The human MUC2 mucin apoprotein appears to dimerize before O-glycosylation and shares epitopes with the ‘insoluble’ mucin of rat small intestine

Noomi ASKER,* Dan BAECKSTRÖM,* Magnus A. B. AXELSSON,* Ingemar CARLSTEDT† and Gunnar C. HANSSON*‡

*Department of Medical Biochemistry, University of Göteborg, Medicinaregatan 9, S-413 90 Gothenburg, and †the Department of Physiological Chemistry, University of Lund, P.O. Box 94, S-221 00 Lund, Sweden

Rabbit antiserum against a synthetic peptide corresponding to a tandemly repeated amino acid sequence in the human intestinal mucin apoprotein MUC2 was used in immunoprecipitation to study the biosynthesis of MUC2 in the colon-carcinoma cell line LS 174T. Under non-reducing conditions, two bands were precipitated, the smaller with an apparent size of about 700 kDa on SDS/PAGE. When analysed by two-dimensional electrophoresis after reduction, the larger band migrated to the same position as the smaller band and was interpreted as a putative disulphide-bond-stabilized dimer. Pulse-chase experiments showed only the monomer after 5 min and the appearance of the putative dimer after 30 min. The MUC2 apoprotein was also precipitated by antisera against the HF-deglycosylated peptides of the two highly glycosylated domains of the ‘insoluble’ mucin complex of rat small intestine [Carlstedt, Herrmann, Karlsson, Sheehan, Fransson and Hansson (1993) *J. Biol. Chem.* **268**,

18771–18781]. Endoprotease Lys-C cleavage of the immunopurified apoprotein gave a large fragment of about 250 kDa that was detected by both the antiserum against the MUC2 tandem repeat and one of the glycopeptide antisera. This supports the view that the ‘insoluble’ mucin of rat small intestine is encoded by the *Muc2* gene, as recently indicated by a partial cDNA sequence [Hansson, Bäckström, Carlstedt and Klinga-Levan (1994) *Biochem. Biophys. Res. Commun.* **198**, 181–190] and that parts of the apoprotein are conserved between the species. A lectin from the snail *Helix pomatia* that detects terminal α -GalNAc residues did not bind to the monomer or putative dimer, suggesting that O-glycosylation starts after dimerization. The results indicate that the biosynthetic pathway of the MUC2 mucin may be similar to that of the von Willebrand factor with which MUC2 shares sequence similarities at its C- and N-termini.

INTRODUCTION

Mucins are a heterogeneous group of glycoproteins characterized by large amounts of O-linked oligosaccharides usually localized to highly glycosylated domains [1,2]. The mucins cover mucosal surfaces, where they are part of the protection of the epithelial cells. The apoproteins of mucins often contain tandemly repeated amino acid regions with a high concentration of threonine, serine and proline. The ‘classical’ mucins are gel-forming and organized into disulphide-bond-stabilized linear oligo- or poly-mers [1]. The small and large intestine is different from stomach, lung and cervical tract in that the intestine contains an ‘insoluble’ mucin complex, not extractable in chaotropic solutions like 6 M guanidinium chloride [3,4] unless disulphide-bond-reducing agents are employed.

Recently, two highly glycosylated domains from the ‘insoluble’ mucin complex of rat small intestine, named glycopeptide A and B, were purified and characterized in detail [4]. They were similarly glycosylated, had similar amino acid compositions and were obtained in about equal molar amounts. The glycopeptide A had a size (both mass and length) about twice that of glycopeptide B. By deglycosylating glycopeptide A with HF and immunization in rabbits, an antiserum was obtained and used for cDNA cloning. The sequence of an obtained cDNA clone (VR-1A) encoded a Ser/Thr-rich region and a Cys-rich region in its

5'-end [5]. The latter showed sequence similarity to a region of the human MUC2 intestinal mucin apoprotein [6] close to the large tandem repeat (Figure 1). This indicated that the ‘insoluble’ mucin probably has an apoprotein based on the *Muc2* sequence.

The human MUC2 mucin amino acid sequence [6,7] reveals three Cys-rich domains, one in the N-terminus, one in the C-terminus and one central domain. Between these Cys-rich domains are two domains with large amounts of Thr, Pro and Ser, making up more than 50 % of the apoprotein. The larger of these domains comprises a repetitive sequence of a well-conserved 23-amino-acid motif and is referred to as the ‘large tandem repeat’ (Figure 1). The smaller Thr/Pro/Ser-rich domain comprises less-well-conserved tandem repeats. Both N- and C-termini show sequence similarity with the respective ends of von Willebrand factor (vWF), especially with respect to the locations of the Cys residues. vWF undergoes a series of post-translational modifications [8], including dimerization in the endoplasmic reticulum.

In the present study the biosynthesis of the human MUC2 apoprotein was investigated in a human colon-carcinoma cell line, as no rat cell line producing *Muc2* was available. We show that an antiserum against the MUC2 tandem repeat as well as antisera against glycopeptide A and B precipitate the same apoprotein, which forms a putative dimer, probably before being O-glycosylated.

Abbreviations used: vWF, von Willebrand factor; α -gpdA, anti-(glycopeptide A); α -gpdB, anti-(glycopeptide B); α -MUC2TR, anti-(MUC2 tandem repeat); DTT, dithiothreitol.

‡ To whom correspondence should be sent.

