The recombinant C-terminus of the human MUC2 mucin forms dimers in Chinese-hamster ovary cells and heterodimers with full-length MUC2 in LS 174T cells

Martin E. LIDELL*, Malin E. V. JOHANSSON*, Matthias MÖRGELIN†, Noomi ASKER*, James R. GUM, Jr.‡, Young S. KIM‡ and Gunnar C. HANSSON*¹

*Department of Medical Biochemistry, Göteborg University, Medicinaregatan 9A, 413 90 Gothenburg, Sweden, †Cell and Molecular Biology, Lund University, BMC C12, 221 84 Lund, Sweden, and ‡Gastrointestinal Research Laboratory, Department of Veterans Affairs Medical Center, University of California, San Francisco, CA 94121, U.S.A.

The entire cDNA corresponding to the C-terminal cysteine-rich domain of the human MUC2 apomucin, after the serine- and threonine-rich tandem repeat, was expressed in Chinese-hamster ovary-K1 cells and in the human colon carcinoma cell line, LS 174T. The C-terminus was expressed as a fusion protein with the green fluorescent protein and mycTag sequences and the murine immunoglobulin κ -chain signal sequence to direct the protein to the secretory pathway. Pulse-chase studies showed a rapid conversion of the C-terminal monomer into a dimer in both Chinese-hamster ovary-K1 and LS 174T cells. Disulphide-bond-stabilized dimers secreted into the media of both cell lines had a higher apparent molecular mass compared with the intracellular forms. The MUC2 C-terminus was purified from the spent culture medium and visualized by molecular electron microscopy. The dimer nature of the molecule was visible clearly and revealed that each monomer was attached to the other by a large globular

domain. Gold-labelled antibodies against the mycTag or green fluorescent protein revealed that these were localized to the ends opposite to the parts responsible for the dimerization. The C-terminus expressed in LS 174T cells formed heterodimers with the full-length wild-type MUC2, but not with the MUC5AC mucin, normally expressed in LS 174T cells. The homodimers of the MUC2 C-termini were secreted continuously from the LS 174T cells, but no wild-type MUC2 secretion has been observed from these cells. This suggests that the information for sorting the MUC2 mucin into the regulated secretory pathway in cells having this ability is present in parts other than the C-terminus of MUC2.

Key words: Chinese-hamster ovary cell, dimerization, electron microscopy, endoplasmic reticulum, mucin.

INTRODUCTION

The MUC2 mucin belongs to the family of gel-forming mucins. These give mucus its gel properties and are the major constituent of the mucus layer that covers all our mucosal surfaces. It is assumed that the major function of mucins is to protect the underlying epithelial cells, illustrated recently by an increase in colon cancer in mice deficient in the MUC2 mucin [1]. Little is known on how this protective effect is achieved, but it is believed that the polymeric nature is a prerequisite.

The human MUC2 mucin shows large similarities in the cysteine positions in the N- and C-termini compared with the von Willebrand factor (vWF) [2]. Biosynthesis of the vWF is well studied and involves dimerization of the protein by disulphide-bond formation between C-termini in the endoplasmic reticulum (ER) [3]. The vWF is then transported via the Golgi apparatus to the *trans*-Golgi network and secretory vesicles, where oligomerization of the protein starts by disulphide bonds involving the N-termini. Alignment of the human MUC2 mucin [2] with the human MUC5AC [4,5] and MUC5B mucins [6], as well as the porcine [7] and bovine [8] submaxillary mucins, shows sequence similarities in the positions of the cysteine residues of the C-termini. The sequence similarities between these mucins and the vWF suggest that these mucins might also assemble in a similar way. We have shown earlier that the biosynthesis

of the human MUC2 and MUC5AC apomucins are similar, forming disulphide-bond-stabilized dimers before O-glycosylation [9]. Studies on the pig submaxillary mucin (PSM) and the rat Muc2 mucin suggest that the C-terminal parts of these proteins are responsible for dimer formation [10–12]. These findings support the hypothesis that dimerization of the human MUC2 takes place via its C-terminus.

The nerve growth factor, transforming growth factor β 2 (TGF β 2) and platelet-derived growth factor-BB all contain a cystine-knot motif [13]. This motif consists of six cysteine residues with a defined distance to each other forming a three-dimensional knot containing three disulphide bonds. The nerve growth factor, TGF β 2 and platelet-derived growth factor-BB form dimers involving the knot, but the interactions between them are different. For TGF β 2, the dimer-forming cystine bond is known to involve the same cysteine residues in both monomer units, as revealed by its crystal structure [14]. The cysteine forming this intermolecular bond is located directly in front of the fourth cysteine residue in the cystine-knot motif. The mucins MUC2, MUC5AC, MUC5B and MUC6 as well as the vWF contain the cystine-knot motif within the last 150 amino acids of their C-termini. This includes one cysteine in the same position as the dimerizing cysteine found in the TGF β 2. This proposes that this cysteine might be involved also in the dimerization of mucins. Studies on rat Muc2 [11,12] point out that this cysteine residue is

Abbreviations used: CHO, Chinese-hamster ovary; endo H, endoglycosidase H; ER, endoplasmic reticulum; α -GalNAcase, α -N-acetylglactosaminidase; GFP, green fluorescent protein; mAb, monoclonal antibody; NEM, N-ethylmaleimide; PNGase F, peptide N-glycosidase F; PSM, pig submaxillary mucin; TGF β 2, transforming growth factor β 2; vWF, von Willebrand factor.

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

essential for dimer formation, whereas studies on vWF [15], PSM [16] and Norrin [17] indicate that additional cysteine residues may be of importance. The mucin genes *MUC2*, *MUC5AC*, *MUC5B* and *MUC6* are clustered on chromosome 11p15.5 [18], all of which contain the cystine-knot motif and show large similarities in the positions of other cysteine residues also outside of the cystine-knot motif, indicating that these mucin genes are part of a family having a common ancestral gene.

The four known gel-forming mucins are produced in goblet cells that have a regulated secretory pathway. Most of these mucins are stored in granulae and released, on stimulation, e.g. with carbachol or ATP [19]. The secretory stimuli known today release only parts of the stored material. It is assumed that other, maybe more local and direct, mechanisms are involved. The mucins must thus be sorted to the regulated secretory pathway, but no study suggests which part of *MUC2* is involved in this sorting.

To address more detailed questions in the assembly and sorting of the *MUC2* mucin, truncated *MUC2* mucin must be expressed as recombinant proteins. In the present study, we have expressed the C-terminal cysteine-rich domain of the human *MUC2* mucin as a fusion protein with green fluorescent protein (GFP) and the mycTag. However, as this is only a minor portion of *MUC2* (981 out of 5179 amino acids), it is important to analyse whether this recombinant protein behaves as predicted from studies of the full-length *MUC2*. For example, is the C-terminus forming dimers? Is the C-terminus involved in the sorting of *MUC2* to the regulated secretory pathway? The first question was addressed by biochemical and electron-microscopy studies of the expressed protein. The second question was addressed by expressing the recombinant protein in cells without [Chinese-hamster ovary (CHO-K1)] and with (LS 174T) a regulated secretory pathway. LS 174T produces both the *MUC2* and *MUC5AC* mucins, thus allowing studies of how the *MUC2* C-terminus is interacting with the full-length mucins. The results suggest that the recombinant *MUC2* C-terminus forms dimers as predicted for the full-length *MUC2* and that the *MUC2* C-terminus alone is not sufficient in directing the *MUC2* mucin to the regulated secretory pathway.

MATERIALS AND METHODS

Antisera and antibodies

The polyclonal antisera α -*MUC2C2* (against amino acids 4995–5013 of the human *MUC2* apomucin) and α -*MUC5ACTR* (against a synthetic peptide based on a repeated threonine-rich sequence in the human *MUC5AC* apoprotein) have been described in [20]. This is also the case for the antisera α -*MUC2TR* (against the *MUC2* tandem repeat) and α -gpdA (against deglycosylated glycopeptide A from rat *Muc2*) [9,21]. The α -myc monoclonal antibody (mAb) was obtained directly from the culture medium of the 9E10.2 hybridoma (CRL-1729; A.T.C.C.). Goat anti-mouse immunoglobulins coupled with alkaline phosphatase (Goat- α -mouse-AP; Dako, Carpinteria, CA, U.S.A.) were used as the secondary antibody in Western-blot experiments.

