Recombinant MUC1 mucin with a breast cancer-like O-glycosylation produced in large amounts in Chinese-hamster ovary cells

Malin BÄCKSTRÖM*, Thomas LINK†, Fredrik J. OLSON*, Hasse KARLSSON*, Rosalind GRAHAM‡, Gianfranco PICCO‡, Joy BURCHELL‡, Joyce TAYLOR-PAPADIMITRIOU‡, Thomas NOLL† and Gunnar C. HANSSON*¹

*Department of Medical Biochemistry, Göteborg University, P.O. Box 440, 405 30 Gothenburg, Sweden, †Institute of Biotechnology, Forschungszentrum Jülich GmbH, D-52425 Jülich, Germany, and ‡Cancer Research UK, Breast Cancer Biology Group, Guy's Hospital, Thomas Guy House, London SE1 9RT, U.K.

We have developed an expression system for the production of large quantities of recombinant MUC1 mucin in CHO-K1 (Chinese-hamster ovary K1) cells. The extracellular part of human MUC1, including 16 MUC1 tandem repeats, was produced as a fusion protein with murine IgG Fc, with an intervening enterokinase cleavage site for the removal of the Fc tail. Stable MUC1–IgG-producing CHO-K1 clones were generated and were found to secrete MUC1–IgG into the culture medium. After adaptation to suspension culture in protein-free medium in a bioreactor, the fusion protein was secreted in large quantities (100 mg/l per day) into the culture supernatant. From there, MUC1 could be purified to homogeneity using a two-step procedure including enterokinase cleavage and ion-exchange chromatography. Capillary liquid chromatography MS of released oligosaccharides from CHO-K1-produced MUC1 identified the main O-glycans as Gal β 1-3GalNAc (core 1) and mono- and di-

sialylated core 1. The glycans occupied on average 4.3 of the five potential O-glycosylation sites in the tandem repeats, as determined by nano-liquid chromatography MS of partially deglycosylated Clostripain-digested protein. A very similar O-glycan profile and site occupancy was found in MUC1–IgG produced in the breast carcinoma cell line T47D, which has O-glycosylation typical for breast cancer. In contrast, MUC1–IgG produced in another breast cancer cell line, MCF-7, showed a more complex pattern with both core 1- and core 2-based O-glycans. This is the first reported production of large quantities of recombinant MUC1 with a breast cancer-like O-glycosylation that could be used for the immunotherapy of breast cancer.

Key words: glycoprotein, glycosylation, oligosaccharide, mass spectrometry (MS), recombinant protein.

INTRODUCTION

The MUC1 mucin (CD227) is a highly O-glycosylated trans-membrane protein, largely built by tandemly repeated mucin domains, each with five O-glycosylation sites. MUC1 was first identified as an epithelial cell protein [1], but has recently also been found at low levels on cells of haematopoietic origin [2–4]. The expression of MUC1 is elevated in many forms of epithelial cancers, including those of the breast, lung, cervix, colon and pancreas [5,6].

MUC1 is aberrantly glycosylated in tumour cells. The O-glycans expressed on MUC1 from normal cells consist of long and branched, mainly core 2-based, oligosaccharides [7,8]. In contrast, MUC1 expressed by breast carcinoma cells has an increased proportion of shorter, non-branched, core 1-based O-glycans with a higher sialic acid content [9–11]. The cancer-derived MUC1 also has, in some cases, a higher occupancy of the O-glycosylation sites in the TRs (tandem repeats) compared with MUC1 from normal cells [12]. The altered O-glycan profile of tumour MUC1 can be partly explained by the enhanced expression of the sialyl-transferase ST3Gal-I in breast cancer cells [13,14], which competes with the core 2-forming enzyme (C2GnT1) for their common core 1 substrate (Gal β 1-3GalNAc).

The under-glycosylation of MUC1 in breast cancer cells results in the exposure of normally cryptic peptide core epitopes and some cancer patients have IgMs and IgGs specific for the MUC1 peptide core in their sera [15,16]. These circulating antibodies

are associated with a good prognosis [17]. Moreover, a report that non-MHC-restricted cytotoxic T-cells could be isolated that recognized an epitope in the TR sequence of MUC1, which is masked in the normal mucin, suggested that cellular immune responses may also be induced [18]. For these reasons, there is considerable interest in using MUC1-based immunogens for immunotherapy, particularly for breast cancer patients. The first clinical trials initiated used formulations based on unglycosylated TR peptides of MUC1 [19–21]. A recombinant MUC1 glycoprotein carrying tumour-associated glycans would be an important form of MUC1 to test as an immunogen, but has so far not been available. In the present study, we report the recombinant production of a MUC1 glycoprotein, in which the extracellular part of human MUC1 was fused with murine IgG Fc and produced in CHO-K1 (Chinese-hamster ovary K1) cells. The protein was secreted into the culture medium from where it could be readily purified. The structures of the O-glycans and the occupancy of O-glycosylation sites in the MUC1 TR were determined and the glycoform was found to closely resemble that of recombinant MUC1–IgG produced by the breast cancer cell line T47D.

MATERIALS AND METHODS

Construction of expression plasmids for MUC1–IgG

cDNA encoding the full-length human MUC1 was subcloned into pcDNA3 (Invitrogen, Baltimore, MD, U.S.A.), as a *Bam*HI/*Xho*I

Abbreviations used: TR, tandem repeat; mAb, monoclonal antibody; LC-ESI MS, liquid-chromatography electrospray-ionization MS; ESI-MS/MS, electrospray-ionization tandem MS; DELFIA, dissociation-enhanced lanthanide fluorescence immunoassay; PNA, peanut agglutinin; HPA, *Helix pomatia* agglutinin; WGA, wheatgerm agglutinin; CHO, Chinese-hamster ovary; i.d., inner diameter.

¹ To whom correspondence should be addressed (e-mail gunnar.hansson@medkem.gu.se).

fragment, using T4 DNA ligase (Invitrogen), giving the plasmid pcDNA3/MUC1. Exon 1–3 of the murine IgG_{2a} Fc domain was amplified by PCR (from an expression plasmid kindly provided by Dr B. Ardman, Department of Medicine, Tupper Research Institute, New England Medical Center Hospitals, Boston, MA, U.S.A.) using the oligonucleotide primers 5'-TT TCA ACC TCA GGC CCA GAT GAT GAT GAT AAA GCA GAG CCC AGA GGG CCC ACA-3' and 5'-TTCT TCT AGA TCA TTT ACC CGG AGT CCG GGA GAA-3', which also encoded the enterokinase-recognition sequence DDDDK (where single-letter amino-acid notation used). The PCR product was digested with *Bsu*36I and *Xba*I and subcloned into pcDNA3/MUC1, replacing all parts of the *MUC1* gene that are downstream of the TR region. The ligated construct was electroporated into *Escherichia coli* XL10 Gold (*RecA*⁻; Stratagene, Cedar Creek, TX, U.S.A.) and correct clones were selected by DNA-restriction analyses. Due to recombination events taking place during propagation of the final plasmid in *E. coli*, the resulting plasmid [named pcDNA3/MUC1(16TR)-IgG; see below] contained fewer than the expected 32 TRs. This was later corrected by replacing a *Bam*HI/*Bsu*36I fragment, encoding MUC1, with the corresponding fragment from the original MUC1 full-length clone, to obtain pcDNA3/MUC1(32TR)-IgG. The sequences of all mutated regions were confirmed by DNA sequencing (MWG Biotech, Ebersberg, Germany). Restriction enzymes were from Invitrogen or New England Biolabs (Beverly, MA, U.S.A.) and were used according to the manufacturers' instructions.

