Evidence for a second peptide cleavage in the C-terminal domain of rodent intestinal mucin Muc3

Ismat A. KHATRI*, Rongquan WANG† and Janet F. FORSTNER*[1]

*Research Institute, Department of Structural Biology and Biochemistry, The Hospital for Sick Children, and Department of Biochemistry, University of Toronto, Toronto, Ontario, Canada, and †Department of Gastroenterolgy, Southwestern Hospital, Third Military Medical University, Chongqing, P.R. of China

Rat intestinal mucin Muc3 (rMuc3), like its human homologue (MUC3) and several other membrane mucins, contains a C-terminally located SEA (sea urchin sperm protein, enterokinase and agrin) module, with an intrinsic proteolytic site sequence G↓SIVV (where G↓S is the glycine serine cleavage site). As shown previously [Wang, Khatri and Forstner (2002) Biochem. J. **366**, 623–631], expression of the C-terminal domain of rMuc3 in COS-1 cells yields a V5 epitope-tagged N-terminal glycopeptide of 30 kDa and a Myc- and His epitope-tagged C-terminal glycopeptide of 49 kDa. The present study shows that the 49 kDa membrane-anchored fragment undergoes a further cleavage reaction which decreases its size to 30 kDa. Western blotting, pulse–chase metabolic incubations, immunoprecipitation and deglycosylation with N-glycosidase F were used to detect and identify the proteolytic products. Both the first and second cleavages are presumed to facilitate solubilization of Muc3 at the apical surface of enterocytes and/or enhance the potential for Muc3 to participate in ligand–receptor and signal transduction events for enterocyte function *in vivo*.

Key words: Muc3, mucus glycoprotein, C-terminus, SEA (sea urchin sperm protein, enterokinase and agrin) module, proteolytic cleavage.

INTRODUCTION

Mucus glycoproteins (mucins) are generally classified into membrane and secretory types. Although both types consist of long central regions composed of variable tandem repeats, and are heavily O-glycosylated, the flanking end regions differ in structure and function. Secretory mucin flanking regions are enriched in cysteine and are structurally similar to the C, B and D domains of von Willebrand factor. Like von Willebrand factor, the secretory mucins tend to form intermolecular disulphide bridges and to assemble into dimers, oligomers and polymers. Once secreted into luminal spaces, the polymers form gels that spread over epithelial surfaces, where they lubricate and protect the epithelium from potentially harmful chemicals, micro-organisms, digestion products and other noxious stimuli [1]. The secretory mucins, together with trefoil factors, also play an important role in the restitution of damaged epithelia [2,3]. Membrane mucins, on the other hand, do not form polymers, but are tethered to apical cell surfaces via transmembrane and cytoplasmic domains located near the C-terminus. The long, O-glycosylated extracellular domains project from the cell surface, where they contribute to protection of the epithelia. Some membrane mucins, such as MUC4, contain EGF (epidermal growth factor)-like motifs near the C-termini and have physiological roles that include regulation of cell growth, differentiation, cell–cell adhesion, uterine embryo implantation, and immune cell recognition [4–10]. Of increasing interest is the recognition that membrane mucins can produce soluble forms, either by alternative splicing or by intracellular proteolysis at sites N-terminal to the transmembrane regions [11–15]. The extended, newly soluble extracellular domains and their shorter membrane-attached cytoplasmic tails appear to take on new functions in the initiation of intracellular signal transduction events via the cytoplasmic tail.

Cleavage of the C-terminal polypeptide core by endoproteolysis has been reported for the membrane mucins MUC1, MUC4 (and its rodent homologue Muc4) and rodent Muc3 [6,10,16–18]. The soluble forms are thought to be of particular importance in metastases of malignant tumours and in decreasing the resistance of tumour cells to immune surveillance mechanisms [5]. Because MUC1 and MUC4 share many structural features in their C-terminal domains with the human membrane mucins MUC3, MUC12, MUC13 and MUC17, the phenomena of splicing and post-translational cleavage to produce soluble forms may be widely applicable, both for normal signal transduction and in tumour pathology.