Construction of pSMG-MUC2C

V5, a PCR-amplified product of the human *MUC2* (bases 12880–15708), was inserted into the *KpnI*–*XbaI* sites of a M13mp18 vector [22]. *XbaI* and *HindIII* sites downstream of the stop codon of the V5 sequence were added artificially.

The V5 fragment was excised with *EcoRI* and *HindIII*, and ligated into a pBluescript SK(–) vector using the same sites. The resulting vector was named pBS-V5. A cDNA clone of the human *MUC2* (bases 12243–13083), SMUC41, was inserted into pBluescript SK(–) [23]. Using PCR, an *XbaI* site was inserted at the N-terminal end using the primers 5'-ATTCTAGAGTAACC-CACACAAGCACAGCAC-3' and 5'-GCCGTTGTACACCTCC-TCACCTG-3' to amplify the 472 bp product. This PCR product was ligated into the pBS-V5 via *XbaI* and *BspI*407 (cleaves at base 13074 in the *MUC2* sequence), creating the pBS-*MUC2C* vector. The pSMG-*MUC2C* plasmid was based on the eukaryotic expression vector pEGFP-C1 (ClonTech, Palo Alto, CA, U.S.A.), where the GFP sequence was replaced with a mouse immunoglobulin κ -chain signal sequence from pSec Tag A (Invitrogen, Carlsbad, CA, U.S.A.) via *NheI* and *EcoRI*. A new multicloning site, including the mycTag (EQKLISEEDL), was inserted into the *HindIII* and *XbaI* sites by inserting the annealed oligonucleotides 5'-AGCTCGAGTGAGCTCACTAGTGAGC-AGAAGCTGATCAGCGAGGAGGACCTGTCTAGAAAGCTT-A-3' and 5'-CTAGTAAGCTTTCTAGACAGGTCCTCCTCGC-TGATCAGCTTCTGCTCACTAGTGAGCTCACTGG-3'. Used cloning sites were destroyed by the flanking sequence of the insert. The resulting vector was designated as pSM. The pBS-*MUC2C* plasmid was digested with *XbaI* and *HindIII*, and the *MUC2* C-terminal domain was inserted into the same sites in the pSM plasmid. The pSM-C was sequenced (ALF Express, Amersham Biosciences) using Cy5-labelled dATP and the Autoread Sequencing Kit (Amersham Biosciences). Seven mutations at the C-terminus of the *MUC2* causing altered amino acid sequence, compared with the consensus of the *MUC2* sequence [24], were found. The pSM-C plasmid was corrected by site-directed mutagenesis (QuickChange™ site-directed mutagenesis kit; Stratagene, La Jolla, CA, U.S.A.). The corrections were determined by sequencing. In-frame ligation of the GFP sequence between the mycTag and the *MUC2* C-terminus was made possible by deleting the base at position 606 of the pEGFP-C1 plasmid (ClonTech) using site-directed mutagenesis (QuickChange™ site-directed mutagenesis kit; Stratagene). The oligonucleotides 5'-GCTAGCGCTACCGGTGCCACCATGGTGAGCAAGG-3' and 5'-CCTTGCTCACCATGGTGGCACCAGGTAGCGCTAGC-3' were used. The mutated GFP sequence was excised with *NheI* and *XbaI* and ligated into the *XbaI* site of the pSM-*MUC2C* plasmid, creating the pSMG-*MUC2C* vector encoding the C-terminal domain of the human *MUC2* (bases 12622–15708) as a C-terminal fusion protein with the mycTag and the GFP, downstream of the immunoglobulin κ -chain signal sequence.

Mutation of the pSMG-MUC2C expression plasmid

The cysteine residue at position 5121 of the *MUC2* sequence [2] was mutated to a serine residue by site-directed mutagenesis (QuickChange™ site-directed mutagenesis kit; Stratagene) using the oligonucleotides 5'-CCTGGACCACAGCTGCTCCAGCT-GCAAAGAGGAGAAAACC-3' and 5'-GGTTTTCTCCTCTTT-GCAGCTGGAGCAGCTGTGGTCCAGG-3'. The sequence of the obtained plasmid, pSMG-*MUC2C* (C5121S), was confirmed by DNA sequencing as described above.

Tissue culture

The colon adenocarcinoma cell line LS 174T (CL-188; A.T.C.C.) and the CHO-K1 cells (CCL-61; A.T.C.C.) were cultured as described previously for LS 174T [9].

DNA transfection to generate stable clones expressing pSMG-MUC2C

Cells were washed in transfection media [7 mM sodium phosphate (pH 7.4)/1 mM MgCl₂/272 mM sucrose] and dissolved at a concentration of 10⁷ cells/500 µl of transfection media containing 10 µg of purified plasmid. Cells were transfected by electroporation with a Gene Pulser II RF module (Bio-Rad Laboratories, Hercules, CA, U.S.A.). Transfections of the CHO-K1 cells were performed using the following parameters: 400 V, 40 kHz and 50 % modulation with 15 bursts during 1 ms each with 1 s intervals. The LS 174T cells were transfected using the following parameters: 300 V, 40 kHz and 70 % modulation with 13 bursts of 3 ms duration each with 0.5 s intervals. Cells were seeded in 56 cm² Petri dishes and cultured as above. Selection with G418 (125 µg/ml for the LS 174T and 250 µg/ml for the CHO-K1 cells respectively) to obtain stable transfectants was started 48 h after transfection. Screening of clones expressing the construct was performed using fluorescence microscopy and immunoprecipitation (results not shown).

Transient transfection with pSMG-MUC2C (C5121S) or pSMG-MUC2C

CHO-K1 cells were grown in 21.5 cm² Petri dishes to approx. 80 % confluency. The cells were then transfected according to the manufacturer's instructions using 16 µg of plasmid [pSMG-MUC2C (C5121S) or pSMG-MUC2C] and 100 µg of DMRIE-C Reagent (Invitrogen, Carlsbad, CA, U.S.A.) per dish. Two days after the transfections were completed, the cells were examined by fluorescence microscopy to verify protein expression, and the transfected cells were then used for further studies.

Metabolic labelling

Confluent cells were labelled metabolically by preincubation in methionine- and cysteine-free minimal essential medium (Gibco), containing 10 % (v/v) foetal bovine serum, 100 i.u./ml penicillin and 100 µg/ml streptomycin for 1–2 h, followed by the addition of 37–115 µCi/ml labelling mix (Redivue Pro-mix L-[³⁵S] *in vitro* labelling mix; Amersham Biosciences). Chases were done by changing to unlabelled regular medium, supplemented with 15 µg/ml methionine and 25 µg/ml cysteine.

Preparation of cell lysates and media

Cells were washed twice with PBS and 1 ml of lysis buffer [50 mM Tris/HCl (pH 7.9)/150 mM NaCl/1 % (v/v) Triton X-100], containing protease inhibitors [1 mM PMSF (Sigma), 100 µM leupeptin (Sigma) and 100 KIU (kallikrein inhibitory units)/ml aprotinin (Trasolyl; Bayer, Leverkusen, Germany)], and 5 mM *N*-ethylmaleimide (NEM) (Sigma) was added to the cells. After sonication (intensity 15) for 3×2 s, on a MSE Soniprep 100 sonifier, the cell debris was removed by centrifugation (16 000 g for 10 min). Protease inhibitors and NEM were added to the media to give final concentrations as for the cell lysates, and centrifuged to remove possible contamination of cells. All the operations were performed on ice.

