High Density O-Glycosylation on Tandem Repeat Peptide from Secretory MUC1 of T47D Breast Cancer Cells*

(Received for publication, November 3, 1998, and in revised form, April 9, 1999)

Stefan Müller‡, Kim Alving§, Jasna Peter-Katalinic§, Natasha Zachara¶, Andrew A. Gooley¶, and Franz-Georg Hanisch‡

From the ‡Institute of Biochemistry, Medical Faculty of the University, Joseph-Stelzmann-Strasse 52, 50931 Köln, Germany, the §Institute of Medical Physics and Biophysics, University of Münster, 48149 Münster, Germany, and the ¶Macquarie University Center for Analytical Biotechnology, Macquarie University, Sydney NSW 2109, Australia

The site-specific O-glycosylation of MUC1 tandem repeat peptides from secretory mucin of T47D breast cancer cells was analyzed. After affinity isolation on immobilized BC3 antibody, MUC1 was partially deglycosylated by enzymatic treatment with α-sialidase/β-galactosidase and fragmented by proteolytic cleavage with the Arg-C-specific endoproteinase clostripain. The PAP20 glycopeptides were isolated by reversed phase high pressure liquid chromatography and subjected to the structural analyses by quadrupole time-of-flight electrospray ionization mass spectrometry and to the sequencing by Edman degradation. All five positions of the repeat peptide were revealed as O-glycosylation targets in the tumor cell, including the Thr within the DTR motif. The degree of substitution was estimated to average 4.8 glycans per repeat, which compares to 2.6 glycosylated sites per repeat for the mucin from milk (Müller, S., Goletz, S., Packer, N., Gooley, A. A., Lawson, A. M., and Hanisch, F.-G. (1997) J. Biol. Chem. 272, 24780–24789). In addition to a modification by glycosylation, the immunodominant DTR motif on T47D-MUC1 is altered by amino acid replacements (PAGSTAPAAHGTSPAESR), which were revealed in about 50% of PAP20 peptides. The high incidence of these replacements and their detection also in other cancer cell lines imply that the conserved tandem repeat domain of MUC1 is polymorphic with respect to the peptide sequence.

Due to the structural complexity of O-linked glycans, this characteristic posttranslational modification of mucin peptides is a polygenic regulated phenomenon and hence is prone to multiple, differentiation-dependent alterations. According to numerous reports, mucin O-glycosylation can now be regarded as a diagnostically relevant indicator of tumor-associated changes that are characterized by 1) the de novo expression of novel glycotopes the ectopic or incompatible expression of carbohydrate blood groups, or 2) by deletion/truncation of glycan chains (1).

Also, the widely distributed epithelial mucin MUC1 has been described to be aberrantly processed in cancer cells (2–4). In breast cancer, the nonexpression of the core2 enzyme, Galβ1–3GalNAc/β-6-N-acetylgalactosaminyltransferase (5), leads to the truncation of polyglactosamine-type chains found on the lactation-associated mucin (6) and to the accumulation of core-type chains (2–4). The preponderance of sialylated core1-trisaccharide on carcinoma-associated MUC1, which can be regarded as a biosynthetic dead end product, has been shown to originate from the simultaneous up-regulation and overexpression of Galβ1–3GalNAc/α-3-sialyltransferase (7). Moreover, not only the chain length of the glycans but also their density has been described to be reduced on breast cancer cell-specific MUC1 (4).

Reduced glycosylation has been assumed to permit the immune system access to the peptide core of the tumor-associated mucin (8). The preferred target site for most peptide-specific mouse antibodies generated to the tumor mucin, to synthetic variable number of tandem repeats (VNTR)3 peptides, or to MUC1 fusion proteins is located at the DTR motif of the repeat peptide (9). In humans, however, the natural response to MUC1 shows a second, even stronger immunodominant motif that comprises the STA sequence within the repeat (10). Recent results have revealed that the majority of murine mAbs generated to MUC1 bind significantly stronger to the DTR motif of VNTR peptide, if this is glycosylated with core-type glycans at threonine (11). Also unexpectedly, the binding activity of these antibodies remained unaffected, if the proximal sites were substituted with core1-disaccharides (11).

These findings were in striking conflict with the experimentally based assumptions that the immunodominant DTR motif had to be nonglycosylated to permit binding of antibodies and that the glycans at proximal sites should negatively influence the DTR antigenicity by steric hindrance. In vitro glycosylation studies had shown that the Ser within VTS and the Thr within DTR were neither target sites for the ppGalNAc-Ts from human cancer cells (12, 13) or from milk (13), nor for the recombinant enzymes rGalNAc-T1 to rGalNAc-T3 (14). Conclusions based on in vivo studies, however, using recombinant glycosylation probes (15) or calculations based on sequence data of glycoproteins (16) were finally confirmed by direct chemical evidence obtained for ex vivo isolated MUC1 from milk (17), and they agreed with the finding that all five putative sites within the tandem repeat of MUC1 were glycosylation targets.