Figure 1 Schematic picture showing the human MUC2 apoprotein

Shown is the human MUC2 apoprotein [7], glycopeptide A (gpdB) [4], the sequenced C-terminal [20] and N-termini [21] of the rat Muc2 apoprotein. Shown is also VR-1A, a partial cDNA clone detected by the α -gpdB antiserum [5]. The proposed positions and estimated sizes of gpdB and gpdB are compared with the sequenced parts of the rat Muc2. Regions of sequence identity between the human MUC2 sequence and the rat Muc2 sequence are indicated by broken lines. The number of cleavage sites for the endoprotease Lys-C are indicated.

MATERIALS AND METHODS

Antibodies and lectins

The rabbit antiserum PH497 (hereafter referred to as ' α -gpdB') obtained by immunization with HF-treated glycopeptide A was described previously [5]. Glycopeptide B was deglycosylated as described for glycopeptide A [5], and the antiserum PH498 (hereafter referred to as ' α -gpdB') against the deglycosylated material was prepared as follows. HF-treated glycopeptide B (80 μ g) was dissolved in 0.9% (w/v) NaCl and injected into New Zealand White rabbits in an emulsion of Freund's complete adjuvant and then twice with the same amount of peptide in Freund's incomplete adjuvant. A synthetic peptide based on the tandem repeat region of MUC2 apoprotein (KYPTTPISTT-TMVTPTPTGTQT) was kindly provided by Dr. Young Kim and Dr. James R. Gum (Veterans Affairs Medical Center, San Francisco, CA, U.S.A.). A polyclonal antiserum, PH897 (hereafter referred to as ' α -MUC2TR'), against this synthetic peptide was prepared by immunizing a New Zealand White rabbit once with 1 mg of peptide and Freund's complete adjuvant and then twice with 0.5 mg of peptide and Freund's incomplete adjuvant. Biotin-labelled snail (*Helix pomatia*) lectin was obtained from Sigma, St. Louis, MO, U.S.A. and rabbit anti-*(H. pomatia* lectin) antiserum was from Serotec, Oxford, U.K.

Tissue culture

Colon-adenocarcinoma cell line LS 174T (A.T.C.C. CL 188) was cultured in Iscove's Modified Dulbecco's Medium (GIBCO, Paisley, Renfrewshire, Scotland, U.K.), containing 10% (v/v) fetal-calf serum and 1% (v/v) supplement (110 mg/l sodium pyruvate, 36 mg/l L-asparagine, 116 mg/l L-arginine, 290 mg/l L-glutamine and 10 mg/l folic acid). Cells were grown at 37 °C in air/CO₂ (19:1) and 95% humidity.

Metabolic labelling

Cells were grown in 6 cm-diameter Petri dishes to about 90–100% confluence and starved for 1 h in minimal essential medium without L-methionine (GIBCO) containing 5% (v/v) fetal-calf serum. Cells were labelled with 150 μ Ci of [³⁵S]Met ('*In vitro* cell labelling mix' with [³⁵S]methionine and [³⁵S]cysteine; Amersham International)/dish. In pulse-chase experiments the cells were pulsed for 2 min with 300 μ Ci of [³⁵S]Met (Amersham)/dish and then chased with 100 μ g of unlabelled Met/dish. In one experiment the pulse-labelling was performed with 250 μ Ci of [³⁵S]Met (the cell labelling mix mentioned above) and chased with 100 μ g of unlabelled Met and 150 μ g of unlabelled Cys. The chase was stopped by placing the dishes on ice. The cells were carefully washed twice with 1 ml of ice-chilled PBS. Cells were lysed in 1 ml PBS and 250 μ l of cell lysate buffer to a final concentration of 50 mM Tris/HCl/150 mM NaCl/5 mM EDTA/1% (v/v) Triton X-100, pH 7.4, with proteinase inhibitors (60 μ g of leupeptin, 20 μ g of aprotinin, 110 μ g of phenylmethanesulphonyl fluoride) and sonicated for 3 $\times$ 2 s (intensity 15) on an MSE Soniprep 100 sonifier.

Immunoprecipitation

The lysate of metabolically labelled cells was incubated with preimmune serum (25 μ l of rabbit serum) for 1 h, followed by a 30 min incubation with 150 μ l of 10% (v/v) immunoprecipitin (formalin-fixed *Staphylococcus aureus*; GIBCO). All incubations were performed at room temperature and under agitation. The supernatant was incubated with 25 μ l of the antisera α -gpdB, α -gpdB or α -MUC2TR for 3 h. Precipitation with 3 μ g of *H. pomatia* lectin for 3 h was used, followed by incubation with α -(*H. pomatia* lectin) antiserum for 3 h. Immunoprecipitation was also performed with antisera preincubated with MUC2 tandem-repeat peptide (10 μ l of α -gpdB, α -gpdB and α -MUC2TR with 200 μ g, 1000 μ g and 200 μ g respectively of the peptide) for 3 h and then added to the cell lysates and incubated for 1 h. After incubation with the specific antisera, 150 μ l of 10% (v/v) immunoprecipitin was added and the mixture was incubated for 30 min. The immune complexes were washed four times in 10 mM Tris/HCl/2 mM EDTA/0.1% (v/v) Triton X-100/0.1% (w/v) SDS, pH 7.4.