Immunoprecipitation

Immunoprecipitation was performed as described, including one or two rounds of precipitation with preimmune serum [9], using 25–40 µl of antiserum, followed by precipitation using 150–

200 µl of 10 % (v/v) Pansorbin (formalin-fixed *Staphylococcus aureus*; Calbiochem, San Diego, CA, U.S.A.). The protein–Pansorbin complexes were washed 4 times in 10 mM Tris/HCl, 2 mM EDTA, 0.1 % (v/v) Triton X-100 and 0.1 % (w/v) SDS (pH 7.4). Precipitation using the α-myc mAb was performed by goat anti-mouse IgG-coupled magnetic beads (Dynabeads; Dynal, Oslo, Norway), precoated with 9E10.2 hybridoma culture supernatant. The proteins coupled with the 'antibody–magnetic bead complexes' were washed 4 times in 50 mM Tris/HCl, 150 mM NaCl and 1 % (v/v) Triton X-100 (pH 7.9).

SDS/PAGE

The bound material was released from the Pansorbin or anti-mouse IgG-coated magnetic beads in sample buffer [25 mM Tris/HCl, 10 % (v/v) glycerol, 2.5 % (w/v) SDS; pH 6.8] with or without 5 % (v/v) 2-mercaptoethanol for 5 min at 95 °C. Samples were analysed by discontinuous SDS/PAGE using 3–5, 3–10 or 3–15 % gradient gels with 3 % stacking gels or 5 % linear gels with 4 % stacking gels [25]. Gels were fixed in 30 % ethanol and 10 % acetic acid, incubated in Amplify (Amersham Biosciences), dried and exposed to Kodak BioMax MS films. The molecular markers High-Range Rainbow [¹⁴C]methylated Protein Molecular Weight Markers (Amersham Biosciences) and Precision Protein Standards (Bio-Rad Laboratories) were used.

Western blotting and detection

Semi-dry electrotransfer of the proteins to PVDF membranes (0.45 µm Immobilon-P; Millipore, Bedford, MA, U.S.A.) was made at 2.5 mA/cm² in a Transfer-Blot SD-Dry Transfer Cell (Bio-Rad Laboratories). The transfer buffer used contained 48 mM Tris, 39 mM glycine, 1.3 mM SDS and 10 % methanol. A blocking solution [PBS containing 5 % (w/v) milk powder/0.1 % Tween 20/0.05 % (w/v) NaN₃] was used both to prevent unspecific binding to the membrane during the detection steps and as solvent for the antibodies [α-myc mAb (hybridoma supernatant) 1:10 and goat anti-mouse IgG coupled with alkaline phosphatase 1:1000 in blocking solution].

Glycosidase treatments

³⁵S-labelled and immunoprecipitated (using α-myc mAb) material from cell lysate and media were treated with different glycosidases or reaction buffer alone, for 16–24 h at 37 °C (on shaker), while still attached to the magnetic beads. For complete deglycosylation, neuraminidase (Roche, Indianapolis, IN, U.S.A.) [50 m-units/ml in 50 mM sodium citrate buffer (pH 6.0), containing 4 mM CaCl₂ and 100 µg/ml BSA], O-glycosidase (Roche) [15 m-units/ml in 50 mM sodium phosphate buffer (pH 7.0) and 100 µg/ml BSA], peptide N-glycosidase F [PNGase F; 50 units/ml in 50 mM sodium phosphate buffer (pH 7.0), 20 mM EDTA, 100 µg/ml BSA] (Roche) and α-N-acetylgalactosaminidase [α-GalNAcase; 500 m-units/ml in citrate/phosphate buffer (pH 4.0) and 100 µg/ml BSA] (Oxford Glycosystems, Abingdon, Oxon., U.K.) were used successively. When only some of the enzymes were used, the omitted glycosidases were replaced by the reaction buffer. After the last treatment, the samples were dried (using a Speed Vac concentrator) before analysis by SDS/PAGE. When endoglycosidase H [endo H; 100 m-units/ml in 50 mM sodium acetate (pH 6.0), 4 mM CaCl₂ and 100 µg/ml BSA] (Roche) was used to elucidate the localization of the intracellular material, the

material was washed once in lysis buffer before elution from the beads and SDS/PAGE.

Anhydrous hydrogen fluoride treatments

³⁵S-labelled and immunoprecipitated (using α -myc mAb) cell lysate and media were released from the beads by the addition of 50 mM Tris/HCl (pH 6.8), 1% (w/v) SDS and heating (95 °C) for 5 min. The supernatants were freeze-dried and all the equipments used in the experiment were dried under nitrogen (gas). Gaseous hydrogen fluoride was frozen under nitrogen (liquified), warmed by hand until melted and then added to the samples. The reactions were allowed to proceed for 1.5 h at 0 °C. The hydrogen fluoride was then removed using nitrogen (gas) and speed-vacuum concentrated successively. Samples were dissolved in 20 μ l of reducing sample buffer [25 mM Tris/HCl (pH 6.8), 10% (v/v) glycerol, 2.5% (w/v) SDS and 5% (v/v) 2-mercaptoethanol] with the addition of 1 μ l of 2 M Tris base before SDS/PAGE.

Purification of SMG-MUC2C

After growing CHO-K1-pSMG-MUC2C cells, 3–4 days after confluence, the cells were washed twice with PBS and Iscove's modified Dulbecco's medium (Gibco) was added. Spent culture media were harvested and new media added every third day. After the addition of protease inhibitors [1 mM PMSF (Sigma), 100 μ M leupeptin (Sigma), 100 KIU/ml aprotinin (Trasylol; Bayer) and 5 mM NEM (Sigma)], the media were centrifuged at 1000 g for 10 min (4 °C) and 0.02% (w/v) of NaN₃ was added. The supernatant was ultrafiltered (YM-100; Millipore), to remove smaller components (<100 kDa) and to exchange the cell-culture media for a 50 mM Tris/HCl buffer (pH 8.0). Ion-exchange chromatography was performed on a MonoQ HR 5/5 column (Amersham Biosciences) using a flow rate of 1.0 ml/min. After washing with 150 mM NaCl in 50 mM Tris/HCl (pH 8.0), bound components were eluted using a linear NaCl gradient (150–600 mM) in 50 mM Tris/HCl (pH 8.0) over 45 min. Column fractions containing the recombinant MUC2 C-terminus were pooled, concentrated using ultrafiltration (Vivaspin 2 ml concentrator, 30000 MWCO RC, 3000 g, 4 °C), loaded directly on a Superose 6 HR 10/30 column (Amersham Biosciences) and eluted with 150 mM ammonium acetate (pH 8.5) at 250 μ l/min. The combined fractions containing the recombinant C-terminus were used for the subsequent studies.

Electron microscopy

The SMG-MUC2C was analysed by negative staining and electron microscopy as described previously [26]. Usual sample concentrations were approx. 20 μ g/ml in 50 mM Tris/HCl and 0.15 M NaCl (pH 7.4). Aliquots (5 μ l) were adsorbed on to carbon-coated grids for 1 min, washed with 2 drops of water, and stained on 2 drops of 0.75% uranyl formate. The grids were rendered hydrophilic by glow discharge at low pressure in air. In some experiments, the N-terminal parts of the protein were identified with the mAb α -myc or the mAb α -GFP (ClonTech) labelled with colloidal thiocyanate gold [27]. Specimens were observed in a Jeol JEM 1230 electron microscope operated at 60 kV accelerating voltage. Images were recorded with a Gatan Multiscan 791 CCD camera.

RESULTS

Construction of the pSMG-MUC2C plasmid

cDNA clones and PCR products, spanning the whole C-terminal sequence of the human MUC2 gene (981 amino acids) starting directly after the region encoding the large serine- and threonine-rich tandem repeat domain, were used to construct the pBS-MUC2C plasmid. Originally, the C-terminal domain was placed directly after the endogenous signal peptide of the human MUC2 by digesting the DNA encoding the N-terminus at its *SacI* site. However, this construct never showed any protein expression after transfer to an expression plasmid and transfection. Instead, the immunoglobulin κ -chain signal sequence was used to direct the C-terminal domain into the secretory pathway. The sequence of the C-terminus was added into the pEGFP-C1 vector containing a neomycin-resistance gene and put into frame after the signal peptide, mycTag and GFP, generating the pSMG-MUC2C plasmid.