We here report on the localization of O-glycosylation sites within the VNTR peptide of tumor-associated MUC1 that is shed into the culture medium by T47D breast cancer cells. This model cell line is representative for breast cancer cells in several respects: (i) by exhibiting aberrant expression of core glycosyltransferases (5), (ii) by expressing mainly core-type sialyloligosaccharides typical for several established breast cancer cell lines (4) and ductal carcinomas of the breast (3), and (iii) by binding of tumor-selective anti-MUC1 antibodies, such as SM-3

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‖ To whom correspondence should be addressed. Tel.: 49-221-478-4493; Fax: 49-221-478-6977; E-mail: franz.hanisch@uni-koeln.de.

1 The abbreviations used are: QTOF, quadrupole time-of-flight; ESI, electrospray ionization; MALDI, matrix-assisted laser desorption ionization; rGalNAc-T, recombinant peptidyl N-acetylgalactosaminyltrnasferase; VNTR, variable number of tandem repeats; RP, reversed phase; HPLC, high pressure liquid chromatography; mAb, monoclonal antibody; PTH, phenylthiohydantoin; GalNAc, N-acetylgalactosamine.
The peptide epitopes of the antibodies were analyzed by three independent groups, and the results were combined in a summary report (9). The epitopes of group 1 (underlined sequence), group 2 (sequence in plain type), and group 3 (sequence in boldface type) do partially overlap. Each antibody shows a distinct dependency on site-specific glycosylation and is, accordingly, grouped into four categories: DTR-specific antibodies of type A to type C (11) and other VNTR-specific antibodies (9), grouped as type D.

Table I

Monoclonal antibodies used in this study

| Antibody | Type | VNTR peptide sequence of MUC1 |
|----------|------|-----------------------------|
| BC3      | C    | -DTR FAPA PPAGVTSA DTR PAP- |
| VA-1     | C    | TR PAP                      |
| SM-3     | A    | APDTR                       |
| VU-3-C6  | B    | APDTR P                     |
| M38      | D    | FAP APAHGVTSA PAP           |

* Type A: antibodies reactive to mono- and oligomeric repeats exhibiting enhancement of binding to VNTR peptide with GalNAc- or Galβ-3GalNAc-substituted DTR motif (11). mAb SM-3 shows reduction of binding after glycosylation of the proximal sites in the VTSA and GSTA motifs (13, 23). However, the positive effect exerted by DTR glycosylation overrides the negative effect by proximal glycosylation on antibody binding (11). Type B: antibodies reactive only to oligorepeats and exhibiting enhancement of binding to VNTR peptide with GalNAc- or Galβ-3GalNAc-substituted DTR motif (11). No negative glycosylation-induced effect on VNTR-C6 binding was observed when analyzing glycopeptides multiply substituted with GalNAc or Galβ-3GalNAc (311). Type C: antibodies reactive to mono- and oligorepeats exhibiting neither positive nor negative effects on their peptide binding activity induced by glycosylation of the VNTR peptide (9, 11, 22). Type D: antibodies defining VNTR epitopes, but not recognizing the DTR motif (9).

Fig. 1. Western blot analysis of secretory T47D-MUC1 prior to and after affinity isolation on immobilized BC3 antibody. Immunostaining of the Western blot was performed with the MUC1-specific antibodies VU-3-C6, M38, and SM3. Lane S, sample prior to application; lane F, flow-through; lane W, wash fractions; lane E, eluate.

Materials and Methods

Cell Culture—The mammary carcinoma cell line T47D was obtained from the American Type Culture Collection and grown in RPMI 1640 containing Glutamax I, 10% fetal calf serum, Pen/Strep (200 IU/ml-200 μg/ml) and insulin (10 μg/ml) (all from Life Technologies, Inc.) at 5% CO2 and 37 °C. Medium (15 ml) was replaced at intervals of 2–3 days. Collected supernatants from confluent cell layers were centrifuged at 500 x g for 10 min, filtered over Miracloth, and stored frozen at −20 °C.

