Glycoconjugate J (1990) 7:525-543

Structures of Acidic O-Linked Polylactosaminoglycans on Human Skim Milk Mucins

FRANZ-GEORG HANISCH[1]*, JASNA PETER-KATALINIC[2], HEINZ EGGE[2], URSULA DABROWSKI[3] and GERHARD UHLENBRUCK[1]

[1] *Institut für Immunbiologie, Universitätskliniken Köln, Kerpener Strasse 15, D-Köln 41, Germany*
[2] *Institut für Physiologische Chemie, Universität Bonn, Nussallee 11, D-5300 Bonn 1, Germany*
[3] *Max-Planck-Institut für Medizinische Forschung, Jahnstrasse 29, D-6900 Heidelberg, Germany*

Received March 9/June 29, 1990.

Key words: O-linked glycans, acidic carbohydrates, mucins, human milk, FAB mass spectrometry, NMR spectroscopy

O-Linked glycans were isolated from human skim milk mucins or mucin-derived high-molecular weight glycopeptides and fractionated by anion exchange chromatography into neutral and acidic alditols. Major oligosaccharides contained in the acidic fraction were purified by high performance liquid chromatography and structurally characterized by a combination of fast atom bombardment mass spectrometry, methylation analysis and 500 MHz ^{1}H-nuclear magnetic resonance spectroscopy. The structural aspects exhibited by these major species in the acidic fraction resemble those established previously for the neutral oligosaccharides from human skim milk mucins:

1) **the size of the alditols varies from tri- to decasaccharides,**
2) **the core structure is of the ubiquitous type 2,**
3) **the backbone sequences are of the poly-*N*-acetyllactosamine type with a particular preponderance of linearly extended GlcNAcβ(1-3)Gal (major) or GlcNAcβ(1-6)Gal units (minor).**

***N*-Acetylneuraminic acid on monosialylated (mucin- or glycopeptide-derived) and disialylated glycans (glycopeptide derived) is linked predominantly to position C-3 of galactose.**

Mucins are high molecular weight glycoproteins, which represent the major components of mucus, the visco-elastic glycocalyx lining the surfaces of glandular epithelial tissues. The clustered carbohydrate chains of mucins (up to 200 per monomeric unit) are linked *O*-glycosidically to the peptide core and exhibit a high degree of structural diversity. According to this particular property mucins as found in human serum and secretions or in epithelial membranes are suitable candidates as tumor diagnostic markers since they may readily indicate differentiation-related variations of the glycosylation status on normal vs. malignant cells. Not surprisingly, since introduction of the hybridoma technique numerous monoclonal antibodies have been generated and shown to define cancer-associated mucin epitopes [1-4]. Hybridoma-defined mucins are regular components of human milk precipitating with the

---

*Author for correspondence.

fraction of caseins [5]. This milk mucin has been termed PAS-O [6] forming an integral membrane component of human milk fat globule membranes (HMFGM), EMA [7], MAM 6 [8] and NPG [9]. Because of the high degree of polymorphism exhibited by the milk mucin at both the protein and DNA levels [10, 11] it is also referred to as PEM (polymorphic epithelial mucin) [12]. Recently, a partial c-DNA of the mucin core peptide has been sequenced leading to the finding that it is built up from repetitive icosapeptides with defined positions of glycosylation [13].

In an attempt to characterize structurally the status of mucin glycosylation on normally differentiated breast epithelial cells during lactation we have previously analyzed the complete primary structures of major neutral O-linked glycans on human skim milk mucins [14]. The carbohydrate structures established belong to the class of poly-*N*-acetyllactosamines comprising up to 16 monosaccharide units and exhibiting a novel type of linearly extended backbones with Galβ(1-4)GlcNAcβ(1-6) repeating units. Here we report the results of structural analyses on the major components contained in the fraction of acidic O-linked glycans from human skim milk mucins.

Materials and Methods

Preparation of Sialyl-oligosaccharide Alditols

Fractionation of human milk and isolation of mucus glycoproteins or Pronase-stable, high molecular weight glycopeptides from mucins contained in the casein fraction were performed according to previously published methods [14]. Sialyl-oligosaccharides were isolated by reductive β-elimination in alkaline borohydride followed by ion exchange chromatography of the released alditols on DEAE-Sephadex A-25 and fractionation on a semi-preparative scale by high performance liquid chromatography on primary amine bonded silica [14]. Carbohydrates were eluted with a gradient of acetonitrile/15 mM KH_2PO_4 in water, pH 4.7, starting from 7/3 up to 1/4 at a flow rate of 1 ml/min and monitored spectrophotometrically at 192 nm. Desalting of HPLC subfractions was performed on 1.3 x 60 cm Sephadex G-25 in water.

Composition and Methylation Analysis

Monosaccharide compositions of carbohydrate samples (10 µg) were analyzed after methanolysis (0.65 N methanolic HCl) for 16 h at 70°C and re-*N*-acetylation according to the method of Chaplin [15]. 1-*O*-Methyl glycosides were trimethylsilylated using a mixture of *N*-methyl-*N*-trimethylsilyltrifluoroacetamide (20 µl) and trifluoroacetic acid (2 µl). The derivatives formed within 10 min at 70°C were analyzed on a fused silica capillary column wall coated with RSL300 heated in a gradient from 100°C to 130°C (16°C/min), followed by a gradient from 130°C to 260°C (4°C/min).

Partially methylated alditol acetates were prepared as described by Stellner *et al.* [16], but hydrolysis was shortened to 6 h. The derivatives were analyzed by gas-liquid chromatography-mass spectrometry on a Hewlett Packard HSD 5970 instrument using a capillary column wall-coated with SE 54 (15 m), which was heated from 100°C to 300°C (10°C/min).

FAB-Mass Spectrometry

FAB-Mass spectrometry in the positive ion mode was performed on a ZAB HF mass spectrometer (VG Analytical, Manchester, UK) using conditions described previously [17]. Briefly, 5 µg of permethylated or peracetylated saccharide alditol solubilized in 1 µl methanol together with 2 µl thioglycerol were applied to the target, which had been preloaded with 1 µl of a 0.1% solution of sodium acetate in methanol. The target was bombarded with xenon atoms having a kinetic energy equivalent to 8.5-9.5 kV.

^{1}H-NMR Spectroscopy

For NMR measurements, ca. 400 µg of MM-A3 or ca. 100 µg of MM-A5 were repeatedly exchanged with ^2H_2O and then dissolved in 0.3 ml of ^2H_2O containing a trace of acetone, which was used as internal reference (δ 2.225 at 307 K).

One dimensional (1 D) spectra were obtained at 500 MHz with a Bruker AH-500 spectrometer. For resolution enhancement. the free induction decays were multiplied by the Lorentzian-Gaussian transformation function. Two-dimensional (2D) scalar shift correlated (COSY) spectra were recorded at 360 MHz with two level $^1H^2HO$ suppression and a relaxation delay of 2 s using a Bruker AM-360 spectrometer. The spectral width was 1300 Hz in both dimensions and the data matrix size in the time domain was 1 K x 256 data points. For each of the 256 t_1 delays, 208 transients were accumulated for MM-A3 and 320 for MM-A5.

The matrix was zero-filled in t_1, multiplied with the sinebell window function in both dimensions, and transformed in the magnitude mode. The resulting frequency-domain spectrum was symmetrized relative to the diagonal.