Cell culture and transfections

CHO-K1 (ATCC CCL-61) and the breast cancer cell lines T47D and MCF-7 were cultured in T flasks in Iscove's modified Dulbecco's medium (Invitrogen) supplemented with 10 % foetal bovine serum, penicillin and streptomycin (Invitrogen). CHO-K1 cells were transfected with pcDNA3/MUC1(16TR)-IgG by electroporation or with pcDNA3/MUC1(32TR)-IgG by DMRIE-C liposomes (Invitrogen). T47D and MCF-7 were transfected with pcDNA3/MUC1(32TR)-IgG by electroporation or using Fugene-6 (Roche, Indianapolis, IN, U.S.A.), respectively. To generate stable clones, 250 mg/l Geneticin (Invitrogen) was added on day 3 after transfection and clones were selected based on high secretion of MUC1-IgG into the culture medium, as determined by anti-mouse Ig-DELFA (dissociation-enhanced lanthanide fluorescence immunoassay; see below) and Western blot using anti-mouse Ig antibodies (Dako, Glostrup, Denmark). For the large-scale production of MUC1(16TR)-IgG, one CHO-K1 clone was transferred to suspension growth in the protein-free medium ProCHO4-CDM (BioWhittaker, Verviers, Belgium) in a 2 litre perfusion culture bioreactor. The culture supernatants containing MUC1-IgG were stored at 4–8 °C during collection and at –20 °C for longer storage.

Southern blotting

Genomic DNA was prepared from the stable MUC1-IgG-secreting CHO-K1 clone using a DnEasy kit (Qiagen). Genomic DNA (10 µg) was digested with *Hinf*I or *Eco*RI, fractionated on a 0.7 % agarose gel and transferred to a Hybond N membrane (Amersham Biosciences, Uppsala, Sweden). The blot was hybridized with a [α -³²P]dCTP-labelled 500 bp probe to the TR region of MUC1, washed to high stringency and then exposed to a phosphor screen overnight at room temperature. The results were visualized using a STORM scanner system.

Purification of MUC1 and MUC1-IgG

Culture supernatant from protein-free perfusion culture of CHO-K1 MUC1-IgG was concentrated through ultrafiltration using an Amicon YM100 filter with a 100 kDa cut-off (Millipore, Bedford, MA, U.S.A.). MUC1-IgG was cleaved by incubating with 100 units of enterokinase (EK_{Max}; Invitrogen) per mg of MUC1-IgG for at least 50 h in EK_{Max} buffer (50 mM Tris/HCl, pH 8.0, 0.1 % Tween 20 and 1 mM CaCl₂) at 37 °C. The mixture of cleaved protein was then loaded on to a HiPrep 16/10 Q FF anion-exchange column connected to an Äkta purifier (both from Amersham Biosciences) at 1 ml/min. After washing out non-bound proteins with a 2 ml/min flow of 50 mM Tris/HCl, pH 8.0, bound proteins were eluted with a 0–0.5 M NaCl gradient for 90 min with a plateau for 20 min at 0.15 M NaCl. Fractions of 3 ml were collected during elution. Uncleaved MUC1-IgG was purified using a similar procedure without prior incubation with EK_{Max}.

MCF-7 and T47D MUC1-IgG was purified by adding anti-mouse Ig-agarose beads (Sigma-Aldrich, St Louis, MO, U.S.A.) directly to the collected culture medium (with 10 % foetal bovine serum) from the respective stable clones and left turning at 4 °C for 16–24 h. Beads were then transferred to empty glass columns and washed with 10 mM sodium phosphate buffer, pH 7.2/0.5 M NaCl. MUC1-IgG was eluted using 0.1 M glycine/HCl, pH 2.4/0.5 M NaCl. Fractions collected were immediately neutralized using 1 M Tris/HCl, pH 9.0. MUC1-IgG-containing fractions were pooled and dialysed against water before analyses.

SDS/PAGE

Proteins were separated under reducing conditions on 4–15 % or 3–10 % polyacrylamide gels with 0.1 % SDS. Proteins were stained by silver staining or with 0.125 % Alcian Blue (Sigma-Aldrich) in 25 % (v/v) ethanol and 10 % (v/v) acetic acid.

Immunoassays

MUC1-IgG concentrations were determined using a DELFIA adapted from Bäckström et al. [22], using rabbit-anti-mouse Ig (Dako) for coating and europium-labelled rabbit anti-mouse IgG (H + L; Jackson ImmunoResearch, West Grove, PA, U.S.A.) for detection. Standards and samples were incubated in the wells for 3 h. Purified MUC1-IgG subjected to quantitative amino acid and monosaccharide analyses was used as a standard.

Reactivities with PNA (peanut agglutinin), WGA (wheatgerm agglutinin) and HPA (*Helix pomatia* agglutinin) were tested in a DELFIA coated with anti-MUC1 mAb (monoclonal antibody) HMFG-1. After blocking with 2 % BSA, sequential incubations were performed with proteins in fractions from HiPrep Q FF, biotinylated lectins (Sigma-Aldrich) at 2 µg/ml and europium-conjugated streptavidin and developed as described in [22].

Reactivities with the MUC1-specific mAbs HMFG-1, HMFG-2 or SM3 were made in wells of ELISA plates coated with 500 ng of pure MUC1-IgG and blocked with 1 % BSA. After sequential incubations with biotinylated mAbs and peroxidase-conjugated streptavidin (Dako), plates were developed using *o*-phenylenediamine (Sigma-Aldrich) and the absorbance at 450 nm was determined using a Dynex MRX II (Jencons).

For slot-blot analysis, equal volumes of relevant fractions from HiPrepQ FF were slot blotted on to Immobilon-P PVDF membrane (Millipore). Glycoproteins bound to the membrane

were deglycosylated in trifluoromethanesulphonic acid (Aldrich, Milwaukee, MI, U.S.A.) for 6 h at 0 °C and then developed with anti-MUC1 HMFG-1 mAb followed by alkaline phosphatase-conjugated goat anti-mouse IgG (Dako) and 5-bromo-4-chloroindol-3-yl phosphate/Nitro Blue Tetrazolium (BCIP/NBT; Promega, Madison, WI, U.S.A.). The intensities of the slots were quantified using video densitometry.

Amino acid analysis

The amino acid composition of 20–50 µg of purified MUC1 was determined by the Amino Acid Analysis Centre, Department of Biochemistry, Uppsala University, Uppsala, Sweden.

Release of O-linked oligosaccharides

O-linked oligosaccharides from 20–25 µg of purified MUC1–IgG were released by reductive β -elimination by a protocol adapted from Schulz et al. [23].

