Glycobiology vol. 10 no. 5 pp. 439–449, 2000

MINI REVIEW

MUC1: the polymorphic appearance of a human mucin

Franz-Georg Hanisch[1] and Stefan Müller

Institute of Biochemistry, Medical Faculty, University of Cologne, Joseph-Stelzmann-Strasse 52, 50931 Köln, Germany

Accepted on December 1, 1999

Key words: mucin/MUC1/O-glycosylation/tumor antigen/cancer

Introduction

The term "mucin" has changed its meaning over the last decades prompted by the impressive progress made in the field of glycobiology. Up to the 1970s researchers exclusively used this term to refer to the major glycoprotein components in secreted mucus lining the surfaces of glandular epithelia. The best characterized species during that time, the mucins from ovine and bovine submaxillary glands, served as structural models for this subclass of glycoproteins, since they exhibited the features regarded as typical for mucins: a high carbohydrate content (exceeding 50% by weight) with a concomitant high buoyant density and a threonine/serine rich peptide core serving as a scaffold for the addition of uniform, simple and mainly acidic oligosaccharides. A dense hydrophilic coat of O-linked negatively charged glycans was a simple structural model, but in accord with the proposed function of mucins which was regarded to lie in the formation of a viscoelastic gel serving in physicochemical protection of epithelial surfaces.

Two important observations, which were made during the 1980s, have changed this view and conferred the mucins increased attraction as a research topic. Using advanced methodology and sophisticated instrumentation (FABMS, 500 MHz H-NMR) several groups could demonstrate by means of structural chemistry that mucins are much more complex glycosylated than expected (Lamblin *et al.*, 1984; Hanisch *et al.*, 1985, 1986; Hounsell *et al.*, 1985, 1989; Mutsaers *et al.*, 1986). A second major point was the identification of tumor-associated epitopes on mucins as immunotargets on malignant epithelial cells and their secretions (Magnani *et al.*, 1983; Hilkens *et al.*, 1984; Burchell *et al.*, 1987). In particular the latter aspect has driven tremendous efforts to characterize distinct mucin species by recombinant technology. In 1990 four groups were able to sequence the first human mucin gene on the DNA level (Gendler *et al.*, 1990; Lan *et al.*, 1990; Ligtenberg *et al.*, 1990; Wreschner *et al.*, 1990). Designated as MUC1 in accord with the Human Genome Mapping conventions, this mucin protein is identical to the polymorphic epithelial mucin (PEM), the polymorphic urinary mucin (PUM), Episialin, DF3 antigen, and several other glycoforms of MUC1 isolated from various sources. MUC1 and meanwhile a series of other human mucins (MUC2 to MUC12) have revealed to exhibit large domains of tandemly repeated peptides as a structural characteristic of the "real" mucins (Porchet *et al.*, 1991; Bobek *et al.*, 1993; Toribara *et al.*, 1993; Gum *et al.*, 1994; Meerzaman *et al.*, 1994; Shankar *et al.*, 1994; Van Klinken *et al.*, 1997; Lapensee *et al.*, 1997; Nollet *et al.*, 1998; Williams *et al.*, 1999). They can be discriminated in this way from mucin-like glycoproteins, like GlyCAM1 or MadCAM1, serving roles in cell adhesion (Shimizu and Shaw, 1993). The currently known MUC species can be subdivided into two groups dependent on their structural aspects and biosynthetic routes. Membrane bound mucins (MUC1, MUC3, MUC4, MUC12) exhibit hydrophobic sequences or "transmembrane domains" responsible for their anchoring in the lipid bilayer and C-terminal peptides enter the cytosol (Table I). With one exception (MUC7) the secretory mucins (MUC2, MUC5AC, MUC5B, MUC6) possess one or several von Willebrandt factor-like D domains (Table I), cystein-rich peptides, which function in the oligomerization of mucin monomers and in the packaging into secretory vesicles (Perez-Vilar and Hill, 1999).

This review will focus on MUC1 by attempting to summarize the actual knowledge on the structural, biosynthetic and functional aspects accumulated over the last decade. While previous reviews have concentrated on the mucin gene (Gendler and Spicer, 1995), the processing of the protein and its function in cell adhesion (Hilkens *et al.*, 1992) or on MUC1 immunology (Apostolopoulos and McKenzie, 1994; Finn *et al.*, 1995), we will emphasize those aspects related to O-glycosylation of MUC1 in health and disease. In this context reference will primarily be made to structural and enzymatic work and only sporadically to evidence from immunochemical studies. Recent findings have led to a series of revisions of previous concepts on site-specific O-glycosylation of MUC1 and glycosylation-induced effects on MUC1 antigenicity. These findings and their implications will be tied into the current concepts of the field and discussed with reference to ongoing developments of tumor defense strategies on the humoral and cellular level.

MUC1 expression and function

MUC1 exhibits a ubiquitous organ distribution (Zotter *et al.*, 1988) including mammary gland acini and ducts, salivary

[1]To whom correspondence should be addressed

© 2000 Oxford University Press

Table I. The MUC series of human mucins

MUC species	secretory membraneous	D domains	Chromosomal localization	Tandem repeats		Reference
				Number	Sequence (residue number)	
MUC1	m,s	-	1q21	pm [a]	AHGVTSAPDTRPAPGSTAPP (20)	Gendler *et al.*, 1990
MUC2	s	+	11p15.5	pm	PTTTPITTTTTVTPTPTPTGTQT (23)	Gum *et al.*, 1994
				16	PPTTTPSPPTTTTTTP (16)	
MUC3	m,s	-	7q22	pm	HSTPSFTSSITTTETTS (17)	Van Klinken *et al.*, 1997
MUC4	m,s	-	3q29	pm	TSSASTGHATPLPVTD (16)	Nollet *et al.*, 1998
MUC5AC	s	+	11p15.5	pm	TTSAPTTS (8)	Guyonnet Duperat *et al.*, 1995
MUC5B	s	+	11p15.5	11	Irregular repeats of 29 aa	Meerzaman *et al.*, 1994
MUC6	s	+	11p15.5	pm	169 aa repeat	Toribara *et al.*, 1993
MUC7	s	-	4q13-21	6	TTAAPPTPSATTPAPPSSSAPPE (23)	Bobek *et al.*, 1993
MUC8	?	?	12q24.3	?	TSCPRPLQEGTRV (13) and	Shankar *et al.*, 1994
					TSCPRPLQEGTPGSRAAHALSRRG	
					HRVHELPTSSPGGDTGF (41)	
MUC9	?	?	1p13	pm	GEKTLTPVGHQSVTP (15)	Lapensee *et al.*, 1997
MUC11[b]			7q22	?	SGLSEESTTSHSSPGSTHTTLSPASTTT (28)	Williams *et al.*, 1999a
MUC12[b]	m,s	?	7q22	?	SGLSQESTTFHSSPGSTETTLSPASTTT (28)	Williams *et al.*, 1999a

[a]pm, genetic polymorphism with respect to the number of tandem repeats.
[b]MUC11 and MUC12 could represent proteins encoded by separate genes or originate from alternative mRNA splicing.

gland ducts and serous acini, but not mucinous acini; squamous epithelium of the esophagus; parietal cells, canaliculi and peptic cells of the stomach; acini and ducts of the pancreas; bile ducts in the liver; enterocytes of the duodenum, but not the large intestine; respiratory and ciliated epithelium of the lungs, serous bronchial glands, but not mucinous ones; distal tubules of the kidney and collecting ducts, but not proximal tubules; bladder urothelium; prostate gland epithelium; resting endometrium of the uterus; rete testis; (activated) mesothelium. Important negative tissues are the skin epithelium and all kinds of mesenchymal tissues.

Nonepithelial expression of the mucin has been described to occur on lymphoid cells and lymphomas, especially on plasma cells and myelomas (Zotter *et al.*, 1988, and references therein; Treon *et al.*, 1999). The same authors described detection of the mucin also on anaplastic large cell lymphomas, some T lymphomas of the Ki-1 (activated) type, and on Reed-Sternberg cells in Hodgkin disease. Recently it was shown that a selected panel of anti-MUC1 antibodies stained positive with a fraction of normal proerythroblasts and erythroblasts in bone marrow (Brugger *et al.*, 1999). An important observation is furthermore that activated T cells, but not resting T lymphocytes switch on MUC1 expression (Agrawal *et al.*, 1998).

The cellular localization of MUC1 in normal ductal epithelia is restricted to the apical surface facing the lumen of the duct. There it may be responsible for the lubrication of the epithelial surfaces, and for the entrapment of particles or cellular debris. A potential protective function mediated by specific binding of ingested bacterial or viral pathogens was demonstrated for MUC1 in human milk (Schroten *et al.*, 1992; Yolken *et al.*, 1992). While protective functions may be common to all mucins including the secreted species, the cytoplasmic tail of the membrane-integrated MUC1 has been proposed to be involved in cell signaling events (Pandey *et al.*, 1995). It contains potential tyrosine phosphorylation sites for binding to SH2 domains, like that on the adaptor protein Grb2. Although the functional role of MUC1 mediated cell signaling is unknown, there are hints that point to a possible involvement in cell adhesion (Yamamoto *et al.*, 1997). A cytoplasmic motif on MUC1 was identified to bind β-catenin, and the interaction with this protein was demonstrated to be induced by adherence of epithelial cells to culture dishes. MUC1 interaction with β-catenin is controlled by glycogen synthase kinase, which phosphorylates serine adjacent to proline in the STDRSPYE motif (Li *et al.*, 1998).