Results

O-Linked glycans on human skim milk mucins or mucin-derived, high molecular weight glycopeptides were released by reductive β-elimination and fractionated by ion-exchange chromatography on DEAE-Sephadex A-25 followed by high performance liquid chromatography on primary amine bonded phase. Elution profiles on HPLC obtained for the respective fractions of acidic oligosaccharides were very similar with regard to the quantitative distribution of the individual components (Fig. 1). The major alditol species contained in the mucin-derived fractions MM-A2, MM-A3, MM-A4, MM-A5, MM-A6 (monosialylated alditols) and those in the glycopeptide-derived fractions MM-A7, MM-A8, MM-A9, MM-A10 and MM-A11 (disialylated alditols) were subjected to structural analyses comprising positive ion FAB-mass spectrometry, methylation analysis and ^{1}H-NMR spectroscopy at 500 MHz.

Fraction MM-A2

The positive ion FAB-mass spectrum registered for the methylated compounds in this fraction revealed pseudo-molecular ions M+H and M+Na at m/z 1118 and 1140, respectively (Table 1). These values in combination with the carbohydrate composition (Table 1)

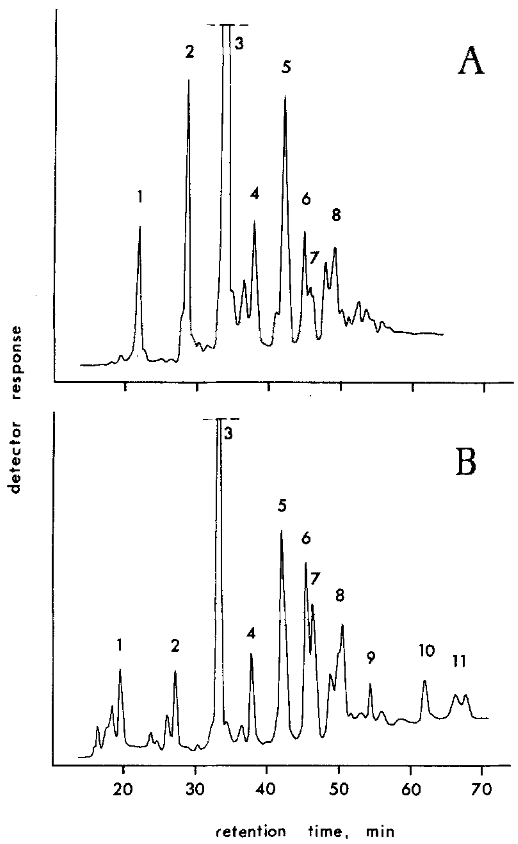


Figure 1. High performance liquid chromatography of a) mucin- or b) glycopeptide-derived oligosaccharide alditols. Sialylated alditols (1 mg; fraction MM-A) were applied to a column of primary amine bonded silica (0.46 x 25 cm). Carbohydrates were eluted at a flow rate of 1 ml/min with acetonitrile/15 mM KH_2PO_4 using a gradient of 7/3 up to 1/4, by vol. Eluting carbohydrates were registered spectrophotometrically at 192 nm. Peak numbers 1-11 refer to fractions MM-A1 to MM-A11.

represent a tetrasaccharide composed of *N*-acetylneuraminic acid (1), galactose (1), *N*-acetylglucosamine (1) and *N*-acetylgalactosaminitol (1). Sequence-dependent primary fragments in FAB-mass spectrometry were observed at m/z 260 ($HexNAc^+$), 376 ($NeuAc^+$), 580 ($NeuAc\text{-}Hex^+$) and 521 ($HexNAc\text{-}HexNAc\text{-}ol^+$) suggesting the presence of two non-reducing termini and a branched core structure. The pattern of substitution was established by GC-MS analysis of partially methylated alditol acetates (Table 2) revealing nonsubstituted *N*-acetylglucosamine, 3-monosubstituted galactose and 3,6-disubstituted *N*-acetylgalactosaminitol as the major derivatives. The data can be summarized in the following structure model for MM-A2 oligosaccharide alditol:

GlcNAc1
6
GalNAc-ol
3
NeuAc2-3Gal1

Table 1. Pseudomolecular ions and monosaccharide compositions of acidic glycans from human skim milk mucins.

HPLC-fraction	Pseudomolecular ions M+H or (M+Na) in FAB-MS		Monosaccharide composition based on pseudomolecular ions (FAB-MS) and GC analysis of trimethylsilylated 1-*O*-methyl glycosides				
			NeuAc	Fuc	Gal	GlcNAc	GalNAc-ol
MM-A2	1118	(1140)	1	0	1	1	1
MM-A3	1322	(1344)	1	0	2	1	1
MM-A4	1496	(1518)	1	1	2	1	1
MM-A5	1771	(1793)	1	0	3	2	1
MM-A6	1945	(1967)	1	1	3	2	1
MM-A7	1234	(1256)	2	0	1	0	1
MM-A8	1683	(1705)	2	0	2	1	1
	2119	(2141)	1	2	3	2	1
	2220	(2242)	1	0	4	3	1
MM-A9	2293	(2315)	1	3	3	2	1
MM-A10	2132	(2154)	2	0	3	2	1
MM-A11	2306	(2328)	2	1	3	2	1
	2581	(2603)	2	0	4	3	1

A minor admixture of a molecule containing a Gal-GlcNAc disaccharide fragment with m/z 464 and being composed of *N*-acetylneuraminic acid (1), galactose (2), *N*-acetylglucosamine (1) and *N*-acetylgalactosaminitol (1) (pseudo-molecular ions M+H and M+Na at 1322 and 1344, respectively) became evident on inspection of the FAB-mass spectrum.

Fraction MM-A3

According to a composition analysis after acidic methanolysis (Table 1) and the pseudomolecular ions M+H or M+Na registered in FAB-MS of the permethylated compound at m/z 1322 or 1344 (Fig. 2), respectively, the pentasaccharide alditol is composed of *N*-acetylneuraminic acid (1), galactose (2), *N*-acetylglucosamine (1) and *N*-acetylgalactosaminitol (1) (Table 1). A sequence analysis based on primary FAB-fragments of the methylated alditol revealed signals at m/z 376 (NeuAc$^+$), 464 (Hex-HexNAc$^+$), 580 (NeuAc-Hex$^+$) and 725 (Hex-HexNAc-HexNAc-ol$^+$). Since signals at m/z 432 or 228, which represent resonance stabilized secondary fragments originating from elimination of the C-3 substituent in Hex1-3/4-HexNAc$^+$, are indicative for the substitution of the HexNAc residue, the mass spectrum supports a Hex1-4HexNAc sequence rather than the isomeric sequence. This assumption is further confirmed by results of a methylation analysis demonstrating the exclusive presence of the partially methylated HexNAc derivative 1,4,5-tri-*O*-acetyl-*N*-acetylglucosaminitol (Table 2). A branched core structure is revealed by the

Table 2. Methylation analysis of oligosaccharide alditols[a].