To obtain standards of the different oligosaccharides, O-glycans from 30 mg of CHO-K1 MUC1(16TR)–IgG were released as described previously [24]. The Gal β 1-3GalNAc, NeuAc α 2-3Gal β 1-3GalNAc and NeuAc α 2-3Gal β 1-3(NeuAc α 2-6)GalNAc oligosaccharides were isolated individually by HPLC on a Hypercarb column (100 mm $\times$ 4.6 mm, 5 µm particles; Thermo-Hypersil, Runcorn, Cheshire, U.K.) connected to an Äkta purifier. A gradient of 0–40 % acetonitrile in 20 mM NH $_4$ HCO $_3$ at 1 %/min was used at a flow of 0.5 ml/min. The amount of each oligosaccharide was determined using monosaccharide composition analysis as described previously [25].

Preparation of MUC1 TR glycopeptides for determination of glycosylation-site occupancy

MUC1 TR glycopeptides were generated by a protocol modified from Müller et al. [12]. MUC1–IgG (30–50 µg) from CHO-K1 or T47D cells was bound to 100 µl of anti-mouse Ig-agarose beads (Sigma-Aldrich) and deglycosylated using sequential incubations at 37 °C with 50 m-units of *Vibrio cholerae* neuraminidase (Roche) in 0.05 M sodium acetate, pH 5.5, with 4 mM CaCl $_2$ and 100 µg/ml BSA for 22 h and 50 m-units of β -galactosidase (from bovine testes; Sigma-Aldrich), in 0.1 M citrate phosphate buffer, pH 5.0, for 22 h. An additional incubation (after the neuraminidase treatment) with α 1-2-fucosidase (1 m-unit) and α 1-3,4-fucosidase (1 m-unit) in 25 mM sodium phosphate, pH 5.0, for 20 h was performed for the T47D-produced protein. The partially deglycosylated MUC1–IgG was released from the agarose beads using 0.1 M glycine/HCl, pH 2.4/0.15 M NaCl, neutralized using 1 M Tris/HCl, pH 9.0, dialysed against water, lyophilized and dissolved in 25 mM sodium phosphate, pH 7.6/0.2 mM CaCl $_2$. Clostripain (3 units; Sigma-Aldrich), which had been pre-activated with 2.5 mM dithiothreitol for 4 h at 20–25 °C, was added and allowed to digest the protein for 20 h at 37 °C.

LC-ESI MS (liquid-chromatography electrospray-ionization MS)

LC-ESI MS and ESI-MS/MS (electrospray-ionization tandem MS) were performed on a Q-ToF mass spectrometer (Micromass, Manchester, U.K.) in negative-ion mode for oligosaccharide alditols and positive-ion mode for glycopeptides. The desalted oligosaccharide alditols (1 µl of 10 µl) were subjected to capillary-LC-ESI MS or LC-ESI MS/MS on a 150 mm $\times$ 0.25 mm i.d. (inner diameter) column packed with 5 µm Hypercarb particles

(Thermo-Hypersil). A HP1100 binary pump (Hewlett-Packard, Palo Alto, CA, U.S.A.) was used and the flow was split in a Valco Tee (VICI, Schenkon, Switzerland) using 40 cm $\times$ 50 µm i.d. fused-silica tubing as a restrictor, down to a flow rate of 5 µl/min of 5 mM NH $_4$ HCO $_3$ through the column. An acetonitrile gradient of 1 % acetonitrile/min was used and the column eluent was transferred via 50 µm i.d. fused-silica tubing into a nano-LC interface (Micromass). Electrospray capillary voltage was –3.5 kV, sample cone voltage was 38 V and sample cone temperature was 100 °C; compressed air was used as the cone gas at a flow rate of 300 l/h and as the nebulizer gas at 0.3 MPa. In LC MS, full scans were performed over the m/z range 200–1500. For LC MS/MS, automatic function switching was used for colliding [M–H] $^-$ ions at intensities higher than 20 counts/s. Argon was used as collision gas with a collision energy of 22 eV for m/z 300–600, 30 eV for m/z 600–900, 45 eV for m/z 900–1100 and 60 eV for m/z 1100–2000. O-glycans were quantified by integrating the individual peaks in mass chromatograms generated by MassLynx 3.4 software (Micromass).

Samples containing known concentrations of the different oligosaccharide standards were analysed by capillary-LC MS as above. The integrated peak areas were calculated and used to obtain correction factors for the signals obtained from neutral, mono-sialylated and di-sialylated oligosaccharides. The ratio between the areas obtained from equal concentrations of Gal β 1-3GalNAc (m/z 384) and NeuAc α 2-3Gal β 1-3GalNAc (m/z 675) were determined to 0.4 ± 0.099 (mean $\pm$ S.D., $n = 11$). The corresponding value when comparing Gal β 1-3GalNAc (m/z 384) and NeuAc α 2-3Gal β 1-3(NeuAc α 2-6)GalNAc (m/z 966) was determined to 0.3 ± 0.10 ($n = 11$). To obtain the relative amounts of the different oligosaccharides, the areas for mono-sialylated species were multiplied by 0.4 and for di-sialylated with 0.3.

MUC1 glycopeptides were analysed by nano-LC-ESI MS; 0.3 µl (of 30 µl) was injected using an M-435 microinjection valve (Upchurch Scientific, Oak Harbor, WA, U.S.A.) into a fused-silica 120 mm $\times$ 0.075 mm i.d. column, packed with Kromasil C $_{18}$ -bonded 3.5 µm porous particles [100 Å (where 1 Å $\equiv$ 0.1 nm); Erka Chemicals, Bohus, Sweden]. PEEK tubing (400 µm i.d.; Upchurch Scientific) was used to seal the fused-silica column in a 1.58 mm Valco through-bore union (Valco Instruments, Houston, TX, U.S.A.) against a 2 µm steel screen (Valco Instruments). A tapered emitter tip made of fused silica (150 µm outer diameter, 20 µm i.d.; Polymicro Technologies, Phoenix, AZ, U.S.A.) was connected to the union by a modified Micromass nano-LC Option using 150 µm i.d. PEEK tubing. The flow rate was split down to 250 nl/min. A gradient of 0–50 % acetonitrile in 0.1 % formic acid (1 % acetonitrile/min) was used. The electrospray capillary voltage was +3.2 kV, the sample cone voltage was 40 V and the sample cone temperature was 80 °C; nitrogen was used as the cone gas at a flow rate of 250 l/h and as the nebulizer gas at 0.2 MPa. Full scans were performed over the range m/z 400–1500. Mass chromatograms for [M + 3H] $^{3+}$ (and [M + 4H] $^{4+}$ when present) for the different MUC1 glycopeptides were generated and the peaks integrated using MassLynx 3.4.

RESULTS

Generation of stable CHO-K1 clones secreting MUC1–IgG

An expression plasmid was constructed where the extracellular part of MUC1, taken from a full-length human MUC1 clone with 32 TRs, was expressed from the cytomegalovirus promoter as a fusion protein with murine IgG Fc (Figure 1). When the plasmid was propagated in *E. coli*, the BamHI/Bsu36I fragment encoding the extracellular part of MUC1 was found to be smaller than

Figure 1 Schematic drawing of the MUC1-IgG construct used for the large-scale production in CHO-K1 cells

CMV, cytomegalovirus; EK, enterokinase.

