Glycoconjugate Journal 15, 823–833 (1998)

Different O-glycosylation of respiratory mucin glycopeptides from a patient with cystic fibrosis

Kristina A. Thomsson[1], Ingemar Carlstedt[2], Niclas G. Karlsson[1], Hasse Karlsson[1] and Gunnar C. Hansson[1]*

[1]*Department of Medical Biochemistry, Göteborg University, Medicinaregatan 9A, 413 90 Gothenburg, Sweden*
[2]*Department of Cell and Molecular Biology, Lund University, PO. Box 94, 221 00 Lund, Sweden*

The O-linked oligosaccharides from three fractions of highly glycosylated mucin glycopeptides obtained from sputum of a patient with cystic fibrosis were characterized and compared regarding size, composition, sequence and when possible linkage positions. Neutral and sialic acid-containing glycans were permethylated and analyzed by high-temperature GC-MS and MALDI-MS, showing more than 60 different oligosaccharides with a size of up to 15 monosaccharide units. Some of the observed oligosaccharides are novel for respiratory secretions, one being a trifucosylated heptasaccharide with the proposed structure: Fuc-Gal-4(Fuc-3)GlcNAc-(Fuc-)Gal-3GalNAcol. The glycosylation of two of the glycopeptide fractions was similar with regard to the neutral and sialylated oligosaccharides despite their different origins from the sol or gel phase. Analysis of the sulfated oligosaccharides by FAB-MS/MS indicated that the gel fraction contained C-6 linked sulfate groups while the two sol fractions also contained C-3 linked sulfate. The results suggest the presence of different glycosylated mucin domains, probably originating from different mucin glycoforms and/or apoproteins in the airway of cystic fibrosis patients.

Keywords: **mucin, cystic fibrosis, oligosaccharides, MALDI, GC-MS**

Abbreviations: **CF, cystic fibrosis; GalNAcol, *N*-acetylgalactosaminitol; Hex, hexose; HexNAc, *N*-acetylhexosamine; HexNAcol, *N*-acetylhexosaminitol; GC, gas chromatography; MALDI, matrix-assisted laser desorption ionization; FAB, fast atom bombardment; MS, mass spectrometry; CID,collision induced dissociation**

Introduction

The major macromolecular components in the protective mucus layer covering the luminal surface of the respiratory tract are mucins, produced by mucous cells in the submucosal glands and goblet cells in the surface epithelia [1]. Mucins are large glycoproteins, O-glycosylated via GalNAc to the amino acids serine and threonine, which together with proline, are enriched into 'domains' of the protein backbone. These 'mucin domains' are resistant to proteolytic digestion due to the presence of densely packed oligosaccharides.

Cystic fibrosis is one of the most common inherited diseases [2]. The genetic defect is a mutation in a membrane-protein (CFTR), that acts as a chloride channel. Patients with cystic fibrosis suffer from an overproduction of highly viscous mucus and persistent infections in the airways with, in particular, *Hemophilus influenzae*, *Staphylococcus aureus* and *Pseudomonas aeruginosa*. Little is known about the relation between the genetic defect and the deranged properties of mucus. Information about structural changes in the respiratory mucins in CF is scarce, although an altered sialylation and sulfation have been observed [3–5]. CF respiratory mucins have macromolecular properties [6] similar to those from normal individuals [7], although being more 'degraded'. In keeping with the increased sulfation, populations with a higher buoyant density and increased 'acidity' were identified, in particular in the sol phase [8]. Oligosaccharides from patients with cystic fibrosis have been analyzed before (see for example [9–11]), and these studies have provided an impressive list of structures. However, the glycosylation of respiratory mucin subpopulations has not been studied and little information concerning the relative abundance of the individual oligosaccharides is available. Here we have analyzed the neutral, sialylated and sulfated oligosaccharides from three respiratory mucin glycopeptide fractions from CF sputum concerning their structure and abundance. Our results

*To whom correspondence should be addressed. + 46-31-773 3488; Fax: + 46-31-416108; E-mail: gunnar.hansson@medkem.gu.se

demonstrate that these populations of mucin domains are indeed differently glycosylated and indicate the extent of variation in glycosylation that exists between glycoforms and/or mucin apoproteins in the respiratory tract. Since carbohydrates are believed to play important roles in the attachment of microbes, the highly glycosylated mucin domains are probably major interaction points and different mucins, possibly with different cellular origin, could very well vary in their potential to interact with microbes.

Material and methods

Isolation of high-molecular mass mucin glycopeptides

Frozen sputum (30 g) from a female CF patient (16 years old, blood group O, Pseudomonas-infected) was thawed with gentle mixing into 30 ml of cold NaCl (0.2 M), sodium phosphate buffer (pH 6.5, 10 mM) containing Na_2EDTA (5 mM) and di-isopropyl phosphofluoridate (1.2 ml, 0.1 M in propanol). The sol and gel phases were separated by ultracentrifugation (Beckman 50.2 Ti rotor, 30 min, 40000 rev min^{-1} at 4 °C) and cold ethanol (4 vols) were added to both samples. The sol was left to precipitate and the gel was gently stirred at 4 °C overnight. After centrifugation (20 min, benchtop centrifuge), the precipitates were dissolved in 4 M guanidinium chloride and reduced, alkylated, nuclease and trypsin digested was described [12]. Four volumes of cold ethanol were added and samples were left to precipitate for 2 h at 4 °C. The precipitates were recovered by centrifugation, dissolved in guanidinium chloride (4.0 M) and chromatographed on a Sephacryl S-200 (Pharmacia Uppsala, Sweden) column (4.5 cm × 88 cm) eluted with guanidinium chloride (4.0 M, pH 7.0) at a flow rate of 50 ml h^{-1}. The high-molecular mass glycopeptides were recovered from the void volume after extensive dialysis against water and lyophilized. The sequence of reduction/alkylation, nuclease- and trypsin digestion as well as the gel chromatography step was repeated once.

Ion-exchange HPLC of high-molecular mass mucin glycopeptides

The high-molecular mass glycopeptides were chromatographed on a Pharmacia Mono Q HR 5/5 ion-exchange column. The column was eluted at a flow rate of 0.5 ml min^{-1}, first isocratically for 10 min with buffer A (piperazine-perchlorate buffer, 10 mM, pH 5.0) and then with a linear gradient for 60 min from 0–100% of buffer B (lithium perchlorate, 0.25 M/piperazine-perchlorate buffer, 10 mM, pH 5.0). Fractions (0.5 ml) were analyzed by automated procedures for neutral sugars and sialic acids [8]. Fractions were pooled and recovered by lyophilization after extensive dialysis against water.

Release and analyses of oligosaccharides

Three glycopeptide fractions obtained after ion-exchange HPLC on Mono Q, Sol-I (5.4 mg), Sol-II (7.2 mg), and Gel-I (8.1 mg), were dissolved in water and oligosaccharides were released with reductive β-elimination and separated into a neutral and an acidic oligosaccharide pool as described [12]. Further fractionation of the acidic oligosaccharides (from 3–4.5 mg glycopeptides) was performed by isolating the neutral pool eluted from a second DEAE-column after esterification (sialylated species) and a pool eluted by high salt (sulfated species) subsequently followed by the conversion of the methyl ester derivative of the sialylated species into *N*-methylamides [13]. The neutral and sialic acid-containing oligosaccharides were permethylated [13] and an aliquot of the sialylated derivatives separated into low and high molecular mass species by gel chromatography [14]. The monosaccharide composition of the neutral and acidic oligosaccharide fractions was determined by GC as their alditol acetates and sialic acids after neuraminidase digestion [12].