Isolation of MUC1—Mouse ascites containing monoclonal antibody BC3 (IgM) was a generous gift from Dr. Ian McKenzie (Austin Research Institute, Heidelberg, Australia). The binding characteristics of this antibody and of all anti-MUC1 mAbs used in the present study were summarized in Table I. mAb BC3 was used for affinity isolation of MUC1 due to its established independence from glycosylation (9). Accordingly, this antibody exhibits high affinity binding to strikingly different glycoforms of the mucin, such as lactation- and tumor-associated MUC1 (9). To prepare an affinity column, the ascites was cleared from debris by a brief centrifugation step and subjected to size exclusion chromatography on Sephacryl S-300, equilibrated in 0.2 M sodium bicarbonate buffer, pH 8.3, 0.5 M sodium chloride. The antibody was collected with the void fraction and coupled to a 1-ml HiTrap NHS activated column (Amersham Pharmacia Biotech) according to the manufacturers instructions. Cell culture supernatant was cleared from cellular debris and fines by a 30-min centrifugation at 10,000 rpm and 4 °C and passage through a short bed of Sephadex G25. Cleared supernatant was supplemented to 0.5 M NaCl, 0.02% NaN3, and 200-ml batches were loaded onto the column. After washing with 20 volumes of 50 mM sodium phosphate, pH 7.2, 0.5 M NaCl, 0.01% Tween 20, 0.02% NaN3, and 10 volumes of 50 mM sodium phosphate, pH 7.2, 0.125 M NaCl, 0.02% NaN3, the column was eluted with 10 volumes of 50 mM glycine/HCl, pH 2.5. The eluted fractions were analyzed by SDS-polyacrylamide gel electrophoresis on 3.5–15% gradient gels. The gels were either stained with silver or blotted onto nitrocellulose membranes. Western blots were developed with anti-MUC1 mouse monoclonal antibodies and anti-mouse POD conjugates using ECL detection. Eluates from five runs, corresponding to 1000 ml of supernatant, were combined, concentrated by ultrafiltration on Amicon XM 300, desalted by gel filtration on Sephadex G25 (PD10), and dried in a speedvac. To prevent selective enrichment of particular glycoforms during the affinity purification, the column was never loaded to saturation, and all flow-through fractions were analyzed for MUC1 in Western blots using a series of anti-MUC1 mAbs with different fine specificities (Table I).
mAbs SM-3 and VU-3-C6 bind to the DTR motif, whereas mAb M38 recognizes the GSTAP sequence within the repeat peptide. Like BC3, all antibodies have been demonstrated to bind to lactation- and tumor-associated glycoforms of the secretory mucin (9).

Preparation of PAP20 Glycopeptides—To reduce the average chain lengths of MUC1 linked glycans and improve thereby the efficiency of clostripain digestion, affinity isolated MUC1 (approximately 10 μg) was partially deglycosylated by sequential treatment with 200 units of neuraminidase (clostridium perfringens, New England Biolabs) in 50 mM sodium citrate, pH 4.5 (37 °C for 16 h), and 25 million units of β-galactosidase (bovine testes, Roche Molecular Biochemicals) in 100 mM sodium phosphate/citrate, pH 4.3, 10% glycerol (37 °C for 16 h) according to the manufacturer’s instructions. To demonstrate the extent of sialic acid and galactose hydrolysis, the enzyme treated mucin was desalted on Sephadex G25 (NAP5) column, and a fraction was subjected to monosaccharide composition analysis as described below. The remaining, partially deglycosylated mucin was dried in a speedvac and digested by addition of 10 μg of activated clostripain (Sigma) as described previously (17). PAP20 glycopeptides were isolated from the crude digest by reversed phase (RP) HPLC on a Beckmann C18 column (2.0 × 150 mm) using previously described conditions (17). PAP20 glycopeptides were detected by enzyme-linked immunosorbent assay using monoclonal antibody M38 (Table I and Ref. 9). Immune-reactive fractions were combined, and the volume was reduced to approximately 50 μl in a speedvac.

Mass Spectrometry—The hybrid quadrupole-time of flight (QTOF) fitted with a Z-spray source (Micromass, Manchester, UK) was used for electrospray ionization (ESI) mass spectrometry in one-dimensional and two-dimensional mass spectrometry experiments in a nanospray mode as described recently (18). Briefly, in one-dimensional mass spectrometry experiments, precursor ions were detected as triply charged ions that correspond to the calculated masses and compositions of PAP20 glycopeptides listed in Table II. The major signal at m/z 1022.1 (mono-isotopic mass of the singly charged ion, 3064.2 mass units) was isolated by RP-HPLC and used for C-terminal sequencing by combining carboxypeptidase Y digestion (Sigma, sequencing grade, 100 ng, enzyme:substrate ratio of 1:1 in 100 mM sodium citrate, pH 5.5) with MALDI mass spectrometry. Edman Degradation—Glycopeptides were covalently bound to arylamine derivatized membranes and sequenced on a Hewlett-Packard G1000A protein sequenator using a modification of Program 3.1 that substitutes methanol for ethyl acetate to transfer the ATZ-amino acid (19). The use of methanol substituted for ethyl acetate enables efficient recovery of glycosylated PTH-derivatives at approximately 75% of the expected yield for a nonglycosylated PTH-derivative. The sequelon AA™ membranes were from the Perseptive Biosystems division of Perkin-Elmer. Quantitation of mono-, di-, and trisaccharide substituted serine/threonine using solid-phase Edman degradation was performed as described (19, 20).

Monosaccharide Composition Analysis—The partially deglycosylated mucin was analyzed for the contents of sialic acid and galactose by quantitative monosaccharide composition analysis using gas chromatography/mass spectrometry on an MD800 (Fisons, Mainz, Germany) according to a previously described protocol (21).

