmitted by the first mass analyzer and directed into the collision cell (with argon as collision gas) with a kinetic energy set around –58 eV.

References

Basa,L.J. and Spellman,M.W. (1990) Analysis of glycoprotein-derived oligosaccharides by high-pH anion-exchange chromatography. *J. Chromatogr.*, **499**, 205–220.

Bergwerff,A.A., Stroop,C.J.M., Murray,B., Holtrof,A.-P., Pluschke,G., van Oostrum,J., Kamerling,J.P. and Vliegenthart,J.F.G. (1995) Variation in N-linked carbohydrate chains in different batches of two chimeric monoclonal IgG1 antibodies produced by different murine SP2/0 transfectoma cell subclones. *Glycoconj. J.*, **12**, 318–330.

Childs,R.A., Gregoriou,M., Scudder,P., Thorpe,S.J., Rees,A.R. and Feizi,T. (1984) Blood group-active carbohydrate chains on the receptor for epidermal growth factor of A431. *EMBO J.*, **3**, 2227–2233.

Cummings,R.D., Mangelsdorf Soderquist,A. and Carpenter,G. (1985) The oligosaccharide moieties of the epidermal growth factor receptor in A-431 cells. *J. Biol. Chem.*, **260**, 11944–11952.

Damm,J.B.L., Kamerling,J.P., van Dedem,G.W.K. and Vliegenthart,J.F.G. (1987) A general strategy for the isolation of carbohydrate chains from N,O-glycoproteins and its application to human chorionic gonadotropin. *Glycoconj. J.*, **4**, 129–144.

Damm,J.B.L., Bergwerff,A.A., Hård,K., Kamerling,J.P. and Vliegenthart,J.F.G. (1989a) Sialic acid patterns in N-linked carbohydrate chains. Structural analysis of the *N*-acetyl-/*N*-glycolyl-neuraminic-acid-containing N-linked carbohydrate chains of bovine fibrinogen. *Recl. Trav. Chim. Pays-Bas*, **108**, 351–359.

Damm,J.B.L., Voshol,H., Hård,K., Kamerling,J.P. and Vliegenthart,J.F.G. (1989b) Analysis of *N*-acetyl-4-*O*-acetylneuraminic-acid-containing N-linked carbohydrate chains released by peptide-N^4-(*N*-acetyl-β-glucosaminyl)asparagine amidase F; application to the structure determination

of the carbohydrate chains of equine fibrinogen. *Eur. J. Biochem.*, **180**, 101–110.

Defize,L.H.K., Arndt Jovin,D.J., Jovin,T.M., Boonstra,J., Meisenhelder,J., Hunter,T., de Hey,H.T. and de Laat,S.W. (1988) A431 cell variants lacking the blood group A antigen display increased high affinity epidermal growth factor-receptor number, protein-tyrosine kinase activity and receptor turnover. *J. Cell Biol.*, **107**, 939–949.

Degenhardt,M., Weber,W., Eschenburg,S., Dierks,K., Rapp,G. and Betzel,C. (1998) Crystallization and preliminary X-ray crystallographic analysis of the EGF receptor ectodomain. *Acta Crystallogr. Sect. D*, **54**, 999–1001.

Dittadi,R., Gion,M., Pagan,V., Brazzale,A., Maschio,O.D., Bargossi,A., Busetto,A. and Bruscagnin,G. (1991) Epidermal growth factor receptor in lung malignancies. Comparison between cancer and normal tissue. *Br. J. Cancer*, **64**, 741–744.

Dua,V.K., Rao,B.N.N., Wu,S.-S., Dube,V.E. and Bush,C.A. (1986) Characterization of the oligosaccharide alditols from ovarian cyst mucin glycoproteins of blood group A using high pressure liquid chromatography (HPLC) and high field ^{1}H NMR spectroscopy. *J. Biol. Chem.*, **261**, 1599–1608.

Engelmann,B. and Schumacher,U. (1993) The emerging role of ABH blood group antigens as modulators of cell membrane function. *Comp. Biochem. Physiol.*, **105A**, 197–203.

Fu,D., Chen,L. and O'Neill,R. (1994) A detailed structural characterization of ribonuclease B oligosaccharides by ^{1}H NMR spectroscopy and mass spectrometry. *Carbohydr. Res.*, **261**, 173–186.