SDS/PAGE and autoradiography

Samples were dissolved from the complex in reducing sample buffer [50 mM Tris/HCl/20% (v/v) glycerol/2% (v/v) mercaptoethanol/5% (w/v) SDS/10 mM dithiothreitol (DTT), pH 6.8] or in non-reducing sample buffer [30 mM Tris/HCl/10% (v/v) glycerol, 2% (w/v) SDS], for 5 min at 95 °C. SDS/PAGE was run on a 1.5 mm-thick gel with a 3–15% gradient gel and a 3% stacking gel [9]. In some experiments the bisacrylamide content of the stacking gel and 3% gradient gel was reduced from 2.67% to 1% of the acrylamide content. Gels were cast on a polyacrylamide-gel support medium (Gel Bond PAG Film from FMC Corp., Rockland, ME, U.S.A.) and run at 25 mA through the stacking gel and 35 mA through the gradient gel. Gels were fixed in 30% (v/v) ethanol and 10% (v/v) acetic acid for more than 3 h, soaked in Amplify (Amersham International) containing 5% (v/v) glycerol for 30 min, dried on a gel dryer for 4 h (or overnight) at 70 °C and exposed (Kodak XAR-5 film) at –80 °C. Rainbow [¹⁴C]methylated protein molecular-mass markers [14.3–200 kDa (Amersham International)] was used. High-molecular-mass calibration kit [67–669 kDa (Pharmacia, Uppsala, Sweden)] and a monoclonal IgM antibody were silver-stained [10].

Agarose-gel electrophoresis

Agarose-gel electrophoresis was performed on 1.5-mm thick gels and cast on an agarose gel support medium (Gel Bond Film, FMC) between two warm glass plates (140 mm × 160 mm). Some modifications to the discontinuous Laemmli buffer system were carried out [11]. The separation gel contained 2% (w/v) agarose [Ultrapure (GIBCO) or SeaPlaque (low gelling temperature; FMC)], and the stacking gel contained 0.8% (w/v) agarose (Ultrapure; GIBCO) and were solidified for 1 h each at +4 °C. The wells were 20 mm from the interface with the separation gel. The gels were run vertically at 15 mA, fixed, dried and exposed to film as described above. For two-dimensional gel electrophoresis, samples were first run under non-reducing conditions on 2% (w/v) agarose gel and the strip was then incubated in upper buffer [0.5 M Tris/HCl/0.4% (w/v) SDS, pH 6.8, containing 50 mM DTT], with constant shaking for 30 min at room temperature. The strip was then placed on top of another 2% (w/v) agarose gel and overlaid with 0.8% (w/v) agarose gel with 0.1 M DTT.

Endoprotease Lys-C treatment

Radioactively labelled immunoprecipitates were eluted from the immunoprecipitin with 2% (w/v) SDS/0.1 M NH_4HCO_3 , pH 9, at 95 °C. The eluates were diluted with an equal volume of 0.1 M NH_4HCO_3 , pH 9, followed by the addition of 0.5 unit of endoprotease Lys-C (analytical grade; Boehringer, Mannheim, Germany). Digestion was performed for 2 h at 37 °C, the samples freeze-dried, dissolved in sample buffer to give a final concentration of 4% (w/v) SDS and subjected to SDS/PAGE.

Fluoroimmunoassay

Cross-reactivity between the antisera was tested by performing a dissociation-enhanced lanthanide fluoroimmunoassay [12]. Microtitre strips were coated overnight at 37 °C with the MUC2 tandem-repeat peptide and HF-treated glycopeptide A and B diluted in PBS starting with 10 µg/ml and then diluted 1:5 (100 µl/well). The amount of HF-treated glycopeptide A and B refers to the mass before HF treatment. Strips were washed twice in washing buffer [5 mM Tris/HCl (pH 8)/0.15 M NaCl/0.005% Tween-20/0.02% NaN_3] and once in 0.15 M NaCl, followed by a 3 h incubation with blocking solution [50 mM Tris/HCl (pH 8)/0.15 M NaCl/6% sorbitol/0.1% BSA/90 µM CaCl_2 /4 µM EDTA/0.02% NaN_3] at 37 °C. The following incubations were performed on a shaking apparatus at room temperature. Antisera α -MUC2TR, α -gpdA or α -gpdB were diluted 1:100 in assay buffer [50 mM Tris/HCl (pH 8)/0.15 M NaCl/1.5% BSA/20 µM diethylenetriamine penta-acetic acid/0.01% Tween-40/0.05% NaN_3] for 2 h, followed by washing of the strips three times in washing buffer. The bound antisera were detected by an europium (Eu^{3+})-labelled anti-rabbit antiserum for 1 h. After the incubation, strips were washed six times in washing buffer and 200 µl of enhancement solution/well was added to release Eu^{3+} , followed by an incubation for 5 min. The fluorescence was measured in an Arcus fluorimeter (Pharmacia).

Inhibition immunoassay was prepared by coating microtitre strips with 1 µg/ml of the synthetic peptide based on the human MUC2 tandem repeat (strips were blocked as described above). Antisera α -MUC2TR, α -gpdA and α -gpdB were diluted in assay buffer, (1:400, 1:400 and 1:200 respectively) and preincubated with a serial dilution of the MUC2 tandem-repeat peptide (starting with 100 µg/ml and then diluted 1:10) for 3 h. The antisera/peptide mixture (100 µl) was added to the strips and

incubated for 1 h before detection of bound antibodies as described above.

RESULTS

Monomer and putative dimer of human MUC2 apoprotein

A polyclonal antiserum directed against the tandem repeat of human MUC2 apoprotein was used to study the primary translation product of the MUC2 mucin in the colon carcinoma cell line LS 174T. An extract of metabolically labelled cells was immunoprecipitated and examined on an agarose gel under reducing and non-reducing conditions. Two major bands appeared under non-reducing conditions, and, after reduction of the sample, the larger band disappeared. This was interpreted as a dimer of the smaller one, formed by disulphide bonds (Figure 2). Samples were also analysed by two-dimensional gel electrophoresis (see below). When analysing the immunoprecipitate on 3–15% (Figure 3) and 3–5% polyacrylamide gel with Rainbow molecular-mass markers, thyroglobulin and IgM antibodies the apparent size of the smaller band was estimated as 700–750 kDa (referred to as '700 kDa monomer') and the apparent size of the larger band to just above 1000 kDa.