Among the numerous functions proposed for MUC1 its potential role in tumor progression and metastasis has to be emphasized. Malignant cells lose their polarity and MUC1 topology changes by expression also at the basolateral surface. Among other factors, this aberrant expression has been claimed to mediate the initial step in the metastatic cascade of tumor cells by the antiadhesive effects exerted by the rod-like, several hundred nanometer extending molecule that shields the cell surface sterically and electrostatically (Hilkens *et al.*, 1992). In this way it is presumed to prevent the formation of cell–cell contacts mediated by E-cadherin (Wesseling *et al.*, 1996) or cell matrix contacts mediated by integrins (Wesseling *et al.*, 1995). In the same way it may also be responsible for the escape of metastasizing tumor cells from the surveillance by the host's immune system. This assumption is based on the finding that mucin-expressing cells are generally resistant to natural killer cells and cytotoxic T cells (Sherblom and Moody, 1986). Moreover, MUC1 on tumor cells was demonstrated to

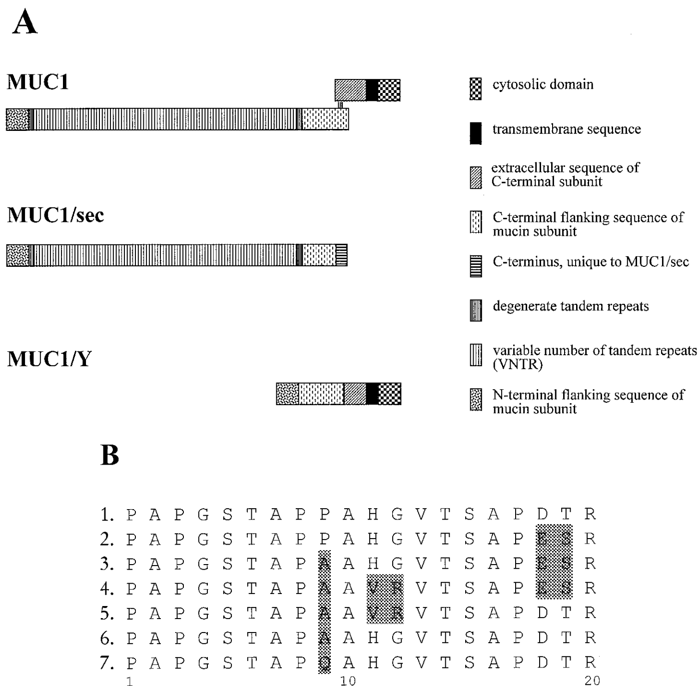

Fig. 1. Primary structure and domain architecture of human MUC1. The primary structures of MUC1 protein splice variants (**A**) and of the MUC1 specific VNTR peptides (**B**) are shown. (**A**) Three isoforms of MUC1 protein are presented, which are derived from the same gene by alternative splicing: MUC1, MUC1/Y, and MUC1/SEC. (**B**) Several variant sequences of the VNTR peptide were reported to occur in human carcinoma cell lines that deviate from the established sequence (1) by a series of amino acid replacements: cancer cell lines T47D (2,3), MCF-7 (4,5), HPAF (6,7).

induce a local T cell anergy (Agrawal *et al.*, 1997), an inhibition of proliferation which was reversible by interleukin-2. Finally, MUC1 has been identified *in vitro* as a potential target in the colonization of metastasizing tumor cells by interacting with ICAM1 (Regimbald *et al.*, 1996) or with E-selectin (Zhang *et al.*, 1996).

In line with these observations is the overexpression of mucin by carcinoma cells making MUC1 a primary target for tumor defense strategies. Since a fraction of the mucin is secreted or shed from the tumor cell surface, MUC1 is detectable in the sera of cancer patients and has meanwhile developed to an established tumor marker (Ca15.3) in the clinical postoperative diagnosis of breast cancer patients (Bon *et al.*, 1990).

Polymorphism of the MUC1 protein

The primary structure and domain architecture of MUC1 protein became accessible by cloning and sequencing of the full-length cDNA derived from different human carcinoma cell lines (Gendler *et al.*, 1990; Lan *et al.*, 1990; Ligtenberg *et al.*, 1990; Wreschner *et al.*, 1990). The deduced organization of exon and intron sequences revealed that among the seven exons the second contains the entire region of multiply repeated 60 bp units. On the protein level this domain can make up about 50–80% of the molecule depending on the number of icosapeptide repeats (Figure 1). This size variation of the domain, which therefor is called a "variable number of tandem repeats" (or VNTR) domain, underlies an inherited polymorphism and comprises about 20–120 repeat units in the northern European population with 40 and 80 repeats being the most frequent (Gendler *et al.*, 1988; Siddiqui *et al.*, 1988). The remainder of the protein consists of 480 amino acids. The VNTR domain in the center of MUC1 is flanked by short regions containing several degenerate repeat peptides with a low percentage of sequence identity to the VNTR peptide (Figure 1). The short N-terminal domain contains a signal peptide and a splice site yielding two alternative products. The more extended C-terminal portions outside the VNTR domain comprise a transmembrane (31aa) and an intracellular domain (69 aa).

While the human MUC1 and its mouse homologue have only limited homology (30% identity) in their repeat domains, the C-terminal cytoplasmic domains are >85% identical (Spicer *et al.*, 1991). This high degree of sequence conservation in conjunction with the cellular topology of the mucin points to a functional relevance of the intracellular peptide. In contrast to the mouse homologue the tandem repeats of the human MUC1 have been reported to be more highly conserved (Gendler *et al.*, 1990; Spicer *et al.*, 1991). In accord with this, the entire domain has been interpreted as a recent expansion of a single repeat (Vos *et al.*, 1991). Unexpectedly, a sequencing study of proteolytic fragments corresponding to the entire VNTR core protein revealed that the established sequence of the repeat may represent only one out of several alternative icosapeptides (Müller *et al.*, 1999). Starting with the AHG motif three positions were reported to be replaced at high incidence: Pro9 → Ala, Asp18 → Glu and Thr19 → Ser (Figure 1). The latter two replacements which were found in about 50% of the repeats occurred always concertedly. Originally reported for a breast cancer cell line (T47D) the same VNTR peptide variations, but also other replacements were detected in MCF-7 and HPAF carcinoma cell lines on the DNA level (Siddiqui *et al.*, 1988; Müller and Hanisch, unpublished observations). The occurrence of variant peptides is not restricted to cancer cells, since MUC1 from individual milk samples showed a similar series of amino acid replacements (Müller and Hanisch, unpublished observations).

The term polymorphism refers also to several isoforms of the mucin originating from alternative splicing events or from processing on the protein level (Figure 1). The membrane-associated MUC1 is expressed as a continuous polypeptide. However, between VNTR and transmembrane domains a proteolytic cleavage site is located which is used during processing of the mucin protein in the Golgi (Ligtenberg *et al.*, 1992). Surface exposure of the cleavage products occurs by formation of a non-covalent heterodimeric complex composed of the mucin subunit and the C-terminal transmembrane unit which anchors MUC1 (Figure 1). This mode of membrane anchorage is not unique for MUC1, but has been reported also for the rat homologue of human MUC4, the ASGP-1/ASGP-2 complex (Sheng *et al.*, 1990). Two variant proteins that are generated by alternative splicing from the MUC1 gene form a similar complex (Baruch *et al.*, 1999). The MUC1/Y isoform, which is devoid of the tandem repeat domain (Zrihan-Licht *et al.*, 1994), spans the membrane and serves as a binding partner for another splice variant of the MUC1 gene, the secreted mucin-like polymorphic MUC1 protein MUC1/SEC (Figure

1), which contains a tandem repeat array (Baruch *et al.*, 1999). MUC1/SEC includes a C-terminal peptide sequence corresponding to parts of intron 2, but no transmembrane domain, and can be detected with sec-peptide specific antibodies in carcinoma cell secretions and in the sera of breast cancer patients.

Polymorphism of MUC1 glycosylation

Cellular biosynthesis of MUC1 glycans

Following the secretory pathway the MUC1 protein core becomes post-translationally modified by extensive glycosylation within, but also outside the VNTR domain. The peptide stretch between VNTR and transmembrane domains contains five potential N-glycosylation sites (Ligtenberg *et al.*, 1990). Within the VNTR domain two serines and three threonines represent five potential O-glycosylation sites per repeat, while numerous putative target sites are distributed irregularly over the remaining extracellular domains. N-Glycosylation by cotranslational transfer of high-mannose glycans occurs in the ER as the initial step (Hilkens and Buijs, 1988). After proteolytic cleavage of the protein core the following processing of N-linked chains to complex-type glycans and the initiation/elongation of O-linked glycans take place in the *cis* to *trans* Golgi. The latter process contributes mainly to the increase in apparent molecular mass of MUC1 and takes about 30 min (Hilkens and Buijs, 1988). However, the final addition of sialic acids, which occurs in the *trans* Golgi, does not reach completion during the first run through, and an immature form of the mucin is exposed on the plasma membrane. The fully mature form, which is detectable after h on the cell surface, is generated by constitutive reinternalization of membrane exposed molecules and their recycling through the *trans* Golgi network (Litvinov and Hilkens, 1993). Several rounds of recycling are necessary to fully sialylate the mucin. Accordingly, the membrane-associated MUC1 comprises a mixture of premature and mature glycoforms, which both are constitutively recycled to achieve or maintain a high degree of sialylation. MUC1 that is released from the cell, on the other hand, corresponds primarily to the mature form of the mucin.

Tissue-specific glycosylation patterns

The structural features of O-linked glycans on MUC1 are common to mucin-type glycosylation, but vary in relation to organ- and differentiation-dependent fluctuations. MUC1 from lactating breast epithelium has been characterized to express primarily long polylactosamine-type chains (Hull *et al.*, 1989; Hanisch *et al.*, 1989, 1990), which are built-up by elongation of the C6-branch of core2 (Figure 2A,B). This core-type, which is formed by branching of the Ser/Thr-linked GalNAc at C3 and C6, is the most frequent of the seven variants described so far on mucins (Table II). A minor fraction of the MUC1 glycans is derived from core 1-disaccharide (F.-G.Hanisch, A.M.Lawson, and T.Feizi, unpublished results), while the occurrence of the less frequent cores 3 and 4 or of the rarely expressed cores 5 to 7 (Table II) could not be demonstrated. The secretory glycoform in human milk is dominated by neutral glycans with linear and branched backbones, which can comprise up to 16 monosaccharide units. The repetitive N-acetyllactosamine units of the linear backbones were found to be linked via C3 or C6 of galactose (Hanisch *et al.*, 1989). Fucose is added to subterminal and internal GlcNAc residues in α4 and α3 linkages resulting in the formation of peripheral and repetitive Lewis type sequences (Figure 2B).

Table II. Core structures of O-linked glycans on mucins

Designation	Structure	Occurrence on	
		MUC1	human mucins
Core 1	Galß1-**3 GalNAc**	**milk, urine**	ubiquitous
Core 2	GlcNAcß1-**6** \\ **GalNAc** / Galß1-**3**	**milk, urine**	ubiquitous
Core 3	GlcNAcß1-**3 GalNAc**	**n.d.**	colon meconium amniotic fluid
Core 4	GlcNAcß1-**6** \\ **GalNAc** / GlcNAcß1-**3**	**n.d.**	respiratory mucus meconium
Core 5	GalNAcα1-**3 GalNAc**	**n.d.**	rectal carcinoma meconium
Core 6	GlcNAcß1-**6 GalNAc**	**n.d.**	seminal fluid meconium ovarian cyst
Core 7	GalNAcα1-**6 GalNAc**	**n.d.**	n.d.

n.d., Not determined.