Figure 2 MUC1-IgG CHO-K1 clone has a gene corresponding to 16TR of MUC1

Southern blot of restriction-enzyme-digested genomic DNA probed with a MUC1 TR-specific DNA probe. Lane 1, molecular-mass markers; lane 2, digestion with *HinfI* (predicted size for 16-TR MUC1 gene is 2.4 kb); lane 3, digestion with *EcoRI* (predicted size for 16-TR MUC1 gene is 6.4 kb).

expected (1.6 kb instead of 2.5 kb) because of a recombination involving the repetitive sequence of the TR domain. To determine the actual number of TRs in the recombined plasmid, genomic DNA isolated from a CHO-K1 clone transfected with the plasmid was digested with different restriction enzymes followed by a Southern blot using a TR-specific probe. The sizes of the genomic DNA fragments obtained after cleavage with *HinfI* or *EcoRI* were 2.4 and 6.4 kb respectively (Figure 2), indicating that the MUC1-IgG produced from the CHO-K1 clone contained 16 TRs. The recombined plasmid was later corrected to contain 32 TRs. Both plasmids (with 16 and 32 TRs) were used to transfect CHO-K1 cells and the MUC1-IgG fusion proteins were secreted from the cells into the culture medium. The two glycoproteins migrated as bands with the apparent molecular masses of 170 kDa (Figure 3A and 3B) and > 300 kDa (results not shown) and reacted with anti-mouse Ig as well as several MUC1-specific mAbs in Western blots (results not shown).

Stably transfected CHO-K1 clones were generated and selected for the secretion of maximum levels of MUC1-IgG. One CHO-K1 MUC1(16TR)-IgG clone was adapted to suspension culture in protein-free medium. When the clone was grown in a perfusion culture bioreactor with controlled pH and pO_2 , very high cell densities, accompanied by high concentrations of MUC1-IgG in the culture medium, were obtained. MUC1-IgG was the main protein component in the medium and, using these conditions, around 100 mg of MUC1-IgG/l per day could be produced. The purity was estimated to be already at least 80% at this stage, as determined by silver-stained SDS/PAGE (Figure 3A, lane 1). One CHO-K1 clone expressing MUC1(32TR)-IgG was adapted to the same suspension culture conditions, but never reached the same high productivity as the CHO-K1 MUC1(16TR)-IgG clone. The CHO-K1 16TR clone was therefore used for large-scale production.

Purification of MUC1-IgG and MUC1 from CHO-K1 cells

To obtain pure MUC1-IgG and MUC1 glycoproteins, the supernatant from CHO-K1 perfusion culture in protein-free medium was first concentrated and the buffer changed by ultrafiltration. From there, MUC1-IgG could be purified in a single step using anion-exchange chromatography (Figure 3A, lane 2). To obtain MUC1 without the IgG tail, enterokinase was added to the re-buffered culture supernatant to cleave off the IgG tail. To ensure that the digestion was complete, a small sample of the protein was subjected to SDS/PAGE before subsequent purification, and visualized by staining with Alcian Blue, which detects sulphated and sialylated glycoproteins. A size shift of around 30 kDa was observed, corresponding to the loss of the IgG tail (Figure 3B). Interestingly, after removal of the IgG tail, the MUC1 glycoprotein could not be visualized by silver staining, probably due to its high glycosylation (Figure 3B). The released MUC1 was then separated from the IgG tail by anion-exchange chromatography. Two main proteins eluted from the column, the first absorbing at both 280 and 215 nm, whereas the second only absorbed at 215 nm (Figure 3C). SDS/PAGE showed that the first peak, eluted at around 100 mM NaCl, contained IgG (30 kDa, silver stained) and the second, eluted at around 200 mM NaCl, contained MUC1 (130 kDa, stained with Alcian Blue; Figure 3D). Some MUC1 co-eluted with the IgG, as detected in DELFIA using HMFG-1 as the capturing antibody and different lectins for detection. PNA and HPA, specific for Gal-GalNAc and GalNAc respectively, detected the early-eluting MUC1. WGA, specific for terminal sialic acids, on the other hand did not, indicating that this MUC1 represented less-sialylated glycoforms (Figure 3E). The MUC1 that eluted in

Figure 3 Purification of MUC1-IgG and MUC1 using anion-exchange chromatography

(A) Purity of MUC1-IgG before (lane 1) and after (lane 2) anion-exchange chromatography using HiPrep Q FF. Proteins were separated by SDS/PAGE (4–15 % gels) under reducing conditions and detected by silver staining. (B) Staining of recombinant MUC1 by Alcian Blue and silver before (–) and after (+) enterokinase cleavage. Proteins were separated by SDS/PAGE (4–15 % gels) under reducing conditions and stained with either Alcian Blue or silver. (C) Chromatogram from HiPrep Q FF. Black and grey lines represent the absorbance at 214 and 280 nm respectively. Broken line represents the conductivity. (D) Fractions from HiPrep Q FF were analysed using SDS/PAGE (4–15 % gels) under reducing conditions and then stained with either Alcian Blue (MUC1) or silver (IgG Fc). (E) Fractions from HiPrep Q FF were analysed in DELFIA using HMFG-1 anti-MUC1 mAb for coating and biotinylated PNA (○), HPA (△) or WGA (■) to detect the different MUC1 glycoforms eluted in the different peaks. (F) Slot blot of equal volumes from each fraction from HiPrep Q FF on to PVDF membrane followed by chemical deglycosylation by trifluoromethanesulphonic acid and detection with HMFG-1 anti-MUC1 mAb. Intensities of the slots were quantified using video densitometry.

the second peak, however, was detected by WGA, but not by PNA and HPA lectins (Figure 3E) and was also stained with Alcian Blue (Figure 3D), indicating higher sialylation. The relative amounts of MUC1 in the different fractions were determined by slot blotting of chemically deglycosylated protein, using the mAb HMFG-1 anti-MUC1 for detection. Deglycosylation was performed as HMFG-1 reacts poorly with highly sialylated MUC1 [26]. An identical slot blot performed without deglycosylation showed a very weak reactivity of HMFG-1 with MUC1 in the second peak (results not shown). In contrast, the slot blot of deglycosylated proteins clearly showed that the majority of MUC1 was found in the second peak (Figure 3F). The pooled MUC1 protein in the second peak (fractions 31–43) was subjected to amino acid analysis and was shown to contain pure MUC1 (Table 1). Using this purification procedure, it was possible to obtain around 80 mg

of pure MUC1 glycoprotein from a single round of culture in a bioreactor using 18 litres of culture medium (Table 2).