Antisera against deglycosylated parts of the rat 'insoluble' intestinal mucin allowed to react with human MUC2 apoprotein

Immunization of rabbits with deglycosylated rat glycopeptide A and B purified from rat 'insoluble' intestinal mucin gave two antisera, namely α -gpdA and α -gpdB respectively. When these antisera were used in immunoprecipitation of metabolically labelled LS 174T cells, two major bands were precipitated. These bands had the same mobilities as the monomer and putative dimer precipitated with the antiserum against the tandem repeat of human MUC2 (α -MUC2TR) as shown in Figure 2.

To confirm that all three antisera precipitate the same protein, immunoprecipitation was performed by extracting the monomer and putative dimer with one antiserum followed by an extra precipitation with immunoprecipitin. The remaining cell lysate was divided and incubated with the other specific antisera. No bands were precipitated when the second specific antiserum was used, indicating that the epitopes of the three different antisera

Figure 2 Immunoprecipitation of a monomer and putative dimer with antisera against deglycosylated rat glycopeptide A and B and the human MUC2 tandem repeat

Colon-carcinoma cell line LS 174T was metabolically labelled for 1 h. Immunoprecipitation on the cell lysate was performed with antisera α -gpdA (lanes 1 and 4), α -gpdB (lanes 2 and 5) and α -MUC2TR (lanes 3 and 6). Samples in lanes 1, 2 and 3 were analysed non-reduced and lanes 4, 5 and 6 reduced by gel electrophoresis using a 2% low-gelling-temperature agarose gel.

Figure 3 Serial immunoprecipitation with antisera against deglycosylated rat glycopeptide A and B and the human MUC2 tandem repeat

LS 174T cells were pulsed for 2 min and chased for 30 min. The cell lysate was immunoprecipitated with α -MUC2TR (lane 1) and then incubated once more with immunoprecipitin. The supernatant was split, followed by immunoprecipitation with the antisera α -gpdA (lane 2) and α -gpdB (lane 3). Immunoprecipitation was also performed with α -gpdB (lane 5), followed by precipitation with α -gpdA (lane 4) and α -MUC2TR (lane 6). The samples were analysed by the modified SDS/PAGE procedure described in the Materials and methods section. Molecular-mass (M) markers indicated are mouse IgM (900 kDa), thyroglobulin (669 kDa), thyroglobulin half-unit (330 kDa) and myosin (200 kDa).

Figure 4 Analysis of the fragments, formed by endoprotease Lys-C cleavage of the immunopurified MUC2 apoprotein

LS 174T cells were metabolically labelled for 1 h and immunoprecipitated with α -MUC2TR. The precipitate was treated with endoprotease Lys-C (lane 2) and the fragments formed were immunoprecipitated again with α -gpdA (lane 3), α -gpdB (lane 4) and α -MUC2TR (lane 5). Samples were reduced and analysed by SDS/3–15% PAGE. Both α -gpdA and α -MUC2TR bound to a fragment of an estimated size of 250 kDa. Immunoprecipitation performed with α -MUC2TR without degradation with Lys-C is shown in lane 1. Abbreviation: M , molecular mass.

are indeed found on the same protein, the human MUC2 apoprotein (Figure 3; precipitation with α -gpdA first is not shown).

Antisera binding to the large tandem-repeat array of human MUC2

To determine the parts of the human MUC2 apoprotein containing the epitopes of the antisera α -gpdA and α -gpdB, the immunopurified and labelled apoprotein was cleaved with endoprotease Lys-C and the fragments formed exposed to the antisera. A fragment of about 250 kDa was found under reducing (Figure 4) and non-reducing conditions (results not shown). This fragment most likely included the 2300-amino-acid-long tandem repeat of the apoprotein as α -MUC2TR antiserum precipitated

Figure 5 Fluoroimmunoassay with the antisera against deglycosylated rat glycopeptide A and B and the human MUC2 tandem repeat

Graphs showing fluoroimmunoassay of the reactivity between antisera α -gpdA (■), α -gpdB (●) and α -MUC2TR (▲) against the three immunization antigens. Microtitre strips were coated with HF-treated glycopeptide A and B (a and b) and a synthetic peptide based on the tandem repeat of human MUC2 (c) starting with 10 μ g/ml and diluted 1:5. All strips were incubated with antisera α -gpdA, α -gpdB and α -MUC2TR diluted 1:100 in assay buffer. Bound antibodies were detected with Eu^{3+} -labelled anti-rabbit antibodies and the fluorescence was measured in an Arcus fluorimeter.

this fragment after cleavage. The α -gpdA found the 250 kDa fragment, whereas α -gpdB did not (Figure 4), which is consistent with the predicted specificities of these antisera based on their immunization antigens (Figure 1).

Cross-reactivity between antisera α -MUC2TR, α -gpdA and α -gpdB

The cross-reactivity between the antisera α -MUC2TR, α -gpdA and α -gpdB was tested with fluoroimmunoassay (Figure 5). The antisera were added to microtitre strips coated with MUC2TR peptide and HF-treated glycopeptides A and B. Cross-reactivity between the antisera α -gpdA and α -gpdB was found. Antisera α -gpdA and α -gpdB also had some affinity for the MUC2TR peptide. The antiserum α -MUC2TR, on the other hand, showed no affinity for the HF-treated glycopeptide A or B.