High-temperature GC and GC-MS of permethylated oligosaccharides

For GC and GC-MS, permethylated oligosaccharides dissolved in ethyl acetate were injected (0.5–1 µl) on-column at 70 °C or 80 °C onto fused silica capillaries (10 m × 0.25 mm i.d, HT-polyimide coated) coated with crosslinked 0.03 µm PS 264 stationary phase (Fluka, Buchs, Switzerland) [13]. Neutral oligosaccharides were analyzed by increasing the temperature 10 °C min^{-1} up to 400 °C and derivatives of the sialic acid-containing oligosaccharides were analyzed by increasing the temperature 50 °C min^{-1} up to 200 °C followed by 10 °C min^{-1} up to 390 °C. For GC, the flame ionization detector was kept at 400 °C. GC-MS was performed on a Hewlett-Packard 5890A-II gas chromatograph interfaced to a JEOL SX-102A mass spectrometer (Jeol, Tokyo, Japan) [13].

MALDI-MS of permethylated oligosaccharides

The matrix-assisted laser desorption time-of-flight mass spectrometry was performed on a VG TofSpec E (Micromass, Manchester, UK) equipped with a nitrogen laser (337 nm, 3 ns pulses) and used in the positive reflectron mode. Acceleration voltage was + 20 or + 22.5 kV and mass spectra were obtained by accumulating 80 laser shots. External multipoint calibration was performed using the lithium adducts of permethylated polyfructanes. The mass accuracy when identifying oligosaccharides was better than 0.05% and the resolution approximately 1500 (full width at half maximum). The mass spectra of the permethylated oligosaccharides showed sodium and potassium adducts of the molecular ions using DHB (2,5-dihydroxybenzoic acid) as matrix, but by adding LiCl to the target mixtures, only lithium adducts were obtained. The samples were analyzed with and without LiCl to ensure that no ion suppression effects occurred. Samples were prepared by using DHB as matrix (4 µl, 50 mM in 1 : 1 H_2O : acetonitrile) mixed with

LiCl (2 μl, 0.05 M in H_2O) and analyte (2 μl, same concentration as for GC, in ethyl acetate). 0.5–1 μl of the mixture was applied to the target and left to dry in air.

Derivatization and analysis of sulfated oligosaccharides by tandem mass spectrometry

Aliquots of the sulfated oligosaccharides were acetylated with 300 μl pyridine and 150 μl acetic acid anhydride or acetic anhydride-d_6. The samples were dissolved in MeOH (100 μl), of which 1–2 μl were mixed with approximately 5 μl of the matrix (triethanolamine) on the FAB target. FAB-MS/MS experiments were performed in the negative mode on a JEOL HX/HX 110A four sector tandem mass spectrometer as described [15].

Results

Isolation of high-molecular mass mucin glycopeptides and preparation of neutral, sialic acid- and sulfate-containing oligosaccharides

Sputum from a blood group O patient with cystic fibrosis was separated into a sol and a gel phase, and trypsin-resistant glycopeptides were isolated following reduction-alkylation, nuclease- and trypsin digestion. The glycopeptides were purified by gel filtration (results not shown) and separated by ion-exchange HPLC (Figure 1). The mucin glycopeptides from the sol were more retarded than those from the gel phase implying a more acidic glycosylation, and only a small part of the material from the gel eluted at the same position as the later part of the sol peak. The results suggest the presence of at least two glycopeptide populations. The peaks shown in Figure 1 represent the major part of the high-molecular mass glycopeptides from the sol and gel phase (measured as total hexose or sialic acid) indicating that their oligosaccharides should represent the vast majority of mucin oligosaccharides from the respiratory tract.

The major part of the material from the gel was pooled as Gel-I. The Sol-I mucin glycopeptide fraction was pooled from the sol phase glycopeptides with the same 'acidity' as the Gel-I, excluding part of the early eluting peak of the sol phase glycopeptides. The late eluting glycopeptides from the sol glycopeptides were combined into the Sol-II fraction. Late eluting material from the gel phase was, due to the small amount of material, not studied further.

The released O-linked oligosaccharide alditols from the Sol-I, Sol-II and Gel-I fractions were separated into neutral and acidic species by anion-exchange chromatography and analyzed for monosaccharide composition (Table 1). The Sol-II glycopeptides contained a higher proportion of acidic oligosaccharides, consistent with the corresponding glycopeptides being more retarded on ion-exchange chromatography. The average chain length of the neutral oligosaccharides was estimated to about 6 residues for the

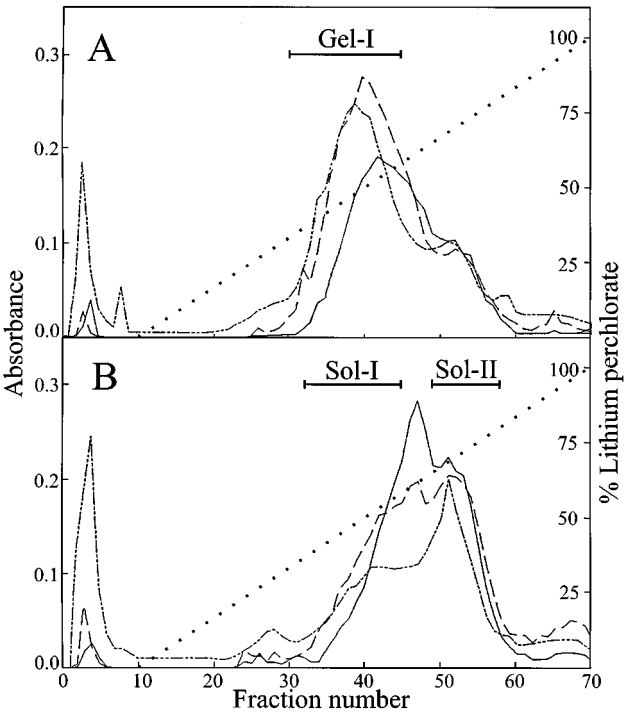

Figure 1. Ion-exchange HPLC of high-molecular mass mucin glycopeptides from (a) the gel and (b) the sol phase of CF sputum. Glycopeptides were chromatographed on a Mono Q HR5/5 column and fractions were analysed for sialic acid (—), hexose (- -) and absorbance ($\times 10^{-1}$) at 280 nm (-···-). The dotted line indicates the percentage of buffer B. Material was pooled as indicated.

Table 1. Monosaccharide composition of neutral and acidic oligosaccharides released from CF respiratory mucin glycopeptides.

	Sol-II	*Sol-I*	*Gel-I*
Neutral oligosaccharides	11%[a]	33%	31%
Fuc	0.4[b]	1.2	1.5
Gal	1.2	2.2	1.9
GlcNAc	0.5	1.4	1.5
GalNAc	1	1	1
Average chain length[c] (No. of residues)	3.1	5.8	5.9
Acidic (sialic acid- and sulfate-containing) oligosaccharides	89%[a]	67%	69%
Fuc	1.3	1.3	0.9
Gal	3.5	2.6	2.6
GlcNAc	1.9	1.7	1.5
GalNAc	1	1	1
NeuAc	1.6	1.1	0.5
Average chain length[c] (No. of residues)	9.3	7.7	6.5

[a] Percentage molar distribution of the sum of monosaccharides between the neutral and acidic glycans. The results are the mean of two separate analyses.
[b] Molar ratios of individual saccharides are given relative to GalNAc.
[c] Assuming that GalNAc only appears as linkage sugar.