Partially methylated alditol	MM-A2	MM-A3	MM-A4	MM-A5	MM-A6	MM-A7	MM-A8	MM-A9	MM-A10	MM-A11
Fucitol										
2,3,4-Tri-*O*-methyl	-	-	+	-	(+)	-	+	+	-	+
Galactitol										
2,3,4,6-Tetra-*O*-methyl	-	+	+	+	+	-	+	-	-	+
2,4,6-Tri-*O*-methyl	+	+	+	+	+	+	+	+	+	+
2,3,4-Tri-*O*-methyl	-	-	-	-	(+)	-	-	-	-	-
3,4,6-Tri-*O*-methyl	-	-	-	-	-	-	-	(+)		
2,4-Di-*O*-methyl	-	-	-	-	-	-	(+)	-	-	+
2-N-Methylacetamido-2-deoxygalactitol										
1,4,5-Tri-*O*-methyl	+	+	+	+	+	+	+	+	+	+
2-N-Methylacetamido-2-deoxyglucitol										
3,4,6-Tri-*O*-methyl	+	-	-	-	-	-	-	-	-	-
3,6-Di-*O*-methyl	(+)	+	-	+	+	-	+	-	+	+
4,6-Di-*O*-methyl	-	-	-	(+)	(+)	-	+	-	-	(+)
6-Mono-*O*-methyl	-	-	+	-	(+)	-	+	+	-	+

[a] Qualitative data presented in brackets refer to partially methylated alditol acetates with a relative abundance of <1 using 1,4,5-tri-*O*-methyl-2-*N*-methylacetamido-2-deoxygalactitol as a reference.

partially methylated HexNAc-ol 3,6-di-*O*-acetyl-*N*-acetylgalactosaminitol. In summary, the data obtained for fraction MM-A3 (refer also to the ^{1}H-NMR data presented below) support the following structure:

Galβ1-4GlcNAcβ1
6
GalNAc-ol
3
NeuAcα2-3Galβ1

Fraction MM-A4

Fraction MM-A4 contains a hexasaccharide alditol with the composition *N*-acetylneuraminic acid (1), fucose (1), galactose (2), *N*-acetylglucosamine (1) and *N*-acetylgalactosaminitol (1) as judged from mass increment calculation (pseudomolecular ions M+H or M+Na at m/z 1496 or 1518, respectively) in conjunction with composition analysis (Table 1). The methylated alditol in FAB-MS gave rise to a set of primary fragments and their presumed daughter ions characterizing the two branches NeuAc-Hex (represented at m/z 376, 580) and Hex-(dHex-)HexNAc (represented at m/z 638). The intense daughter ion at m/z 432,

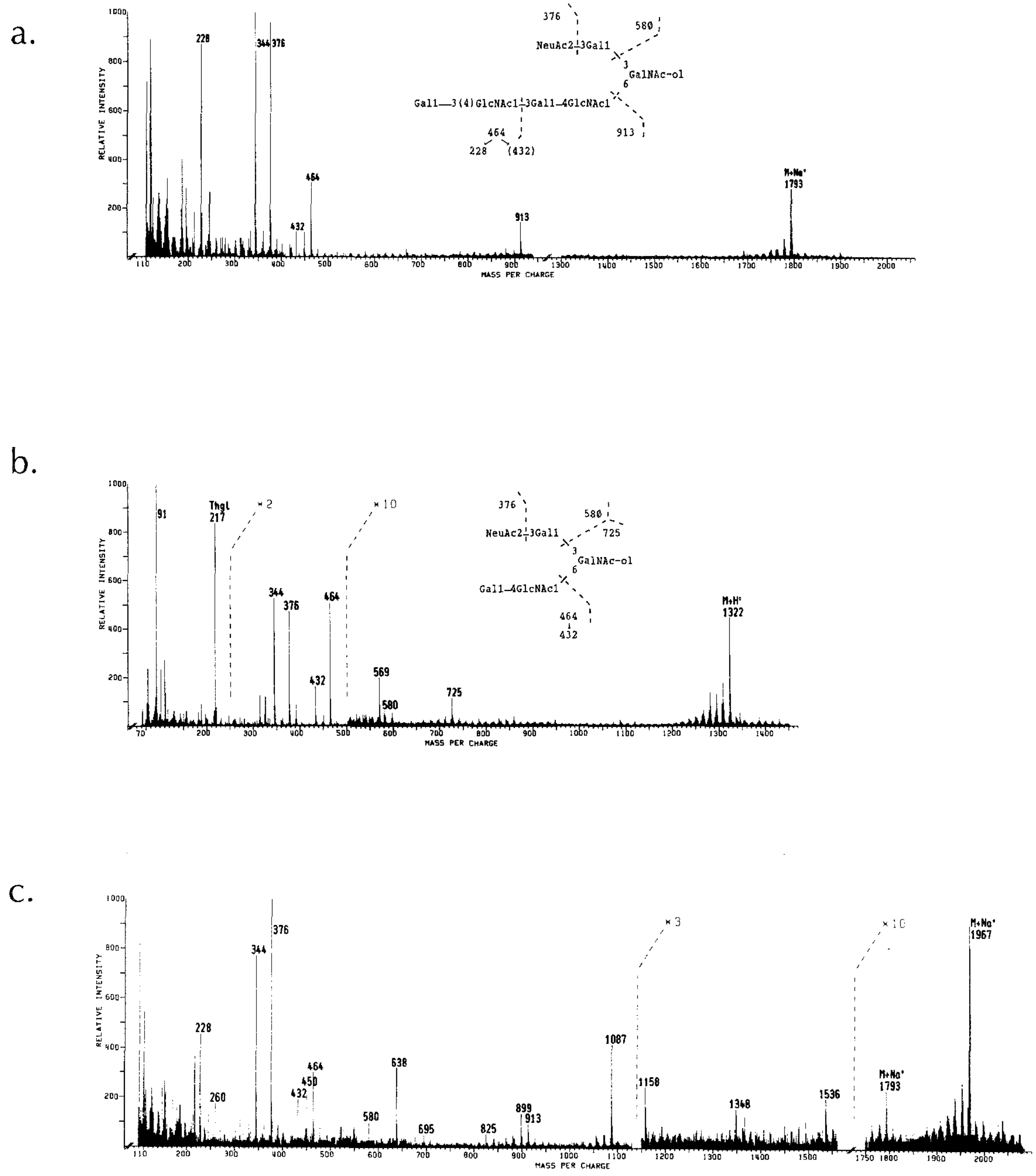


Figure 2. Positive ion FAB-mass spectra of permethylated saccharide alditols.
a, fraction MM-A3; b, fraction MM-A5; c, fraction MM-A6. For details refer to the Materials and Methods section. Structural assignments of fragment ions are given in the Figures.

which is presumably formed by elimination of dHex from the primary fragment at m/z 638, strongly suggests that fucose is linked to position C-3 of the subterminal *N*-acetylglucosamine. In conjunction with results of a methylation analysis (Table 2) the following structure model for the alditol in fraction MM-A4 may be postulated:

Gal1-4GlcNAc1
3 6
Fuc1 GalNAc-ol
3
NeuAc2-3Gal1

Fraction MM-A5

The permethylated alditol in fraction MM-A5 is represented by the pseudomolecular ions M+H or M+Na at m/z 1771 or 1793 in FAB-MS and, accordingly, is calculated to be composed of *N*-acetylneuraminic acid (1), galactose (3), *N*-acetylglucosamine (2) and *N*-acetylgalactosaminitol (1) (Table 1). Two branches with the proposed sequences NeuAc-Hex and Hex-HexNAc-Hex-HexNAc can be assumed on the basis of ion series in FAB-MS at m/z 376 (NeuAc$^+$), 580 (NeuAc-Hex$^+$) and m/z 464 (Hex-HexNAc$^+$), 913 (Hex$_2$-HexNAc$_2$$^+$) (Fig. 2). The substitution of the internal galactosyl residue within the sequence Gal-GlcNAc-Gal-GlcNAc was established by two independent lines of evidence: GC-MS-identification of partially methylated 1,3,5 tri-*O*-acetyl galactitol (Table 2) as the only derivative of mono-substituted galactose and chemical shifts registered for the C3 proton (H3) in 500 MHz ^{1}H-NMR (see below).