CHO-K1-produced MUC1 reacts with mAbs specific for tumour MUC1

We analysed the ability of CHO-K1 MUC1-IgG to react with the MUC1-specific mAbs HMFG-1, HMFG-2 and SM3, all recognizing epitopes in the PDTRP sequence of the TR, but showing different glycosylation dependence. The glycoprotein reacted with all three mAbs, the weakest reaction being that of mAb HMFG-1 (Figure 4), which is negatively affected by the presence of sialic acids [26]. The binding of the mAbs SM3 and HMFG-2 is inhibited instead by long O-glycan chains, for mAb SM3 more

Table 1 Amino acid composition of purified recombinant MUC1

The theoretical value is based on 16 identical TRs with the sequence HGVTSAPDTRPAP-GSTAPPA. Amino acids in the TR are shown in bold. n.d., not determined.

Amino acid	Theoretical value (% by frequency)	Observed value (mol %)
Asx	6.21	7.24
Thr	14.6	13.3
Ser	12.6	12.4
Glx	1.71	3.55
Pro	21.2	18.0
Gly	9.85	9.50
Ala	17.3	16.0
Val	5.57	6.44
Met	0.21	0.24
Ile	0	0.65
Leu	1.07	2.14
Tyr	0	0.51
Phe	0	0.56
His	4.50	3.99
Lys	0.86	1.46
Arg	4.07	4.04
Trp	0.21	n.d.
Cys	0	n.d.

specifically by the presence of GlcNAc in the β 1-6 position of the core GalNAc [14]. Both SM3 and HMFG-2 were found to react well with the CHO-K1-produced MUC1. Taken together, this indicated that CHO-K1 MUC1 had O-glycans that are short and sialylated, as in breast cancer cells.

MUC1-IgG produced in CHO-K1 and T47D cells have almost identical O-glycan structures

To determine if the O-glycans present on CHO-K1-produced MUC1 had any similarity with those found on breast carcinoma cells, we generated stable MUC1-IgG-producing clones of two breast carcinoma cell lines, T47D and MCF-7. MUC1-IgG glycoproteins secreted from these clones were purified to high purity using affinity chromatography (Figure 5A).

The O-glycans on MUC1-IgG secreted from the different cell lines were released from the purified proteins using reductive β -elimination and analysed directly by capillary-LC-ESI MS and LC-ESI MS/MS. It was found that the main oligosaccharide from CHO-K1 MUC1-IgG was NeuAc α 2-3Gal β 1-3GalNAc (sialyl-T), making up 84 % of the O-glycans. Smaller amounts of NeuAc α 2-3Gal β 1-3(NeuAc α 2-6)GalNAc (disialyl-T, 13 %) and non-sialylated Gal β 1-3GalNAc (core 1 or T antigen, 1.9 %) were

Figure 4 Reactivity of CHO-K1-produced MUC1-IgG with MUC1-specific mAbs

Equal amounts of recombinant MUC1-IgG was coated in wells of microtitre plates and the three biotinylated MUC1-specific mAbs HMFG-2 (●), SM3 (■) and HMFG-1 (▲) were titrated in the wells and detected with streptavidin-horseradish peroxidase conjugate using o-phenylenediamine as the substrate.

also found (Figure 5B and Table 3). A similar composition of the O-glycans was found in MUC1-IgG from T47D cells, but with a higher proportion of the non-sialylated core 1 structure (32 %). For both CHO-K1 and T47D MUC1-IgG, all oligosaccharides had less than five monosaccharide units and all were core 1-based.

The O-glycans released from MUC1-IgG from MCF-7 cells were more heterogeneous than those found in CHO-K1 and T47D. A large number of oligosaccharides could be detected in the LC-MS chromatogram and of these, 17 could be completely or partially structurally determined using LC-MS/MS. Seven of the oligosaccharides were neutral and 10 were sialylated, and the size reached up to eight monosaccharide units. Only seven of the structures were core 1-based, together making up approx. 16 % of the O-glycans, the rest being core 2-based (Figure 5B and Table 3).

MUC1-IgG produced in CHO-K1 and T47D cells have similar O-glycan site occupancies

The number of O-glycosylation sites occupied in the MUC1 glycoproteins produced in CHO-K1 and T47D was also determined. Purified MUC1-IgG was enzymically deglycosylated, leaving only the core GalNAc residues directly attached to the polypeptide chain. This allowed the endoprotease Clostripain to cleave the proteins (at every Arg) into 20-amino-acid-long

Table 2 Yield of purified MUC1 from perfusion culture supernatant of CHO-K1 MUC1-IgG

The concentration of MUC1-IgG was determined by anti-mouse Ig DELFIA.

Sample	Volume (l)	Concentration MUC1-IgG (mg/l)	MUC1-IgG (mg)	MUC1 (mg)	Yield of MUC1 (% of starting material)
MUC1-IgG-containing culture supernatant	18	40*	720	540†	100 %
MUC1-IgG in concentrate after ultrafiltration	1.3	200	260	200†	37 %
Pure MUC1 after enterokinase cleavage and anion-exchange purification	—	—	—	83‡	15 %

* This value may be overestimated because the free IgG tail, which may be present in low amounts, is also detected by the anti-mouse Ig DELFIA.

† Calculated from the amount of MUC1-IgG, using the calculated molecular masses 73 kDa for MUC1(16TR)-IgG and 46 kDa for MUC1(16TR) and adjusting for the carbohydrate content (40 % in MUC1-IgG and 50 % in MUC1).

‡ Sum of the yields from several preparations, which were determined by amino acid analyses. The carbohydrate content, 50 %, was added to the mass.

Figure 5 Capillary-LC-ESI MS of released oligosaccharides from MUC1-IgG

(A) SDS/PAGE (3–10% gels) of MUC1(32TR)-IgG from T47D and MCF-7 breast cancer cell lines. Proteins were detected by silver stain. Lanes 1, spent culture medium from MUC1-IgG-secreting stable clones; lanes 2, non-bound proteins after incubation with anti-mouse Ig-agarose beads; lanes 3, purified MUC1-IgG (> 250 kDa) eluted from anti-mouse Ig-agarose beads. The additional band in purified T47D MUC1-IgG corresponds to the non-O-glycosylated precursor, as it is detected by anti-MUC1 mAbs in Western blot (results not shown). (B) Capillary-LC-ESI MS chromatograms of oligosaccharide alditols released from MUC1-IgG produced in CHO-K1, T47D and MCF-7. Peaks are annotated with mass and charge of the molecular ions corresponding to each component. For structural and quantitative data, see Table 3.

glycopeptides, corresponding to a single TR. The masses of the glycopeptides were then analysed directly by nano-LC-ESI MS and reflected the number of GalNAcs on the glycopeptides.

To investigate the potential influence of the number of TRs on the O-glycosylation site occupancy, MUC1-IgG glycoproteins containing either 16 or 32 TRs produced in CHO-K1 cells were analysed. The mass spectrum of glycopeptides from CHO-K1-produced MUC1-IgG (with both 16 and 32 TRs) showed TR peptides with three to five sites occupied by GalNAcs, the largest peak being that of peptides with all five sites occupied (Figure 6, Table 4). No large difference in the average occupancy of glycosylation sites between the proteins with different number of TRs could be identified. The corresponding mass spectrum of MUC1 glycopeptides from T47D showed a very similar occupancy of O-glycosylation sites as that of CHO-K1 cells, with the largest peak being that of all five sites occupied. In addition, a small peak corresponding to glycopeptides with two GalNAcs was also found (Table 4). The average occupancy of sites in MUC1-IgG from the different sources was calculated from the integrated peaks of MUC1 glycopeptides, and were determined to be 4.3 for CHO-K1 MUC1(16TR)-IgG, 4.6 for CHO-K1 MUC1(32TR)-IgG and 4.2 for T47D MUC1(32TR)-IgG.