Structural studies on MUC1-derived glycans from other organs are rare. As a component of human gallbladder epithelium MUC1 (epitectin) has been purified from urine and analyzed with respect to the glycan sizes and structures expressed in this organ (Bhavanandan *et al.*, 1998). Deviating from the patterns of lactating breast epithelium the sizes of the major neutral glycans do not exceed a tetrasaccharide (Figure 2A). The predominant species Galβ1–3GalNAc, Galβ1–3(GlcNAcβ1–6)GalNAc, and Galβ1–3(Galβ1–4GlcNAcβ1–6)GalNAc were also identified as their sialylated derivatives.

Tumor-associated alterations of MUC1 O-glycosylation

In breast cancer the chain lengths of the O-linked glycans on MUC1 are restricted to the core-type level (Hanisch *et al.*, 1996; Lloyd *et al.*, 1996), and sialylated glycans predominate over neutral ones (Figure 3). Enzymatic studies revealed that some mammary carcinoma cell lines and individual breast tumors do not express functional β-glucosaminyltransferases (Brockhausen *et al.*, 1995), which is in accord with the glycan structures found on tumor-associated MUC1 (Figure 3). The low expression or even lack of the core2 forming β6-glucosaminyltransferase in breast cancer cells (Brockhausen *et*

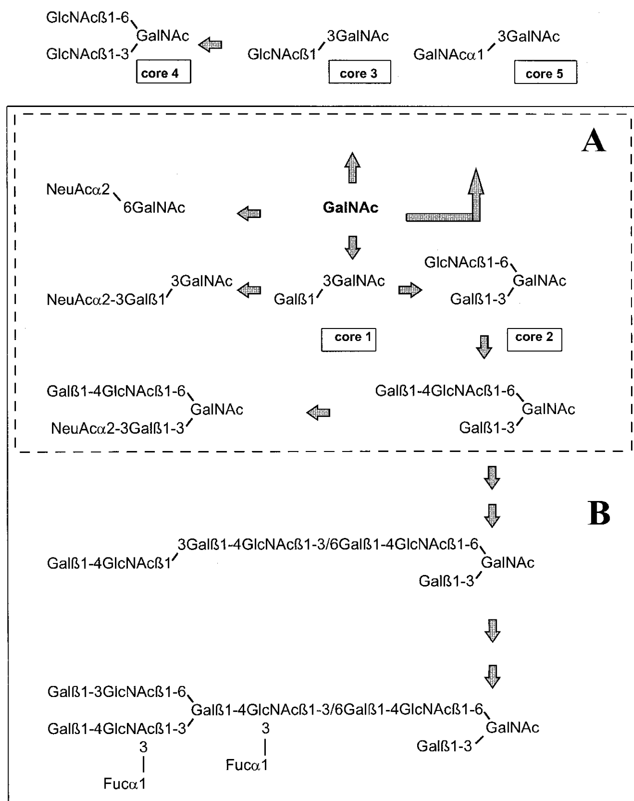

Fig. 2. Structurally defined O-linked glycans on tissue-specific glycoforms of human MUC1. Major glycans representing the glycosylation patterns of the urinary MUC1 (PUM, epitectin) from human bladder epithelium (**A**) and of human milk (**B**). Other core-types were reported to be characteristic for normal human colon (core 3) or to occur in bronchial mucins (core 4). Core 5-based glycans were not described as constituents of normal human mucins.

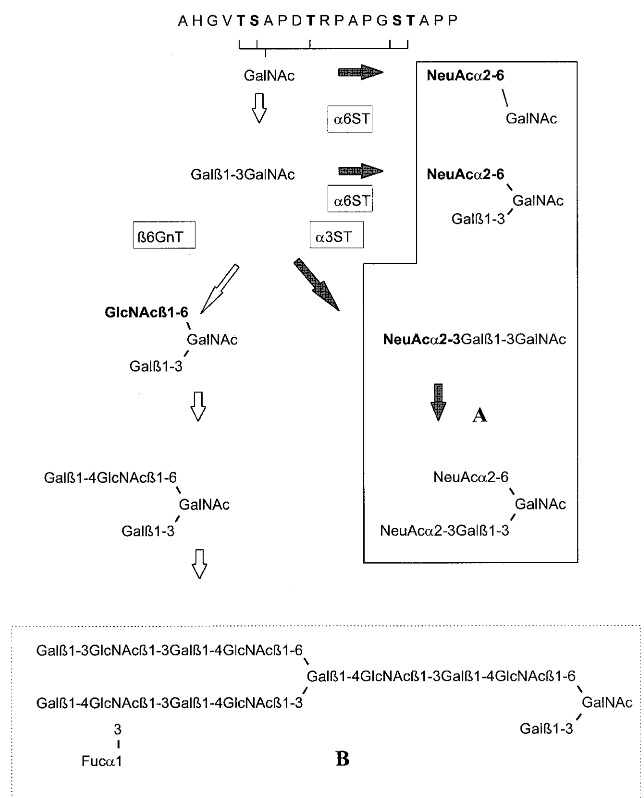

Fig. 3. Cancer-associated O-glycosylation of MUC1. A characteristic feature of breast carcinoma cells is the reduced expression or lack of core 2-specific β6-N-acetylglucosaminyltransferase (β6GnT) and the concomitant overexpression of the α3-sialyltransferase (α3ST) terminating the elongation of core 1 disaccharide by introducing NeuAc at C3 of galactose. This results in the formation of short sialylated oligosaccharides (**A**) instead of long polylactosamine-type chains synthesized by normal glandular cells during lactation (**B**, representative structure). A further cancer-associated pathway is initiated by α6-sialylation of the GalNAc-specific α6-sialyltransferase (α6ST), which competes with β6GnT for core1-disaccharide substrate.

al., 1995) leads to an accumulation of core1-disaccharide, which serves as a substrate of Gal-specific α3-sialyltransferase and of GalNAc-specific β6-sialyltransferases (Figure 3). There is evidence for overexpression of CMP-NeuAc : Galβ1–3GalNAc α3-sialyltransferase in these cells, which may compete with the core 2 enzyme for substrate (Whitehouse *et al.*, 1997). In accord with this, the most prominent glycans found on MUC1 from T47D breast cancer cells were the trisaccharides NeuAcα2–3Galβ1–3GalNAc and NeuAcα2–6(Galβ1–3)GalNAc, and no Lewis antigens were detectable (Hanisch *et al.*, 1996). It has to be emphasized, however, that the outlined features of cancer-associated MUC1 are not shared by all breast cancer cell lines analyzed so far on the chemical level. An exceptionally high degree of complex glycosylation has been found on recombinant MUC1 glycosylation probes expressed in MCF-7 cells, which are able to synthetize GlcNAc-containing oligosaccharides with fucosylated Lewis-type epitopes (Lloyd *et al.*, 1996; S.Müller and F.-G.Hanisch, unpublished observations). On the other hand, T47D cell line can serve as a representative model, reflecting the alterations of O-glycosylation patterns in primary breast tumors in an extreme state. Another alteration of O-glycosylation, which is restricted to cancer-associated MUC1 and, hence, exhibits the qualities of a specific tumor marker in breast cancer, is the expression of the Hanganutsiu Deicher (HD) antigen. HD antigen is identical to an N-glycolyl sialic acid variant (NeuGc) that is not found in birds and man. However, gangliosides of oncofetal origin (Kawai *et al.*, 1991) and MUC1 from breast carcinoma cells (cell lines and solid tumors) were demonstrated immunochemically (Devine *et al.*, 1991) and chemically to contain this variant in a fraction of their sialic acids (Hanisch *et al.*, 1996).

Less specific information is available on MUC1 in colorectal carcinomas, where it is coexpressed with the major intestinal mucin, MUC2. Earlier structural studies on mucin glycosylation in rectal adenocarcinomas had revealed that the glycan patterns were restricted to short core-type sequences derived from the cores 1, 3 and 5 (Table II), which contained as a common structural unit the disaccharide NeuAcα2–6GalNAc (Kurosaka *et al.*, 1983). Enzymatic studies suggested a general shift of core synthesis in colorectal carcinomas from core 3, which is predominant in the normal colon (Podolski, 1985a,b), to core 2 (Brockhausen *et al.*, 1991; Yang *et al.*, 1994; Corfield *et al.*, 1995; Vavasseur *et al.*, 1995).

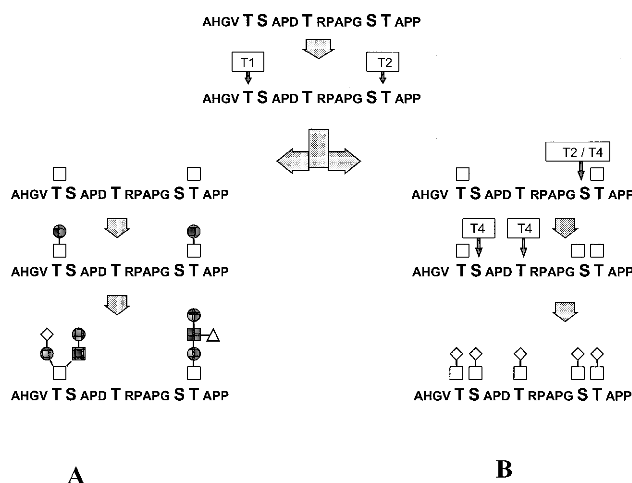

Fig. 4. Site-specific O-glycosylation of MUC1 tandem repeat peptide *in vitro*. The sequential order of GalNAc addition to the five potential sites is shown. T1 to T4 correspond to the recombinant enzymes rGalNAc-T1 to rGalNAc-T4 acting in a site-selective and sequential manner. (**A**) Formation of core structures by addition of Gal or GlcNAc to the Ser/Thr-linked GalNAc presumably prevents initial glycosylation at proximal sites resulting in the low-density glycosylation characteristic for MUC1 in milk. (**B**) If there is no competition with core-specific enzymes, the peptide scaffold can be fully glycosylated at each putative site resulting in the high-density glycoform characteristic for some carcinoma cell lines. Open squares, GalNAc; solid circles, Gal; solid squares, GlcNAc; diamonds, NeuAc; open triangles, Fuc.