The combined data of composition analyses, FAB-MS, methylation analysis and ^{1}H-NMR are summarized in the proposed structure:

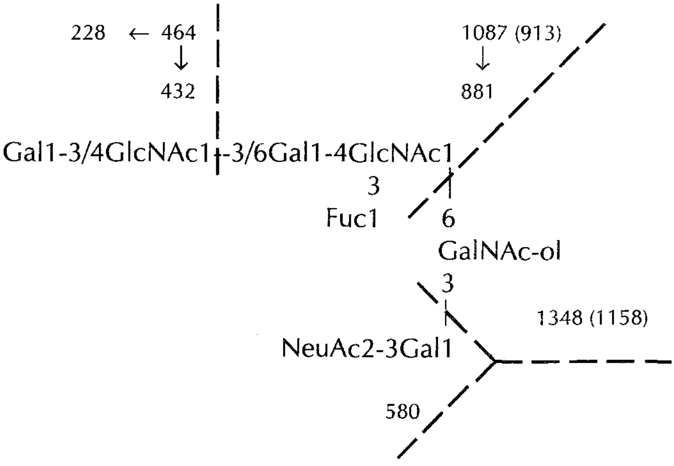
Galβ1-3/4GlcNAcβ1-3Galβ1-4GlcNAcβ1
6
GalNAc-ol
3
NeuAcα2-3Galβ1

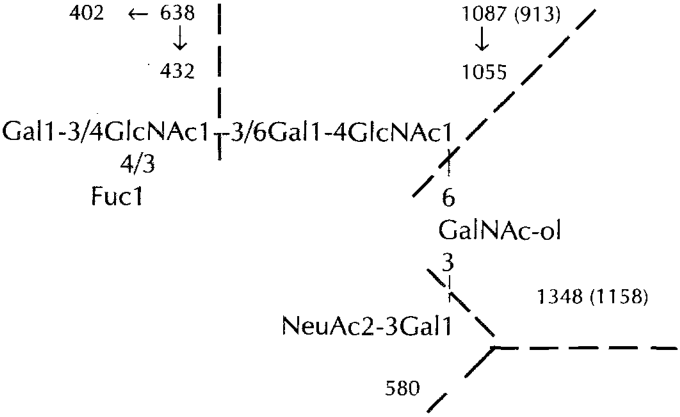
Fraction MM-A6

Pseudomolecular ions M+H or M+Na in FAB-MS at m/z 1945 or 1967, respectively, indicate that the major alditol in fraction MM-A6 is an octasaccharide with the composition *N*-acetylneuraminic acid (1), fucose (1), galactose (3), *N*-acetylglucosamine (2) and *N*-acetylgalactosaminitol (1) (Table 1). A minor heptasaccharide NeuAc-Hex$_3$-HexNAc$_3$-ol represented by its M+Na at m/z 1793 may have originated from the major component by cleavage of fucose during derivatization and work-up, or could represent a linkage isomer of the alditol in fraction MM-A5.

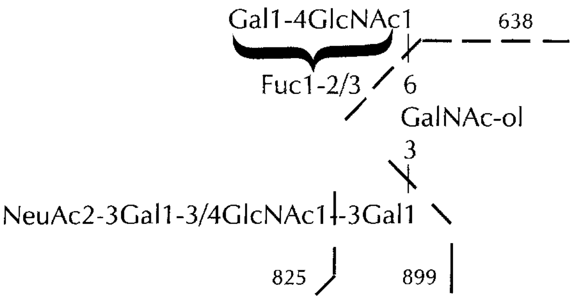
Sequence-dependent ions registered at m/z 638 (dHex-Hex-HexNAc$^+$) or 1087 (dHex-Hex$_2$-HexNAc$_2^+$) and presence of the ion at m/z 464 (Hex-HexNAc$^+$) support the assumption that fraction MM-A6 is composed of two isomeric oligosaccharides with alternative positions of fucosylation at the subterminal or internal *N*-acetylglucosamine residues (Fig. 2). Elimination of deoxyhexose from dHex-Hex$_2$-HexNAc$_2^+$ with formation of the corresponding resonance stabilized secondary ion at m/z 881 (not very pronounced in the spectrum presented) is indicative for a C-3 rather than C-4-linkage of dHex to the internal *N*-acetylglucosamine residue, while fucose of the fragment dHex-Hex-HexNAc$^+$ presumably is linked to both C-3 (major) or C-4 (minor) of the subterminal *N*-acetylglucosamine residue as indicated by secondary ions at m/z 432 or 402.

A methylation analysis (Table 2) confirmed the presence of terminal fucose, C-3- or C-6-monosubstituted galactose, C-3- or C-4 monosubstituted and C-3-, C-4-disubstituted *N*-acetylglucosamine and C-3-, C-6-disubstituted *N*-acetylgalactosamine. Results from FAB-MS and methylation analysis are summarized in the tentative structure models of two isomeric octasaccharides and a minor defucosylated heptasaccharide (primary fragment ions in brackets):





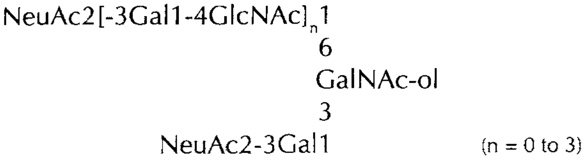
The existence of an alternative sequence can be predicted on the basis of sequence ions at m/z 825, 899 and 638 which fit with the structure model shown below:



Minor fragment ions in FAB-MS at m/z 260 (HexNAc$^+$) and 1536 (dHex-Hex$_3$-HexNAc$_3$$^+$) point to further admixtures of isomeric oligosaccharide alditols which are presumably not represented by the above formulated structure models, and should at least in part correspond to molecular species with different monomeric compositions.