Sol-I and Gel-I and 3 for the Sol-II assuming that all GalNAc residues were linked to the peptide backbone (blood group O patient). The acidic oligosaccharides in Sol-I and Gel-I were, on average, of the same size as the corresponding neutral ones, whereas Sol-II contained acidic oligosaccharides that were significantly longer, having an average length of approximately nine residues.

Analysis of the neutral oligosaccharides

MALDI mass spectra of the three neutral permethylated oligosaccharide fractions (Sol-I, Sol-II and Gel-I) are shown in Figure 2. The addition of LiCl to the matrix was shown to quantitatively suppress sodium and potassium adducts. Monosaccharide composition showed that the oligosaccharides consist of Gal, GlcNAc, Fuc, and GalNAc. The majority of GalNAc residues within the samples are assumed to be directly linked to the protein backbone and thus converted to GalNAcol during the release under reductive conditions. This assumption was strengthened by the fact that the analyzed patient was blood group A negative. Assuming that all GalNAc residues were converted into GalNAcol, and that the monosaccharide residues found by the compositional analysis are the only ones present, the pseudomolecular ions obtained by MALDI-MS were assessed (Table 2). Only oligosaccharides with up to 15 residues were observed, with weak intensities of the larger pseudomolecular ions, in agreement with the majority of the oligosaccharides being within the low mass region, as predicted from the monosaccharide composition. Also predicted from the monosaccharide composition is the finding that the larger oligosaccharides are present in the Sol-I and Gel-I fractions, in line with the observation that the largest oligosaccharides detected by MALDI-MS were observed in these fractions (Figure 2). The larger neutral species contain approximately one Fuc residue per Gal-GlcNAc unit, indicating the presence of fucosylated *N*-acetyllactosamine structures. The Sol-II fraction revealed a glycosylation profile different from those of Sol-I and Gel-I with less abundant higher mass pseudomolecular ions (Figure 2 and Table 1). The variation of the glycosylation detected by MALDI-MS was also reflected in the gas chromatograms (Figure 3) of the permethylated neutral oligosaccharides, where Sol-II could be distinguished from Sol-I and Gel-I.

Twenty oligosaccharide sequences were characterized from the mass spectra obtained by GC-MS of the three fractions (Table 2). A high degree of fucosylation of the oligosaccharides were observed as exemplified by the component 7.a, interpreted as Fuc-Gal-4(Fuc-3)GlcNAc-(Fuc-)Gal-3GalNAcol (Table 2) from its mass spectrum shown in Figure 4. In this and the other mass spectra fragment ions are mainly oxonium- and inductive ions [12–14]. The

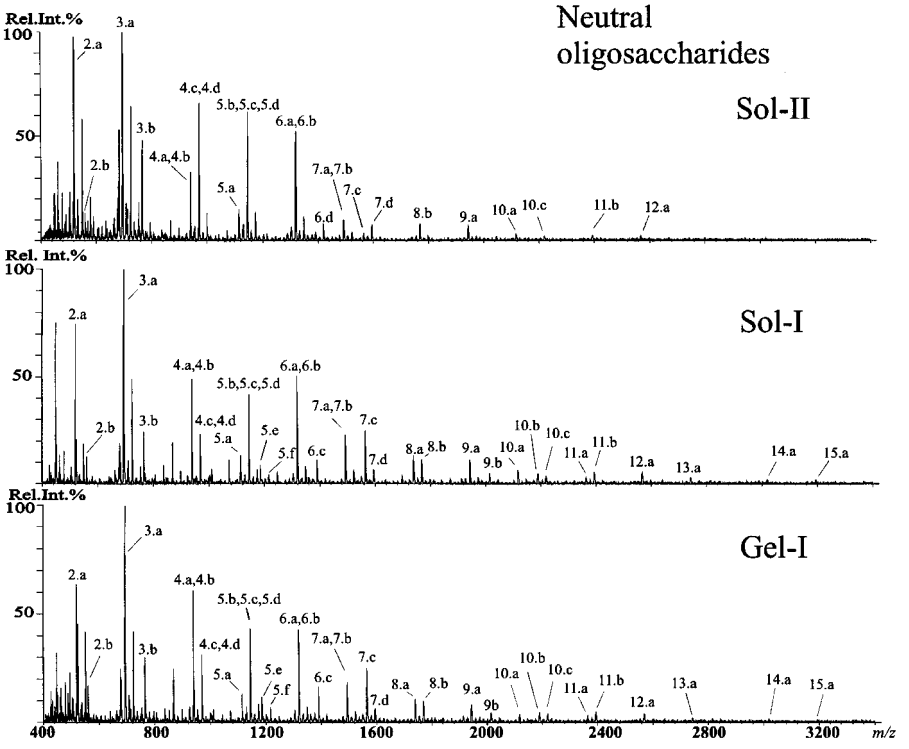

Figure 2. MALDI mass spectra of permethylated neutral oligosaccharides from the glycopeptide fractions Sol-II, Sol-I and Gel-I. Pseudomolecular ions occur as lithium adducts $[M + Li]^+$. The deduced oligosaccharide compositions are listed in Table 2. Conditions for MALDI are described in Materials and Methods.

Table 2. Neutral oligosaccharides of mucin glycopeptides from sputum of a patient with cystic fibrosis. Molar ratios of neutral oligosaccharides in Sol-I, Sol-II and Gel-I were calculated relative to *GlcNAc-3GalNAcol* analysed by GC. The sequences are listed together with the oligosaccharides observed by MALDI-MS and their proposed composition.