No structural or enzymatic evidence is currently available for tumor-associated MUC1 from other organ sites.

Site-specific O-glycosylation of MUC1

Site-specificity and regulation of initial O-glycosylation

Initial GalNAc transfer to serine and threonine by polypeptidyl N-acetylgalactosaminyltransferases (ppGalNAc-Ts) is not ruled by a consensus sequence of the peptide substrate. This renders the prediction of O-glycosylated sites more difficult. On the other hand, much progress has been made with respect to our knowledge on the initial events regulating site-specific incorporation of O-linked chains into mucin-type peptides. This is partly due to approaches based on structural identification of O-glycosylation sites on *ex vivo*–derived glycoproteins (Gooley and Williams, 1994) enabling the training of neural network algorithms with information on the actual glycosylation sites *in vivo*. Another trait that has been followed over the last years is based on *in vitro* glycosylation of synthetic peptide substrates using a panel of recombinant polypeptidyl N-acetylgalactosaminyl-transferases (rGalNAc-Ts). Up to seven rGalNAc-Ts have been cloned and functionally expressed so far (Clausen and Bennett, 1996; Bennett *et al.*, 1998; Hagen *et al.*, 1998; Bennett *et al.*, 1999; Ten Hagen *et al.*, 1999) and rGalNAc-T1 to -T3 were demonstrated to exhibit distinct, but partially overlapping site specificities (Wandall *et al.*, 1997; Hanisch *et al.*, 1999). While the rGalNAc-T1 and rGalNAc-T2 show nearly ubiquitous or broader expression, other enzyme variants, like rGalNAc-T3 and rGalNAc-T6, are much more restricted (Clausen and Bennett, 1996; Bennett *et al.*, 1999).

In the MUC1 model the site-specific GalNAc transfer to the five potential positions within the repeat peptide has been extensively studied *in vitro* (Nishimori *et al.*, 1994; Stadie *et al.*, 1995). rGalNAc-T1 prefers the Thr within VTSA (Figure 4), but is able to glycosylate also the Ser and Thr residues in the GSTA motif (Wandall *et al.*, 1997; Hanisch *et al.*, 1999). On the other hand, rGalNAc-T2 transfers the sugar more rapidly to Thr within GSTA, while Thr in VTSA and Ser in GSTA are less efficient substrate positions. The remaining sites, Ser in VTSA and Thr in PDTR, cannot be glycosylated by these two enzymes, but it has been shown that the “fill-up” reactions are catalyzed by rGalNAc-T4 (Bennett *et al.*, 1998), an enzyme exhibiting a strict dependency on previous glycosylation. The concerted action of rGalNAc-T1, -T2, and -T4, and possibly also other rGalNAc-Ts, would expectedly yield a fully glycosylated MUC1 tandem repeat peptide (Figure 4). Whether a state of full glycosylation is reached or not may depend on several parameters: the cellular repertoire and activity of ppGalNAc-Ts and a postulated dynamic regulation of initial O-glycosylation considering epigenetic effects of previous glycosylation on the GalNAc transfer to other vicinal or proximal sites. Such glycosylation-induced effects have been demonstrated to exist in two *in vitro* studies (Brockhausen *et al.*, 1996; Hanisch *et al.*, 1999). While positive effects were revealed on the monosaccharide level of glycosylated MUC1 peptides, only negative influences were exerted by core 1-disaccharide on glycosylation at adjacent and distant sites. These antagonistic effects of mono- and di(oligo)saccharides could underlie a postulated regulatory mechanism, which assumes early competition of initial O-glycosylation with core glycan synthesis in the *cis* Golgi. A competition of ppGalNAc-Ts with the core-specific glycosyltransferases involved in core 1 and core 2 synthesis could finally determine the density of O-glycosylation (Figure 4).

Localization of O-glycosylation sites on in vivo modified MUC1

The predictions from the recent *in vitro* data are in accord with the actual glycosylation patterns revealed for *in vivo* glycosylated MUC1 from lactating breast epithelium or from breast cancer cells (Figure 5). Proteolytic fragments generated by cleavage of partially deglycosylated MUC1 from milk fat membranes were demonstrated in a combined mass spectrometric and Edman sequencing approach to represent a mixture of mono- to pentasubstituted repeat peptides (Müller *et al.*, 1997). Each of the five putative positions was identified as a target site for GalNAc transfer, but on the average only 50% of the sites were actually glycosylated (Figure 5). There is some preference of particular sites (Thr in VTSA and Ser-Thr in GSTA), but no regular site-specific pattern of GalNAc substitution suggesting a random distribution of the glycans along the peptide scaffold that yields multiple glycoforms of the repeat peptide (Müller *et al.*, 1997). In light of the above outlined mechanistic model, the intermediate density of O-glycosylation found on the lactation-associated MUC1 could be explained by a rapid elongation of core-GalNAc at the initially substituted sites and a concomitant inhibition of further initiation events in proximal positions (Figure 4A).

These results from structural chemistry are in striking conflict with assumptions based on immunological approaches. Most antibodies reactive to peptide motifs local-

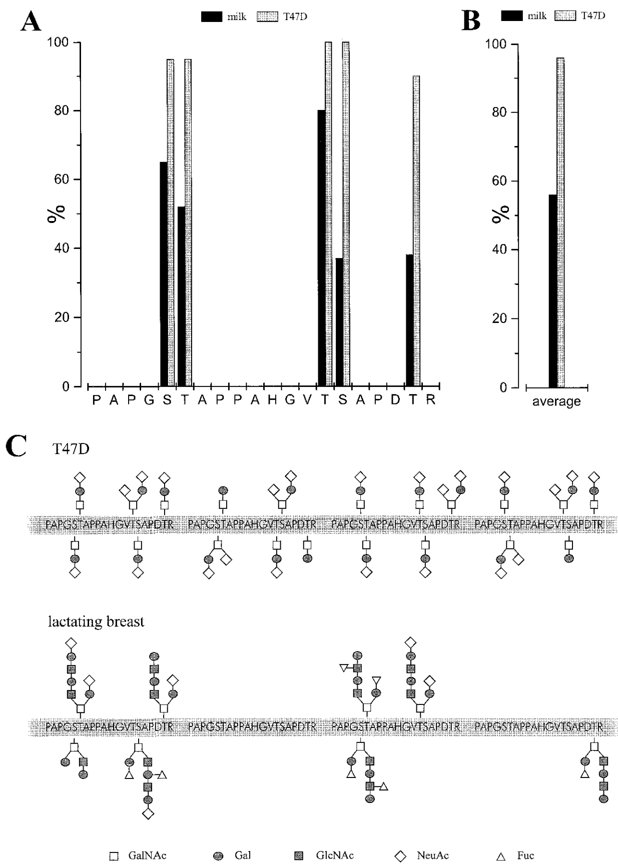

Fig. 5. Site-specific O-glycosylation of MUC1 tandem repeats *in vivo*. The quantitative substitution patterns of the five potential O-glycosylation sites on VNTR peptide are shown as revealed by combined mass spectrometric and Edman sequencing. The PAP20 glycopeptide fragments were derived from MUC1 on human milk fat globules or from secretory MUC1 isolated from the supernatant of T47D breast cancer cells. (A) The degree of glycosylation (%) at each of the five putative sites is shown. (B) The average degree of glycosylation per repeat is given in percent of full substitution (five glycans per repeat). Open squares, GalNAc; solid circles, Gal; solid squares, GlcNAc; diamonds, NeuAc; open triangles, Fuc.

ized within the tandem repeats are negatively influenced in their binding efficiency by GalNAc incorporation into proximal positions (Stadie *et al.*, 1995; Hanisch, 1998). This glycosylation-induced effect on antibody binding can be assumed to increase with the chain lengths of the glycans. In accord with this was the observation that neither MUC1 from normal breast epithelium nor the glycoform expressed by the lactating mammary gland were bound by most of the peptide-specific antibodies. As outlined above, both MUC1 glycoforms are characterized by long polylactosamine-type chains, which presumably by way of steric hindrance prevent access of the antibody to the mucin peptide. Expectedly, the MUC1 glycoform on cancer cells exhibits strong binding activity with peptide-specific antibodies, due to its glycosylation with shorter glycan chains. However, it was also assumed that the higher antigenicity of the tumor-associated glycoform should result from under- or nonglycosylation of the immunodominant DTR motif within the repeats (Burchell *et al.*, 1993). In striking conflict with this assumption is the recent finding that in a breast cancer cell line the motif is glycosylated in more than 90% of the repeats (Müller *et al.*, 1999). A high degree of substitution at DTR is in line with the demonstration of nearly full glycosylation at each of the five putative sites (Figure 5). Whether such high density O-glycosylation revealed for the secretory mature glycoform of a single cell line is a general characteristic of breast cancer cells remains, however, to be established. Earlier evidence from other laboratories had suggested that MUC1 on breast cancer cells is underglycosylated compared with normal epithelial cells in two ways, by reduction of the glycan chain lengths and by reduced density of O-glycosylation (Lloyd *et al.*, 1996).

MUC1 O-glycosylation and immunogenicity/antigenicity

MUC1 as B cell immunotarget

Antibodies to MUC1 could serve functions in immunotherapeutic approaches. While exerting a limited effect against the primary tumor, they may be effective in eradicating circulating single tumor cells or tumor cell embolies. Besides their established roles in antibody-dependent cell-mediated cytotoxicity and complement-mediated lysis, anti-MUC1 could exert effects by uncovering cell surface receptors and restoring in this way cell adhesion or allowing immune recognition of tumor cells. In support of these considerations a recent study on breast cancer patients revealed that naturally occurring anti-MUC1 in their sera favorably influenced overall survival in stage I and II patients (von Mensdorff-Pouilly *et al.*, 1996, and in press).