Fractions MM-A7 to MM-A11

Alditols chromatographing in fractions MM-A7 to MM-A11 (Fig. 1b) are more prominent in the glycopeptide-derived fractions and mainly represent disialylated or polyfucosylated species (Table 1). The major components contained in fractions A7, A8, A10 and A11 are identical with a homologous series of disialylated di- to octasaccharides exhibiting 6-GalNAc-ol branches of the poly-*N*-acetyllactosamine-type with n units (n = 0 to 3). Positive ion FAB-mass spectrometry of the permethylated alditols revealed a homologous series of pseudomolecular ions M+H (or M+Na) at m/z 1234 (1256), 1683 (1705), 2132 (2154) and 2581 (2603) (Table 1) which differ by the mass increment m/z 449 corresponding to a Hex-HexNAc unit. Primary fragment ions arising from cleavage of sialylated polylactosamine sequences with increasing chain lengths are registered at m/z 825 (NeuAc-Hex-HexNAc$^+$), 1274 (NeuAc-Hex$_2$-HexNAc$_2$$^+$) or 1723 (NeuAc-Hex$_3$-HexNAc$_3$$^+$) (Table 3). The branching pattern of the oligosaccharide alditols could be established by GC-MS identification of the partially methylated alditol acetates (Table 2). All fractions (A7 to A11) yield the 3,6-di-*O*-acetylated *N*-acetylgalactosaminitol derivative indicating a branched core structure, the 1,3,5-tri-*O*-acetylated derivative of galactose and (with the exception of fraction A7) the 1,4,5-tri-*O*-acetyl derivative of *N*-acetylglucosamine indicating a preponderance of linear -3Gal1-4GlcNAc sequences. The structural aspects of the major disialylated alditols in fractions A7 to A11 may tentatively be summarized in the general formula:



$$\begin{array}{ll} \text{NeuAc2[-3Gal1-4GlcNAc]}_n\text{1} & \\ \quad\text{6} & \\ \quad\text{GalNAc-ol} & \\ \quad\text{3} & \\ \text{NeuAc2-3Gal1} & \text{(n = 0 to 3)} \end{array}$$

Fraction A7: n = 0 A9: n = 2
A8: n = 1 A11: n = 3

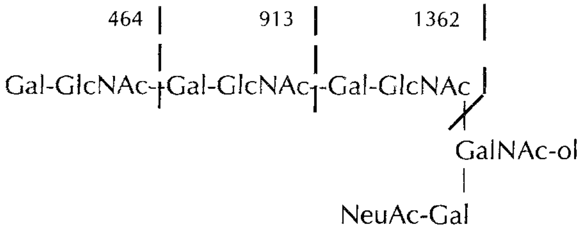
Table 3. Primary fragment ions registered in positive ion FAB-mass spectrometry of permethylated alditols.

Primary fragment ion at m/z	Corresponding fragment compositions	Detected in HPLC subfractions MM-A
260	HexNAc$^+$	2,6
376	NeuAc$^+$	2,3,4,5,6,7,8,9,10,11
464	Hex-HexNAc$^+$	2,3,5,6,8,11
521	HexNAc-HexNAc$^+$	2
580	NeuAc-Hex$^+$	2,3,4,5,6,7,8,9,10,11
638	dHex-Hex-HexNAc$^+$	4,6,8,11
725	Hex-HexNAc-HexNAc-ol$^+$	3
812	dHex$_2$-Hex-HexNAc$^+$	8,9
825	NeuAc-Hex-HexNAc$^+$	6,8,10,11
913	Hex$_2$-HexNAc$_2^+$	5,6,8,11
1087	dHex-Hex$_2$-HexNAc$_2^+$	6,8,11
1274	NeuAc-Hex$_2$-HexNAc$_2^+$	11
1360	NeuAc$_2$-dHex-Hex-HexNAc$^+$	11
1362	Hex$_3$-HexNAc$_3^+$	8
1435	dHex$_3$-Hex-HexNAc$^+$	9
1448	NeuAc-dHex-Hex$_2$-HexNAc$_2^+$	11
1536	dHex-Hex$_3$-HexNAc$_3^+$	6
1723	NeuAc-Hex$_3$-HexNAc$_3^+$	11

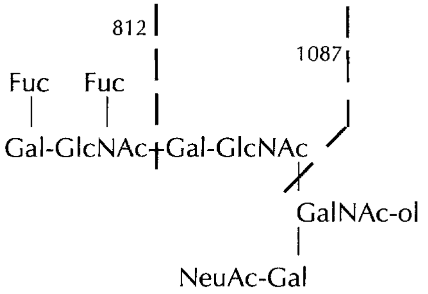
Small amounts of the 1,3,5,6-tetra-*O*-acetylated galactose derivative in fraction A8 and a major portion of this alditol derivative in fraction A11 indicate that a subfraction of these polylactosamine-type alditols may exhibit branched backbone sequences (Table 2).

Fractions MM-A8 and MM-A9 contain monosialylated alditols some of which exhibit further substitution by fucose (Table 1). Due to the heterogeneity of these fractions structural assignments with regard to linkage position could not unambiguously be made and, accordingly, the tentative structure models presented below are based merely on sequence data from FAB-MS and compositional data.

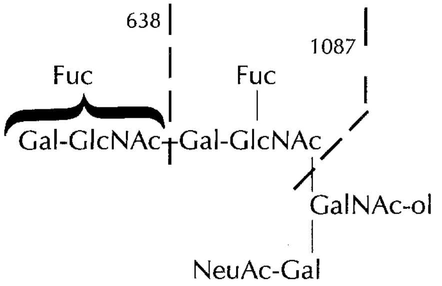
A monosialylated alditol identified in fraction MM-A8 (corresponding to the pseudomolecular ion M+H at m/z 2220 in Table 1) exhibits chain elongation of a polylactosamine sequence as indicated by the ion series at m/z 464, 913 and 1362. Composition and sequence data fit with the formula:



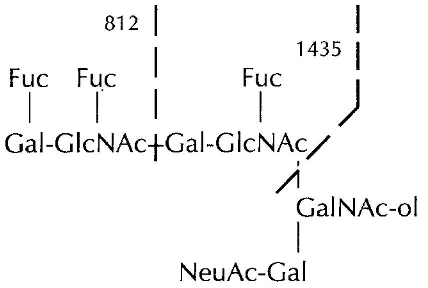
Fucosylated species contained in fractions MM-A8 and MM-A9 are presumably derived from the basic structure in fraction MM-A5 by further substitution with two or three fucosyl residues in various positions:



MM-A8



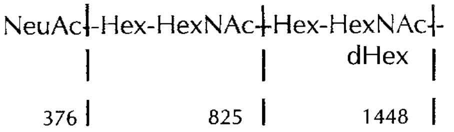
MM-A8



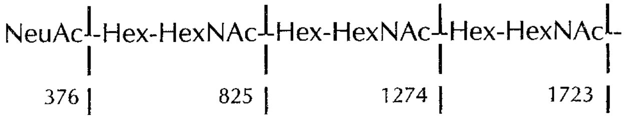
MM-A9

Alditols in fraction MM-A11 are of particular interest with regard to a structural heterogeneity of carbohydrate sequences at the non-reducing termini. Although only two molecular species could be identified in FAB-MS by their respective pseudomolecular ions M+H at m/z 2306 and 2581 (corresponding to the compositions NeuAc$_2$-dHex-Hex$_3$-HexNAc$_2$-HexNAc-ol and NeuAc$_2$-Hex$_4$-HexNAc$_3$-HexNAc-ol) as many as 11 primary sequence ions could be observed representing a series of alternative terminal structures:

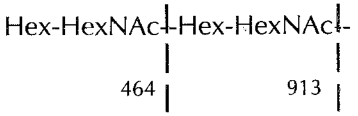
(1)



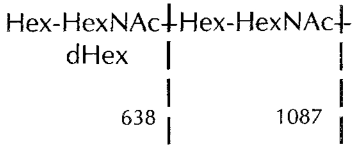
(2)



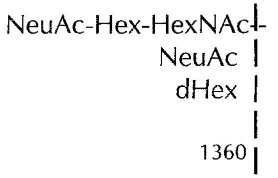
(3)