Geyer,R., Geyer,H., Kühnhardt,S., Mink,W. and Stirm,S. (1983) Methylation analysis of complex carbohydrates in small amounts: capillary gas chromatography-mass fragmentography of methylalditol acetates obtained from N-glycosidically linked glycoprotein oligosaccharides. *Anal. Biochem.*, **133**, 197–207.

Günther,N., Betzel,C. and Weber,W. (1990) The secreted form of the epidermal growth factor receptor. *J. Biol. Chem.*, **265**, 22082–22085.

Hård,K., Mekking,A., Damm,J.B.L., Kamerling,J.P., de Boer,W., Wijnands,R.A. and Vliegenthart,J.F.G. (1990) Isolation and structure determination of the intact sialylated N-linked carbohydrate chains of recombinant human follitropin expressed in Chinese hamster ovary cells. *Eur. J. Biochem.*, **193**, 263–271.

Hård,K., Mekking,A., Kamerling,J.P., Dacremont,G.A.A. and Vliegenthart,J.F.G. (1991) Different oligosaccharides accumulate in the brain and urine of a cat with α-mannosidosis: structure determination of five brain-derived and seventeen urinary oligosaccharides. *Glycoconj. J.*, **8**, 17–28.

Hård,K., van Zadelhoff,G., Moonen,P., Kamerling,J.P. and Vliegenthart,J.F.G. (1992) The Asn-linked carbohydrate chains of human Tamm-Horsfall glycoprotein; novel sulfated and novel GalNAc-containing N-linked carbohydrate chains. *Eur. J. Biochem.*, **209**, 895–915.

Hardy,M.R. and Townsend,R.R. (1988) Separation of positional isomers of oligosaccharides and glycopeptides by high-performance anion-exchange chromatography with pulsed amperometric detection. *Proc. Natl Acad. Sci. USA*, **85**, 3289–3293.

Hermentin,P., Witzel,R., Vliegenthart,J.F.G., Kamerling,J.P., Nimtz,M. and Conradt,H.S. (1992) A strategy for the mapping of N-glycans by high-pH anion-exchange chromatography with pulsed amperometric detection. *Anal. Biochem.*, **203**, 281–289.

Hokke,C.H., van Dedem,G.W.K., Kamerling,J.P. and Vliegenthart,J.F.G. (1991) Determination of branch location of extra *N*-acetyllactosamine units in sialo N-linked tetraantennary oligosaccharides. *FEBS Lett.*, **286**, 18–24.

Hokke,C.H., Bergwerff,A.A., van Dedem,G.W.K., Kamerling,J.P. and Vliegenthart,J.F.G. (1995) Structural analysis of the sialylated N- and O-linked carbohydrate chains of recombinant human erythropoietin expressed in Chinese hamster ovary cells. *Eur. J. Biochem.*, **228**, 981-1008.

Hurwitz,D.R., Emanuel,S.L., Nathan,M.H., Sarver,N., Ullrich,A., Felder,S., Lax,I. and Schlessinger,J. (1991) EGF induces increased ligand binding affinity and dimerization of soluble epidermal growth factor receptor extracellular domain. *J. Biol. Chem.*, **266**, 22035–22043.

Jay,A. (1996) The methylation reaction in carbohydrate analysis. *J. Carbohydr. Chem.*, **15**, 897–923.

Kamerling,J.P. and Vliegenthart,J.F.G. (1989) Carbohydrates. In Lawson,A.M. (ed.), *Clinical Biochemistry—Principles, Methods, Applications,* **1**. *Mass spectrometry.* Walter de Gruyter, Berlin, pp. 176–263.

Kamerling,J.P. and Vliegenthart,J.F.G. (1992) High-resolution ^{1}H-nuclear magnetic resonance spectroscopy of oligosaccharide-alditols released from mucin-type *O*-glycoproteins. In Berliner,L.J. and Reuben,J. (eds.), *Biological Magnetic Resonance*, **10**. *Carbohydrates and nucleic acids.* Plenum Press, New York, pp. 1–194.

Kobata,A., Totani,K., Endo,T., Kuroki,M., Matsuoka,Y. and Yamashita,K. (1989) Comparative study of the sugar chains of carcinoembryonic antigen and its crossreacting antigen purified from meconium. In Yachi,A. and Shively,J.E. (eds.), *The carcinoembryonic antigen gene family.* Elsevier Biomedical, Amsterdam, pp. 12–22.