No.	*Sequences of neutral oligosaccharides*[a]	*Molecular mass*[b] *(Da)*	*GC*[c]			*MALDI-TOF MS*		
			Sol-II	*Sol-I*	*Gel-I*	*Sol-II*	*Sol-I*	*Gel-I*
1.a	GalNAcol	307.2	9.1	5.3	2.4	NA	NA	NA
2.a	Gal-3GalNAcol	511.3	4.7	2.7	1.8	x	x	x
2.b	*GlcNAc-3GalNAcol (reference)*	*552.3*	*1.0*	*1.0*	*1.0*	x	x	x
3.a	Fuc-Gal-3GalNAcol	685.4	5.6	7.0	4.4	x	x	x
3.b	Gal-4GlcNAc-3GalNAcol	756.4	1.4	1.2	1.0	x	x	x
4.a	Fuc-Gal-3GlcNAc-3GalNAcol	930.5	0.26	1.6	2.4	x	x	x
4.b	Fuc-Gal-4GlcNAc-3GalNAcol	930.5	0.51	1.1	0.39	x	x	x
4.c	Gal-3**(Gal-4GlcNAc-6)**GalNAcol	960.5	0.95	0.62	0.65	x	x	x
4.d	Gal-4GlcNAc-Gal-3GalNAcol	960.5	1.1	0.22	0.20	x	x	x
5.a	Fuc-Gal-(Fuc-)GlcNAc-3GalNAcol	1104.6	0.45	0.44	0.35	x	x	x
5.b	Gal-3**(Fuc-Gal-GlcNAc-6)** GalNAcol	1134.6	0.41	0.40	0.40	x	x	x
5.c	Fuc-Gal-GlcNAc-Gal-3GalNAcol	1134.6	0.58	0.47	0.48	x	x	x
5.d	Fuc-Gal-3**(Gal-GlcNAc-6)**GalNAcol	1134.6	0.52	0.38	0.32	x	x	x
5.e	GlcNAc-3**(Fuc-Gal-GlcNAc-6)**GalNAcol	1175.6	ND	0.34	0.41		x	x
5.f	Gal_2, $GlcNAc_2$, GalNAcol	1205.6	ND	ND	ND		x	x
6.a	Fuc-Gal-(Fuc-)GlcNAc-Gal-3GalNAcol	1308.7	0.44	0.25	0.17	x	x	x
6.b	Fuc-Gal-3 **(Fuc-Gal-4GlcNAc-6)**GalNAcol	1308.7	0.41	0.69	0.64	x	x	x
6.c	Gal-GlcNAc-3**(Fuc-Gal-GlcNAc-6)**GalNAcol	1379.7	ND	0.29	0.38		x	x
6.d	Gal_3, $GlcNAc_2$, GalNAcol	1409.7	ND	ND	ND	x		
7.a	Fuc-Gal-4(Fuc-3)GlcNAc-(Fuc-)Gal-3GalNAcol	1482.8	ND	0.15	0.12	x	x	x
7.b	Fuc-Gal-3**(Fuc-Gal-(Fuc-)GlcNAc-6)**GalNAcol	1482.8	ND	0.23	0.20	x	x	x
7.c	Fuc-Gal-GlcNAc-3**(Fuc-Gal-GlcNAc-6)**GalNAcol	1553.8	< 0.10	0.57	0.58	x	x	x
7.d	Fuc, Gal_3, $GlcNAc_2$, GalNAcol	1583.8				x	x	x
8.a	Fuc_3, Gal_2, $GlcNAc_2$, GalNAcol	1727.9					x	x
8.b	Fuc_2, Gal_3, $GlcNAc_2$, GalNAcol	1757.9				x	x	x
9.a	Fuc_3, Gal_3, $GlcNAc_2$, GalNAcol	1932.0				x	x	x
9.b	Fuc_2, Gal_3, $GlcNAc_3$, GalNAcol	2003.0					x	x
10.a	Fuc_4, Gal_3, $GlcNAc_2$, GalNAcol	2106.1				x	x	x
10.b	Fuc_3, Gal_3, $GlcNAc_3$, GalNAcol	2177.1					x	x
10.c	Fuc_2, Gal_4, $GlcNAc_3$, GalNAcol	2207.1				x	x	x
11.a	Fuc_4, Gal_3, $GlcNAc_3$, GalNAcol	2351.2					x	x
11.b	Fuc_3, Gal_4, $GlcNAc_3$, GalNAcol	2381.2				x	x	x
12.a	Fuc_4, Gal_4, $GlcNAc_3$, GalNAcol	2555.3				x	x	x
13.a	Fuc_5, Gal_4, $GlcNAc_3$, GalNAcol	2729.4					x	x
14.a	Fuc_4, Gal_5, $GlcNAc_4$, GalNAcol	3004.5					x	x
15.a	Fuc_5, Gal_5, $GlcNAc_4$, GalNAcol	3178.6					x	x

NA = not analyzed, ND = not detected, x = present.

[a] The saccharides marked in bold are located on C-6 of GalNAcol. The following assumptions have been made: hexose residues are Gal, *N*-acetylhexosamine residues are GlcNAc, deoxyhexose residues are Fuc and *N*-acetyl hexosaminitol residues are GalNAcol.

[b] Permethylated oligosaccharides, calculated monoisotopic mass. Mass accuracy was better than 0.05% when analyzed by MALDI-MS.

[c] The molar ratio of oligosaccharides are given relative to *GlcNAc-3GalNAcol*. The signal area response of each oligosaccharide was divided with its molecular mass. The larger oligosaccharides were not analyzable by GC and GC-MS.

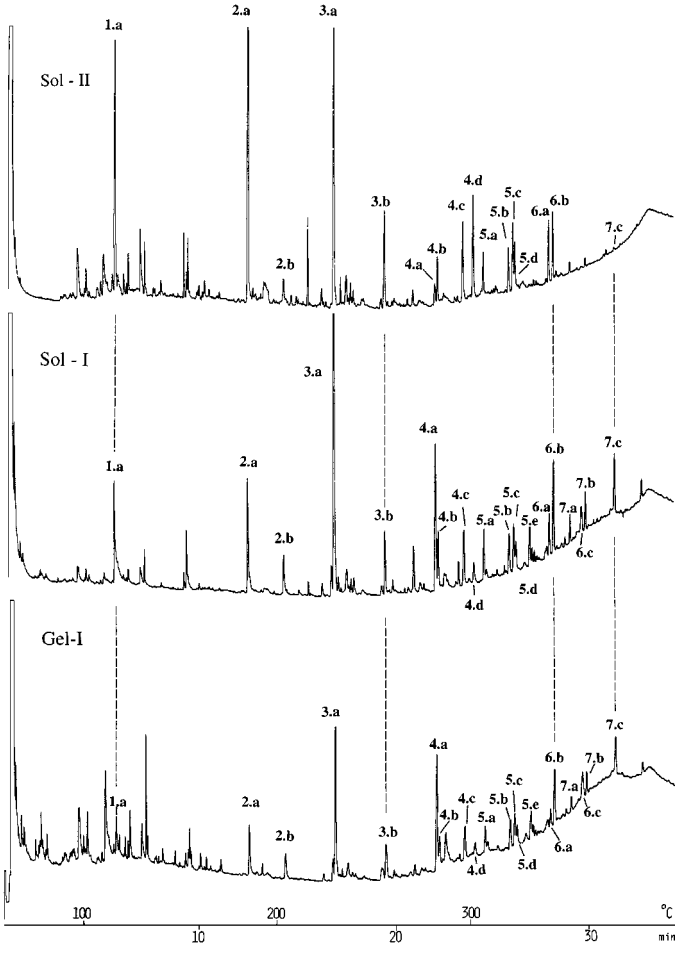

Figure 3. Gas chromatograms of permethylated neutral oligosaccharides from the glycopeptide fractions Sol-II, Sol-I and Gel-I. The numbers above the peaks correspond to the sequences listed in Table 2. Chromatographic details are given in Materials and Methods.