Expression of the human MUC2 C-terminal domain in CHO-K1 cells

To study the biosynthesis of the recombinant C-terminus of the human MUC2 in a cell line lacking a regulated secretory pathway, the pSMG-MUC2C plasmid was transfected into CHO-K1 cells. After transfection, the expression of the GFP was studied by fluorescence microscopy and used as guidance for the following selection of stable clones. One stable expressing cell line was used for further studies. That the recombinant protein was inserted correctly into the ER, and that it followed the secretory pathway, was evident from the presence of the C-terminus in the media of transfected cells. The protein produced was analysed by both SDS/PAGE and Western-blot analysis (Figure 1A) and by metabolic labelling, immunoprecipitation by an α -mycTag mAb and SDS/PAGE (Figure 1B). When the spent culture media were analysed under non-reducing conditions (Figure 1A, lane 5; Figure 1B, lane 4), a high-molecular-mass band migrating with an estimated apparent molecular mass of approx. 470 kDa was observed. When the same media were analysed under reducing conditions, only one band that migrated with an estimated mass of approx. 250 kDa was found (Figure 1A, lane 3; Figure 1B, lane 2). This suggests that the dimer is made up of SMG-MUC2C only. The mass of the reduced secreted C-terminus is larger than the predicted mass of 158 kDa for the recombinant protein with 18 N-glycans. However, the C-terminus found in the cell lysate was smaller than the secreted form and more in line with the predicted mass. It showed an apparent molecular mass of approx. 190 kDa under reducing conditions, and approx. 330 kDa under non-reducing conditions (Figures 1A and 1B). The estimated masses of the non-reduced and reduced secreted C-termini suggest that the C-terminus forms disulphide-bond-stabilized dimers.

From Figure 1, it is evident that the secreted C-terminus has a higher apparent molecular mass compared with the intracellular form. Exo- and endoglycosidase digestions and analysis by SDS/PAGE further explored the cause for this change in migration. The intracellular form of the C-terminus was sensitive to endo H, whereas the secreted form was resistant (Figure 2C). This suggests that most of the intracellular C-terminus localized to the ER. This is also supported by the observation that this form was unaffected by both neuraminidase and O-glycosidase treatments, whereas the secreted C-terminus was affected by these treatments (Figure 2A). Sensitivity to these two enzymes is in line with the CHO-K1 cells producing mainly mono- and disialylated Gal-GalNAc-O-glycans (G.C. Hansson

Figure 1 Expression and secretion of the C-terminus of the human MUC2 in CHO-K1 cells

Media and cell lysates of CHO-K1 cells stably expressing pSMG-MUC2C were analysed by Western blotting (**A**) using a SDS/PAGE [3–10% (w/v) gradient gel] and detection with the α -mycTag mAb. Cell lysate from a stable expressing cell line (lanes 2 and 4) and its spent culture media (lanes 3 and 5) were analysed under reducing (lanes 2 and 3) and non-reducing conditions (lanes 4 and 5). A reduced cell lysate from LS 174T cells stably expressing pSMG-MUC2C is shown for comparison (lane 1). (**B**) The stably expressing CHO-K1 cell line was labelled metabolically for 6 h with [35 S]Promix. Immunoprecipitation using the α -mycTag mAb coupled with magnetic beads was performed on cell lysates (lanes 1 and 3) and media (lanes 2 and 4). The samples were analysed under reducing (lanes 1 and 2) and non-reducing (lanes 3 and 4) conditions on a 3–10% gradient SDS/PAGE gel followed by fluorography. R, reduced samples; NR, non-reduced samples. Positions of molecular-mass standards are indicated on the left-hand side, and the monomeric and dimeric forms of SMG-MUC2C are indicated on the right-hand side.

and F.J. Olson, unpublished work). The mass of the secreted C-terminus was still larger than the intracellular form after sequential treatment with neuraminidase, O-glycosidase, PNGase F and α -GalNAcase (Figure 2A). This mass difference was further explored by deglycosylation using anhydrous hydrogen fluoride treatments and analysis by SDS/PAGE (Figure 2B). After this treatment, which is supposed to remove all glycans, a mass difference was still observed. Thus the main part of the explanation for the larger mass of the secreted C-terminus is due to glycosylation, but there are probably additional modifications of the secreted molecule causing the increased mass.

To analyse further the assembly of the C-terminus, pulse–chase studies were performed and analysed by SDS/PAGE (Figure 3). The monomeric C-terminus was detectable in the cell lysate already after the pulse and reached a maximum level after 15 min. An intracellular dimer was first observed after a 15 min chase, and this component showed a maximum level after 60 min. These time data are in accordance with the interpretation that the dimerization occurs in the ER. Analysing the cell lysates under non-reducing conditions sometimes revealed diffuse high-molecular-mass bands (Figure 1B, lane 3; Figure 3B, lane 30 min), possibly due to the formation of aggregates in the ER. The higher intensity of the band for reduced monomer when compared with

Figure 2 Glycosidase and hydrogen fluoride treatments of the recombinant C-terminus of the human MUC2 mucin expressed in CHO-K1 cells

The CHO-K1 cell line expressing the C-terminus of the MUC2 mucin was labelled metabolically with [35 S]Promix (**A**, 16 h; **B**, 6 h; **C**, 8 h). Immunoprecipitations using the α -mycTag mAb coupled with magnetic beads were performed on cell lysates and media. After glycosidase (**A**, **C**) or anhydrous hydrogen fluoride treatments (**B**), the samples were analysed by SDS/PAGE on 3–5% (**A**), 3–10% (**B**) or 5% (**C**) gels under non-reducing (**A**) or reducing (**B**, **C**) conditions and analysed by fluorography. (**A**) Cell lysates and media treated sequentially with neuraminidase, O-glycosidase, PNGase F and α -GalNAcase; when enzymes were left out, they were treated with their reaction buffer only. (**B**) Cell lysates and media treated and untreated with anhydrous hydrogen fluoride. (**C**) Cell lysates and media treated with endo H or reaction buffer only.

the non-reduced dimer suggested that some of the intracellular material is probably present in other forms, instead of dimers (cf. lanes 1 and 3 in Figure 1B).

The fully glycosylated, high-molecular-mass species of the SMG-MUC2C dimer was detected in the medium after 180 min (Figure 3). However, this band could be observed already after 60 min if the films were exposed for longer times (results not shown). Interestingly, this band has never been detected in cell lysates (Figure 3) [28]. Thus the passage from the ER, as revealed by the endo H-sensitive SMG-MUC2C, out into the medium is fast. This can only be accounted for if one assumes that the C-terminal dimer is secreted via the constitutive pathway.

Mutation of a cysteine residue that is suggested to be involved in dimerization

Based on the known cysteine residues involved in the dimerization of the TGF β 2 and vWF, site-directed mutagenesis was performed. A pSMG-MUC2C plasmid with a C5121S mutation was created and transiently transfected into CHO-K1 cells. Immunoprecipitations using the α -mycTag mAb and analysis by non-reducing SDS/PAGE revealed that the mutants produced an intracellular disulphide-bond-stabilized dimer (Figure 4, lane 3), as it was resistant to boiling in SDS and was possible

Figure 3 Biosynthesis of the recombinant MUC2 C-terminus in CHO-K1 cells

The CHO-K1 cell line expressing the C-terminus of the MUC2 mucin was labelled metabolically with [35 S]Promix for 15 min and chased in unlabelled medium with excess of methionine and cysteine for 0, 15, 30, 60 and 180 min. Cell lysates and media were prepared and immunoprecipitations, using the α -mycTag mAb coupled with magnetic beads, were performed. Samples were analysed by SDS/PAGE (3–10 % gel) under reducing (**A**) and non-reducing (**B**) conditions and analysed by fluorography. Non-transfected CHO-K1 cells (chased for 180 min) are referred to as control. Positions of molecular-mass standards are indicated on the left-hand side in (**A**) and on the right-hand side in (**B**). The monomeric and dimeric forms of SMG-MUC2C are indicated on the right-hand side.

to reduce to a monomer by dithiothreitol (results not shown). The intracellular form of the mutant C-terminus was sensitive to endo H (Figure 4B), suggesting that it was localized to the ER. Media of CHO-K1 cells transiently transfected with the non-mutated C-terminus gave a band identical with the one in the stable expressing CHO-K1 cells. However, when the media from the transiently transfected CHO-K1 cells with the mutated plasmid were analysed (Figure 4A), no mutated SMG-MUC2C was detected, showing that it was not secreted. This could indicate that the mutated cysteine residue is important for folding and may also play a significant role in the formation of correct dimers accepted by the ER exit-control system.