Enzyme Immunoassays—Immunochemical detection of PAP20 glycopeptides in eluates from RP-HPLC was performed by enzyme-linked immunosorbent assay as described previously (17, 22). mAbs M38 and VA-1 were available as pure antibodies at 1 mg/ml from the ISOBM TD-4 International Workshop on Monoclonal Antibodies against MUC1 (see Ref. 9 and Table I) and used at 0.5 μg/ml. The epitope of mAb M38, which was localized to the GSTAP sequence of the MUC1 tandem repeat peptide, is not affected by clostripain cleavage of the mucin core. By contrast, one of the postulated sequences recognized by mAb VA-1 is destroyed by enzymatic hydrolysis at Arg-C (Table I).

RESULTS

Isolation of MUC1—Secretory MUC1 was isolated from supernatants of the human mammary carcinoma cell line T47D by affinity chromatography on immobilized monoclonal antibody BC3. The antibody is directed against an epitope within the VNTR peptide and has been shown to react with highly glycosylated milk MUC1, urinary MUC1, and with unglycosylated synthetic peptides covering the DTR motif, indicating a low selectivity for particular MUC1 glycoforms (9). However, to
From the measured triply charged ions (M + 3H)\(^{-}\), the deconvoluted values of singly charged ions (M + H)\(^{-}\) were calculated by multiplication with a factor of 3 and subtraction of the mass of 2 protons. These values were compared with the corresponding pseudomolecular ions calculated for glycosylated PAP20 on the basis of monoisotopic mass increments. Mass values matched with a series of differently glycosylated PAP20 peptides and sequence variants of PAP20 peptides exhibiting Pro → Ala and/or Asp-Thr → Glu-Ser replacements according to the Edman degradation data. The positions of the amino acid exchanges were defined by Edman degradation (Fig. 6) and by PSD MALDI mass spectrometry (Fig. 4). Signals with relative intensities exceeding 20% were regarded as major, and the corresponding triply charged ions are given in boldface.

| (M + 3H)\(^{-}\) measured | (M + H)\(^{-}\) deconvoluted | Proposed structure of glycopeptides |
|--------------------------|-----------------------------|----------------------------------|
| 765.1                    | 2293.3                      | HexNAc\(_2\)                      |
| 832.7                    | 2496.1                      | HexNAc\(_3\)                      |
| 900.4                    | 2699.2                      | Hex HexNAc\(_3\)                  |
| 954.1                    | 2863.3                      | Hex HexHexNAc\(_4\)               |
| 988.1                    | 2902.3                      | Hex HexHexHexNAc\(_5\)            |
| 1008.4                   | 3023.2                      | Hex HexHexHexHexNAc\(_6\)         |
| 1022.1                   | 3064.2                      | Hex HexHexHexHexHexNAc\(_7\)      |
| 1076.1                   | 3226.3                      | Hex HexHexHexHexHexHexNAc\(_8\)   |
| 1119.1                   | 3355.4                      | NeuAc Hex HexHexNAc\(_4\)         |
| 1130.1                   | 3388.4                      | NeuAc Hex HexHexHexNAc\(_5\)      |
| 1173.2                   | 3517.5                      | NeuAc Hex\(_2\) HexHexNAc\(_3\)   |
| 1227.2                   | 3679.5                      | NeuAc Hex\(_2\) HexHexHexNAc\(_4\) |
| 1184.1                   | 3590.4                      | Hex HexHexNAc\(_5\)               |
| 1281.1                   | 3841.4                      | NeuAc Hex\(_2\) HexHexHexHexNAc\(_6\) |

\(^{a}\) Asp-Thr → Glu-Ser.
\(^{b}\) Pro → Ala.
\(^{c}\) Pro → Ala and Asp-Thr → Glu-Ser.

confirm that all MUC1 glycoforms were bound to the column, the loaded sample, all fractions of the flow-through, and the wash fraction and the eluate were analyzed for MUC1 in Western blots (Fig. 1) using monoclonal antibodies SM-3, M38, and VU-3-C6 (Table I). According to previous analyses (9), the binding characteristics of these three anti-MUC1 antibodies cover lactation- and tumor-associated glycoforms of MUC1. Moreover, epitope repetition on individual MUC1 molecules ensures that among a multitude of glycoforms of the repeat peptide at least a subfraction on each mucin molecule will be recognized by any of the antibodies used. A strong increase of the MUC1 signal in the eluted fraction as compared with the loaded sample indicates the enrichment of the mucin. Due to the high carbohydrate content of the mucin, the band is smeared and only faintly stained by silver. The stain further revealed the presence of three minor protein contaminants, migrating with apparent molecular masses of 40, 70, and 90 kDa. No attempt was made to further purify the affinity-isolated sample, because a final selection was provided by the specific proteolytic cleavage with clostripain in conjunction with RP-HPLC of the MUC1 derived glycopeptides. In accordance with this, the RP-HPLC profiles registered for the digested sample were dominated by mucin-derived glycopeptides (see Fig. 2 and below).