An inhibition immunoassay was performed by preincubating antisera α -gpdA, α -gpdB and α -MUC2TR with serial dilutions

Figure 6 Inhibition fluoroimmunoassay with antisera against deglycosylated rat glycopeptide A and B and the human MUC2 tandem-repeat antisera by the MUC2 tandem-repeat peptide

Graphs showing inhibition fluoroimmunoassay with MUC2 tandem-repeat peptide and antisera α -gpdA, α -gpdB and α -MUC2TR. Microtitre strips were coated with 1 μ g/ml of synthetic peptide based on the tandem repeat of human MUC2. Antisera α -gpdA (1:400), α -gpdB (1:200) and α -MUC2TR (1:400) were preincubated for 3 h with the serially diluted MUC2 tandem-repeat peptide, starting with 100 μ g/ml peptide and diluted 1:10 in assay buffer, followed by incubation of the mixture to the microtitre strips for 1 h. Bound antibodies were detected with Eu^{3+} -labelled anti-rabbit antiserum and the fluorescence was measured in an Arcus fluorimeter.

Figure 7 Immunoprecipitation with antisera against deglycosylated rat glycopeptide A and B and the human MUC2 tandem repeat inhibited by the MUC2 tandem-repeat peptide

Preincubation for 3 h with 10 μ l each of antisera α -gpdA, α -gpdB and α -MUC2TR, with 200 μ g, 1000 μ g and 200 μ g of the synthetic peptide of the MUC2 tandem repeat respectively was performed. Metabolically labelled lysate from LS 174T cells was immunoprecipitated with the peptide-inhibited antisera α -gpdA (lane 2), α -gpdB (lane 4) and α -MUC2TR (lane 6) for 1 h. As a control, immunoprecipitation was performed with 10 μ l of non-inhibited antisera of α -gpdA (lane 1), α -gpdB (lane 3) and α -MUC2TR (lane 5). Samples were run under reducing condition by modified SDS/3–15%-PAGE.

of the MUC2TR peptide and the mixture incubated with strips coated with MUC2 tandem-repeats peptide (Figure 6). The bound antibodies were detected with Eu^{3+} -labelled anti-rabbit antiserum. The binding of α -gpdA and α -gpdB to the MUC2

Figure 8 Two-dimensional gel electrophoresis of immunopurified monomer and dimer of MUC2

Precipitate from 1 h-labelled cells after reaction with the antiserum α -gpdA was run on a 2% low-gelling-temperature agarose gel under non-reducing conditions (right to left), followed by another 2% low-gelling-temperature agarose gel under reduced conditions (top to bottom). In the second dimension, the putative dimeric form of the human MUC2 apoprotein migrated to the same position as the 700 kDa monomer. An additional band with an estimated size of about 100 kDa smaller than the monomer could be detected. A weak band migrating to a position between the monomer and the putative dimer could be seen in the first dimension that, after reduction, migrated to the same position as the additional band.

tandem-repeat peptide was inhibited by the peptide itself. To determine whether the antisera against the glycopeptide A and B contained antibodies reacting with parts of the human MUC2 mucin other than the tandem repeat, immunoprecipitation was performed. From data in Figure 6, the amount of peptide sufficient to inhibit the binding of the antibodies reacting with tandem repeat of the MUC2 was determined. Preincubation of 10 μ l of the antisera α -gpdA, α -gpdB and α -MUC2TR with 200, 1000 and 200 μ g of peptide respectively was performed for 3 h. LS 174T cells were metabolically labelled for 1 h and incubated with the antisera/peptide mixture. Results show that both α -gpdA and α -gpdB antisera still reacted with the human MUC2 mucin, even after inhibition of the antisera with the MUC2TR peptide (Figure 7). On the other hand, no band was precipitated with antiserum α -MUC2TR after inhibition of the antiserum with the peptide.

Two-dimensional gel electrophoresis

To additionally confirm that the larger band is the putative dimeric form of the 700 kDa band, immunoprecipitation with α -gpdA was analysed by two-dimensional gel electrophoresis. Samples were first separated on a non-reducing 2%-agarose gel and then analysed for a second time under reducing conditions (Figure 8). As is indicated, the putative dimeric form of the human MUC2 apoprotein migrated to the same position as the 700 kDa monomer.

When the apoprotein of MUC2 in LS 174T was examined by immunoprecipitation with all the three antisera, two other bands of variable intensity appeared on the autoradiograms most of the time. One of the bands had a size about 100 kDa smaller than the monomer apoprotein of MUC2 and the other had a size between that of the monomer and dimer of MUC2. The antiserum α -gpdA reacted more strongly with these additional bands (Figure 2) than the other two antisera. The two-dimensional gel revealed that the larger additional band migrated to a position equal to that of the smaller additional band under reducing conditions (Figure 8).

Figure 9 Assembly of the MUC2 apoprotein into a dimer analysed by pulse-chase experiments

LS 174T cells were metabolically labelled for 2 min in [35 S]Met and chased for 5, 30, 60 and 180 min. Immunoprecipitation was performed with α -MUC2TR and samples analysed under non-reducing conditions by modified SDS/3–15%-PAGE (as described in the Materials and methods section). The monomer could be detected after 5 min chase, and the putative dimer was visible after a 30 min chase.