Electron microscopy of the secreted human MUC2 C-terminal dimers

To analyse the structure and oligomeric status of the SMG-MUC2C secreted from CHO-K1 cells, the purified non-reduced recombinant protein was subjected to electron microscopy after negative staining. As a guide in the interpretation of the structure, α -GFP and α -myc antibodies were gold-labelled and used for immunostaining the SMG-MUC2C. The general structure revealed dimers that could be labelled with one gold particle at each outer end of the dimer (Figure 5). This picture was identical for both antibodies (α -GFP in Figures 5A–5C and α -

Figure 4 Expression of recombinant MUC2 C-terminus with the mutation C5121S in CHO-K1 cells and endo H treatment of the product

(**A**) CHO-K1 cells expressing pSMG-MUC2C (lanes 1 and 5, stable transfectants; lanes 2 and 6, transient transfectants) or pSMG-MUC2C (C5121S) (lanes 3 and 7, transient transfectants) as well as non-transfected CHO-K1 cells (lanes 4 and 8) were labelled metabolically with [35 S]Promix for 16 h. Immunoprecipitation of cell lysates and media were performed using the α -mycTag mAb coupled with magnetic beads. The proteins were then analysed by SDS/PAGE (3–10 % gel) under non-reducing conditions and analysed by fluorography. (**B**) CHO-K1 cells expressing the mutated C-terminus (transient transfectants) were labelled metabolically with [35 S]Promix for 17 h and immunoprecipitation, using the α -mycTag mAb coupled with magnetic beads, was performed on cell lysates. After endo H treatment (or reaction buffer only) the samples were analysed by SDS/PAGE (5 % gel) under reducing conditions and analysed by fluorography. Lane 1, sample treated with reaction buffer only; lane 2, endo H-treated sample.

myc in Figures 5D–5E), suggesting that the tagged N-terminus of this construct was farthest away from the central part where the two monomers were held together. This is consistent with the localization (see above) of the cystine knot and potential cysteine residues responsible for dimerization to the C-terminal end. The central part of the structure must thus contain at least the C-CK domains [29], as suggested in Figure 5(G). This appears white on the micrographs, suggesting a denser and more condensed protein. A similar phenomenon was observed recently for the oligomerization domains of the MUC2 N-terminus [30]. The central part is attached to the gold-labelled ends via a single flexible thread, which may be a single peptide chain. The nature and position of this peptide in the sequence of MUC2 is currently not possible to predict, but there is a mucin-like sequence in the N-terminal end of the MUC2 C-terminal construct that probably should give an extended structure. The D4 domain is situated between the C-CK domains and the mucin-like sequence, something that could suggest that also the D4 domain is located in the central part together with the C-CK domains. However, the precise localization of the D4 domain in the electron microscopy pictures is presently not possible to predict. The outer part labelled with the mAb must contain the GFP and mycTag sequences.

Expression of the human-MUC2 C-terminal domain in a MUC2-producing cell line

To analyse if the recombinantly expressed MUC2 C-terminus is assembled correctly in a cell line normally expressing the MUC2

Figure 5 Electron microscopy of the purified MUC2 C-terminus dimer stained with gold-labelled antibodies

The purified SMG-MUC2C protein was immunostained on the grid with gold-labelled α -GFP (A–C) and α -myc (D, E) antibodies. The grid was stained negatively and analysed by electron microscopy. A schematic representation of the SMG-MUC2C dimer and potential interpretation of the pictures are given in (G). The small images shown are representative of the structures captured in an extended form. The scale bars represent 50 nm in (A) and 25 nm in (B–F) (shown in F).

Figure 6 Secretion of the recombinant C-terminus of the human MUC2 from LS 174T and CHO-K1 cells

Media of LS 174T and CHO-K1 cells stably expressing pSMG-MUC2C were separated by SDS/PAGE on a 3–10% gradient gel under reducing (lanes 1–3) and non-reducing (lanes 4–6) conditions. After Western blotting, the α -mycTag mAb was used for detection. Media from CHO-pSMG-MUC2C (lanes 1 and 4), non-transfected LS 174T (lanes 2 and 5) and LS 174T-pSMG-MUC2C cells (lanes 3 and 6). R, reduced samples; NR, non-reduced samples. Positions of molecular-mass standards are indicated on the left-hand side, and the monomeric and dimeric forms of SMG-MUC2C are indicated on the right-hand side.

mucin, the pSMG-MUC2C plasmid was transfected into LS 174T cells, stable expressing clones were selected, and one of them was studied further. The spent culture media and cell lysates were analysed by SDS/PAGE and Western blotting using an α -mycTag antibody. The reduced and non-reduced secreted forms of the C-terminus from LS 174T cells migrated as slightly larger

Figure 7 Biosynthesis of the MUC2 C-terminus in LS 174T cells and heterodimerization of the C-terminus with the wild-type MUC2

(A) LS 174T cells stably expressing pSMG-MUC2C were labelled metabolically with [35 S]Promix for 15 min and chased in unlabelled medium with excess of methionine and cysteine for 0 min (lane 1), 15 min (lane 2), 30 min (lane 3) and 60 min (lane 4). Sequential immunoprecipitation was performed with α -MUC2TR antisera (results not shown) and α -myc mAb (lanes 1–4). Samples were analysed by SDS/PAGE (3–15% gel) under non-reducing conditions. The positions of the SMG-MUC2C monomer and dimer as well as the 220-kDa mass marker are indicated. (B) Non-transfected LS 174T cells (lane 1) expressing the MUC2 mucin and stable transfectants of the same cells with pSMG-MUC2C (lane 2) were labelled metabolically for 2 h with [35 S]Promix. Immunoprecipitation using antiserum α -MUC2TR was performed on the cell lysate. Samples were analysed under non-reducing conditions by SDS/PAGE (3–5% gel). Positions of the MUC2 monomer, the MUC2 dimer and the MUC2+SMG-MUC2C heterodimer are marked. The nature of the diffuse band with a mobility lower than the MUC2+SMG-MUC2C heterodimer is not known. The SMG-MUC2C dimer is too small to be observed on this gel.

molecules when compared with the ones from CHO-K1 cells (Figure 6). The non-reduced MUC2 C-terminus from LS 174T cells was also approx. 2-fold larger when compared with the reduced one. The LS 174T cell line is derived from the intestine and is normally synthesizing longer oligosaccharides compared with CHO-K1 cells, suggesting that the mass difference is due to differences in glycosylation.

To analyse further the processing of the recombinant C-terminus in the stable LS 174T cell line expressing the construct, metabolic labelling was performed for 15 min and was followed by different chase times. The C-terminus was immunoprecipitated from the cell lysates using the α -mycTag mAb and analysed by non-reducing SDS/PAGE (Figure 7A). The monomeric C-terminus was visible after the 15 min pulse as well as a weak band that migrated as a dimer. This dimer band became strong after 30 min chase and the monomeric band almost disappeared when chased for 60 min. At this time additional weak and larger bands appeared, possibly reflecting formation of higher oligomers.