In the case of MUC1, the protein core, although heavily glycosylated, seems to elicit the stronger immune response in mice, since most hybridoma antibodies generated to this mucin bind to peptide epitopes within the VNTR domain (Price *et al.*, 1997). The preferred target for most peptide-specific antibodies generated to the tumor mucin is located at the DTR motif within the repeat peptide (Price *et al.*, 1997). In humans, however, the natural B cell response to MUC1 shows a second immunogenic motif (Petrarca *et al.*, 1996). The most frequent minimal epitope of natural IgM and IgG in healthy subjects as well as in cancer patients was recently identified as RPAPGS, followed by PPAHGVT and PDTRP (von Mensdorff-Poilly *et al.,* unpublished observations). Unlike the vaccine-induced antibodies the natural anti-MUC1 reacts more strongly to the glycosylated (GalNAc substituted) peptides. One would expect that glycosylation of the DTR motif, as in the case of a tumor-associated glycoform of the mucin, should reduce or abolish binding of antibodies directed to this peptide sequence. Recent evidence, however, suggests that glycosylation of the motif with core-type glycans increases its antigenicity rather than to reduce it (Karsten *et al.*, 1998). Among a series of DTR-specific hybridoma antibodies the majority showed this unexpected behavior (Table III). While recognition of the nonsubstituted DTR motif was demonstrated to be affected by proximal glycosylation (Stadie *et al.*, 1995), the fully glycosylated repeat peptide with GalNAc or Gal-GalNAc in each of the five positions exhibited similar binding activity as the singly, DTR-glycosylated TR peptide (Karsten *et al.*, 1998).

Overall these observations could be explained on the basis of established interactions between the carbohydrate substituent (GalNAc) and the peptide core. Glycosylation at DTR could stabilize in this way a particular conformation of the peptide,

Table III. Grouping of MUC1 specific antibodies directed to the DTR motif according to their reactivities to DTR glycosylated and nonglycosylated repeat peptides

Type A monomeric VNTR	Type B oligomeric VNTR	Type C
Mc5	Ma552	BC2
GP1.4	VU-3-C6	BC3
VU-3-D1	A76-C/A7	VA1
E29	VU-11-D1	VA2
SM-3	VU-11-E1	HMFG1
VU-12-E1	VU-11-E2	
MF06		
214D1		
BCP8		
B27.29		

Type A antibodies are reactive to monomeric repeat peptides and exhibit strong enhancement of binding to glycosylated DTR. Type B antibodies bind to non-glycosylated oligorepeats and to DTR glycosylated monorepeats. Type C antibodies bind to monorepeats irrespective of DTR glycosylation, but do not show a site-specific enhancement of binding.

which favorably interacts with the antibody binding site. This would be in accord with the finding that substitution of the motif with GalNAc is sufficient to induce the observed effect (Karsten *et al.*, 1998). DTR-specific antibodies, which remain unaffected by glycosylation of the motif or which show even enhancement of binding, should define a peptide epitope at the backside of the glycosylated motif, where interaction with the side chain of the Thr residue and sterical hindrance by the sugar are unlikely to occur. In support of this hypothesis are results of a recent x-ray crystallographic study on DTR-glycosylated TR peptide in complex with antibody (SM-3) (Freemont *et al.*, Workshop on Carbohydrates and the Immune Response, Lesbos, 1999). The conformation of non-glycosylated TR peptide in complex with SM-3 and the sites of interaction with the antibody had previously been reported (Dokurno *et al.*, 1998). Reactivity of this and possibly other antibodies seems not to be directed to a "knob-like" or convex epitope as had been postulated (Fontenot *et al.*, 1993). The constraints inherent to this motif result more likely in a turn-like, self-stabilizing, and more compact structure compared to the flanking stretches (Kirnarsky *et al.*, 1998). Two structural models had been previously postulated for the immunodominant motif on the basis of NMR data, a type I reverse β-turn formed by the PDTR sequence (Scanlon *et al.*, 1992) and a type II β-turn formed by the APDT sequence (Fontenot *et al.*, 1995). The flanking regions of the DTR motif seem to adopt more extended conformations with polyproline II helical elements (Kirnarsky *et al.*, 1998), which can be stabilized by GalNAc substitution.

MUC1 as T cell immunotarget

It remains to be established whether the glycosylation-induced effects on antigenicity of the DTR motif have impact also on its immunogenicity. B- and T-cell responses are dependent on the proteolytic processing and MHC presentation of the protein fragments, and currently only limited information is available on these processes with regard to normal mature or cancer-associated glycoforms of MUC1. Model studies with a nonimmunogenic peptide from hemoglobin that had been substituted with GalNAc and other short glycan moieties have indicated that core-type glycans on MHC class II presented peptides can be recognized by T cell receptors (Galli-Stampino *et al.*, 1997; Jensen *et al.*, 1997). In accord with this, core 1 glycosylated TR peptides of MUC1 were able to induce strong, polyclonal proliferative responses of CD4+ T cells from both healthy donors and breast cancer patients (Burchell *et al.*, Workshop on Carbohydrates and the Immune Response, Lesbos, 1999). The Th1 type response was independent of a presentation of the MUC1 glycopeptides by dendritic cells, but showed a strict dependency on the site of glycan substitution within the repeat peptide.

The first evidence that MUC1 can serve as a target antigen for tumor specific cytotoxic T cells (CTL) came from studies on breast cancer patients (Jerome *et al.*, 1991). It could be demonstrated that tumor-reactive T cells from peritumoral lymph nodes of these patients were able to kill MUC1 positive cancer cells. The epitopes recognized by these T cells were localized within the VNTR domain of MUC1 and identified as the SM-3 defined PDTRP peptide motif (Jerome *et al.*, 1991). The mode of CTL activation was unusual, since it was independent of MUC1 processing and presentation. A likely explanation of these unexpected findings was the assumption that cross-linking of T cell receptors was mediated by their simultaneous binding to the repetitive epitopes within the VNTR domain of the mucin. Unfortunately, such a mode of CTL activation circumventing the participation of T helper cells results in responses of low efficacy. The induction of CD4+ T cell responses, however, seems to be critically dependent on the glycosylation state of MUC1. Secretory MUC1 glycoforms from tumor ascites exhibiting a high degree of complex O-glycosylation are unable to prime strong helper T cell responses when administered to dendritic cells *in vitro* (Hiltbold *et al.*, 1998). On the other hand, a synthetic peptide corresponding to five unglycosylated repeats primed CD4+ T cells from healthy donors when presented by dendritic cells (Hiltbold *et al.*, 1998). The naturally processed class II epitope was identified as the dodecapeptide PGSTAPPAHGVT, which was restricted to presentation by HLA-DR3. Other peptide epitopes of the VNTR domain had previously been shown to be restricted to HLA-A2 (Apostolopoulos *et al.*, 1997), and to HLA-A11 (Domenech *et al.*, 1995), but the responses were of low efficacy and T helper cell independent. In summary, the above findings clearly indicate that the healthy human peripheral T cell repertoire contains T helper cells capable of recognizing MUC1 epitopes on the nonglycosylated mucin peptide. Fully glycosylated MUC1 as found in tumor ascites is presumably not properly processed by the antigen presenting cells due to proteolytic resistance of the peptide core. This resistance is likely to be mediated by high density O-glycosylation of the tandem repeat peptides. Recent evidence suggests, on the other hand, that also glycosylated forms of MUC1 can be processed by dendritic cells. However, proteolytic cleavage and presentation are restricted to MHC class I and the efficiency of processing was shown to be inversely correlated with the degree of MUC1 peptide glycosylation (Hiltbold *et al.*, 1999).

Concluding remarks and perspectives

The recent evidence summarized above suggests that the striking effects mediated by O-linked glycans on the antigenicity and immunogenicity of MUC1 will have to be considered in future designs of MUC1-based vaccines. Site-specific O-glycosylation has impact on VNTR peptide antigenicity and on the strength of the VNTR glycopeptide-mediated effect on human T cell proliferation. Structural approaches aiming at the identification of O-glycosylation sites have proven to be superior compared to predictions based on *in vitro* glycosylation data. Similar approaches could be used to learn more about the processing of glycosylated VNTR peptides in APCs, in particular with reference to the glycan structures and to the site-specific substitution patterns required for efficient proteolytic cleavage and binding of the fragments to MHC II protein. Structural work on MHC II presented glycopeptide fragments could support the design of tumor vaccines that induce strong T helper cell responses, a prerequisite for the induction of MUC1-specific IgG and ADCC and/or effective T cell mediated cytotoxicity.

Acknowledgments

The authors acknowledge support of their work by the Deutsche Forschungsgemeinschaft Grants Ha2092/4–1, Ha2092/4–2, and Ha2092/4–3. We also thank Dr. Jo Hilgers for valuable suggestions and discussions during preparation of the manuscript.

References

Agrawal,B., Krantz,M.J., Reddish,M.A. and Longenecker,B.M. (1997) *Cancer*-associated MUC1 mucin inhibits human T-cell proliferation, which is reversible by IL-2. *Nature Med.*, **4**, 43–49.

Agrawal,B., Krantz,M.J., Parker,J. and Longenecker,B.M. (1998) Expression of MUC1 mucin on activated human T cells: implications for a role of MUC1 in normal immune regulation. *Cancer Res.*, **58**, 4079–4081.

Apostolopoulos,V. and McKenzie,I.F.C. (1994) Cellular mucins: targets for immunotherapy. *Crit. Rev. Immunol.* **14**, 293–309.

Apostolopoulos,V., Karanikas,V., Haurum,J.S. and McKenzie,I.F. (1997) Induction of HLA-A2-restricted CTLs to the mucin 1 human breast cancer antigen. *J. Immunol.*, **159**, 5211–5218.

Baruch,A., Hartmann,M., Yoeli,M., Adereth,Y., Greenstein,S., Stadler,Y., Skornik,Y., Zaretzky,J., Smorodinsky,N.I., Keydar,I. and Wreschner,D.H. (1999) The breast cancer-associated MUC1 gene generates both receptor and its cognate binding protein. *Cancer Res.*, **59**, 1552–1561.

Bennett,E.P., Hassan,H., Mandel,U., Mirgorodskaya,E., Roepstorff,P., Burchell,J., Taylor-Papadimitriou,J., Hollingsworth,M.A., Merkx,G., Van Kessel,A.G., Eiberg,H., Steffensen,R., Clausen,H. (1998) Cloning of a human UDP-GalNAc: polypeptide N-acetylgalactosaminyl-transferase that complements other GalNAc-transferases in complete O-glycosylation of the MUC1 tandem repeat. *J. Biol. Chem.*, **273**, 30472–30481.