The first C-terminal cleavage of rodent Muc3 takes place within minutes of protein translation. The cleavage site was identified by expressing the C-terminal domain (380 amino acids) in COS-1 and CaCo-2 cells, and using epitope tags at each end of the construct to detect the products. An N-terminal V5 epitope immunoreactive fragment of approx. 30 kDa and a C-terminal Myc immunoreactive glycosylated fragment of approx. 49 kDa were detected. The cleavage occurred within a SEA (sea urchin sperm protein, enterokinase and agrin) module located between the first and second EGF-like motifs. The cleavage sequence (Leu-Ser-Lys-Gly↓Ser-Ile-Val-Val, where the arrow indicates the cleavage site) is similar to a SEA module cleavage sequence reported in MUC1 [18] and several other proteins [19]. Both the extracellular (soluble) and membrane fragments of Muc3 remain associated throughout intracellular transport via non-covalent interactions, and both fragments appear at the plasma membrane. It is assumed that the soluble fragment can then be secreted or shed from the epithelial surface. The fragments can be irreversibly dissociated *in vitro* if cell lysates are treated with SDS or boiling. Both fragments of Muc3 contain N-glycans, but the N-glycans did not strongly influence the cleavage or reassociation phenomena [16,20].

Abbreviations used: ASGF, ascites sialoglycoprotein; EGF, epidermal growth factor; ER, endoplasmic reticulum; rMuc3, rat intestinal mucin Muc3; SEA, sea urchin sperm protein, enterokinase and agrin.

[1] To whom correspondence should be addressed: The Hospital for Sick Children, 555 University Avenue, Toronto, ON, Canada M5G 1X8 (e-mail jfforst@sickkids.on.ca).

During the course of our experiments, it was noted that the 30 kDa N-terminal fragment gave a very broad band on SDS/PAGE, and was reduced in size to 22 kDa after incubation of cell lysates with N-glycosidase F. In Western blots the 30 (and 22) kDa bands reacted intensely with anti-V5 antibody. In some experiments a faint (and unexplained) reaction with anti-Myc antibody was also observed at 30 kDa (for example Figure 6, lane b in [16]). This finding prompted us to investigate the 30 kDa fragment more thoroughly. We have now explored the possibility that more than one peptide might be present at 30 kDa, with the implication that there could be a second cleavage step involved in the processing and solubilization of Muc3.

MATERIALS AND METHODS

Reagents

Cell labelling ^{35}S-Pro-mix containing L-[^{35}S]methionine and L-[^{35}S]cysteine was purchased from Amersham Corp. (Oakville, Ontario, Canada). Protein A–agarose and dithiothreitol were purchased from Sigma Chemical Co. (St. Louis, MO, U.S.A.). N-glycosidase F, protease inhibitor cocktail tablets, RIPA buffer kit and Nitro Blue Tetrazolium/5-bromo-4-chloroindol-3-yl phosphate were purchased from Roche Molecular Biochemicals (Laval, Quebec, Canada). His-Bind® resin and buffer kit were purchased from Novagen Inc. (Madison, WI, U.S.A.).

Antibodies

Monoclonal anti-Myc and anti-(V5 epitope) antibodies were obtained from Invitrogen (Groningen, The Netherlands). Alkaline phosphatase-conjugated goat anti-(mouse IgG) and alkaline phosphatase-conjugated goat anti-(rabbit IgG) antibodies were purchased from Bio-Rad Laboratories (Hercules, CA, U.S.A). Polyclonal anti-6279 antibody, raised in rabbits to a peptide sequence within the C-terminus of rMuc3 (rat intestinal mucin Muc3), has been described previously [21].

C-terminal domain constructs of rMuc3 (p20 and p20s/a)

Constructs p20, p20s/a and the pSec expression vector used in these studies have been described previously [16].