Preparation of PAP20 Glycopeptides—Exoglycosidase-treated mucin was demonstrated by monosaccharide analysis to have lost about 70% of its galactose and more than 90% of its sialic acid. This partially deglycosylated MUC1 was effectively cleaved with clostripain and yielded a pool of glycosylated icosapeptides from the MUC1 tandem repeat domain, referred to as PAP20-glycopeptides. The fragments were isolated from the crude digest by RP-HPLC, as shown in Fig. 2. A heterogenous peak coeluted with a synthetic MUC1 glycopeptide, a 21-mer carrying three core1-disaccharides (Table I, Fig. 2), had been demonstrated previously (4). Accordingly, the amounts of MUC1 tandem repeat domain, referred to as PAP20-glycopeptides, were found by QTOF(+)ESI mass spectrometry and proposed structure of glycopeptides.
HexNAc found on individual PAP20 glycopeptides should reflect the number of peptide linked GalNAc. The QTOF-ESI spectrum of combined, PAP20 containing HPLC-fractions (Fig. 3) was dominated by triple charged ions ($m/z$ 1022.1, 1076.1, 1130.1, 1173.2, and 1227.2) corresponding to a series of PAP20 glycoforms, originating from PAP20-HexNAc$_5$ ($m/z$ 968.1). The deconvoluted masses of these ions are related to each other by mass increments corresponding to Hex (162.2 mass units) and NeuAc (291.3 mass units) residues (Table II), which were not completely removed by glycosidase digestion of the mucin sample. A similar series ($m/z$ 954.4 and 1008.4) was also present for PAP20-HexNAc$_4$ ($m/z$ 900.4), but the signal intensities were significantly lower. In addition, a weak ion corresponding to PAP20-HexNAc$_3$ was detected at $m/z$ 832.7. Because the intensity of this ion was low and because there was no series of related ions, PAP20-HexNAc$_3$ represents obviously only a minor component in the PAP20 glycopeptide mixture.

The presence of the PAP20-HexNAc$_5$ based ion series was the most important result of the described mass spectrometric analysis, because it demonstrates that (i) all of the five potential glycosylation sites are substituted in the average glycoform of MUC1 tandem repeat peptide in T47D cells, and (ii) that all five positions can principally be glycosylation targets for the various ppGalNAc-Ts. Moreover, the high signal intensities of the PAP20-HexNAc$_5$ series, as compared with the PAP20-HexNAc$_4$ and PAP20-HexNAc$_3$ series leads to the conclusion that MUC1 expressed in T47D is almost completely O-glycosylated and that this glycoform of the tandem repeat is characteristic for these tumor cells.

Referring to Table II and to the deconvoluted ion masses, the major signals are accompanied by an ion series at $m/z$ 26 mass units. These match the masses of PAP20 glycopeptides exhibiting a replacement of Pro by Ala, which was identified by Edman degradation (see below). A concerted replacement of Thr-19 by Ser and of Asp-18 by Glu was revealed by MALDI mass spectrometry of deglycosylated PAP20 fragments after sequential digestion with carboxypeptidase Y. While the mass differences of the $M+Na$ ion series at $m/z$ 1082, 1181, 1282, 1389, 1652, 1753, and 1909 correspond to the C-terminal sequence VTSAPDTR, the intense signal at $m/z$ 1666 indicates liberation of Ser from the precursor at $m/z$ 1753 and corresponds to the Glu-containing residual peptide PAPGSTAPPAHGVTSAE (Fig. 4).

A sequencing of glycopeptides was performed by MS/MS analysis in QTOF(+)-ESI mass spectrometry to localize the $O$-glycosylation sites and to identify amino acid replacements at particular sites. A signal corresponding to PAP20 (glyco)peptides with two HexNAc residues was detectable with very low intensity at $m/z$ 765.1 (Fig. 3). MS/MS analysis of this ion revealed the formation of y-series ions from the C-terminal end
TABLE III

**MS/MS analysis by QTOF-ESI(+) mass spectrometry of the triply charged ion at m/z 1022.1 corresponding to PAP20-HexNAc**

A quadrupole analyzer was used to select the precursor ion for fragmentation in the hexapole collision cell. The collision gas was argon at a pressure of approximately $5 \times 10^{-3}$ mbar and a collision energy of 30 V. The C-terminal fragments of the y-series were more intense than the N-terminal fragments of the b-series. High mass fragments generally appeared as doubly or triply charged ions. Also weak fragment ions corresponding to the Thr → Ser and Asp → Glu replacement were included in the list or explicitly indicated as not detected (ND).

| b<sub>n</sub> | GalNAc | b<sub>n</sub> | GalNAc | b<sub>n</sub> | GalNAc | b<sub>n</sub> | GalNAc |
|----------------|--------|----------------|--------|----------------|--------|----------------|--------|
| 1 P | 276.1 | 479.1/465.1 | 2 P | T/S | 465.1 | 3 P | D/E | 3 |
| 4 G | 323.1 | 488.1 | 5 P | A | 488.1 | 6 T | S | 6 |
| 7 A | 582.1 | 1153.3 | 1153.3 | 12 T | 1711.8 | 12 T | 1711.8 |
| 8 P | 903.2 | 1512.7 | 9 G | H | 1512.7 | 10 A | 9 G |
| 11 H | A | 1512.7 | 12 G | P | 1512.7 | 13 V | P | 1512.7 |
| 14 T | A | 1711.8 | 13 T | A | 1711.8 | 15 S | T | 1711.8 |
| 16 A | T | 1914.9 | 15 T | T | 1914.9 | 16 A | T | 1914.9 |
| 17 P | 1496.1 | 1719.7 | 1719.7 | 18 T | 2125.9 | 18 T | 2125.9 |
| 18 D/E | 1611.1/ND | 1922.8 | 1922.8 | 19 T/S | 2329.0 | 19 T/S | 2329.0 |
| 20 R | 2523.2 | 2735.2 | 20 R | P | 2735.2 | 21 R/S | 2902.4 | 21 R/S |