The recombinant C-terminus is able to form heterodimers with the wild-type MUC2

If the recombinant C-terminus was forming dimers with itself, it might also interact with the full-length wild-type MUC2 mucin. To analyse if this was the case, the LS 174T cells

stably expressing the C-terminus construct were metabolically labelled and cell lysates immunoprecipitated with an antiserum (α -MUC2TR) reacting with the non-O-glycosylated tandem repeat of the MUC2 mucin. As these epitopes are absent from the recombinant C-terminus, only wild-type MUC2 should be precipitated. When the immunoprecipitate was analysed by SDS/PAGE, the expected MUC2 monomer, as found in wild-type LS 174T cells, was found (Figure 7B, lanes 1 and 2 respectively). However, this was considerably weaker in the cells also expressing the recombinant C-terminus. Instead, an intense band migrating with a slightly lower mobility when compared with the full-length MUC2 monomer was observed. This band was sensitive to reduction (results not shown) and was therefore interpreted as a disulphide-bond-stabilized heterodimer of full-length MUC2 and the recombinant C-terminus, as it migrated at about the expected position for assumed masses of 600 and 1200 kDa for the MUC2 monomers and dimers. That this band contained the recombinant C-terminus was confirmed by Western-blot analysis showing the presence of the mycTag in this band (results not shown). Attempts to verify further this interpretation by immunoprecipitation using the α -mycTag mAb failed, as no bands migrating in the gel were observed. This suggests that the mycTag epitope could be buried in the case of the heterodimer. The band for the full-length MUC2 dimer was very weak, suggesting that the SMG-MUC2C protein competed effectively with the wild-type MUC2 for dimerization. No secreted heterodimers were observed (results not shown).

Does the recombinant C-terminus form heterodimers with the MUC5AC mucin?

The cell line LS 174T expresses not only the MUC2 mucin, but also the MUC5AC mucin. Previous studies have shown that both these mucins form dimers in similar ways, but despite their high sequence similarities in the positions of the cysteine residues in their C-termini they do not form any heterodimers [20]. The observation that the recombinant C-terminus formed heterodimers with the full-length MUC2 in LS 174T cells suggested that the recombinant C-terminus was folded correctly and behaved as endogenous MUC2. To support this further, the probable formation of heterodimers of the C-terminus and the MUC5AC mucin was studied. As a control, lysates of the wild-type LS 174T cells were immunoprecipitated with an antiserum (α -gpdA) against glycopeptide A from rat Muc2, detecting both non-O-glycosylated MUC2 and MUC5AC, and analysed by non-reducing SDS/PAGE. As shown in a previous study [20], bands for the monomers and dimers of both MUC2 and MUC5AC were found (Figure 8, lane 6). The LS 174T cell line stably expressing the recombinant C-terminus was labelled metabolically and immunoprecipitated with the α -MUC2TR and α -gpdA antisera (Figure 8, lanes 3 and 4). The heterodimer of the full-length MUC2 and the C-terminus was the major band precipitated, but also MUC2 as well as MUC5AC monomers and dimers were detected as weak bands. Precipitation with the α -MUC5ACTR antiserum revealed only the MUC5AC monomer and dimer (Figure 8, lane 1), although weaker when compared with the α -gpdA antiserum as observed before [20]. For unknown reasons, it has never been possible to precipitate the MUC2+SMG-MUC2C heterodimers using the α -myc antibody. Lack of additional bands in between the MUC5AC monomer and dimer (lane 1) as well as between the MUC5AC and MUC2 monomers in the lanes precipitated by the α -gpdA antiserum suggests that the C-terminus does not form heterodimers with the MUC5AC mucin. This further supports the idea that the C-terminus is folded correctly

Figure 8 The MUC2 C-terminus heterodimerizes with the MUC2, but not with the MUC5AC mucin in LS 174T cells

LS 174T cells (lane 6) and LS 174T cells stably expressing pSMG-MUC2C (lanes 1–5) were incubated with [35 S]Promix for 90 min and the cell lysates were subjected to immunoprecipitation using antiserum α -MUC5ACTR (lanes 1 and 2), α -gpdA (lanes 3 and 6), α -MUC2TR (lane 4) and mAb α -myc (lane 5). Proteins were analysed by non-reducing SDS/PAGE (3–5% gel). The migration positions of the MUC2 and MUC5AC monomers and dimers and of the MUC2+SMG-MUC2C heterodimer are marked. Lanes 1 and 2 are originally the same lane, but in lane 1 the film was exposed for a longer time.

and behaves like the wild-type MUC2 in the LS 174T cell line.

DISCUSSION

The human MUC2 mucin contains two mucin domains rich in the amino acids serine and threonine. The larger mucin domain is made up of a 23-amino-acid sequence, which is repeated tandemly approx. 100 times. These mucin domains are flanked by the cysteine-rich domains, of which the N- and C-termini are believed to be involved in the oligomerization of MUC2. Previously, we have studied the biosynthesis of the human MUC2 mucin and showed that the apomucin forms dimers in the ER [31]. However, which terminus and cysteine residues are involved has not been determined. Detailed studies on the assembly of the MUC2 mucin have been difficult due to the large mass of the monomer, 540 kDa without glycans, and the heterogeneous glycosylation of the mature mucin giving a mass of up to 3 MDa for its monomer.

In constructing an expression vector for the C-terminus of the human MUC2 mucin, the assembled corresponding cDNA was first ligated to the sequence encoding the endogenous signal peptide of MUC2. However, no expression of the C-terminus was observed by immunoprecipitation and analysis by SDS/PAGE. The reason for this is not known, but when the signal peptide was exchanged with the murine immunoglobulin κ -chain signal sequence, the C-terminus was expressed in the transfected cells. The GFP and mycTag sequences were also added in front of the sequence encoding the full-length MUC2 C-terminus to allow easier detection. In Western-blot analysis, and after metabolic labelling and analysis by SDS/PAGE (both under non-reducing conditions), the recombinant C-terminus was detected as a disulphide-bond-stabilized dimer with an apparent molecular mass of approx. 470 kDa in spent medium from CHO-K1 cells. When the same media were analysed under reducing conditions,

a monomer with a mass of approx. 250 kDa was found. This was larger than the corresponding reduced C-termini found inside the cells. The larger mass of the secreted form is mainly due to glycosylation, but this is probably not the only explanation.

The presence of the GFP protein in the MUC2 C-terminus fusion protein was advantageous in several ways. During the cloning procedure, fluorescent clones were selected, limiting the need for screening by other procedures, since they are destructive to the cells. To be handled correctly by the cell, the C-terminus should be passing through the secretory pathway. This was indicated by the ER-like distribution of GFP (results not shown) and proven by the fact that the recombinant C-terminus was glycosylated and secreted from the cells. The formation of heterodimers between the recombinant C-terminus and the wild-type MUC2 in LS 174T cells should also make it possible to study later stages of the MUC2 assembly and secretion. However, the LS 174 cells grow in large clumps, and it has so far been impossible to obtain useful pictures using conventional confocal microscopy. The pSMG-MUC2C plasmid was constructed with a 10-amino-acid motif detected by the α -myc mAb to allow the specific precipitation of this fusion protein separately from the wild-type MUC2 protein. The mAb was detecting both the monomer and dimer of SMG-MUC2C, but not when linked to the full-length MUC2 monomer. The reason for this discrepancy is not understood currently, but could be due to other domains of the full-length MUC2 covering the myc-binding epitope.