(4)



(5)



^{1}H-NMR Spectroscopy

Fractions MM-A3 and -A5 were analyzed by 1D and 2D ^{1}H-NMR spectroscopy, and the results are presented in Table 4 and in Fig. 3. The GalNAc-ol-I, GlcNAc-II and Gal-III resonances for MM-A3 are practically identical with those found in the previous work [14] for the neutral MM-N4 component isolated from the same source. In contrast, all resonances for the side-chain Gal-VI residue, in particular the H3 resonance, are low field shifted (δH3 = 0.44 ppm), as expected for a 3-sialylated galactose [18-20]. The same refers to residues I, II and VI of the MM-A5 component, whereas the chemical shifts for its now glycosylated Gal-III residue differ strongly. The glycosylation site at its C-3 can readily be deduced with the aid of the previously established rules [19, 21], which indicate that glycosylation generates marked low field shifts for the aglycone protons at the substitution site and at both vicinal sites (sometimes even stronger for the latter), but leaves other proton resonances virtually unaffected: thus, the middle of the three strongly perturbed protons, here obviously H3 (δH2 = 0.059 ppm, δH3 = 0.069 ppm and δH4 = 0.220 ppm), indicates the site of the glycosidic linkage. This inference is convincingly confirmed by the identical chemical shifts, within experimental error, obtained by Strecker *et al.* [22] for galactose H2-H4 protons in a segment [GlcNAcβ1-3Galβ-] of lacto-*N*-tetraose. The IV1-3III linkage is also in full accord with the observation that chemical shifts for anomeric protons located at glycosidic bridges to ring carbons (C-2, C-3, C-4) occur towards low field, if compared with the situation found for glycosidic bonds to C-6 hydroxymethyl groups [21, 23]. Thus, the 4.742 ppm GlcNAc-IV H1 chemical shift additionally confirms the different substitution patterns with respect to the corresponding (1-6)-linked segments of the neutral MM-N6 species [14] for which 4.697 ppm was found. Again, our GlcNAc-IV H1 and H2 chemical shifts are practically the same

Table 4. ^{1}H Chemical shifts for the oligosaccharide alditols MM-A3 and -A5, as measured at 307 K in ^2H_2O relative to acetone set at 2.225 ppm.

		V	IV	III	II	VII	VI	I
Substance	Proton	Galβ1-3/4[a]	GlcNAcβ1-3	Galβ1-4	GlcNAcβ1—	NeuAcα2—	3Galβ1-3)	—6GalNAc-ol
MM-A3	H1			4.468	4.564		4.530	3.742;3.811
	H2			3.538	3.747		3.602	4.385
	H3			3.667		1.800;2.778	4.108	4.064
	H4			3.923		3.695	3.932	3.446
	H5							4.263
	$H6_A$							3.692
	$H6_B$							3.916
MM-A5	H1	4.437	4.742	4.460	4.562		4.530	3.770;3.800
	H2	3.527	3.895	3.597	3.740		3.602	4.379
	H3	3.601		3.736		1.800;2.776	4.107	4.064
	H4	3.885		4.143		3.698	3.931	3.442
	H5							4.260
	$H6_A$							3.693
	$H6_B$							3.913

[a] The dominant component is (1-3)-linked; the Gal-V H1 signal for a small admixture of the (1-4)-linked component is seen in the 1D spectrum, but no cross peaks are observed in the 2D spectrum.

as found by Strecker *et al.* [22]. The H2 resonance deserves special attention since its low field shift with respect to the H2 resonance for the 4-glycosylated GlcNAc-II residue (3.895 vs 3.740 ppm) indicates that the former H2 is vicinal to the glycosylation site, i.e. that GlcNAc-IV is 3-glycosylated. Finally, the terminal position of the Gal-V residue is evidenced by the high field position of all of its assigned signals, as compared with those for the substituted internal Gal-III and -VI residues. Here too, the agreement with the galactose chemical shifts obtained for the terminal Galβ1-3GlcNAcβ- segment of lacto-*N*-tetraose [22] is excellent.

Discussion

Biochemical and immunological evidence suggests that human mammary epithelial cells secrete and express on their plasma membranes a polymorphic family of sialomucins. Expression of the mucin linked carbohydrates is developmentally regulated and, accordingly, their molecular structures vary with differentiation stages in the normal gland and in malignancy. Since mucins in human milk serum and on human milk fat globule membranes are derived from the normal lactating epithelium of the mammary gland, they may be regarded as molecular models representing the glycosylation status of normally-differentiated