Lax,I., Mitra,A.K., Ravera,C., Hurwitz,D.R., Rubinstein,M., Ullrich,A., Stroud,R.M. and Schlessinger,J. (1991) Epidermal growth factor (EGF) induces oligomerization of soluble, extracellular, ligand-binding domain of EGF receptor. *J. Biol. Chem.*, **266**, 13828–13833.

Lemmon,M.A., Bu,Z., Ladbury,J.E., Zhou,M., Pinchasi,D., Lax,I., Engelman,D.M. and Schlessinger,J. (1997) Two EGF molecules contribute additively to stabilization of the EGFR dimer. *EMBO J.*, **16**, 281–294.

Libermann,T.A., Razon,N., Bartal,A.D., Yarden,Y., Schlessinger,J. and Soreq,H. (1984) Expression of epidermal growth factor receptors in human brain tumors. *Cancer Res.*, **44**, 753–760.

Maaheimo,H., Renkonen,R., Turunen,J.P., Pentillä,L. and Renkonen,O. (1995) Synthesis of a divalent sialyl Lewis x O-glycan, a potent inhibitor of lymphocyte-endothelium adhesion. *Eur. J. Biochem.*, **234**, 616–625.

Mangelsdorf Soderquist,A., Stoscheck,C. and Carpenter,G. (1988) Similarities in glycosylation and transport between the secreted and plasma membrane forms of the epidermal growth factor receptor in A-431 cells. *J. Cell Physiol.*, **136**, 447–454.

Mayes,E.L.V. and Waterfield,M.D. (1984) Biosynthesis of the epidermal growth factor receptor in A431 cells. *EMBO J.*, **3**, 531–537.

Michalski,J.-C., Wieruszeski,J.-M., Alonso,C., Cache,P., Montreuil,J. and Strecker,G. (1991) Characterization and 400-MHz ^{1}H-NMR analysis of urinary fucosyl glycoasparagines in fucosidosis. *Eur. J. Biochem.*, **201**, 439–458.

Slieker,L.J. and Lane,M.D. (1985) Post-translational processing of the epidermal growth factor receptor. *J. Biol. Chem.*, **260**, 687–690.

Strecker,G., Wieruszeski,J.-M., Michalski,J.-C., Alonso,C., Boilly,B. and Montreuil,J. (1992) Characterization of Le^x, Le^y and ALe^y antigen determinants in KDN-containing O-linked glycan chains from *Pleurodeles waltlii* jelly coat eggs. *FEBS Lett.*, **298**, 39–43.

Todderud,G. and Carpenter,G. (1988) Presence of mannose phosphate on the epidermal growth factor receptor in A-431 cells. *J. Biol. Chem.*, **263**, 17893–17896.

Toida,T., Vlahov,I.R., Smith,A.E., Hileman,R.E. and Linhardt,R.J. (1996) C-2 Epimerization of *N*-acetylglucosamine in an oligosaccharide derived from heparan sulfate. *J. Carbohydr. Chem.,* **15**, 351–360.

Ullrich,A., Coussens,L., Hayflick,J.S., Dull,T.J., Gray,A., Tam,A.W., Lee,J., Yarden,Y., Libermann,T.A., Schlessinger,J. and others. (1984) Human epidermal growth factor receptor cDNA sequence and aberrant expression of the amplified gene in A431 epidermoid carcinoma cells. *Nature*, **309**, 418–425.

Ullrich,A. and Schlessinger,J. (1990) Signal transduction by receptors with tyrosine kinase activity. *Cell*, **61**, 203–212.

Van Rooijen,J.J.M., Kamerling,J.P. and Vliegenthart,J.F.G. (1998) Sulfated di-, tri- and tetraantennary N-glycans in Tamm-Horsfall glycoprotein. *Eur. J. Biochem.*, **256**, 471–487.

Vliegenthart,J.F.G., Dorland,L. and van Halbeek,H. (1983) High-resolution, ^{1}H-nuclear magnetic resonance spectroscopy as a tool in the structural analysis of carbohydrates related to glycoproteins. *Adv. Carbohydr. Chem. Biochem.*, **41**, 209–374.

Weber,W., Gill,G.N. and Spiess,J. (1984) Production of an epidermal growth factor receptor-related protein. *Science*, **224**, 294–297.

Weber,W., Wenisch,E., Günther,N., Marnitz,U., Betzel,C. and Righetti,P.G. (1994) Protein microheterogeneity and crystal habits: the case of epidermal growth factor receptor isoforms as isolated in a multicompartiment electrolyzer with isoelectric membranes. *J. Chromatogr. A,* **679**, 181–189.