DISCUSSION

We have generated a mammalian expression system for the production of large quantities of recombinant MUC1 in CHO-K1 cells. The protein was produced as a fusion protein with murine IgG_{2a} Fc in order to obtain active secretion from the cells, which simplified the purification. The produced MUC1 had an O-glycosylation pattern very similar to that of MUC1 from breast cancer cells, as represented in the extreme case by the metastatic breast cancer cell line T47D. The O-glycans occupied on average 4.3 of the five potential sites in the MUC1 TR and were structurally determined by capillary-LC MS to be the core 1-structure Galβ1-3GalNAc (T antigen) and mono- and di-sialylated Galβ1-3GalNAc.

One CHO-K1 MUC1-IgG clone was adapted to growth in suspension in protein-free medium in a perfusion culture bioreactor. The amounts of MUC1-IgG accumulating in the culture medium were then increased, giving an excellent production system for recombinant MUC1 that produced around 100 mg of MUC1-IgG/l per day. The MUC1-IgG in the culture medium was already quite pure and could easily be purified further with or without removal of the Fc tail, using the same anion-exchange protocol. From a single round of perfusion culture over 80 mg of pure MUC1 could be obtained, a figure which can be increased by optimizing the purification procedure. During the first step in the purification, concentrating the MUC1-IgG-containing culture medium by ultrafiltration, a substantial amount of MUC1-IgG was lost (Table 2). However, it is possible that the concentration of MUC1-IgG in the culture supernatant was over-estimated, due to the presence of low levels of free IgG Fc tail in the medium, caused by proteolytic cleavage during culture or storage. We have found that the sandwich immunoassay used to quantify MUC1-IgG detects free IgG Fc with a higher affinity than MUC1-IgG, therefore leading to an over-estimation of MUC1-IgG. The free IgG Fc would pass through the 100 kDa membrane used in the ultrafiltration and thus could be responsible for some of the losses at this step.

CHO-K1 cells are often the expression system of choice to obtain recombinant mammalian glycoproteins with heterogeneous glycosylation. They lack important core 2-glycosyltransferases and produce mainly core 1-based O-glycans. Detailed structural analyses of the O-glycans on recombinant human erythropoietin and human thrombomodulin expressed in CHO-K1 cells have previously been performed and the glycans were, in both cases,

Table 3 Structure of O-glycans on MUC1–IgG produced in different cell lines

O-glycans were released from purified MUC1–IgG from the different cell lines by reductive β -elimination and were subjected to capillary LC-ESI MS and LC-ESI MS/MS in negative-ion mode for structural determination. Integrated peak areas in mass chromatograms made for each oligosaccharide separately were used to calculate the relative amounts of each structure. Areas were corrected for the different signal intensities obtained by neutral, mono-sialylated (correction factor 0.4) and di-sialylated (correction factor 0.3) oligosaccharides, as described in the Materials and methods section. CHO-K1 shows the mean of two determinations from one CHO-K1 MUC1(16TR)–IgG clone; T47D shows the mean of two determinations of combined protein from two T47D MUC1–IgG clones; MCF-7 shows the mean of two determinations of two different MCF-7 MUC1–IgG clones. Bold text indicates 6-linked branches. (i) and (ii) are used to distinguish two structures with the same molecular ion; the designations are also used in Figure 5. GalNAcol means reduced GalNAc as formed during release from the protein.

Structure/composition of oligosaccharide alditols	Molecular ion [M-H] [−]	Relative amount (%)		
		CHO-K1	T47D	MCF-7
Neutral				
Galβ1-3GalNAcol	384 [−]	1.9	32.3	6.9
Fuc-Galβ1-3GalNAcol	530 [−]	−	8.2	< 0.5
Gal-Galβ1-3GalNAcol	546 [−]	−	−	3.1
Galβ1-3(GlcNAcβ1-6)GalNAcol	587 [−]	−	−	3.7
Galβ1-3(Gal-GlcNAcβ1-6)GalNAcol	749 [−]	−	−	20.3
Fuc,Gal ₂ ,GlcNAc,GalNAcol	895 [−]	−	−	3.2
Gal ₃ ,GlcNAc ₂ ,GalNAcol	1114 [−]	−	−	2.6
Sialylated				
Galβ1-3(NeuAcα2-6)GalNAcol	675 [−] (i)	0.5	4.4	1.0
NeuAcα2-3Galβ1-3GalNAcol	675 [−] (ii)	84.3	46.9	4.4
Fuc-Galβ1-3(NeuAcα2-6)GalNAcol	821 [−]	−	−	< 0.5
NeuAcα2-3Gal1β1-3(GlcNAcβ1-6)GalNAcol	878 [−]	−	−	1.6
NeuAcα2-3Galβ1-3(NeuAcα2-6)GalNAcol	966 [−] , 483 ^{2−}	13.3	8.1	0.8
NeuAcα2-3Galβ1-3(Gal-GlcNAcβ1-6)GalNAcol	1040 [−] (i)	−	−	11.7
Galβ1-3(NeuAc-Gal-GlcNAcβ1-6)GalNAcol	1040 [−] (ii)	−	−	8.9
NeuAcα2-3Galβ1-3(NeuAc-Gal-GlcNAcβ1-6)GalNAcol	665 ^{2−}	−	−	27.9
NeuAc,Gal ₃ ,GlcNAc ₂ ,GalNAcol	1405 [−]	−	−	< 0.5
NeuAc ₂ ,Gal ₃ ,GlcNAc ₂ ,GalNAcol	847 ^{2−}	−	−	3.7

Figure 6 O-glycan site occupancy in CHO-K1 MUC1–IgG

Combined nano-LC-ESI mass spectra of MUC1 glycopeptides obtained after enzymic deglycosylation and Clostripain digestion of MUC1(16TR)–IgG from CHO-K1 cells. Peaks correspond to MUC1 TR peptides with three to five GalNAcs.

identified as the sialylated core 1 structures NeuAc α 2-3Gal β 1-3GalNAc and NeuAc α 2-3Gal β 1-3(NeuAc α 2-6)GalNAc, the main component being the di-sialylated form [27–29]. We found the same two sialylated core 1-based structures on MUC1 produced in CHO-K1, but the mono-sialylated form was more abundant than the di-sialylated form. In addition, we found a small proportion of non-sialylated Gal β 1-3GalNAc. As there is only one O-glycan on erythropoietin and five on thrombomodulin, one

possible explanation for the discrepancy in sialylation could be the incomplete action of the sialylation machinery due to the abundance of O-glycan chains on the recombinant MUC1 protein (approx. 70 in the 16-TR MUC1 protein). However, we observed the same ratio of mono- and di-sialylated Gal β 1-3GalNAc on a smaller MUC1–IgG reporter protein with only 1.7 TRs (results not shown). More likely, different subtypes of CHO cells produce different glycoforms.