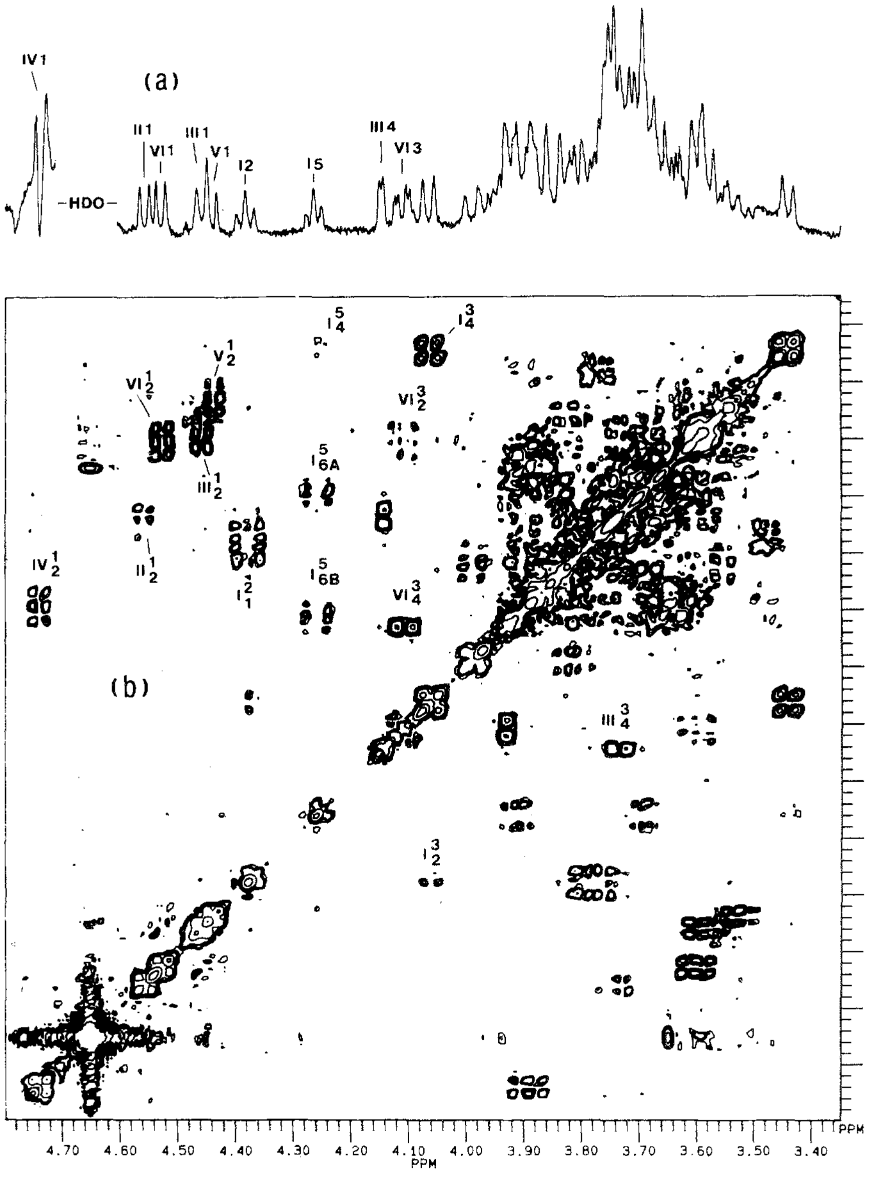
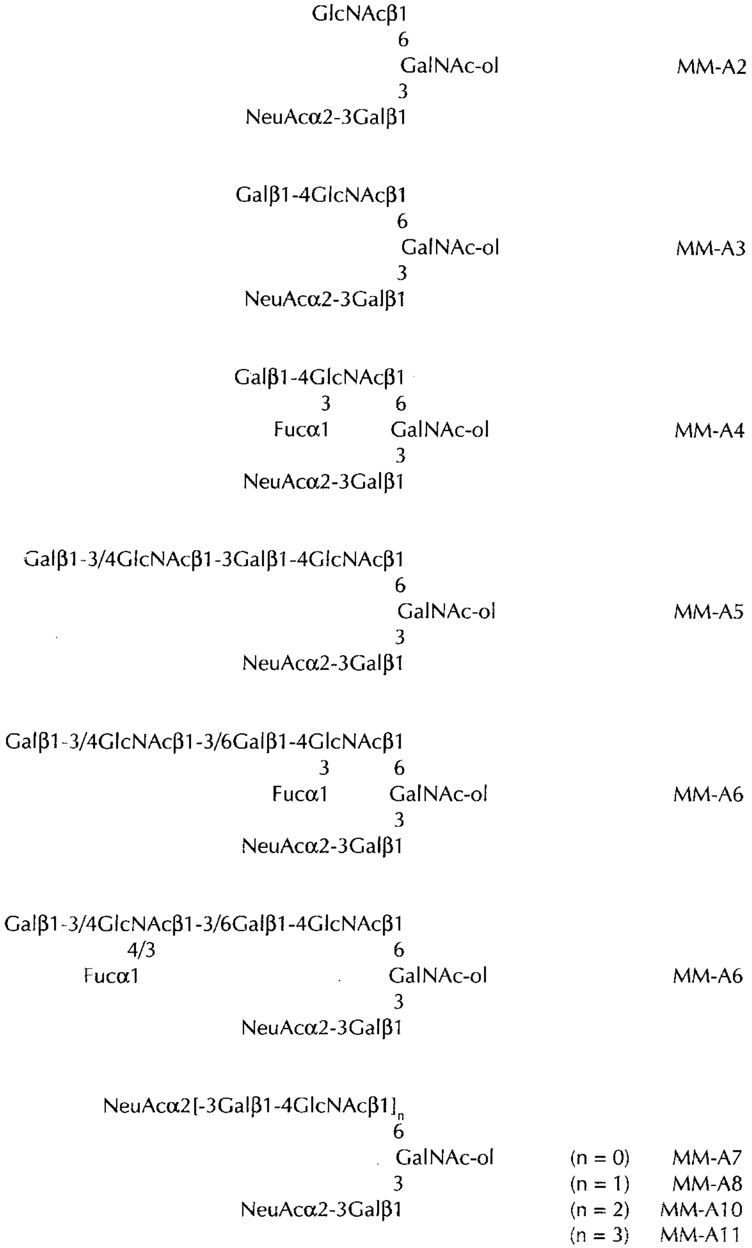


Figure 3. Region of characteristic sugar ring proton resonances of the ^{1}H-NMR spectra of MM-A5 (for structure see Table 4).

a) 500 MHz one dimensional spectrum. Arabic numerals refer to the protons in the sugar residue denoted by roman numerals (denotation of the sugar units as in Table 4),

b) COSY spectrum at 360 MHz. The correlations should be read as in the following example: a contour labelled IV$^1_2$ is located at the chemical shift of the IV-1 proton (horizontal scale) and the chemical shift of the IV-2 proton (vertical scale).



GlcNAcβ1
6
GalNAc-ol MM-A2
3
NeuAcα2-3Galβ1

Galβ1-4GlcNAcβ1
6
GalNAc-ol MM-A3
3
NeuAcα2-3Galβ1

Galβ1-4GlcNAcβ1
3 6
Fucα1 GalNAc-ol MM-A4
3
NeuAcα2-3Galβ1

Galβ1-3/4GlcNAcβ1-3Galβ1-4GlcNAcβ1
6
GalNAc-ol MM-A5
3
NeuAcα2-3Galβ1

Galβ1-3/4GlcNAcβ1-3/6Galβ1-4GlcNAcβ1
3 6
Fucα1 GalNAc-ol MM-A6
3
NeuAcα2-3Galβ1

Galβ1-3/4GlcNAcβ1-3/6Galβ1-4GlcNAcβ1
4/3 6
Fucα1 GalNAc-ol MM-A6
3
NeuAcα2-3Galβ1

NeuAcα2[-3Galβ1-4GlcNAcβ1]$_n$
6
GalNAc-ol (n = 0) MM-A7
3 (n = 1) MM-A8
NeuAcα2-3Galβ1 (n = 2) MM-A10
(n = 3) MM-A11

Figure 4. Proposed primary structures of the major sialyl-oligosaccharide alditols from human skim milk mucins.

epithelial cells in this organ. Accordingly, we have investigated in a series of studies the primary structures of neutral [14] and acidic glycans (this contribution) derived from skim milk mucins or from the membrane mucin on milk fat globules (unpublished results).

The results revealed an extraordinary structural heterogeneity, particularly within the fraction of neutral glycans. Neutral O-linked glycans are of the polylactosamine-type, comprise up to 16 monosaccharide units and exhibit unique structural features with regard to their backbone sequences [14]. While a significant portion of these backbone sequences is branched at internal galactose residues, the major species are characterized by linear polylactosamine chains containing the unusual structural element GlcNAcβ1-6Gal rather than GlcNAcβ1-3Gal. The occurrence of this type of linkage conflicts with a largely accepted biosynthetic rule, the "3 before 6 rule" [24], which states that position C-3 of Gal/GalNAc must be substituted prior to addition of *N*-acetylglucosamine in position C-6 of the target sugar. On the other hand, structural evidence from other investigators [25, 26] and the existence of a β6-*N*-acetylglucosaminyltransferase in Novikoff ascites tumor cells [27] support the assumption that human tissues may contain an enzyme variant which is independent from C-3-substitution of its target sugar.

The structural features of sialic acid containing glycans may he summarized as follows (see also Fig. 4): The core structures are of the general type 2 exhibiting C-6 branches of the polylactosamine type with GlcNAcβ1-3Gal as major linkage unit. The unusual structural element GlcNAcβ1-6Gal has also been found in a minor fraction of acidic glycans (MM-A6). Sialic acid (*N*-acetylneuraminic acid) is exclusively linked to position C-3 of galactose (with the exception of the glycan in MM-A7). Mono-sialylated glycans predominate in the fraction of mucin-derived alditols, while glycopeptide-derived glycans contain also significant portions of disialylated species.