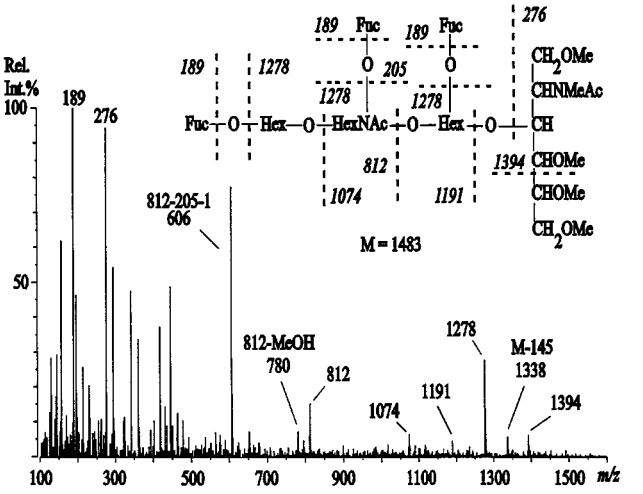

Figure 4. Mass spectrum (background subtracted) from GC-MS of the permethylated neutral trifucosylated heptasaccharide component 7.a (Table 1) recorded from the analysis of the oligosaccharides from the glycopeptide fraction Sol-I.

intense fragment ion at *m/z* 189 indicated several terminal Fuc residues in this molecule. The fragment ion at *m/z* 1394 (cleavage between C-4 and C-5 of the HexNAcol) showed that the single saccharide chain Fuc-Hex-(Fuc-)HexNAc-(Fuc-)Hex-, identified by the fragmentation, is located at C-3 of GalNAcol. The presence of the intense signal at *m/z* 606 actually enables some assignment of linkage position since it is typical for difucosylated type 2 chains [16]. All the deduced oligosaccharides were semiquantified by GC (Figure 3) and the relative molar amounts of the individual components are compiled in Table 2. Oligosaccharides identified with GC-MS were supported by the presence of the corresponding $[M + Li]^+$ ions as observed by MALDI-MS.

Analysis of the sialic acid-containing oligosaccharides

The sialylated oligosaccharides were further separated from the sulfate-containing ones using on-column derivatization followed by conversion into *N*-methyl amides [13]. The neutral and sialic acid-containing oligosaccharides were permethylated and analyzed by MALDI-MS and high-temperature GC and GC-MS. MALDI-MS of the three permethylated *N*,*N*-dimethyl amide sialic acid-containing oligosaccharide fractions from Sol-I, Sol-II and Gel-I showed $[M + Li]^+$ ions corresponding to oligosaccharides with up to 10 residues (Figure 5). As for the neutral oligosaccharides, the largest components were observed in the Sol-I and Gel-I fractions. The most obvious difference when comparing the gas chromatograms (Figure 6) was the larger amount of the trisaccharide 3.2 (NeuAc-Gal-3GalNAcol) in the Sol-II fraction compared to Sol-I and Gel-I.

Sixteen sialic acid-containing oligosaccharide sequences were deduced by GC-MS and quantified by GC (Table 3). Sialic acid was found attached to Gal and C-6 of GalNAcol. To permit analysis of the three fractions by GC, oligosaccharides with 1–7 residues were isolated by gel filtration. The relative amounts of the late eluting components, indicated between brackets in Table 3, may not truly reflect the actual situation due to an overlap in the separation of the smaller and larger oligosaccharides by gel filtration.

Analysis of the sulfated oligosaccharides

The acetylated and perdeuteroacetylated sulfated oligosaccharides from the three fractions were analyzed by FAB-MS and detected as $[M–H]^-$ ions (Table 4). In the fractions Gel-I and Sol-I, 5 and 6 different molecular ions of sulfated oligosaccharides were observed, respectively, whereas only two were found in Sol-II. The $[M–H]^-$ ions were collided and analyzed by FAB-MS/MS. The fragment ion *m/z* 139 (^{0,4}A) is characteristic of a peracetylated hexose or *N*-acetylhexosamine with sulfate on the C-6 [15], and *m/z* 181 (^{1,3}A) of a peracetylated hexose with sulfate on C-3. In the latter

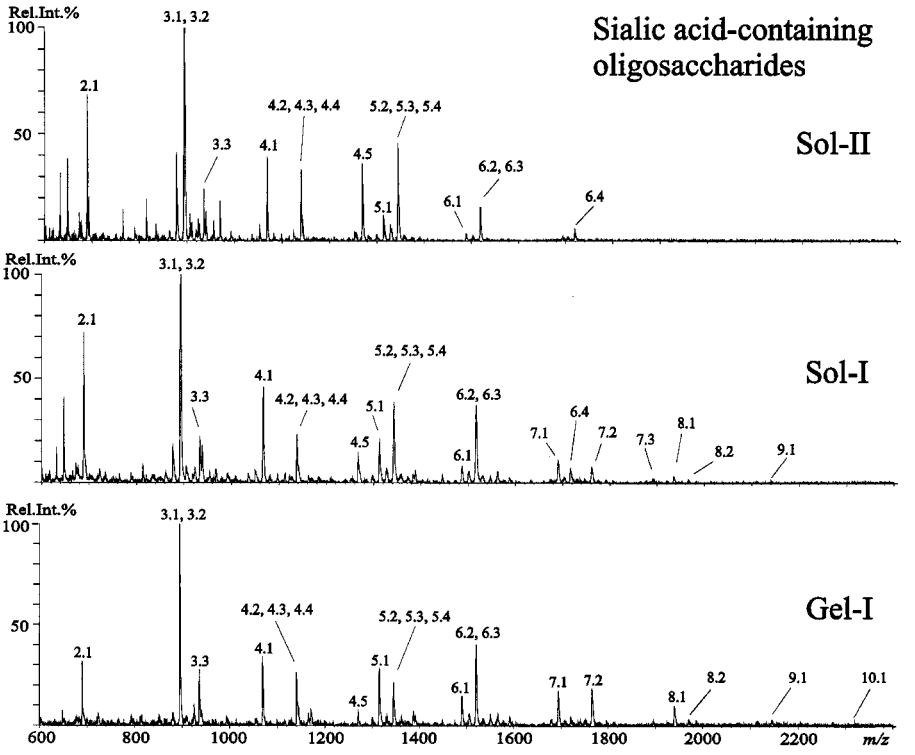

Figure 5. MALDI mass spectra of permethylated *N,N*-dimethyl amide derivatives of the sialic acid-containing oligosaccharides from Sol-II, Sol-I and Gel-I. Pseudomolecular ions occur as lithium adducts $[M + Li]^+$. The predicted oligosaccharide compositions based on the observed masses are listed in Table 3. Conditions for MALDI are described in Materials and Methods.

case, the fragmentation was confirmed by analyses of standard sulfated oligosaccharides. Based on these ions, the sulfate group was found either on C-3 of hexose or on C-6 of *N*-acetylhexosamine residues in the Sol-I and Sol-II fractions (Table 4). In contrast, in the Gel-I fraction, sulfate was only detected on C-6 and not on C-3. As the two Sol fractions contained a mixture of C-3 and C-6 linked sulfate, most of the molecular ions observed (Table 4) contained a mixture of more than one component making the MS/MS spectra difficult to interpret. However, the CID spectra of the oligosaccharides in the Gel-I fraction were less complex due to the single type of sulfation, and four structures could be deduced (Table 5). The fragmentation patterns obtained with high energy collision of sulfated oligosaccharides using FAB-MS/MS have been described [15]. Three of these contained a sulfated blood group H-type 1 or 2 epitope on the C-6 branch of GalNAcol. The CID-spectrum of one of the components with the sequence Fuc-Gal-3**(Fuc-Gal-4(SO_3^--6)GlcNAc-6)**GalNAcol is shown in Figure 7. The intense ions *m/z* 1549 and 1261 (^{1,5}X) were due to the terminal Fuc-Gal- epitope. The presence of ions due to cleavage in the *N*-acetylgalactosaminitol (*m/z* 944, $^4A_{0\alpha}$-Ac; *m/z* 1016, $^3A_{0\alpha}$-Ac), ions at *m/z* 1157 ($Z_{1\beta}$-OAc) and at *m/z* 1445 ($Z_{2\beta}$-OAc) together with the ion *m/z* 713 ($^{3,5}A_{1\alpha}$), the latter originating from a cleavage in the *N*-acetylhexosamine residue, localized the branches.