The occupancy of the potential O-glycosylation sites in the MUC1–IgG was the same irrespective of the number of TRs in the protein. We observed no difference between the proteins with 32, 16 or 1.7 TRs, indicating that the occupancy was determined by the cells rather than by the number of potential O-glycan sites. The average occupancy obtained for MUC1–IgG from T47D, 4.2 glycans per TR, was similar to that obtained from CHO-K1 cells, but somewhat lower than that previously reported for endogenous MUC1 from T47D (4.8 per TR) [12]. It is difficult to know whether this difference is statistically significant, but it is possible that the IgG tail in the present construct prevents the complete action of the polypeptide-GalNAc-transferases that initiates the synthesis of the O-glycan chains, maybe by moving the protein faster through the Golgi apparatus.

There was a large structural difference between the O-glycans released from MUC1–IgG from the two breast cancer cell lines investigated. In T47D, only core 1-based structures were found, consistent with the earlier finding that T47D have lost the ability to make core 2 glycans, as they lack important core 2 β 6-GlcNAc-transferases [9]. In MCF-7, this enzyme is still present, explaining the synthesis of core 2-based structures, but the presence of the α 2,3-sialyltransferase ST3Gal-I, which adds sialic acid in the α 2,3 position of the Gal in Gal β 1,3GalNAc, prevents to some extent the conversion into core 2 structures [9,14]. The MCF-7 cell line is representative of a differentiated breast carcinoma in that it shows polarized expression of MUC1

Table 4 O-glycosylation site occupancy of MUC1-IgG from CHO-K1 and T47D

MUC1 TR glycopeptides were generated by partial deglycosylation of purified MUC1-IgG glycoproteins and subsequent cleavage with the endoprotease Clostripain. Glycopeptides were then analysed by nano-LC-ESI MS in positive-ion mode. Peaks in the LC MS chromatogram representing MUC1 TR glycopeptides with *n* GalNAcs were integrated and expressed as a percentage of the area from all MUC1 TR peptide peaks. Sites/TR were calculated using the formula [$\Sigma n \times (\% \text{ of peptides with } n \text{ sites})$]/100, where *n* = 2–5.

Number of sites (<i>n</i>)	Molecular ion (carbohydrate residues)	<i>m/z</i> _{calc} (¹² C)	% Peptides with <i>n</i> sites‡		
			CHO-K1 MUC1(16TR)*	CHO-K1 MUC1(32TR)*	T47D MUC1(32TR)†
5	[M + 3H] ³⁺ (5 GalNAc)	968.12, 959.51‡	52	72	57
	[M + 4H] ⁴⁺ (5 GalNAc)	726.34			
4	[M + 3H] ³⁺ (5 GalNAc + 1 Gal)	1022.13			
	[M + 3H] ³⁺ (4 GalNAc)	900.42, 891.75‡	30	18	19
	[M + 4H] ⁴⁺ (4 GalNAc)	675.57			
	[M + 3H] ³⁺ (4 GalNAc + 1 Gal)	954.44			
3	[M + 3H] ³⁺ (3 GalNAc)	832.73, 824.06‡	18	10	14
	[M + 3H] ³⁺ (3 GalNAc + 1 Gal)	886.75			
2	[M + 3H] ³⁺ (2 GalNAc)	765.04, 756.36‡	0	0	11
Sites/TR			4.3	4.6	4.2

* Mean from two determinations.

† One determination of combined protein from two clones.

‡ *m/z* corresponding to MUC1 TR peptides with the non-typical sequences PAPGSTAPAAHGVTSAPDTR or PAPGSTAPAAHGVTSAPESR, which were found at lower frequencies than the typical sequence PAPGSTAPAAHGVTSAPDTR.

to some degree. Interestingly, the level of ST3Gal-I expression has been shown to be related to the grade of the tumour [13]. A difference in the MUC1 O-glycans from different breast carcinoma cell lines, including T47D and MCF-7, has been observed before [11]. That study also found predominantly core 1-based structures in T47D, whereas in MCF-7 many oligosaccharides were core 2-based. However, there is a large number of discrepancies in the structures from MCF-7 MUC1 to the ones described here. This may be explained by variations in the MCF-7 cell line or by differences in the MUC1 constructs used, but the explanation is more likely to be found in the methods used for structural identification. These authors based their interpretation on retention times of fluorescently labelled oligosaccharides with and without prior specific glycosidase treatments. This approach is difficult to interpret in complex mixtures like these. Instead, capillary-LC MS and capillary-LC MS/MS, as used in the present study, provide more precise structural information at a detection level sufficient for analysis of small amounts of cell-culture supernatants, as shown here.

In conclusion, we have generated an efficient expression system in CHO-K1 cells for the production of recombinant MUC1 glycoprotein with a breast cancer-like O-glycosylation. The produced MUC1 reacts well with the MUC1-specific mAb SM3, which specifically binds MUC1 in breast cancer cells, including primary breast cancers [13]. Using perfusion culture in protein-free medium, we obtained around 100 mg/l per day of secreted MUC1-IgG glycoprotein, a value which could be increased another two or three times, without observable changes in the O-glycosylation pattern, by optimizing different culture parameters. The ability to produce recombinant MUC1 with a tumour-relevant glycoform in such large amounts is unique and will be extremely valuable for preclinical, immunological and tumour-protection studies and possibly also for clinical studies. This approach also makes it possible to specifically modify the expression system to obtain other glycoforms of MUC1, by using variant CHO cells with mutated glycosylation enzymes or by the introduction of specific glycosyltransferases. This will prove

to be very useful for investigating which MUC1 glycoforms can provide protective immunity against different cancers, which may show differences in the MUC1 glycoform produced.

This work was supported by the European Union, grant nos QLK3-CT-1999-027 and QLK3-CT-2002-02010, the Swedish Research Council, the Swedish Cancer Foundation and IngaBritt and Arne Lundberg's Foundation.