Bennett,E.P., Hassan,H., Mandel,U., Hollingsworth,M.A., Akisawa,N., Ikematsu,Y., Merkx,G.,van Kessel,A.G., Olofsson,S. and Clausen,H. (1999) Cloning and characterization of a close homologue of human UDP-GalNAc: polypeptide N-acetylgalactosaminyl-transferase-T3, designated GalNAc-T6. Evidence for genetic but not functional redundancy. *J. Biol. Chem.*, **274**, 25362–25370.

Bhavanandan,V.P., Zhu,Q., Yamakami,K., Dilulio,N.A., Nair,S., Capon,C., Lemoine,J. and Fournet,B. (1998) Purification and characterization of the MUC1 mucin-type glycoprotein, epitectin, from human urine: structures of the major oligosaccharide alditols. *Glycoconj. J.*, **15**, 37–49.

Bobek,L.A., Tsai,H., Biesbrock,A.R. and Levine,M.J. (1993) Molecular cloning, sequence and specificity of expression of the gene encoding the low molecular weight human salivary mucin (MUC7). *J. Biol. Chem.*, **268**, 20563–20569.

Bon,G.G., Kenemans,P., van Kamp,G.J., Yedema,C.A. and Hilgers,J. (1990) Review on the clinical value of polymorphic epithelial mucin markers for the management of carcinoma patients. *J. Nucl. Allied Sci.*, **34**, 151–162.

Brockhausen,I., Romero,P.A., Herscovics,A. (1991) Glycosyltransferase changes upon differentiation of CaCo-2 human colonic adenocarcinoma cells. *Cancer Res.* **51**, 3136–3142.

Brockhausen,I., Yang,J.M., Burchell,J., Whitehouse,C. and Taylor-Papadimitriou,J. (1995) Mechanisms underlying aberrant glycosylation of MUC1 mucin in breast cancer cells. *Eur. J. Biochem.*, **233**, 607–617.

Brockhausen,I., Toki,D., Brockhausen,J., Peters,S., Bielfeldt,T., Kleen,A., Paulsen,H., Meldal,M., Hagen,F. and Tabak,L.A. (1996) Specificity of O-glycosylation by colostrum UDP-GalNAc: polypeptide N-acetylgalactosaminyl-transferase using synthetic glycopeptide substrates. *Glycoconj. J.*, **13**, 849–856.

Brugger,W., Bühring,H.-J., Grünebach,F., Vogel,W., Kaul,S., Müller,R., Brümmendorf,T.H., Ziegler,B.L., Rappold,I., Brossart,P., Scheding,S. and Kanz,L. (1999) Expression of MUC1 epitopes on normal bone marrow: implications for the detection of micrometastatic tumor cells. *J. Clin. Oncol.*, **17**, 1535–1544.

Burchell,J., Taylor-Papadimitriou,J. (1993) Effect of modification of carbohydrate side chains on the reactivity of antibodies with core-protein epitopes of the MUC1 gene product. *Epith. Cell Biol.*, **2**, 155–162.

Burchell,J., Gendler,S., Taylor-Papadimitriou,J., Girling,A., Lewis,A., Millis,R. and Lamport,D. (1987) *Development* and characterization of breast cancer reactive monoclonal antibodies directed to the core protein of the human milk mucin. *Cancer Res.*, **47**, 5476–5482.

Clausen,H. and Bennett,E.P. (1996) A family of UDP-GalNAc: polypeptide N-acetyl- galactosaminyl-transferases control the initiaition of mucin-type O-linked glycosylation. *Glycobiology*, **6**, 635–646.

Corfield,A.P., Myerscough,N., Gough,M., Brockhausen,I., Schauer,R. and Paraskeva,C. (1995) Glycosylation patterns of mucins in colonic disease. *Biochem. Soc. Trans.*, **23**, 840–845.

Devine,P.L., Clark,B.A., Birrel,G.W., Layton,G.T., Ward,B.G., Alewood,P.F. and McKenzie,I.F. (1991) The breast tumor-associated epitope defined by monoclonal antibody 3E1.2 is an O-linked mucin carbohydrate containing N-glycolylneuraminic acid. *Cancer Res.*, **51**, 5826–5836.

Dokurno,P., Bates,P.A., Band,H.A., Steward,L.M.D., Lally,J.M., Burchell,J.M., Taylor-Papadimitriou,J., Snary,D., Sternberg,M.J.E. and Freemont,P.S. (1998) Crystal structure at 1.95 A resolution of the breast tumor-specific antibody SM-3 complexed with its peptide epitope reveals novel hypervariable loop recognition. *J. Mol. Biol.*, **284**, 713–728.

Domenech,N., Henderson,R.A. and Finn,O.J. (1995) Identification of an HLA-A11-restricted epitope from the tandem repeat domain of the epithelial tumor antigen mucin. *J. Immunol.*, **155**, 4766–4774.

Finn,O.J., Jerome,K.R., Henderson,R.A., Pecher,G., Domenech,N., Magarian-Blander,J. and Barratt-Boyes,S.M. (1995) MUC1 epithelial tumor mucin-based immunity and cancer vaccines. *Immunol. Rev.*, **145**, 61–89.

Fontenot,J.D., Tjandra,N., Bu,D., Ho,C., Montelaro,R.C. and Finn,O.J. (1993) Biophysical characterization of one-, two- and three-tandem repeats of human mucin (MUC1) protein core. *Cancer Res.*, **53**, 5386–5394.

Fontenot,J.D., Mariappan,S.V., Catasti,P., Domenech,N., Finn,O.J. and Gupta,G. (1995) Structure of a tumor associated antigen containing a tandemly repeated immunodominant epitope. *J. Biomol. Struct. Dyn.*, **13**, 245–260.

Galli-Stampino,L., Meinjohanns,E., Frische,K., Meldal,M., Jensen,T., Werdelin,O. and Mouritsen,S. (1997) T-cell recognition of tumor-associated carbohydrates: the nature of the glycan moiety plays a decisive role in determining glycopeptide immunogenicity. *Cancer Res.*, **57**, 3214–3222.

Gendler,S.J. and Spicer,A.P. (1995) Epithelial mucin genes. *Annu. Rev. Physiol.*, **57**, 607–634.

Gendler,S.J., Taylor-Papadimitriou,J., Duhig,T., Rothbard,J. and Burchell,J. (1988) A highly immunogenic region of a human epithelial mucin expressed by carcinomas is made up of tandem repeats. *J. Biol. Chem.*, **263**, 12820–12823.

Gendler,S.J., Lancaster,C.A., Taylor-Papadimitriou,J., Duhig,T., Peat,N., Burchell,J., Pemberton,L., Lalani,E. and Wilson,D. (1990) Molecuilar cloning and expression of the human tumor-associated polymorphic epithelial mucin PEM. *J. Biol. Chem.*, **265**, 15286–15293.

Gooley,A.A. and Williams,K.L. (1994) Towards characterizing O-glycans: the relative merits of *in vivo* and *in vitro* approaches in seeking peptide motifs specifying O-glycosylation sites. *Glycobiology*, **4**, 413–417.

Gum,J.R., Hicks,J.B., Toribara,N.W., Siddiki,B. and Kim,Y.S. (1994) Molecular cloning of human intestinal mucin (MUC2) cDNA. Identification of

the amino terminus and overall sequence similarity to prepro-von Willebrand factor. *J. Biol. Chem.*, **269**, 2440–2446.

Guyonnet Duperat,V., Audie, J-P., Debailleul,V., Laine,A., Buisine, M-P., Galiegue-Zouitina,S., Pigny,P., Degand,P., Aubert, J-P. and Porchet,N. (1995) Characterization of the human mucin gene MUC5AC: a consensus cysteine-rich domain for 11p15 mucin genes? *Biochem. J.* **305**, 211–219.

Hagen,K.G., Hagen,F.K., Balys,M.M., Beres,T.M., Van Wuyckhuyse,B. and Tabak,L.A. (1998) Cloning and expression of a novel, tissue specifically expressed member of the UDP-GalNAc:polypeptide N-acetylgalactosaminyl-transferase family. *J. Biol. Chem.*, **273**, 27749–27754.

Hanisch,F.-G. (1998) Specificity clusters of MUC1-reactive mouse monoclonal antibodies. *Tumor Biol.*, **19**, 111–117.

Hanisch,F.-G., Egge,H., Peter-Katalinic,J. and Uhlenbruck,G. (1986) Primary structure of a major sialyl-saccharide alditol from human amniotic mucins expressing the tumor-associated sialyl-Lewis x antigenic determinant. *FEBS Lett.*, **200**, 42–46.

Hanisch,F.-G., Egge,H., Peter-Katalinic,J., Uhlenbruck,G., Dienst,C. and Fangmann,R. (1985) Primary structures and Lewis blood-group-dependent expression of major sialylated saccharides from mucus glycoproteins of human seminal plasma. *Eur. J. Biochem.*, **152**, 343–351.

Hanisch,F.-G., Müller,S., Hassan,H., Clausen,H., Zachara,N., Gooley,A.A., Paulsen,H., Alving,K. and Peter-Katalinic,J. (1999) Dynamic epigenetic regulation of initial O-glycosylation by UDP-GalNAc:peptide N-acetylgalactosaminyltransferases. *J. Biol. Chem.*, **274**, 9946–9954.

Hanisch,F.-G., Peter-Katalinic,J., Egge,H., Dabrowski,U. and Uhlenbruck,G. (1990) Structures of acidic O-linked polylactosaminoglycans on human skim milk mucins. *Glycoconj. J.*, **7**, 525–543.

Hanisch,F.-G., Stadie,T.R.E., Deutzmann,F. and Peter-Katalinic,J. (1996) MUC1 glycoforms in breast cancer: cell line T47D as a model for cancer-associated alterations of O-glycosylation. *Eur. J. Biochem.*, **236**, 318–327.

Hanisch,F.-G., Uhlenbruck,G., Peter-Katalinic,J., Egge,H., Dabrowski,J. and Dabrowski,U. (1989) Structures of neutral O-linked polylactosaminoglycans on human skim milk mucins. *J. Biol. Chem.*, **264**, 872–883.

Hilkens,J. and Buijs,F. (1988) Biosynthesis of MAM-**6**, an epithelial sialomucin. Evidence for involvement of a rare proteolytic cleavage step in the endoplasmic reticulum. *J. Biol. Chem.*, **263**, 4215–4222.