Discussion

In the present work, the glycosylation profiles of three glycopeptide fractions from the gel and sol phases of sputum from a patient with cystic fibrosis have been analyzed. The structures of twenty neutral and sixteen sialic acid-containing oligosaccharides from the three isolated glycopeptide fractions were elucidated using GC-MS. Core 1, core 2, core 3 and core 4 structures were identified and the branches on these consist of fucosylated and nonfucosylated Gal-GlcNAc units. The single GalNAcol together with the disaccharide Gal-3GalNAcol and trisaccharide Fuc-Gal-3GalNAcol were the most abundant neutral species, while Gal-3**(NeuAc-6)**GalNAcol and NeuAc-Gal-3GalNAcol were the most prominent sialic acid-containing ones. Most sequences were identical to oligosaccharides previously found in sputum from patients suffering from cystic fibrosis or bronchiectasis [9–11, 17–20]. However, two of the neutral (no. 7.a and 5.c) and two of the sialic acid-containing ones (no. 4.3 and 4.4) have not been identified earlier in human sputum. The neutral saccharide 7.a (Fuc-Gal-4(Fuc-3)GlcNAc-(Fuc-)Gal-3GalNAcol) seems to be a novel structure (CarbBank version 3.0). Among the sialic acid-containing oligosaccharides, 4.4 (NeuAc-Gal-GlcNAc-3GalNAcol) has only been observed when sulfated on C-6 of the GlcNAc residue in tracheobronchial mucins from an

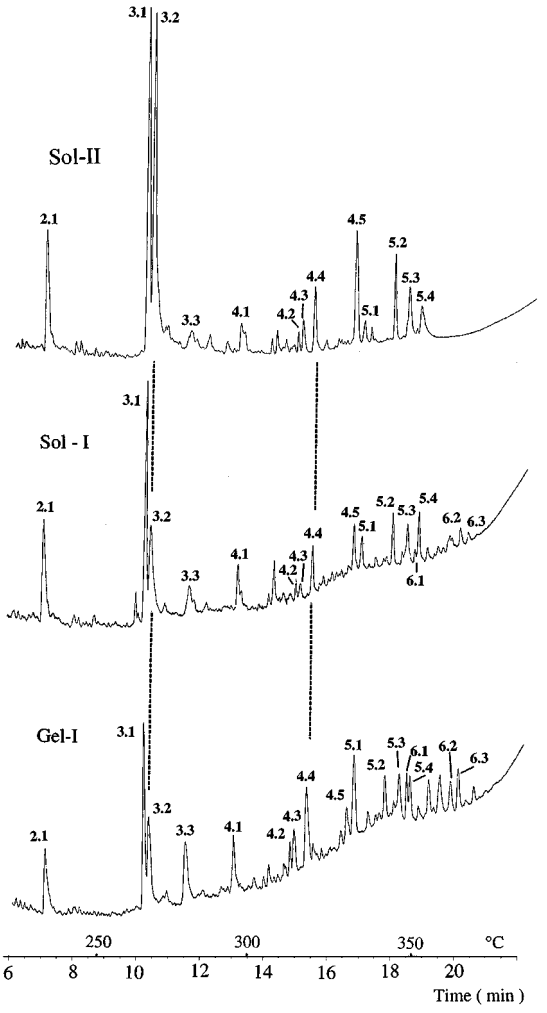

Figure 6. Gas chromatograms of the permethylated *N,N*-dimethyl amide derivatives of the smaller sialic acid-containing oligosaccharides isolated by gel filtration from the glycopeptide fractions Sol-II, Sol-I and Gel-I. The numbers above the peaks correspond to the sequences listed in Table 3. Chromatographic details are given in Materials and Methods.

individual suffering from cystic fibrosis [21]. Sulfated oligosaccharides similar to those described here have been identified in pig intestine using FAB-MS/MS [15], except for the novel sequence: GlcNAc-Gal-3**(Fuc-Gal-(SO$_3^-$-6)GlcNAc-6)**GalNAcol.

Identification of individual oligosaccharides is important, but a better molecular understanding of mucins and mucin domains also requires knowledge of the relative abundance of the individual components. GC is believed to give reliable quantitative information and to evaluate the use of MALDI-MS in this context, the latter technique was compared with GC for the quantification of oligosaccharides with up to seven residues. The estimates agree fairly well (data not shown), supporting the use of MALDI-MS for 'profiling' mixtures of oligosaccharides and for assessing the relative abundance of the various glycans. However, a trend with a gradual decrease in the amounts as estimated by GC compared to MALDI-MS was observed when quantifying the largest oligosaccharides analyzable with GC. The apparent underestimation of large oligosaccharides with GC has been noticed before and is probably due to thermal degradation on the capillary column at higher temperatures [14].

The results from MALDI-MS and GC indicate a shorter average length of both neutral and sialic acid-containing oligosaccharides in the Sol-II fraction compared to Sol-I and Gel-I. In contrast, the monosaccharide analyses predict significantly larger acidic oligosaccharides in Sol-II implying that this fraction may also contain significantly longer sulfated structures. The presence of high-molecular mass sulfated oligosaccharides from CF sputum (up to 160–200 residues) has been indicated [22].

The relative higher amounts of oligosaccharides with terminal NeuAc-Gal- versus **NeuAc-6**GalNAcol in Sol-II than Sol-I and Gel-I together with the less frequent occurrence of blood group H-type determinants within Sol-II (Tables 2 and 3) suggest that the Sol-I and Gel-I glycopeptides have a similar origin, different from the Sol-II glycopeptides. This is probably an oversimplification, as analysis of the sulfated oligosaccharides indicated that the Gel-I fraction contained sulfate groups linked mainly to C-6 of *N*-acetylhexosamine residues, whereas Sol-I and Sol-II also contained C-3 linked sulfate. Nevertheless, it has been observed that different glycoforms of the MUC5B apomucin do exist within the respiratory tract [23] and together with MUC5AC these apomucins appear to be the major contributor to respiratory mucus [24, 25].

The size of respiratory mucin domains is in the order of 300–700 kDa [23]. Using an average size of 400 kDa and assuming that oligosaccharides comprise 90% of the mass, the total number of saccharides per domain is estimated to be 240–300 chains based on the average number of residues per oligosaccharide (Table 1) (average mass of each chain is 1200–1700 Da). About 60 neutral oligosaccharides would be found on such a mucin domain of Sol-II, and 90–100 neutral oligosaccharides on Sol-I and Gel-I (assuming that the oligosaccharides observed by GC represent 90% of the structures). Out of these chains, 20–30 would be either component 1.a or 2.a, and 10–30 component 3.a. With these assumptions, only the oligosaccharides that occur in lowest amounts, (6.a, 7.a or 7.b) would appear as single copies on such a mucin domain. Although the major part of the mucin domains appears to be covered by epitopes presented by only a few different glycans, epitopes present on longer and less frequent oligosaccharides should be readily exposed and may also shield epitopes present on the shorter ones.