The C-terminus of the human MUC2 apomucin contains 86 cysteine residues. The positions of the cysteine residues are very similar to the C-terminus of the vWF [3], but also to the human mucins MUC5AC [4,5] and MUC5B [6] as well as the PSM [7]. The cysteine residues in the far C-terminal end are part of the cystine-knot motif known to be involved in the dimerization of some human growth factors [2,23]. TGF β 2 dimerizes via a cysteine residue located directly at the N-terminus of the fourth cysteine residue in the cystine-knot motif [14]. Perez-Vilar and Hill [16] have performed a thorough mutagenesis study of the cysteine residues in the C-terminus of PSM and analysed these constructs in transfected COS-7 cells. Although they show a relatively high background 'activity' in the stacking/separation gel interface, they found three cysteine residues in the cystine-knot motif that were likely to be involved in the dimer formation. When one of the corresponding cysteine residues (position 5121) in the MUC2 C-terminus was mutated to a serine, expressed in CHO-K1 cells and analysed in cell lysates and media, it was shown that dimers were still produced inside the cells. However, these dimers were trapped in the ER and consequently not secreted. This could indicate that the mutated cysteine residue is involved directly in the formation of a correct disulphide-bond linkage between the monomers. However, additional cysteine residues are probably involved as well, and therefore the incorrectly formed dimer might not be able to pass out of the ER. It is also possible that this cysteine residue is not involved directly in dimer formation and that the exchange of this residue for a serine results in an altered three-dimensional structure, preventing its transport out of the ER. This result shows that single amino acid substitutions also affect the exit of MUC2 from the ER, giving limitations to mutagenesis experiments for the study of MUC2 assembly.

When the recombinant C-terminus of the PSM was expressed in COS-7 cells, disulphide-bond-stabilized dimers were observed [10]. However, larger material (aggregates?) not entering the gel was found, as well as an independence of N-glycans for dimerization. This is in contrast with our studies of the wild-type MUC2 in LS 174T cells, where MUC2 lacking N-glycans did not

exit the ER [31]. By expressing the human MUC2 C-terminus in COS-7 cells followed by immunoprecipitation, we observed that the MUC2 C-terminus formed dimers in the cells, but that these were not secreted (results not shown). Also, larger disulphide-bond-stabilized forms, maybe non-specific aggregates, were observed in the cells, and it was evident by pulse-chase studies that the larger complexes formed relatively quickly and that these and the monomer did not disappear with time. Together, the results suggest that the C-terminus has difficulties in exiting the ER of the COS-7 cells, possibly due to an abnormal folding and intermolecular disulphide-bond formation. In contrast with the COS-7 cells, the CHO-K1 cells secreted only one band migrating as a dimer, suggesting that certain cell lines can process the recombinant MUC2 C-terminus in a correct way. This was also the case when expressed in the LS 174T cells, as the C-terminus behaved as expected. It quickly formed dimers without any staining in the stacking/separation gel interface and application wells. It also formed heterodimers with the full-length, wild-type MUC2. The amount of recombinant C-terminus expressed was probably higher than the wild-type MUC2, as suggested by the secretion of recombinant C-terminal dimers. These high amounts of C-termini caused an almost complete prevention of the formation of wild-type MUC2 dimers.

It is known that the wild-type MUC2 does not form heterodimers with the MUC5AC mucin also expressed in the same cells, despite the large similarities in the localization of the cysteine residues [18,20]. The recombinant C-terminus behaved in the same way as the full-length MUC2 in this respect. These observations support the recombinant C-terminus as a model for studying the role of the MUC2 C-terminus in the assembly and function of the MUC2 mucin.

There is limited information available on the structure of the MUC2 C-terminal dimer. The electron-microscopy pictures presented (Figure 5), however, give a first general impression and outline of the MUC2 C-terminal structure. The pictures demonstrate unequivocally that this part of MUC2 forms dimers. These are shown as symmetrical dimers possible to label with a maximum of two mAbs, one at each end of the dimer. The pictures also reveal that the protein is dense and compact in the domains holding the dimer together. In a recent study [30], we described also that the MUC2 N-terminal trimer is also held together within a compact and dense domain. This interpretation was supported further by the observation that this part of the N-terminus was resistant to trypsin digestion. The dimer-forming part of the C-terminus is linked to the remaining molecule by a thin and highly flexible thread, probably a single peptide chain. In the full-length mucin, the large mucin domain is attached at the present position for the GFP and mycTags in the recombinant protein. The presence of this thread within the C-terminus will probably give the full-length MUC2 mucin a high flexibility at this end of the molecule. This is of importance probably, as other parts of MUC2, especially the mucin domains, are suggested to be relatively rigid.

The CHO-K1 cells lack a regulated secretory pathway. Thus proteins from this cell line are secreted via the constitutive pathway, as was suggested by the fast and continuous secretion of the recombinant MUC2 C-terminus in the present study. It can be concluded that this part of the MUC2 protein is assembled correctly and is capable of passing the secretory pathway. In addition to a constitutive secretion, the LS 174T cells also have a regulated secretory pathway. This latter cell line also synthesizes normal mucins, namely the MUC2 and MUC5AC mucins. When the recombinant SMG-MUC2C protein was expressed in this cell line, it formed SMG-MUC2C homodimers and heterodimers with MUC2, but not with MUC5AC mucin. This is similar to

previous observations for full-length mucins [20,31], suggesting that the recombinant MUC2 C-terminus behaves as the wild-type MUC2 and takes part in a correct assembly in the ER. When the spent culture media from the LS 174T-pSMG-MUC2C cells were analysed, disulphide-bond-stabilized homodimers of the recombinant C-terminus was found. More detailed pulse-chase studies in the LS 174T-pSMG-MUC2C cells also suggest that the intracellular, fully glycosylated form is not accumulated inside the cell [28]. The secretion occurred without stimulation, and all these results suggest that the SMG-MUC2C dimers pass through these cells via the constitutive pathway. It has been suggested previously that GFP by itself could provide sufficient targeting information for sorting into the regulated secretory pathway [32]. However, this does not seem to be the case for the GFP in the SMG-MUC2C protein in the tested cells. In contrast with the constitutive secretion of SMG-MUC2C dimers, no wild-type MUC2+SMG-MUC2C heterodimers were detected in the medium, despite their formation in the cell. This is not surprising, as the LS 174T cells do not secrete any appreciable amount of full-length MUC2 [31]. The observation of intracellular heterodimers, but only secretion of the recombinant MUC2 C-terminal homodimers, suggests that the MUC2 C-terminus does not contain sorting information targeting it into the regulated secretory pathway. This is in analogy with the vWF, where the C-terminus has no role in targeting the protein to the regulated secretory pathway [33]. The information for vWF targeting to the regulated secretory pathway has instead been localized to the N-terminal part of vWF. The observations made here suggest that this could also be the case for the MUC2 mucin. This interpretation is supported by the observation that no secretion of the recombinant MUC2 N-terminus [30] was observed when expressed in the LS 174T cells (M.E.V. Johansson and G.C. Hansson, unpublished work).

In summary, the recombinant MUC2 C-terminus forms dimers and can replace the full-length MUC2 in the dimerization process occurring in the ER. It is also suggested that the targeting information for sorting MUC2 into the regulated secretory pathway is located in parts, other than the C-terminus.

Ms. Karin Axelsson is acknowledged for skilful technical assistance. The work was supported by the Swedish Science Foundation (No. 7461) and by IngaBritt and Arne Lundbergs Foundation. J.R.G. and Y.S.K. are supported by the VA Medical Research Service.