Hilkens,J., Buijs,F., Hilgers,J., Hageman,P., Calafat,J., Sonnenberg,A. and van der Valk,M. (1984) Monoclonal antibodies against human milk-fat globule membranes detecting differentiation antigens of the mammary gland and its tumors. *Int. J. Cancer*, **34**, 197–206.

Hilkens,J., Ligtenberg,M.J-L., Vos,H.L. and Litvinov,S.V. (1992) Cell membrane associated mucins and their adhesion modulating property. *Trends Biochem. Sci.*, **17**, 359–363.

Hiltbold,E.M., Alter,M.D., Ciborowski,P. and Finn,O.J. (1999) Presentation of MUC1 tumor antigen by class I MHC and CTL function correlate with the glycosylation state of the protein taken up by dendritic cells. *Cell. Immunol.*, **194**, 143–149.

Hiltbold,E.M., Ciborowski,P. and Finn,O.J. (1998) Naturally processed class II epitope from the tumor antigen MUC1 primes human CD4+ T cells. *Cancer Res.*, **58**, 5066–5070.

Hounsell,E.F., Lawson,A.M., Feeney,J., Gooi,H.C., Pickering,N.J., Stoll,M.S., Lui,S.C. and Feizi,T. (1985) Structural analysis of the O-glycosidically linked core-region oligosaccharides of human meconium glycoproteins which express oncofetal antigens. *Eur. J. Biochem.*, **148**, 367–377.

Hounsell,E.F., Lawson,A.M., Stoll,M.S., Kane,D.P., Cashmore,G.C., Carruthers,R.A., Feeney,J. and Feizi,T. (1989) Characterization by mass spectrometry and 500-MHz proton nuclear magnetic resonance spectroscopy of penta- and hexasaccharide chains of human fetal gastrointestinal mucins (meconium glycoproteins). *Eur. J. Biochem.*, **186**, 597–610.

Hull,S.R., Bright,A., Carraway,K.L., Abe,M., Hayes,D.F. and Kufe,D.W. (1989) Oligosaccharide differences in the DF3 sialomucin antigen from normal human milk and the BT20 human breast carcinoma cell line. *Cancer Commun.*, **1**, 261–267.

Jensen,T., Hansen,P., Galli-Stampino,L., Mouritsen,S., Frische,K., Meinjohanns,E., Meldal,M. and Werdelin,O. (1997) Carbohydrate and peptide specificity of MHC class II-restricted hybridomas raised against an O-glycosylated self peptide. *J. Immunol.*, **158**, 3769–3778.

Jerome,K.R., Barud,D.L., Bendt,K.M., Boyer,C.M., Taylor-Papdimitriou,J., McKenzie,I.F.C., Bast,R.C. and Finn,O.J. (1991) Cytotoxic T-lymphocytes derived from patients with breast adenocarcinoma recognize an epitope present on the protein core of a mucin molecule preferentially expressed by malignant cells. *Cancer Res.*, **51**, 2908–2916.

Karsten,U., Diotel,C., Klich,G., Paulsen,H., Goletz,S., Müller,S. and Hanisch,F.-G. (1998) Enhanced binding of antibodies to the DTR motif of MUC1 tandem repeat peptide is mediated by site-specific glycosylation. *Cancer Res.*, **58**, 2541–2549.

Kawai,T., Kato,A., Higashi,H., Kato,S. and Naiki,M. (1991) Quantitative determination of N-glycolylneuraminic acid expression in human cancerous tissues and avian lymphoma cell lines as a tumor-associated sialic acid by gas chromatography-mass spectrometry. *Cancer Res.*, **51**, 1242–1246.

Kirnarsky,L., Nomoto,M., Ikematsu,Y., Hassan,H., Bennett,E.P., Cerny,R.L., Clausen,H., Hollingsworth,M.A. and Sherman,S. (1998) Structural analysis of peptide substrates for mucin-type O-glycosylation. *Biochemistry*, **37**, 12811–12817.

Kurosaka,A., Nakajima,H., Funakoshi,I., Matsuyama,M., Nagayo,T. and Yamashina,I. (1983) Structures of the major oligosaccharides from human rectal adenocarcinoma glycoprotein. *J. Biol. Chem.*, **258**, 11594–11598.

Lamblin,G., Boersma,A., Lhermitte,M., Roussel,P., Mutsaers,J.H., van Halbeek,H. and Vliegenthardt,J.F. (1984) Further characterization, by combined high-performance liquid chromatography/^{1}H-NMR approach, of the heterogeneity displayed by the neutral carbohydrate chains of human bronchial mucins. *Eur. J. Biochem.*, **143**, 227–236.

Lan,M.S., Batra,S.K., Qui,W.-N, Metzgar,R.S. and Hollingsworth,M.A. (1990) Cloning and sequencing of a human pancreatic tumor mucin cDNA. *J. Biol. Chem.*, **265**, 15294–15299.

Lapensee,L., Paquette,Y. and Bleau,G. (1997) Allelic polymorphism and chromosomal localization of the human oviductin gene (MUC9). *Fert. Ster.*, **68**, 702–708.

Li,Y., Bharti,A., Chen,D., Gong,J. and Kufe,D. (1998) Interaction of glycogen synthase kinase 3β with the DF3/MUC1 carcinoma-associated antigen and β-catenin. *Mol. Cell Biol.*, **18**, 7216–7224.

Ligtenberg,M.J., Kruijshaar,L., Buijs,F., van Meijer,M., Litvinov,S.V. and Hilkens,J. (1992) cell-associated episialin is a complex containing two proteins derived from a common precursor. *J. Biol. Chem.*, **267**, 6171–6177.

Ligtenberg,M.J.L., Vos,H.L., Gennissen,A.M.C. and Hilkens,J. (1990) Episialin, a carcinoma associated mucin, is generated by a polymorphic gene encoding splice variants with alternative amino termini. *J. Biol. Chem.*, **265**, 5573–5578.

Litvinov,S.V. and Hilkens,J. (1993) The epithelial sialomucin, episialin, is sialylated during recycling. *J. Biol. Chem.*, **268**, 21364–21371.

Lloyd,K., Burchell,J., Kudryashov,V., Yin,B.W.T. and Taylor-Papadimitriou,J. (1996) Comparison of O-linked carbohydrate chains in MUC1 mucin from normal breast epithelial cell lines and breast carcinoma cell lines. Demonstration of simpler and fewer glycan chains in tumor cells. *J. Biol. Chem.*, **271**, 33325–33334.

Magnani,J.L., Steplewski,Z., Koprowski,H. and Ginsburg,V. (1983) Identification of the gastrointestinal and pancreatic cancer-associated antigen detected by monoclonal antibody 19–9 in the sera of patients as a mucin. *Cancer Res.*, **43**, 5489–5492.

Meerzaman,D., Charles,P., Daskal,E., Polymeropoulos,M.H., Martin,B.M. and Rose,M.C. (1994) Cloning and analysis of cDNA encoding a major airway glycoprotein, human tracheobronchial mucin (MUC5). *J. Biol. Chem.*, **269**, 12932–12939.

Müller,S., Goletz,S., Packer,N., Gooley,A., Lawson,A.M. and Hanisch,F.-G. (1997) Localization of O-glycosylation sites on glycopeptide fragments from lactation-associated MUC1. *J. Biol. Chem.*, **272**, 24780–24793.

Müller,S., Alving,K., Peter-Katalinic,J., Zachara,N., Gooley,A.A. and Hanisch,F.-G. (1999) High density O-glycosylation on tandem repeat peptide from secretory MUC1 of T47D breast cancer cells. *J. Biol. Chem.*, **274**, 18165–18172.

Mutsaers,J.H., van Halbeek,H., Vliegenthardt,J.F., Wu,A.M. and Kabat,E.A. (1986) Typing of core and backbone domains of mucin-type oligosaccharides from human ovarian-cyst glycoproteins by 500-MHz ^{1}H-NMR spectroscopy. *Eur. J. Biochem.*, **157**, 139–146.

Nishimori,I., Johnson,N.R., Sanderson,S.D., Perini,R., Mountjoy,K., Cernag,R.L., Gross,M.L. and Hollingsworth,M.A. (1994) Influence of acceptor substrate primary amino acid sequence on the activity of human UDP-GalNAc: polypeptide N-acetylgalactosaminyltransferase. *J. Biol. Chem.*, **269**, 16123–16130.

Nollet,S., Moniaux,N., Maury,J., Petitprez,D., Degand,P., Laine,A., Porchet,N. and Aubert,J.P. (1998) Human mucin gene MUC4: organization of its 5′-region and polymorphism of its central tandem repeat array. *Biochem. J.*, **332**, 739–748.

Pandey,P., Kharbanda,S. and Kufe,D. (1995) Association of the DF3/MUC1 breast cancer antigen with Grb2 and the Sos/Ras exchange protein. *Cancer Res.*, **55**, 4000–4003.

Perez-Vilar,J. and Hill,R.L. (1999) The structure and assembly of secreted mucins. *J. Biol. Chem.*, **274**, 31751–31754.

Petrarca,C., Rughetti,A., Rahimi,H., Dagostini,F., Turchi,V., Ghetti,C.A., Scambia,G., Frati,L. and Nuti,M. (1996) Human antibodies against the polymorphic epithelial mucin in ovarian cancer patients recognize a novel sequence in the tandem repeat region. *Eur. J. Cancer*, **32A**, 2155–2163.

Podolski,D.K. (1985) Oligosaccharide structures of human colonic mucin. *J. Biol. Chem.*, **260**, 8262–8271.

Podolsky,D.K., (1985) Oligosaccharide structures of isolated human colonic mucin species. *J. Biol. Chem.*, **260**, 15510–15515.

Porchet,N., Nguyen,V.C., Dufosse,J., Audie,J.P., Guyonnet-Duperat,V., Gross,M.S., Denis,C., Degand,P., Bernheim,A. and Aubert,J.P. (1991) Molecular cloning and chromosomal localization of a novel human tracheo-bronchial mucin cDNA containing tandemly repeated sequences of 48 base pairs. *Biochem. Biophys. Res. Commun.*, **175**, 414–422.

Price,M.R. and 59 coauthors (1997) Summary report on the ISOBM TD-4 Workshop: analysis of 56 monoclonal antibodies against the MUC1mucin. *Tumor Biol.*, **19 (Suppl. 1)**, 1–20.