REFERENCES

- Gendler, S. J., Lancaster, C. A., Taylor-Papadimitriou, J., Duhig, T., Peat, N., Burchell, J., Pemberton, L., Lalani, E. N. and Wilson, D. (1990) Molecular cloning and expression of human tumor-associated polymorphic epithelial mucin. *J. Biol. Chem.* **265**, 15286–15293
- 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
- Wykes, M., MacDonald, K. P., Tran, M., Quin, R. J., Xing, P. X., Gendler, S. J., Hart, D. N. and McGuckin, M. A. (2002) MUC1 epithelial mucin (CD227) is expressed by activated dendritic cells. *J. Leukocyte Biol.* **72**, 692–701
- Correa, I., Plunkett, T., Vlad, A., Mungul, A., Candelora-Kettel, J., Burchell, J. M., Taylor-Papadimitriou, J. and Finn, O. J. (2003) Form and pattern of MUC1 expression on T cells activated *in vivo* or *in vitro* suggests a function in T-cell migration. *Immunology* **108**, 32–41
- Girling, A., Bartkova, J., Burchell, J., Gendler, S., Gillett, C. and Taylor-Papadimitriou, J. (1989) A core protein epitope of the polymorphic epithelial mucin detected by the monoclonal antibody SM-3 is selectively exposed in a range of primary carcinomas. *Int. J. Cancer* **43**, 1072–1076
- Gendler, S. J. (2001) MUC1, the renaissance molecule. *J. Mammary Gland Biol. Neoplasia* **6**, 339–353
- Hanisch, F.-G., Uhlenbruck, G., Peter-Katalinic, J., Egge, H., Dabrowski, J. and Dabrowski, U. (1989) Structures of neutral O-linked poly(lactosaminoglycans) on human skim milk mucins. A novel type of linearly extended poly-N-acetylglucosamine backbones with Galβ(1–4)GlcNAcβ(1–6) repeating units. *J. Biol. Chem.* **264**, 872–877
- Lloyd, K., Burchell, J., Kudryashov, V., Yin, B. W. T. and Taylor-Papadimitriou, J. (1996) Comparison of O-linked carbohydrate chains in MUC-1 mucin from normal breast epithelial cell lines and breast carcinoma cell lines. *J. Biol. Chem.* **271**, 33325–33334
- 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

- 10 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 carcinoma-associated alterations of O-glycosylation. *Eur. J. Biochem.* **236**, 318–327
- 11 Müller, S. and Hanisch, F.-G. (2002) Recombinant MUC1 probe authentically reflects cell-specific O-glycosylation profiles of endogenous breast cancer mucin. High density and prevalent core 2-based glycosylation. *J. Biol. Chem.* **277**, 26103–26112
- 12 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
- 13 Burchell, J., Poulsom, R., Hanby, A., Whitehouse, C., Cooper, L., Clausen, H., Miles, D. and Taylor-Papadimitriou, J. (1999) An α 2,3 sialyltransferase (ST3Gal) is elevated in primary breast carcinomas. *Glycobiology* **9**, 1307–1311
- 14 Dalziel, M., Whitehouse, C., McFarlane, I., Brockhausen, I., Gschmeissner, S., Schwientek, T., Clausen, H., Burchell, J. M. and Taylor-Papadimitriou, J. (2001) The relative activities of the C2GnT1 and ST3Gal-I glycosyltransferases determine O-glycan structure and expression of a tumor-associated epitope on MUC1. *J. Biol. Chem.* **276**, 11007–11015
- 15 Kotera, Y., Fontenot, J. D., Pescher, G., Metzgar, R. S. and Finn, O. J. (1994) Humoral immunity against a tandem repeat epitope of human mucin MUC-1 in sera from breast, pancreatic, and colon cancer patients. *Cancer Res.* **54**, 2856–2860
- 16 Hamaoka, Y., Suehiro, Y., Fukui, M., Shikichi, K., Imai, K. and Hinoda, Y. (2003) Circulating anti-MUC1 IgG antibodies as a favourable prognostic factor for pancreatic cancer. *Int. J. Cancer* **103**, 97–100
- 17 von Mensdorff-Pouilly, S., Verstraeten, A. A., Kenemans, P., Snijdwint, F. G. M., Kok, A., Van Kamp, G. J., Paul, M. A., van Diest, P. J., Meijer, S. and Hilgers, J. (2000) Survival in early breast cancer patients is favourably influenced by a natural humoral immune response to polymorphic epithelial mucin. *J. Clin. Oncol.* **18**, 574–583
- 18 Jerome, K. R., Barnd, D. L., Bendt, K. M., Boyer, C. M., Taylor-Papadimitriou, J., McKenzie, I. F. C., Bast, Jr, 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
- 19 Goydos, J. S., Elder, E., Whiteside, T. L., Finn, O. J. and Lotze, M. T. (1996) A phase I trial of a synthetic mucin peptide vaccine. Induction of specific immune reactivity in patients with adenocarcinoma. *J. Surg. Res.* **63**, 298–304
- 20 Karanikas, V., Hwang, L. A., Pearson, J., Ong, C. S., Apostolopoulos, V., Vaughan, H., Xing, P. X., Jamieson, G., Pietersz, G., Tait, B. et al. (1997) Antibody and T cell responses of patients with adenocarcinoma immunized with mannan-MUC1 fusion protein. *J. Clin. Invest.* **100**, 2783–2792
- 21 Reddish, M., MacLean, G. D., Koganty, R. R., Kan-Mitchell, J., Jones, V., Mitchell, M. S. and Longenecker, B. M. (1998) Anti-MUC1 class I restricted CTLs in metastatic breast cancer patients immunized with a synthetic MUC1 peptide. *Int. J. Cancer* **76**, 817–823
- 22 Bäckström, D., Hansson, G. C., Nilsson, O., Johansson, C., Gendler, S. and Lindholm, L. (1991) Purification and characterization of a membrane-bound and a secreted mucin-type glycoprotein carrying the carcinoma-associated sialyl-Le^x epitope on distinct core proteins. *J. Biol. Chem.* **266**, 21537–21547
- 23 Schulz, B. L., Oxley, D., Packer, N. H. and Karlsson, N. G. (2002) Identification of two highly sialylated human tear-fluid DMBT1 isoforms: the major high-molecular-mass glycoproteins in human tears. *Biochem. J.* **366**, 511–520
- 24 Karlsson, H., Carlstedt, I., and Hansson, G. C. (1989) The use of gas chromatography and gas chromatography-mass spectrometry for the characterization of permethylated oligosaccharides with molecular mass up to 2300. *Anal. Biochem.* **182**, 438–446
- 25 Karlsson, N. G. and Hansson, G. C. (1995) Analysis of monosaccharide composition of mucin oligosaccharide alditols by high-performance anion-exchange chromatography. *Anal. Biochem.* **224**, 538–541
- 26 Burchell, J. and 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. *Epithelial Cell Biol.* **2**, 155–162
- 27 Sasaki, H., Bothner, B., Dell, A. and Fukuda, M. (1987) Carbohydrate structure of erythropoietin expressed in Chinese hamster ovary cells by a human erythropoietin cDNA. *J. Biol. Chem.* **262**, 12059–12076
- 28 Hokke, C. H., Bergwerf, A. A., van Dedem, G. W., Kamerling, J. P. and Vliegthart, J. F. (1995) Structural analysis of the sialylated N- and O-linked carbohydrate chains of recombinant human erythropoietin expressed in Chinese hamster ovary cells. Sialylation patterns and branch location of dimeric N-acetylglucosamine units. *Eur. J. Biochem.* **228**, 981–1008
- 29 Itoh, S., Kawasaki, N., Ohta, M. and Hayakawa, T. (2002) Structural analysis of a glycoprotein by liquid chromatography-mass spectrometry and liquid chromatography with tandem mass spectrometry. Application to recombinant human thrombomodulin. *J. Chromatogr. A* **978**, 141–152

Received 25 July 2003/22 August 2003; accepted 2 September 2003

Published as BJ Immediate Publication 2 September 2003, DOI 10.1042/BJ20031130