| y<sub>n</sub> | R | T/S |
|----------------|----|-----|
| 1 | 1 | 2 |
| 2 | 3 | 4 |

*ND, not determined

For each mass signal, the corresponding y<sub>2</sub> and y<sub>4</sub> ions were obtained for the major b- and y-series. The ions at m/z 479.1 (y<sub>2</sub>) and 691.2 (y<sub>4</sub>) were used as representative examples in Table III. The ions of the b- and y-series were in agreement with the known PAP20 sequence, and for Thr-19 the substitution with HexNAc could be demonstrated by registration of the y<sub>4</sub> ion at m/z 577.1, probably due to the Asp → Glu replacement, which compensates the mass difference of 14 units.

**High Density of O-Glycosylation on Peptide Variants as Revealed by Edman Degradation**

To identify the glycosylated sites and to quantify the degree of glycosylation for each individual site, the PAP20 glycopeptide mixture was subjected to Edman sequencing (Figs. 6 and 7). It has been shown previously that glycosylated phenylthiohydantoin derivatives of serine or threonine are recovered in a pulsed liquid-phase sequenator and can be quantified, using appropriate correction factors (19). The determined sequence was identical to that expected for PAP20 peptides (Figs. 6 and 7), and no unrelated sequences were detectable. However, Edman degradation revealed clear evidence for variant sequences of the repeat peptide (see below) that accounted for a series of mass signals at m/z 479.1 (y<sub>2</sub>) and 691.2 (y<sub>4</sub>) demonstrating that within the fraction of PAP20-GalNAc<sub>2</sub> at least one substitution isomer carries GalNAc at Thr-19 (Fig. 5). This finding implies that the DTR motif is a glycosylation target on VNTR-peptides with no Ser-5 glycosylation and substitution with GalNAc. The other sites did not show preferential glycosylation with GalNAc or Gal-GalNAc.

About one-third of Pro-9 was replaced by Ala, and as much as half of Thr-19 by Ser (Figs. 6 and 7). Information on Asp-18 could not be obtained due to the carbodiimide used during coupling of the glycopeptide. Referring to Table II, consider...
ation of the above amino acid replacements gives rise to additional series of mass matches for PAP20 glycopeptides with variant sequences. Due to the coincidence of Asp → Glu and Thr → Ser exchanges revealed by mass spectrometry, several combinations of amino acid replacements in the variant PAP20 glycopeptides can be ignored.

**DISCUSSION**

Both findings reported in this contribution can be regarded to have implications of biological relevance, because the structural modifications concern the antigenicity of an immunodominant target site on MUC1, a mucin with high impact in tumor biology. Earlier evidence from other laboratories had suggested that MUC1 on breast cancer cells is underglycosylated compared with normal epithelial cells (4) and that this reduced substitution of the peptide scaffold should unmask the DTR motif for the binding of antibodies (8). We demonstrate in this study that the breast cancer cell line T47D glycosylates the MUC1 tandem repeat peptide at higher density than lactating breast epithelia and that even the DTR motif is substituted in >90% of the repeats. Expectedly, glycosylation of the motif should reduce or abolish binding of DTR-specific antibodies, such as SM-3. However, we were previously able to demonstrate a strong enhancement of DTR antigenicity, if the motif in a series of synthetic glycopeptides was substituted with core-type glycans (11). DTR glycosylated peptides remained effective immunotargets after glycosylation of all proximal sites (11). In the light of the present report, these previous findings from *in vitro* binding studies now appear to have biological relevance. The results suggest that tumor cells bind DTR-specific antibodies more effectively, because modification of the motif with short glycans stabilizes the epitope in a favorable way. Antigenicity of the MUC1 tandem repeat peptide should also be altered by the reported amino acid replacements, particularly because two sites of the DTR motif in T47D-MUC1 are exchanged at high incidence. Further studies should reveal whether the existence of these variant sequences is a general feature of carcinoma cells. This finding would have a strong impact on cancer immunology of the mucin.