Figure 10 Serial immunoprecipitation with *H. pomatia* and α -MUC2TR

Immunoprecipitation of LS 174T cells labelled for 3 min with [35 S]Met and [35 S]Cys, chased for 10 min (lanes 1 and 3) and for 1 h (lanes 2 and 4). Samples were analysed on an agarose gel (Ultrapure; GIBCO). Lanes 1 and 2 shows samples incubated in 3 μ g/ml of *H. pomatia* lectin. This complex was precipitated and the supernatant further incubated with α -MUC2TR (lanes 3 and 4). 'a' indicates the non-glycosylated putative dimer and 'b' the non-glycosylated 700 kDa monomer.

Dimer formation

The dimerization of MUC2 was studied by pulse-chase experiments with LS 174T cells with a 2 min pulse and a chase for 5, 30, 60 and 120 min. The MUC2 apoprotein was purified by immunoprecipitation with α -MUC2TR antiserum and analysed non-reduced on SDS/3–15%-PAGE. After a 5 min chase, only the 700 kDa monomer was visible, and after 30 min chase the putative dimer started to appear (Figure 9). The intensity of the monomer and dimer bands started to fade after a 1 h chase, but even after a 24 h chase there was some monomer and putative dimer left (results not shown).

O-glycosylation occurs after dimerization

To determine whether the MUC2 700 kDa monomer and its putative dimer were O-glycosylated, *H. pomatia* lectin, which binds to terminal α -GalNAc, was used for precipitation. Confluent cells were labelled for 3 min, followed by a chase of 10 min and 1 h. The precipitate with *H. pomatia* was removed with immunoprecipitin and the supernatant was then incubated with α -MUC2TR antiserum. Samples were eluted and analysed under non-reducing conditions on agarose gel (Figure 10). Neither the 700 kDa monomer nor its dimer were precipitated with the *H. pomatia* lectin. However, the lectin precipitated heterogeneous material of low mobility after a 1 h chase, but not after 10 min.

DISCUSSION

By analysing the human colon-carcinoma cell line LS 174T with an antiserum directed against the tandem-repeat motif of the human MUC2 mucin, it has been possible to show the early appearance of a primary translation product. Within 30 min a larger band appeared on the gels, which, upon reduction of disulphide bonds, gave a band of the same size as the primary translation product (Figures 2 and 8). From this we conclude that the primary product was the MUC2 monomer which later forms a disulphide-bond-stabilized polymer, a putative dimer. Even though the amount of dimer increased during the first 1 h of chase, a substantial amount of monomer remained. After 1 h the intensity of the monomeric and dimeric bands started to fade away, but even after 24 h of chase there were some monomers and dimers left. Currently we have no explanation for this, but one may speculate that the monomer and dimer not properly folded are retained in the endoplasmic reticulum and slowly degraded.

The entire cDNA of the human *MUC2* has recently been sequenced by Gum et al. [6,7]. The apoprotein contains a 23-amino-acid tandem repeat with many Thr and Pro residues and a smaller repeated domain, rich in Thr, Pro and Ser. The repeated regions are flanked by Cys-rich domains (Figure 1). The size of the MUC2 apoprotein determined by analysis of the cDNA sequence of the *Muc2* gene, including the signal peptide, assuming 100 tandem repeats and glycosylation of all 30 possible N-glycosylation sites, would give a size of about 600 kDa. When analysing the MUC2 apoprotein precipitated from LS 174T cells by SDS/PAGE, the apparent size of the monomer was estimated to be 700–750 kDa and that of the putative dimer to be just slightly larger than 1000 kDa. However, the estimated sizes of these large proteins on SDS-containing gels can only be described as apparent. The relationship between molecular mass and mobility on these gels is not always linear. Parameters such as binding of SDS and unfolding of the mucin apoprotein compared with globular proteins have to be taken into consideration. This could be the reason for a dimer, composed of subunits with the same size as the monomer (Figure 8) not having a size twice that of the 700 kDa monomer.

The sequence of the *MUC2* gene is similar to that of vWF in both the C- and N-terminal domains, especially in respect of the positions of the Cys residues [7]. Previous studies performed on the biosynthesis of the vWF show that the dimerization occurs in the endoplasmic reticulum via disulphide-bond formation and that the dimers are transported to the Golgi apparatus where they are glycosylated and later multimerized. Both the C- and N-terminals are involved in the dimerization and polymerization of vWF [13], which suggests similarities in the biosynthesis of the vWF and the human MUC2. These similarities also argue for the 1000 kDa band actually being a

dimer. The dimer of vWF is formed via joining the C-termini of two monomers [14], but whether the dimer of the human MUC2 apoprotein is linked C-terminus to C-terminus, N-terminus to N-terminus, or N-terminus to C-terminus is still not known.

The recent results obtained by Ohara et al. [15], who studied cell-bound and secreted mucins in the MUC2-producing cell line HM3 (subclone of LS 174T cells), reported that only the secreted, but not the cell-bound, mucin showed a decrease in size on disulphide-bond reduction. The discrepancy between our results and theirs could be due to their use of higher-concentration polyacrylamide gels, which would not allow the dimers to enter the gel.

Studies on the biosynthesis of mucins in rat stomach explants showed the presence of a 300 kDa monomer that was converted into disulphide-bond-stabilized oligomers in the endoplasmic reticulum before being glycosylated [16]. Bands interpreted as mono-, di-, tri- and tetra-mers were separated by SDS/PAGE. This contrasts with our results, where only two non-glycosylated bands were detected, the monomer and the dimer. However, the gastric mucins could be formed by different routes, as the MUC2 mucin is not expressed in the stomach [17].