Mucus secretions from CF patients appear to differ from those from normal individuals both regarding the presence of apoproteins and oligosaccharide structures. Adding to the complexity of the disease, Li *et al.* [26] have shown that *Pseudomonas aeruginosa* lipopolysaccharide upregulates

Table 3. Sialic acid-containing oligosaccharides of mucin glycopeptides from sputum of a patient with cystic fibrosis. Molar ratios of sialic acid-containing oligosaccharides were calculated relative to *NeuAc-6GalNAcol* analyzed by GC. The sequences are listed together with the oligosaccharides observed by MALDI-MS and their proposed composition.

No.	*Sequences of neutral oligosaccharides*[a]	*Molecular mass*[b] *(Da)*	*GC*[c]			*MALDI-TOF MS*		
			Sol-II	*Sol-I*	*Gel-I*	*Sol-II*	*Sol-I*	*Gel-I*
2.1	*NeuAc-6GalNAcol (reference)*	*681.4*	*1.0*	*1.0*	*1.0*	x	x	x
3.1	Gal-3**(NeuAc-6)**GalNAcol	885.5	2.1	1.1	1.3	x	x	x
3.2	NeuAc-Gal-3GalNAcol	885.5	3.0	0.9	1.0	x	x	x
3.3	GlcNAc-3**(NeuAc-6)**GalNAcol	926.5	0.19	0.24	0.79	x	x	x
4.1	Fuc-Gal-3**(NeuAc-6)**GalNAcol	1059.6	0.14	0.20	0.41	x	x	x
4.2	Gal-GlcNAc-3**(NeuAc-6)**GalNAcol	1130.6	0.05	0.05	0.15	x	x	x
4.3	NeuAc-Gal-3**(GlcNAc-6)**GalNAcol	1130.6	0.14	0.06	0.29	x	x	x
4.4	NeuAc-Gal-GlcNAc-3GalNAcol	1130.6	0.28	0.21	0.62	x	x	x
4.5	NeuAc-Gal-3**(NeuAc-6)**GalNAcol	1259.7	0.50	0.19	0.23	x	x	x
5.1	Fuc-Gal-GlcNAc-3**(NeuAc-6)**GalNAcol	1304.7	0.08	0.14	0.51	x	x	x
5.2	NeuAc-Gal-3**(Gal-GlcNAc-6)**GalNAcol	1334.7	0.22	0.15	0.21	x	x	x
5.3	Gal-3**(NeuAc-Gal-GlcNAc-6)**GalNAcol	1334.7	0.26	0.18	0.29	x	x	x
5.4	NeuAc-Gal-GlcNAc-Gal-3GalNAcol	1334.7	0.20	0.17	0.20	x	x	x
6.1	Fuc-Gal-(Fuc-)GlcNAc-3**(NeuAc-6)**GalNAcol	1478.8	ND	0.04	0.17	x	x	x
6.2	NeuAc-Gal-(Fuc-)GlcNAc-Gal-3GalNAcol	1508.8	(ND)[d]	($>$ 0.06)	($>$ 0.17)	x	x	x
6.3	Fuc-Gal-3**(NeuAc-Gal-GlcNAc-6)**GalNAcol	1508.8	(ND)	($>$ 0.03)	($>$ 0.18)	x	x	x
6.4	$NeuAc_2$, Gal_2, GlcNAc, GalNAcol	1708.9				x	x	
7.1	NeuAc, Fuc_2, Gal_2, GlcNAc, GalNAcol	1682.9					x	x
7.2	NeuAc, Fuc, Gal_2, $GlcNAc_2$, GalNAcol	1753.9					x	x
7.3	$NeuAc_2$, Fuc, Gal_2, GlcNAc, GalNAcol	1883.0					x	
8.1	NeuAc, Fuc_2, Gal_2, $GlcNAc_2$, GalNAcol	1928.0					x	x
8.2	NeuAc, Fuc, Gal_3, $GlcNAc_2$, GalNAcol	1958.0					x	x
9.1	NeuAc, Fuc_2, Gal_3, $GlcNAc_2$, GalNAcol	2132.1					x	x
10.1	NeuAc, Fuc_3, Gal_3, $GlcNAc_2$, GalNAcol	2306.2						x

ND = not detected, x = present.
[a] The saccharides marked in bold are located on C-6 of GalNAcol. The following assumptions have been made: hexose residues are Gal, *N*-acetylhexosamine residues are GlcNAc, deoxyhexose residues are Fuc and *N*-acetyl hexosaminitol residues are GalNAcol.
[b] Permethylated oligosaccharides, calculated monoisotopic mass. The mass accuracy was better than 0.05% when analyzed by MALDI-MS.
[c] The molar ratio of oligosaccharides are given relative to *NeuAc-6GalNAcol*. The signal area response of each oligosaccharide was divided with its molecular mass. The larger oligosaccharides are not analyzable by GC and GC-MS.
[d] The amounts given between brackets are underestimated as oligosaccharides of this size and larger were separated from the shorter ones by gel filtration before analysis by GC and GC-MS.

Table 4. Sulfated oligosaccharides of mucin glycopeptides from sputum of a patient with cystic fibrosis analyzed by FAB-MS/MS.

Composition[a]	*Molecular mass*[b] *(Da)*	*Presence of oligosaccharides*[c]		
		Sol-II	*Sol-I*	*Gel-I*
Fuc, Gal, GlcNAc, GalNAcol, SO_3^-	1275.4	−	+	+ (C-6)
Gal_2, GlcNAc, GalNAcol, SO_3^-	1333.4	+ (C-3, C-6)	+	−
Fuc, Gal_2, GlcNAc, GalNAcol, SO_3^-	1563.4	+ (C-3, C-6)	+ (C-3, C-6)	+ (C-6)
Fuc_2 ,Gal_2, GlcNAc, GalNAcol, SO_3^-	1793.5	−	+ (C-3, C-6)	+ (C-6)
Fuc, Gal_2, $GlcNAc_2$, GalNAcol, SO_3^-	1850.5	−	+ (C-6)	+ (C-6)
Fuc_2, Gal_2, $GlcNAc_2$, GalNAcol, SO_3^-	2080.6	−	+ (C-3, C-6)	+ (C-6)

+ present, − not present
[a] The following assumptions have been made: hexose residues are Gal, *N*-acetylhexosamine residues are GlcNAc, deoxyhexose residues are Fuc and *N*-acetyl hexosaminitol residues are GalNAcol.
[b] Acetylated oligosaccharides, calculated monoisotopic mass.
[c] Where collision experiments were performed, the identified sulfate linkages are indicated in brackets as C-3 when linked to C-3 of hexose and C-6 when linked to C-6 of hexose or *N*-acetylhexosamine residues.

Table 5. Structures of sulfated mucin oligosaccharides from Gel-I deduced by FAB-MS/MS.

Molecular mass[a]	*Structures*[b]
1275.4	Fuc-Gal-3**(SO_3^--GlcNAc-6)**GalNAcol
1793.5	Fuc-Gal-3**(Fuc-Gal-(SO_3^--6)GlcNAc-6)**GalNAcol
1850.5	GlcNAc-Gal-3**(Fuc-Gal-(SO_3^--6)GlcNAc-6)**GalNAcol
2080.6	Fuc-Gal-GlcNAc-3**(Fuc-Gal-(SO_3^--6)GlcNAc-6)**GalNAcol

[a] Acetylated oligosaccharides, calculated monoisotopic mass.
[b] The saccharides marked in bold are located on C-6 of GalNAcol. The following assumptions have been made: hexose residues are Gal, *N*-acetylhexosamine residues are GlcNAc, deoxyhexose residues are Fuc and *N*-acetyl hexosaminitol residues are GalNAcol.