Regimbald,L.H., Pilarski,L.M., Longenecker,B.M., Reddish,M.A., Zimmermann,G. and Hugh,J.C. (1996) The breast mucin MUC1 as a novel adhesion ligand for endothelial intercellular adhesion molecule 1 in breast cancer. *Cancer Res.*, **56**, 4244–4249.

Scanlon,M.J., Morley,S.D., Jackson,D.E., Price,M.R. and Tendler,S.J. (1992) Structural and computational investigations of the conformation of antigenic peptide fragments of human polymorphic epithelial mucin. *Biochem. J.*, **284**, 137–144.

Schroten,H., Hanisch,F.-G., Plogmann,R., Hacker,J., Uhlenbruck,G., Nobis-Bosch,R. and Wahn,V. (1992) Inhibition of adhesion of S-fimbriated *Escherichia coli* to buccal epithelial cells by human milk fat globule membrane components: a novel aspect of the protective function of mucins in the nonimmunoglobulin fraction. *Infect. Immun.*, **60**, 2893–2899.

Shankar,V., Gilmore,M.S., Elkins,R.C. and Sachdev,G.P. (1994) A novel human airway mucin cDNA encodes a protein with unique tandem-repeat organization. *Biochem. J.*, **300**, 295–298.

Sheng,Z.Q., Hull,S.R. and Carraway,K.L. (1990) Biosynthesis of the cell surface sialomucin complex of ascites 13762 rat mammary adenocarcinoma cells from a high molecular weight precursor. *J. Biol. Chem.*, **265**, 8505–8510.

Sherblom,A.P. and Moody,C.E. (1986) Cell surface sialomucin and resistance to natural cell-mediated cytotoxicity of rat mammary tumor ascites cells. *Cancer Res.*, **46**, 4543–4546.

Shimizu,Y. and Shaw,S. (1993) Cell adhesion. Mucins in the main stream. *Nature*, **366**, 630–631.

Siddiqui,J., Abe,M., Hayes,D., Shani,E., Yunis,E., Kufe,D. (1988) Isolation and sequencing of a cDNA coding for the human DF3 breast carcinoma-associated antigen. *Proc. Natl. Acad. Sci. USA*, **85**, 2320–2323.

Spicer,A.P., Parry,G., Patton,S. and Gendler,S.J. (1991) Molecular cloning and analysis of the mouse homologue of the tumor-associated mucin, MUC1, reveals conservation of potential O-glycosylation sites, transmembrane and cytoplasmic domains and a loss of minisatellite-like polymorphism. *J. Biol. Chem.*, **266**, 15099–15109.

Stadie,T., Chai,W., Lawson,A.M., Byfield,P. and Hanisch,F.-G. (1995) Studies on the order and site-specificity of GalNAc-transfer to MUC1 tandem repeats by UDP-GalNAc: polypeptide N-acetylgalactosaminyltransferase from milk or mammary carcinoma cells. *Eur. J. Biochem.*, **229**, 140–147.

Ten Hagen,K.G., Tetaert,D., Hagen,F.K., Richet,C., Beres,T.M., Gagnon,J., Balys,M.M., Van Wuyckhuyse,B., Bedi,G.S., Degand,P. and Tabak,L.A. (1999) Characterization of a UDP-GalNAc: polypeptidyl N-acetylgalactosaminyl-transferase that displays glycopeptide N-acetylgalactosyltransferase activity. *J. Biol. Chem.*, **274**, 27867–27874.

Toribara,N.W., Roberton,A.M., Ho,S.B., Kuo,W.L., Gum,E., Hicks,J.W., Gum,J.R., Byrd,J.C., Siddiki,B. and Kim,Y.S (1993) Human gastric mucin. Identification of a unique species by expression cloning. *J. Biol. Chem.*, **268**, 5879–5885.

Treon,S.P., Mollick,J.A., Urashima,M., Teoh,G., Chauhan,D., Ogata,A., Raje,N., Hilgers,J., Nadler,L., Belch,A.R., Pilarski,L.M. and Anderson,K.C. (1999) MUC1 core protein is expressed on multiple myeloma cells and is induced by dexamethasone. *Blood*, **93**, 1287–1298.

Van Klinken,B.J., Van Dijken,T.C., Oussoren,E., Buller,H.A., Dekker,J. and Einerhand,A.W. (1997) Molecular cloning of human MUC3 cDNA reveals a novel 59 amino acid tandem repeat region. *Biochem. Biophys. Res. Commun.*, **238**, 143–148.

Vavasseur,F., Yang,J.M., Dole,K., Paulsen,H. and Brockhausen,I. (1995) Synthesis of O-glycan core 3: characterization of UDP-GlcNAc: GalNAc-R β3-N-acetylglucosaminyltransferase activity from colonic mucosal tissues and lack of the activity in human cancer cell lines. *Glycobiology*, **5**, 351–357.

Vinall,L.E., Hill,A.S., Pigny,P., Pratt,W.S., Toribara,N., Gum,J.R., Kim,Y.S., Porchet,N., Aubert,J.-P. and Swallow,D.M. (1998) Variable number tandem repeat polymorphism of the mucin genes located in the complex on 11p15.5. *Hum. Genet.*, **102**, 357–366.

von Mensdorff-Pouilly, Gourevitch,M.M., Kenemans,P., Verstraeten,A.A., Litvinov,S.V., van Kamp,G.J., Meijer,S., Vermorken,J. and Hilgers,J. (1996) Humoral immune response to polymorphic epithelial mucin (MUC1) in patients with benign and malignant breast tumors. *Eur. J. Cancer*, **32A**, 1325–1331.

von Mensdorff-Pouilly,S., Verstraeten,A.A., Kenemans,P., Snijdewint,F.G.M., Kok,A., van Kamp,G.J., Paul,M.A., van Diest,P.J., Meijer,S. and Hilgers,J. (1999) Survival in early breast cancer patients is favourably influenced by a natural humoral immune response to polymorphic epithelial mucin (MUC1). *J. Clin. Oncol.*, in press.

Vos,H.L., de Vries,Y. and Hilkens,J. (1991) The mouse episialin (Muc1) gene and its promoter: rapid evolution of the repetitive domain in the protein. *Biochem. Biophys. Res. Commun.*, **181**, 121–130.

Wandall,H.H., Hassan,M., Mirgorodskaya,E., Kristensen,A.K., Roepstorff,P., Bennett,E.P., Nielsen,P.A., Hollingsworth,M.A., Burchell,J., Taylor-Papadimitriou,J. and Clausen,H. (1997) Substrate specificities of three members of the human UDP-GalNAc: polypeptide N-acetylgalactosaminyltransferase family, GalNAc-T1, -T2 and -T3. *J. Biol. Chem.*, **272**, 23503–23514.

Wesseling,J., van der Valk,S., Sonnenberg,A. and Hilkens,J. (1995) Episialin (MUC1) overexpression inhibits integrin-mediated cell adhesion to extracellular matrix components. *J. Cell Biol.*, **129**, 255–265.

Wesseling,J., van der Valk,S.W. and Hilkens,J. (1996) A mechanism for inhibition of E-cadherin-mediated cell–cell adhesion by the membraneassociated mucin episialin/MUC1. *Mol. Biol. Cell*, **7**, 565–577.

Whitehouse,C., Burchell,J., Gschmeissner,S., Brockhausen,I., Lloyd,K.O., Taylor-Papadimitriou,J. (1997) A transfected sialyltransferase that is elevated in breast cancer and localizes to the medial/trans-Golgi apparatus inhibits the development of core-2-based O-glycans. *J. Cell Biol.*, **137**, 1229–1241.

Williams,S.J., McGuckin,M.A., Gotley,D.C., Eyre,H.J., Sutherland,G.R. and Antalis,T.M. (1999a) Two novel mucin genes down-regulated in colorectal cancer identified by differential display. *Cancer Res.*, **59**, 4083–4089.

Williams,S.J., Munster,D.J., Quin,R.J., Gotley,D.C. and McGuckin,M.A. (1999b) The MUC3 gene encodes a transmembrane mucin and is alternatively spliced. *Biochem. Biophys. Res. Commun.*, **261**, 83–89.

Wreschner,D.H., Hareuveni,M., Tsarfaty,I., Smorodinsky,N., Horev,J., Zaretzky,J., Kotkes,P., Weiss,M., Lathe,R., Dion,A. and Keydar,I. (1990) Human epithelial tumor antigen cDNA sequences. Differential splicing may generate multiple protein forms. *Eur. J. Biochem.*, **189**, 463–473.

Yamamoto,M., Bharti,A., Li,Y. and Kufe,D. (1997) Interaction of the DF3 / MUC1 breast carcinoma-associated antigen and β-catenin in cell adhesion. *J. Biol. Chem.*, **272**, 12492–12494.

Yang,J.M., Byrd,J.C., Siddiki,B.B., Chung,Y.S., Okuno,M., Sowa,M., Kim,Y.S., Matta,K.L. and Brockhausen,I. (1994) Alterations of O-glycan biosynthesis in human colon cancer tissues. *Glycobiology*, **4**, 873–884.

Yolken,R.H., Peterson,J.A., Vonderfecht,S.L., Fouts,E.T., Midthun,K. and Newburg,D.S. (1992) Human milk mucin inhibits rotavirus replication and prevents experimental gastroenteritis. *J. Clin. Invest.*, **90**, 1984–1991.

Zhang,K., Baeckstrom,D., Brevinge,H. and Hansson,G.C. (1996) Secreted MUC1 mucins lacking their cytoplasmic part and carrying sialyl-Lewis a and×epitopes from a tumor cell line and sera of colon carcinoma patients can inhibit HL-60 leukocyte adhesion to E-selectin-expressing endothelial cells. *J. Cell Biochem.*, **60**, 538–549.

Zotter,S., Hageman,P.C., Lossnitzer,A., Mooi,W.J. and Hilgers,J. (1988) Tissue and tumor distribution of human polymorphic epithelial mucin. In Hilgers,J. and Zotter,S. (eds.), *Polymorphic Epithelial Mucin and CA125-Bearing Glycoprotein As Tumor-Associated Antigens. Cancer Rev.*, Vol. **11/12**, pp. 55–101.

Zrihan-Licht,S., Vos,H.L., Baruch,A., Elroy-Stein,O., Sagiv,D., Keydar,I., Hilkens,J. and Wreschner,D. (1994) Characterization and molecular cloning of a novel MUC1 protein, devoid of tandem repeats, expressed in human breast cancer tissue. *Eur. J. Biochem.*, **224**, 787–795.