REFERENCES

- 1 Velcich, A., Yang, W. C., Heyer, J., Fragale, A., Nicholas, C., Viani, S., Kucherlapati, R., Lipkin, M., Yang, K. and Augenlicht, L. (2002) Colorectal cancer in mice genetically deficient in the mucin Muc2. *Science* **295**, 1726–1729.
- 2 Gum, Jr, J. R., Hicks, J. W., Toribara, N. W., Siddiki, B. and Kim, Y. S. (1994) Molecular cloning of human intestinal mucin (MUC2) cDNA. Identification of the amino terminus and overall sequence similarity to prepro-von Willebrand factor. *J. Biol. Chem.* **269**, 2440–2446.
- 3 Sadler, J. E. (1998) Biochemistry and genetics of von Willebrand factor. *Annu. Rev. Biochem.* **67**, 395–424.
- 4 Meerzaman, D., Charles, P., Daskal, E., Polymeropoulos, M. H., Martin, B. M. and Rose, M. C. (1994) Cloning and analysis of cDNA encoding a major airway glycoprotein, human tracheobronchial mucin (Muc5). *J. Biol. Chem.* **269**, 12932–12939.
- 5 Lesuffleur, T., Roche, F., Hill, A. S., Lacasa, M., Fox, M., Swallow, D. M., Zweibaum, A. and Real, F. X. (1995) Characterization of a mucin cDNA clone isolated from HT-29 mucus-secreting cells – the 3' end of MUC5AC? *J. Biol. Chem.* **270**, 13665–13673.
- 6 Desseyn, J. L., Aubert, J. P., van Seuningen, I., Porchet, N. and Laine, A. (1997) Genomic organization of the 3' region of the human mucin gene MUC5B. *J. Biol. Chem.* **272**, 16873–16883.
- 7 Eckhardt, A. E., Timpte, C. S., Abernethy, J. L., Zhao, Y. and Hill, R. L. (1991) Porcine submaxillary mucin contains a cysteine-rich, carboxyl-terminal domain in addition to a highly repetitive, glycosylated domain. *J. Biol. Chem.* **266**, 9678–9686.
- 8 Bhargava, A. K., Woitach, J. T., Davidson, E. A. and Bhavanandan, V. P. (1990) Cloning and cDNA sequence of a bovine submaxillary gland mucin-like protein containing two distinct domains. *Proc. Natl. Acad. Sci. U.S.A.* **87**, 6798–6802.
- 9 Asker, N., Baekstrom, D., Axelsson, M. A. B., Carlstedt, I. and Hansson, G. C. (1995) The human MUC2 mucin apoprotein appears to dimerize before O-glycosylation and shares epitopes with the 'insoluble' mucin of rat small intestine. *Biochem. J.* **308**, 873–880.
- 10 Perez-Vilar, J., Eckhardt, A. E. and Hill, R. L. (1996) Porcine submaxillary mucin forms disulfide-bonded dimers between its carboxyl-terminal domains. *J. Biol. Chem.* **271**, 9845–9850.
- 11 Bell, S. L., Khatiri, I. A., Xu, G. Q. and Forstner, J. F. (1998) Evidence that a peptide corresponding to the rat Muc2 C-terminus undergoes disulphide-mediated dimerization. *Eur. J. Biochem.* **253**, 123–131.
- 12 Bell, S. L., Xu, G. and Forstner, J. F. (2001) Role of the cystine-knot motif at the C-terminus of rat mucin protein Muc2 in dimer formation and secretion. *Biochem. J.* **357**, 203–209.
- 13 McDonald, N. Q. and Hendrickson, W. A. (1993) A structural superfamily of growth factors containing a cystine knot motif. *Cell (Cambridge, Mass.)* **73**, 421–424.
- 14 Daopin, S., Piez, K. A., Ogawa, Y. and Davies, D. R. (1992) Crystal structure of transforming growth factor- β 2: an unusual fold for the superfamily. *Science* **257**, 369–373.
- 15 Katsumi, A., Tuley, E. A., Bodó, I. and Sadler, J. E. (2000) Localization of disulfide bonds in the cystine knot domain of human von Willebrand factor. *J. Biol. Chem.* **275**, 25585–25594.
- 16 Perez-Vilar, J. and Hill, R. L. (1998) The carboxyl-terminal 90 residues of porcine submaxillary mucin are sufficient for forming disulfide-bonded dimers. *J. Biol. Chem.* **273**, 6982–6988.
- 17 Perez-Vilar, J. and Hill, R. L. (1997) Norrie disease protein (norrin) forms disulfide-linked oligomers associated with the extracellular matrix. *J. Biol. Chem.* **272**, 33410–33415.
- 18 Pigny, P., Guyonnet-Duperat, V., Hill, A. S., Pratt, W. S., Galiegue-Zouitina, S., d'Hooge, M. C., Laine, A., van Seuningen, I., Degand, P., Gum, J. R. et al. (1996) Human mucin genes assigned to 11p15.5: identification and organization of a cluster of genes. *Genomics* **38**, 340–352.
- 19 Jackson, A. D. (2001) Airway goblet-cell mucus secretion. *Trends Pharmacol. Sci.* **22**, 39–45.
- 20 Asker, N., Axelsson, M. A. B., Olofsson, S. O. and Hansson, G. C. (1998) Human MUC5AC mucin dimerizes in the rough endoplasmic reticulum, similarly to the MUC2 mucin. *Biochem. J.* **335**, 381–387.
- 21 Hansson, G. C., Baekstrom, D., Carlstedt, I. and Klinga-Levan, K. (1994) Molecular cloning of a cDNA coding for a region of an apoprotein from the insoluble mucin complex of rat small intestine. *Biochem. Biophys. Res. Commun.* **198**, 181–190.
- 22 Gum, Jr, J. R. (1992) Mucin genes and the proteins they encode: structure, diversity, and regulation. *Am. J. Respir. Cell Mol. Biol.* **7**, 557–564.
- 23 Gum, Jr, J. R., Hicks, J. W., Toribara, N. W., Rothe, E. M., Lagace, R. E. and Kim, Y. S. (1992) The human MUC2 intestinal mucin has cysteine-rich subdomains located both upstream and downstream of its central repetitive region. *J. Biol. Chem.* **267**, 21375–21383.
- 24 Gum, J. R., Byrd, J. C., Hicks, J. W., Toribara, N. W., Lampion, D. T. A. and Kim, Y. S. (1989) Molecular cloning of human intestinal mucin cDNAs sequence analysis and evidence for genetic polymorphism. *J. Biol. Chem.* **264**, 6480–6487.
- 25 Laemmli, U. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature (London)* **227**, 680–685.
- 26 Engel, J. and Furthmayr, H. (1984) Electron microscopy and other physical methods for the characterization of extracellular matrix components: laminin, fibronectin, collagen IV, collagen VI, and proteoglycans. *Methods Cell Biol.* **145**, 3–78.
- 27 Baschong, W., Lucocq, J. M. and Roth, J. (1985) Thiocyanate gold: small (2–3 nm) colloidal gold for affinity cytochemical labeling in electron microscopy. *Histochemistry* **83**, 409–411.
- 28 Lidell, M. E., Johansson, M. E. V. and Hansson, G. C. (2003) An autocatalytic cleavage in the C-terminus of the human MUC2 mucin occurs at the low pH of the late secretory pathway. *J. Biol. Chem.*, doi 10.1074/jbc.M210069200
- 29 Perez-Vilar, J. and Hill, R. L. (1999) The structure and assembly of secreted mucins. *J. Biol. Chem.* **274**, 31751–31754.
- 30 Godt, K., Johansson, M. E. V., Lidell, M. E., Morgelin, M., Karlsson, H., Olson, F. J., Gum, Jr, J. R., Kim, Y. S. and Hansson, G. C. (2002) The N-termini of the MUC2 mucin form trimers that are held together within a trypsin-resistant core fragment. *J. Biol. Chem.* **277**, 47248–47256.
- 31 Asker, N., Axelsson, M. A. B., Olofsson, S. O. and Hansson, G. C. (1998) Dimerization of the human MUC2 mucin in the endoplasmic reticulum is followed by a N-glycosylation-dependent transfer of the mono- and dimers to the Golgi apparatus. *J. Biol. Chem.* **273**, 18857–18863.

-
- 32 Molinete, M., Lilla, V., Jain, R., Joyce, P. B. M., Gorr, S.-U., Ravazzola, M. and Halban, P. A. (2000) Trafficking of non-regulated secretory proteins in insulin secreting (INS-1) cells. *Diabetologia* **43**, 1157–1164
- 33 Wagner, D. D., Saffari-pour, S., Bonfanti, R., Sadler, J. E., Cramer, E. M., Chapman, B. and Mayadas, T. N. (1991) Induction of specific storage organelles by von Willebrand factor propolypeptide. *Cell* (Cambridge, Mass.) **64**, 403–413

Received 2 January 2003/20 February 2003; accepted 21 February 2003

Published as BJ Immediate Publication 21 February 2003, DOI 10.1042/BJ20030003