Sialylated O-linked glycans on human skim milk mucins structurally resemble the O-glycosidically linked carbohydrate chains on human secretory IgA [28]. As revealed by immunochemical analyses with anti-human sIgA, no cross-reactivity could be found in the mucin fraction (unpublished results).

According to structural analyses of mucin-glycans on human mammary carcinoma cell lines BT-20 [29] and MDA-MB231 (unpublished results) and to epitope analyses of hybridoma antibodies generated against a mucin-type antigen (TAG 72) on metastasized mammary carcinoma cells [30, 31], the glycosylation of the epithelial mucins changes dramatically in several ways.

In summary, the overall glycosylation of the carcinoma-associated mucins seems to be reduced, as is the average chain length of the O-linked carbohydrates. Short-chain sequences of the general structure NeuAcα2-6[R-3]GalNAc (R: H or galactose) as prominent immunogenic glycan species on TAG72 [30, 31] point to incomplete glycosylation with concomitant accumulation of biosynthetic dead-end products. The latter may be due to enhanced activities of an α6-sialyltransferase which adds *N*-acetylneuraminic acid to the C-6 position of *N*-acetylgalactosamine, thus introducing a biosynthetic stop signal in early stages of the glycan processing [24]. This enzyme species is probably less prominent in normally differentiated epithelial tissues of the mammary gland which favours elongation of the glycan chains to the fully processed structures exhibited by skim milk mucins.

Also a reduced level of β-glucosaminyltransferase activities could explain the accumulation of precursor substrates. Particularly, the core-disaccharide Galβ1-3GalNAc, which forms the immunodominant group of the TF-antigen, is only weakly expressed in an immunologically accessible form on normal, lactating epithelia of the mammary gland [32], but becomes a prominent marker on carcinoma tissue of the breast during tumor progression [33, 34].

Further investigations are needed to characterize in detail the structural changes associated with malignant transformation of mammary epithelial cells on a chemical basis.

Acknowledgements

This investigation was supported by Deutsche Forschungsgemeinschaft grant Uh 8/14-2 and Uh 8/14-3 and the Fritz Thyssen Stiftung (to U.D.). We are obliged to Prof. Dr. A. Bolte, University Clinic of Gynecology and Obstetrics, Cologne, for providing the samples of human milk, Prof. Dr. M. Donike and Dr. W. Schänzer for kind support of our GC-MS analyses, and Mrs. C. Böttinger for her skilful technical assistance.

References

1 Hilkens J, Buijs F, Hilgers J, Hagemann, Ph, Calafat J, Sonnenberg A, Van der Valk N (1984) Int J Cancer 34:179-206.
2 Kufe D, Inghirami G, Abe M, Hayes D, Justi-Wheeler H, Schlom J (1984) Hybridoma 3:223-32.
3 Johnson VG, Schlom J, Paterson AJ, Bennett J, Magnani JL, Colcher D (1986) Cancer Res 46:850-57.
4 Stähli C, Caravatti M, Aeschbacher M, Kocyba C, Takacs B, Carmann H (1988) Cancer Res 48:6799-802.
5 Hanisch F-G, Uhlenbruck G, Dienst C, Stottrop M, Hippauf E (1985) Eur J Biochem 149:323-30.
6 Shimizu M, Yamauchi K (1982) J Biochem (Tokyo) 91:515-19.
7 Ormerod MG, Steele K, Westwood JH, Mazzini MN (1983) Br J Cancer 48:533-541.
8 Hilkens J, Kroezen V, Buijs F, Hilgers J, Van Vliet M, De Voogd W, Bonfer J, Bruning PF (1985) in Monoclonal Antibodies and Breast Cancer, ed. Ceriani RL, Martinus Nijhoff, The Hague, p 28-42.
9 Ceriani RL, Peterson JA, Lee JJ, Moncada R, Blank EW (1983) Somatic Cell Genet 9:415-27.
10 Swallow DM, Gendler S, Griffiths B, Corney G, Taylor-Papadimitriou J, Bramwell HE (1987) Nature 328:82-84.
11 Gendler SJ, Burchell JM, Duhig T, Lamport D, White R, Parker M, Taylor-Papadimitriou J (1987) Proc Nat Acad Sci USA 84:6060-64.
12 Burchell J, Taylor-Papadimitriou J (1989) Cancer Invest 7:53-61.
13 Gendler S, Taylor-Papadimitriou J, Duhig T, Rothbard J, Burchell J (1988) J Biol Chem 263:12820-28.
14 Hanisch F-G, Uhlenbruck G, Peter-Katalinic J, Egge H, Dabrowski J, Dabrowski U (1989) J Biol Chem 264:872-83.
15 Chaplin MF (1982) Anal Biochem 123:336-41.
16 Stellner K, Saito H, Hakomori S (1973) Arch Biochem Biophys 155:464-72.

17 Egge H, Peter-Katalinic J, Paz-Parente J, Strecker G, Montreuil J, Fournet B (1983) FEBS Lett 156:357-62.

18 Geyer R, Dabrowski J, Dabrowski U, Lindner D, Schlüter M, Schott H-H, Stirm S (1990) Eur J Biochem 187:95-110.

19 Berman E, Dabrowski U, Dabrowski J (1988) Carbohydr Res 176:1-15.

20 Koerner TAW, Prestegard JH, Demou PC, Yu RK (1983) Biochemistry 22:2676-87.

21 Dabrowski J, Hanfland P, Egge H (1982) Methods Enzymol 83:69-86.

22 Strecker G, Wieruszeski J-M, Michalski J-C, Montreuil J (1989) Glycoconjugate J 6:67-83.

23 Vliegenthart JFG, Dorland L, van Halbeek H (1983) Adv Carbohydr Chem Biochem 41:209-374.

24 Schachter H (1986) Biochem Cell Biol 64:163-81.

25 Van Halbeek H, Vliegenthart JFG, Fiat AM, Jolles P (1985) FEBS Lett 187:81-88.

26 Hounsell EF, Lawson AM, Feeney J, Gooi HC, Pickering NJ, Stoll MS, Lui SC, Feizi T (1985) Eur J Biochem 148:367-77.

27 Koenderman AHL, Koppen PL, Van den Eijnden DH (1987) Eur J Biochem 166:199-208.

28 Pierce-Cretel A, Decottignies J-P, Wieruszeski JM, Strecker G, Montreuil J, Spik G (1989) Eur J Biochem 182:457-76.

29 Hull SR, Bright A, Carraway KL, Abe M, Kufe D (1988) J Cell Biochem Suppl 12E:130.

30 Kjeldsen T, Clausen H, Hirohashi S, Ogawa T, Jijima H, Hakomori S (1988) Cancer Res 48:2214-20.

31 Hanisch F-G, Uhlenbruck G, Egge H, Peter-Katalinic J (1989) Biol Chem Hoppe-Seyler 370:21-26.

32 Newman RA, Uhlenbruck G (1977) Eur J Biochem 76:149-55.

33 Klein PJ, Newman RA, Müller P, Uhlenbruck G, Citoler P, Schaefer HE, Lennartz KJ, Fischer R (1979) J Cancer Res Clin Oncol 93:205-14.

34 Wolf MF, Ludwig A, Fritz P, Schumacher K (1988) Tumor Biol 9:190-94.