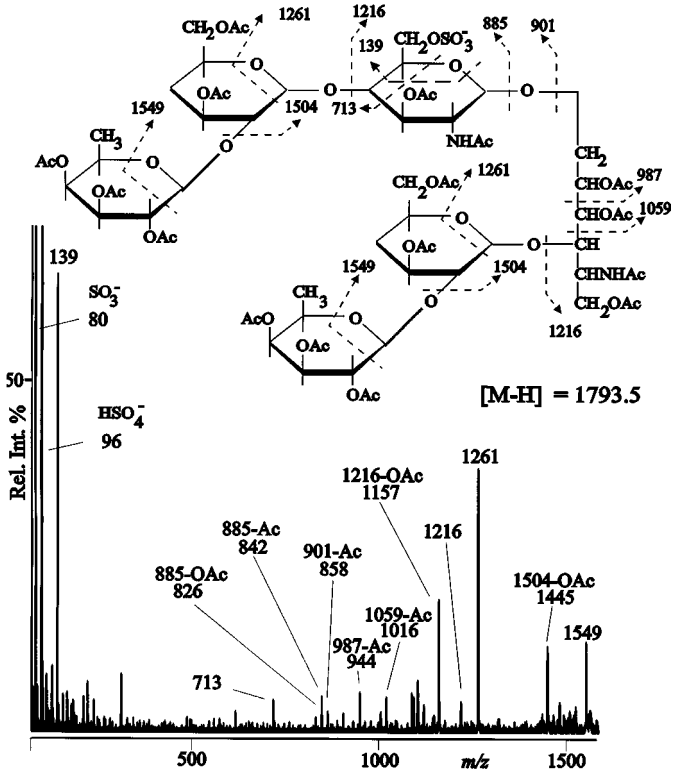

Figure 7. Negative ion FAB-MS/MS showing the CID-spectrum of the precursor ion $[M–H]^- = 1274$ of a peracetylated monosulfated hexasaccharide Fuc-Gal-3**(Fuc-Gal-4(SO_3^--6)GlcNAc-6)**GalNAcol. Fragment ions up to approximately 30 amu below the parent ion were collected.

MUC2 mRNA transcription, thereby suggesting that CF respiratory secretions and their properties could be altered due to appearance of a mucin not usually present in the airways. Infections could also modify the relative abundance of mucin oligosaccharides on mucins already present. An example is the observation of an altered glycosylation of Muc2 in the small intestine in rats infected with the parasite *Nippostrongylus brasiliensis* (Karlsson *et al*., unpublished results). Multiple alterations in CF mucins are thus likely to occur and more detailed studies addressing the glycosylation of individual mucins in normal individuals and CF patients infected with different pathogens are needed.

Acknowledgments

This work was supported by the Swedish Medical Research Council (7461, 7902 and 10446); the Medical Faculty, University of Lund, Sweden; Ingabritt and Arne Lundbergs Stiftelse; Greta och Johan Kocks Stiftelser, Trelleborg, Sweden, Alfred Österlunds Stiftelse and the Swedish Cystic Fibrosis Foundation. The mass spectrometers were obtained from FRN and the Knut and Alice Wallenberg foundation and supported by the Swedish Medical Research Council (3967).

References

1 Rose MC (1992) *Am J Physiol* **263**: L413–L429.
2 Davis PB, Drumm M, Konstan MW (1996) *Am J Resp Crit Care Med* **154**: 1229–56.
3 Carnoy C, Ramphal R, Scharfman A, Lo-Guidice JM, Houdret N, Klein A, Galabert C, Lamblin G, Roussel P (1993) *Am J Resp Cell Mol Biol* **9**: 323–34.
4 Cheng PW, Boat TF, Cranfill K, Yankaskas JR, Boucher RC (1989) *J Clin Invest* **84**: 68–72.
5 Zhang YL, Doranz B, Yankaskas JR, Engelhardt JF (1995) *J Clin Invest* **96**: 2997–3004.
6 Thornton DJ, Sheehan JK, Lindgren H, Carlstedt I (1991) *Biochem J* **276**: 667–75.
7 Thornton DJ, Davies JR, Kraayenbrink M, Richardson PS, Sheehan JK, Carlstedt I (1990) *Biochem J* **265**: 179–86.
8 Thornton DJ, Sheehan JK, Carlstedt I (1991) *Biochem J* **276**: 677–82.
9 Van Halbeek H, Darland L, Vliegenhart JFG, Hull WE, Lamblin G, Lhermitte M, Boersma A, Roussel P (1982) *Eur J Biochem* **127**: 7–20.
10 Breg J, Van Halbeek H, Vliegenhart JFG, Lamblin G, Houvenaghel M-C, Roussel P (1987) *Eur J Biochem* **168**: 57–68.
11 Lo-Guidice J-M, Wieruszeski J-M, Lemoine J, Verbert A, Roussel P, Lamblin G (1994) *J Biol Chem* **269**: 18794–813.
12 Carlstedt I, Herrmann A, Karlsson H, Sheehan J, Fransson L-Å, Hansson GC (1993) *J Biol Chem* **268**: 18771–81
13 Karlsson NG, Karlsson H, Hansson GC (1995) *Glycoconjugate J* **12**: 69–76.
14 Karlsson H, Carlstedt I, Hansson GC (1989) *Anal Biochem* **182**: 438–46.
15 Karlsson NG, Karlsson H, Hansson GC (1996) *J Mass Spectrom* **31**: 560–72.
16 Smith EL, McKibbin JM, Karlsson K-A, Pascher I, Samuelsson BE (1975) *Biochim Biophys Acta* **388**: 171–79.
17 Breg J, Van Halbeek H, Vliegenhart JFG, Klein A, Lamblin G, Roussel P (1988) *Eur J Biochem* **171**: 643–54.
18 Lhermitte M, Rahmoune H, Lamblin G, Roussel P, Strang A-M, Van Halbeek H (1991) *Glycobiology* **1**: 277–93.

19 Klein A, Lamblin G, Lhermitte M, Roussel P, Breg J, Van Halbeek H, Vliegenhart JFG (1988) *Eur J Biochem* **171**: 631–42.
20 Van Halbeek H, Breg J, Vliegenhart JFG, Klein A, Lamblin G, Roussel P (1988) *Eur J Biochem* **177**: 443–60.
21 Mawhinney TP, Landrum DC, Gayer DA, Barbero GJ (1992) *Carbohydr Res* **235**: 179–97.
22 Sangadala S, Bhat UR, Mendicino J (1993) *Mol Cell Biochem* **126**: 37–47.
23 Thornton DJ, Howard M, Khan N, Sheehan JK (1997) *J Biol Chem* **272**: 9561–6.
24 Thornton DJ, Carlstedt I, Howard M, Devine PL, Price MR, Sheehan JK (1996) *Biochem J* **316**: 967–75.
25 Hovenberg HW, Davies JR, Carlstedt I (1996) *Biochem J* **318**: 319–24.
26 Li J-D, Dohrman AF, Gallup M, Miyata S, Gum RJ, Kim YS, Nadel JA, Prince A, Basbaum C (1997) *Proc Natl Acad Sci* **94**: 967–72.

Received 18 November 1997, revised 3 March 1998, accepted 6 March 1998