Two trypsin-resistant highly glycosylated domains, glycopeptide A and B, were isolated previously from the 'insoluble' mucin complex of rat small intestine [4]. Antisera against the deglycosylated glycopeptides A and B were also used in immunoprecipitation of LS 174T cells, as they were known to produce the MUC2 mucin [18,19]. Both antisera precipitated the same monomer and dimer as the antiserum against the tandem repeat of the human MUC2 apoprotein. The antiserum against glycopeptide A was recently used in screening a rat cDNA library giving a clone, VR-1A, containing a Cys-rich and a Ser/Thr-rich region. The Cys-rich sequence showed similarity to a Cys-rich region close to the tandem repeat of the human MUC2 [5], but the Ser/Thr-rich region was not similar to the human tandemly repeated sequence (Figure 1). At that time it could only be suggested that the glycopeptide A and B from the 'insoluble' mucin complex was encoded by the *Muc2* gene, but the present results further support this assumption. The proposed positions and sizes of the rat glycopeptide A and B as compared with the sequenced parts of rat *Muc2* [5,20,21] and the full sequence of human MUC2 are shown in Figure 1.

Sequence comparisons of both the small and large tandem repeat from human MUC2 with the corresponding regions of the three partial sequences of the rat *Muc2* sequence [5,20,21] showed a low degree of sequence identity. Despite this, both α -gpdA and α -gpdB antisera reacted with the synthetic peptide based on the large tandem repeat of the human MUC2 mucin, as shown by fluoroimmunoassay (Figure 5 and 6). These results could indicate a non-specific binding of the antisera α -gpdA and α -gpdB against the human MUC2 mucin. The part of the antisera reacting with the peptide was inhibited before immunoprecipitation to establish whether the antisera still contained antibodies against other parts of the human MUC2. As shown in Figure 7, both antisera α -gpdA and α -gpdB still precipitated the human MUC2 mucin from LS 174T cell line, even if the band precipitated by α -gpdB was slightly weaker than the control. This indicates that the glycopeptides contained conserved unique amino acid sequences remaining in both peptides, giving rise to antibodies reacting with identical regions of the rat and human MUC2 sequences.

By using the known amino acid sequence of the MUC2 mucin, possible cleavage sites for the Lys-C enzyme can be predicted. These cleavage sites are shown in the schematic presentation of MUC2 (Figure 1). The large tandem repeat is not fully sequenced, but a lack of Lys residues [6,19] would give a fragment of 254 kDa assuming 100 tandem repeats and including parts of the

central and C-terminal Cys-rich regions [6,22]. This size is in accordance with the 250 kDa Lys-C fragment precipitated with the α -MUC2TR antiserum. The cells were metabolically labelled with radioactive Met and, as indicated by Gum et al. [19], a few positions in the large tandem repeat of human MUC2 did contain Met, making incorporation of label possible. The size of the large fragment was not influenced by reduction and alkylation before cleavage (results not shown), which further support the notion that the 250 kDa fragment is the tandem repeat of MUC2. The antiserum α -gpdA also reacted with this fragment, a finding which is consistent with the fluoroimmunoassay results (Figures 5 and 6) and the fact that not only the repeat is included in this 250 kDa fragment. The absence of a band when precipitating the Lys-C cleaved MUC2 with the α -gpdB antiserum was probably due to the destruction or fragmentation on the epitopes by the Lys-C cleavage. This indicates that the small Ser/Thr/Pro region of MUC2 (Figure 1), predicted to give a fragment of 40 kDa with four Met residues, does not contain epitopes for the α -gpdB antiserum.

The existence of a putative 'link' protein involved in the formation of mucins has been debated [23]. A 118 kDa glycopeptide of 689 amino acids cleaved from the C-terminal at the amino acid sequence GPH of the rat intestinal MUC2 homologue was reported previously [20]. This C-terminus was later shown to have considerable similarity to the C-terminus of the human MUC2 apoprotein [6]. The fragment cleaved off at the same position in human MUC2 apoprotein has a mass of about 81 kDa, and glycosylation of all 12 possible N-glycosylation sites should give a size of about 100 kDa. The 'link' peptide is believed to remain attached to the mucin via disulphide bonds and released upon reduction [20]. When the monomer and dimer of the human MUC2 were studied, a band with an apparent size estimated to about 600 kDa could sometimes be visible on the reduced as well as the non-reduced gels. The size of the fragment cleaved off is similar to the glycosylated 'link' peptide. If the 600 kDa protein is the monomer with its C-terminal 'link' peptide cleaved off, then these are not attached via disulphide bonds, as boiling in SDS can separate them.

When analysing immunoprecipitated samples of the 700 kDa monomer and its putative dimer on a two-dimensional gel, a possible dimeric form of the additional 600 kDa band was revealed. The putative dimer of this additional band was found in smaller amounts than was the monomeric form. The antisera α -gpdA, α -gpdB and α -MUC2TR reacted with these additional bands to different extents. The antiserum α -gpdA seems to be the most reactive of the three, almost always precipitating the 600 kDa monomer. One possible explanation could be that the monomer with the C-terminus cleaved off has changed its conformation and that α -gpdA had higher affinity for this than the α -gpdB and α -MUC2TR antisera.