The material under study can be regarded as representative, because 1) the cell line T47D is a well established model with respect to O-glycosylation of MUC1 (3–5) and to the binding of MUC1-specific antibodies (25), and 2) it comprised all MUC1 glycoforms secreted by T47D cells. The latter statement is founded on consideration of the binding characteristics of anti-MUC1 antibody BC3, which recognizes a broad spectrum of glycoforms of the tandem repeat peptide (9, 23). It is also supported by the previous finding that glycosylation of the repeat peptides within individual MUC1 molecules is not uniform (17) but gives rise to statistical variations. Hence, the broad reactivity pattern of BC3 covers also strikingly different glycoforms of the mucin, such as lactation- and tumor-associated MUC1 (9). Moreover, BC3 does not belong to the group of antibodies with enhanced binding to glycosylated DTR (11), which could have led to the affinity enrichment of particular glycoforms of the mucin with high density of O-glycosylation. The efficiency of affinity chromatography was confirmed by immunochromatography analysis of all flow-through and wash fractions using a panel of anti-MUC1 antibodies with distinct fine specificities. The possibility that highly glycosylated peptides were enriched due to proteolytic digestion of poorly glycosylated peptide regions during exoglycosidase treatments can be excluded, because synthetic tandem repeat peptides were stable under the same conditions.

There should be a mechanism ruling the density of site occupation during initial O-glycosylation. We recently obtained evidence that glycosylated peptides may represent better substrates for the ppGalNAc-Ts than nonglycosylated peptides (26). This holds particularly true for vicinal glycosylation of Ser/Thr diads within the MUC1 repeat peptide, when using GalNAc-substituted peptide substrates. Negative effects were observed with glycopeptide substrates carrying core1-disaccharides in vicinal or proximal positions of the glycosylation target site (26). The antagonistic effects on the mono- and disaccharide levels point to a competition between ppGalNAc-Ts and the core enzymes αGalNAc-R:β-3-galactosyltransferase (core1 enzyme) and Galβ1–3GalNAc-R:β-6-glucosaminyltransferase (core2 enzyme). Such a competition could underlie the reported differences in glycosylation densities, when comparing lactation- and tumor-associated glycoforms of the mucin (26). In T47D cells, no functional core2 enzyme is expressed (6), and initial O-glycosylation can proceed to completion before inhibitory glycan substituents (sialylated core1) are formed in the trans-Golgi. One of the enzymes involved in initial O-glycosylation, the recently described rGalNAc-T4, which adds GalNAc site-specifically to the DTR motif, was found to prefer substrates carrying GalNAc at the proximal sites (24). However, we could demonstrate that the DTR motif is glycosylated *in vivo* to a significant extent even in the fraction of PAP20-GalNAc<sub>α</sub> pointing to the existence of other enzymes exhibiting similar site specificity, but no dependence on prior glycosylation.

The conserved VNTR domain of MUC1 expressed by T47D...
cells contains variants of the accepted sequence of tandem repeat peptide exhibiting amino acid replacements at three sites (boldface): PAPGSTPAAHGTVTASPESR. Interestingly, neither of the replacements seems to have influence on the degree of glycosylation of the repeat peptide. In particular, the DTR → ESR exchange does not result in decreased glycosylation at Ser-19, as could have been predicted from in vitro glycosylation studies (27). Also the nonconservative exchange of Pro by Ala in the critical position +3 (28) did not influence glycosylation at Thr within the STA motif. In conclusion, the data obtained for in vivo derived material are in obvious contradiction to predictions from in vitro studies as revealed previously for the milk mucin (17). The relative abundance of the sequence variants can be estimated to sum up for about 50% of the total repeats, indicating that they should not result from cell-specific mutations or from cleavage of degenerated repeats flanking the conserved domain. In previous sequencing studies on the DNA level the variant peptides were missed, because direct sequence information was obtained only for a few peripheral repeats of the conserved VNTR domain (29–32). Interestingly, a partial DNA corresponding to internal regions of the MUC1 VNTR domain from MCF-7 cells agrees with respect to the replacements described in our study (33). Moreover, cDNA clones derived from pancreatic tumor cells (HPAF) revealed the same pattern of replacements as detected in T47D-MUC1. A similar polymorphism had not been observed on analysis of MUC1 repeat peptides from pooled human milk samples (17).

Because our study was performed on a model cell line, there is a need to expand this work on other cell lines derived from the same organ site and tumor type. A comparative study is in progress using recombinant glycosylation probes, which are expressed as secretory fusion proteins in a panel of human breast cancer cells.

REFERENCES

1. Ha, S. B., and Kim, Y. S. (1991) Semini. Cancer Biol. 2, 389–400
2. Hull, S. R., Bright, A., Carraway, K. L., Abe, M., Hayes, D. F., and Kufe, D. W. (1989) Cancer Commun. 1, 261–267
3. Hanisch, F.-G., Stadie, T. R. E., Deutzmann, F., and Peter-Katalinic, J. (1996) Eur. J. Biochem. 236, 318–327
4. Lloyd, K., Burchell, J., Kudryashov, V., Yin, B. W. T., and Taylor-Papadimitriou, J. (1996) J. Biol. Chem. 271, 33325–33334
5. Brockhausen, I., Yang, J. M., Burchell, J. M., Whithouse, C., and Taylor-

2 S. Müller and F.-G. Hanisch, unpublished and 5th Workshop on Carcinoma-associated Mucins, Cambridge, United Kingdom, 1998.