A weak band of variable intensity, larger in size than the MUC2 dimer, could sometimes be detected using the α -MUC2TR, α -gpdA and α -gpdB antisera. This band was usually stronger using the α -MUC2TR antiserum and was never present in reduced samples (Figure 2). It could not be extracted with the *H. pomatia* lectin, suggesting that it is a non-glycosylated component (Figure 10). The size of this band was smaller than expected for a non-glycosylated tetramer and, as we have been unable to trace it on two-dimensional gels, its composition remains unknown.

Larger bands, corresponding to glycosylated dimer and oligomers, were never observed with the three antisera, probably reflecting that their epitopes became hidden by glycosylation. The *H. pomatia* lectin did not precipitate the 700 kDa monomer or its putative dimer suggesting that they were not yet O-

glycosylated. Instead it precipitated heterogeneous material with a larger size than the dimer, which could correspond to the dimer being glycosylated with an *N*-acetylgalactosaminyltransferase in the *cis*-Golgi compartment. Other large glycoproteins produced by the LS 174T cells could also account for this diffuse band. One such candidate is the MUC3 mucin synthesized by this cell line. However, the amount as estimated from the transcripts is more than 3000 times lower (D. Baeckström and G. C. Hansson, unpublished work).

The sequence similarity between human MUC2 and thus probably also rat Muc2 with vWF suggest that these form both oligomers and polymers. Long polymers could be the basis for the formation of 'insoluble' mucin complexes found in the small and large intestine of both rat and human [3,4,24].

Note added in proof (received 2 March 1995)

After submission of this paper, McCool et al. [25] demonstrated the formation of disulphide-bond-stabilized oligomers of the MUC2 mucin.

We are indebted to Dr. Young S. Kim and James R. Gum for the gift of the MUC2 tandem-repeat peptide. This work was supported by the Swedish Medical Research Council (grants 7461, 7902 and 10443) and by Ingabritt and Arne Lundberg's Fund.

REFERENCES

- 1 Carlstedt, I., Sheehan, J. K., Corfield, A. P. and Gallagher, J. T. (1985) *Essays Biochem.* **20**, 40–76
- 2 Strous, G. J. and Dekker, J. (1992) *Crit. Rev. Biochem. Mol. Biol.* **27**, 57–92
- 3 Carlstedt, I., Karlsson, H., Sundler, F. and Fransson, L.-Å. (1982) *Adv. Exp. Med. Biol.* **144**, 155–157
- 4 Carlstedt, I., Herrmann, A., Karlsson, H. et al. (1993) *J. Biol. Chem.* **268**, 18771–18781
- 5 Hansson, G. C., Baeckström, D., Carlstedt, I. and Klinga-Levan, K. (1994) *Biochem. Biophys. Res. Commun.* **198**, 181–190
- 6 Gum, J. R., Hicks, J. W., Toribara, N. W., Rothe, E.-M., Lagace, R. E. and Kim, Y. S. (1992) *J. Biol. Chem.* **267**, 21375–21383
- 7 Gum, J. R., Hicks, J. W., Toribara, N. W., Siddiki, B. and Kim, Y. S. (1994) *J. Biol. Chem.* **269**, 2440–2446
- 8 Ruggeri, Z. M. and Ware, J. (1993) *FASEB J.* **7**, 308–316
- 9 Laemmli, U. K. (1970) *Nature (London)* **227**, 680–685
- 10 Heukeshoven, J. and Dernick, R. (1985) *Electrophoresis* **6**, 103–112
- 11 Ruggeri, Z. M. and Zimmerman, T. S. (1981) *Blood* **57**, 1140–1143
- 12 Hemmälä, I., Dakubu, S., Mukkala, V.-M., Siitari, H. and Lövgren, T. (1984) *Anal. Biochem.* **137**, 335–343
- 13 Wagner, D. D. (1990) *Annu. Rev. Cell Biol.* **6**, 217–246
- 14 Wagner, D. D., Lawrence, S. O., Ohlsson-Wilhelm, B. M., Fay, P. J. and Marder, V. J. (1987) *Blood* **69**, 27–32
- 15 Ohara, S., Byrd, J. C., Gum, J. R. and Kim, Y. S. (1994) *Biochem. J.* **297**, 509–516
- 16 Dekker, J., Van Der Ende, A., Aelmans, P. H. and Strous, G. J. (1991) *Biochem. J.* **279**, 251–256
- 17 Ho, S. B., Niehans, G. A., Lyftogt, C. et al. (1993) *Cancer Res.* **53**, 641–651
- 18 Byrd, J. C., Lamport, D. T. A., Siddiqui, B. et al. (1989) *Biochem. J.* **261**, 617–625
- 19 Gum, J. R., Byrd, J. C., Hicks, J. W., Toribara, N. W., Lamport, D. T. A. and Kim, Y. S. (1989) *J. Biol. Chem.* **264**, 6480–6487
- 20 Xu, G., Huan, J.-L., Khatri, I. A. et al. (1992) *J. Biol. Chem.* **267**, 5401–5407
- 21 Ohmori, H., Dohrman, A. F., Gallup, M. et al. (1994) *J. Biol. Chem.* **269**, 17833–17840
- 22 Toribara, N. W., Gum, J. R., Culhane, P. J. et al. (1991) *J. Clin. Invest.* **88**, 1005–1013
- 23 Robertson, A. M., Mantle, M., Fahim, R. E. F. et al. (1989) *Biochem. J.* **261**, 637–647
- 24 Mårtensson, S., Lundblad, A., Hansson, G. C. and Carlstedt, I. (1988) *Scand. J. Clin. Lab. Invest.* **48**, 633–640
- 25 McCool, D. J., Forstner, J. F. and Forstner, G. G. (1994) *Biochem. J.* **302**, 111–118