Papadimitriou, J. (1999) Eur. J. Biochem. 233, 607–617
6. Hanisch, F.-G., Uhlenbruck, G., Peter-Katalinic, J., Egge, H., Dabrowski, J., and Dabrowski, U. (1989) J. Biol. Chem. 264, 872–883
7. Whitehouse, C., Burchell, J., Geschmiesser, S., Brockhausen, I., Lloyd, K. O., and Taylor-Papadimitriou, J. (1997) J. Cell Biol. 137, 1229–1241
8. Burchell, J., and Taylor-Papadimitriou, J. (1993) Epith. Cell Biol. 2, 155–162
9. M. R., et al. (1997) Tumor Biol. 19, 1–15
10. Petrarca, C., Huang, R., Rahimi, H., Dugostini, F., Turchi, V., Gheti, C. A., Scambia, G., Frati, L., and Nuti, M. (1996) Eur. J. Cancer 32, 2155–2163
11. Karsten, U., Dietel, C., Klich, G., Paulsen, H., Goletz, S., Müller, S., and Hanisch, F.-G. (1998) Cancer Res. 58, 2541–2549
12. Nishimori, I., Johnson, N. R., Sanderson, S. D., Perini, R., Mountjoy, K. Cerny, R. L., Gross, M. L., and Hollingsworth, M. A. (1994) J. Biol. Chem. 269, 16120–16130
13. Stadie, T. R. E, Chai, W., Lawson, A. M., Byfield, P. G. H., and Hanisch, F.-G. (1995) Eur. J. Biochem. 229, 140–147
14. Wandall, H. H., Hassan, H., 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) J. Biol. Chem. 272, 23503–23514
15. Nehrke, K., Hagen, F. K., and Tabak, L. A. (1996) J. Biol. Chem. 271, 7061–7065
16. Hansen, J. E., Lund, O., Engelbrecht, J., Bohr, H., Nielsen, J. O., Hansen, J.-E. S., and Brunak, S. (1995) Biochem. J. 308, 801–813
17. Müller, S., Goletz, S., Packer, N., Gooley, A. A., Lawson, A. M., and Hanisch, F.-G. (1997) J. Biol. Chem. 272, 24780–24793
18. Avling, K., Körner, R., Paulsen, H., and Peter-Katalinic, J. (1998) J. Mass Spectrom. 33, 1124–1133
19. Zachara, N. E., and Gooley, A. A. (1998) in Mucin Methods and Protocols. Methods in Molecular Biology (Fearfield, T., ed) Humana Press, Totowa, NJ, in press
20. Harrison, M. J., Nouwens, A. S., Jardine, D. R., Zachara, N. E., Gole, A. A., Nevalainen, H., and Packer, N. H. (1998) Eur. J. Biochem. 256, 119–127
21. Murre, E. R., and Poppe, I. (1994) Methods Enzymol. 230, 1–15
22. Hanisch, F.-G., Hanski, C., and Hasegawa, A. (1992) Cancer Res. 52, 3138–3144
23. Hanisch, F.-G. (1998) Tumor Biol. 19, 111–117
24. Bennett, E. P., Hassan, H., Mandel, U., Mirgorodskaya, E., Roepstorff, P., Burchell, J., Taylor-Papadimitriou, J., Hollingsworth, M. A., Merks, G., van Kessel, A. G., Eiberg, H., Stelfens, R., and Clausen, H. (1998) J. Biol. Chem. 273, 30472–30481
25. Cao, Y., Karsten, U., and Hilgers, J. (1998) Tumor Biology 19, 88–99
26. Hanisch, F.-G., Müller, S., Hassan, H., Clausen, H., Zachara, N., Gole, A. A., Paulsen, H., Avling, K., and Peter-Katalinic, J. (1999) J. Biol. Chem. 274, 9446–9454
27. Wang, Y., Abernethy, J. L., Eckhardt, A. E., and Hill, R. L. (1992) J. Biol. Chem. 267, 12709–12716
28. Wilson, I. B. H., Gavel, Y., and van Heijne, G. (1991) Biochem. J. 275, 529–534
29. Gendler, S. J., Lancaster, C. A., Taylor-Papadimitriou, J., Dubig, T., Peat, N., Burchell, J., Pemberton, L., Lalani, E., and Wilson, D. (1990) J. Biol. Chem. 265, 15286–15293
30. Lichtenberg, M. J. L., Vas, H. L., Gennissen, A. M. C., and Hilkens, J.-E. S., and Brunak, S. (1995) Eur. J. Biochem. 229, 1–15
31. Riggs, R., and Taylor-Papadimitriou, J., Hollingsworth, M. A., Burchell, J., Taylor-Papadimitriou, J., and Clausen, H. (1997) J. Biol. Chem. 273, 30472–30481
32. 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) Eur. J. Biochem. 189, 463–473
33. Siddiqui, J., Abe, M., Hayes, D., Shani, E., Yunis, E., and Kufe, D. (1988) Proc. Natl. Acad. Sci. U. S. A. 85, 2520–2523