Transfection experiments

DNA transfections in COS-1 cells and Caco-2 cells, and SDS/PAGE and Western blotting of lysates and spent media were performed as described previously [16]. Since the products of p20 and p20s/a were poorly secreted, assays of cell lysates rather than media were chosen for presentation. Purification and deglycosylation of His-tagged proteins with N-glycosidase F has been described previously [16].

Metabolic labelling and immunoprecipitation of expressed proteins

At the end of the transfection period (40 h), COS-1 or Caco-2 cells were pulsed (15 min) with ^{35}S-Pro-mix and chased in unlabelled media for 0, 2 and 4 h as described previously [16]. Cells were lysed in RIPA buffer, centrifuged and supernatants subjected to immunoprecipitation with anti-V5 antibody essentially as described previously [16]. Immunoprecipitates were subjected to SDS/PAGE and gels were dried and exposed to X-ray film. In some experiments, immunoprecipitates were resuspended in RIPA buffer and incubated with N-glycosidase F (10 units) overnight at 37 °C prior to SDS/PAGE.

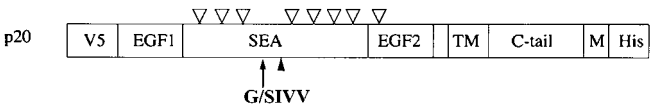

Figure 1 Schematic representation of the product of the p20 construct of rMuc3

cDNA (p20) for the C-terminal 380 amino acids of rMuc3 was cloned into the expression vector pSec. EGF1, EGF2, SEA and TM represent two EGF-like domains, the SEA module and the transmembrane region respectively. V5, M and His refer to V5 epitope, Myc and polyhistidine tags. Open arrowheads indicate N-glycan consensus sites. The closed arrowhead indicates the position of the sequence designated 6279. The arrow points to the G↓S cleavage site in the LSKGSIVV sequence at amino acid position 125 [16].

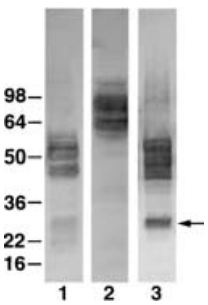

Figure 2 Expression of p20 and p20s/a in COS-1 cells

Products of p20 (lane 1), p20 s/a (lane 2) and HisBind-purified p20 (lane 3) in transfected COS-1 cell lysates were separated by SDS/PAGE and Western blotted using anti-Myc antibody. The arrow shows the position of the 30 kDa band. Molecular mass markers (kDa) are shown on the left.

Denaturation by SDS or heating

Prior to immunoprecipitation, ^{35}S-labelled cell lysates of p20 were subjected to 2 % (w/v) SDS or boiling for 5 min. Samples containing 2 % SDS were diluted with RIPA buffer and the SDS concentration was decreased to 0.4 %. Boiled samples were cooled to room temperature. Immunoprecipitation was then carried out essentially as described previously [16] using anti-V5 antibody.

RESULTS

Cleavage at the G↓SIVV sequence

Figure 1 is a schematic representation of the rMuc3 product of construct p20, with N-terminal V5 and C-terminal Myc and poly(His) epitope tags. As published previously [16,20], post-translational cleavage at the G↓SIVV site in the SEA module generates a 30 kDa V5-positive glycopeptide and a C-terminal 49 kDa glycopeptide bearing the poly(His), Myc and 6279 immunoreactive epitopes. The 6279 sequence has been described previously [21] and refers to a 14-residue synthetic peptide corresponding to a sequence located 10 amino acids C-terminal of the GSIVV site. Transfection of COS-1 (or Caco-2) cells with the p20 construct, followed by SDS/PAGE and Western blotting with anti-Myc antibody, gave a broad or laddered band at 49–55 kDa (Figure 2, lane 1). Transfection with p20s/a, a construct in which the serine of the G/S site was mutated to alanine, prevented cleavage, and the only Myc-positive band appeared as glycoforms

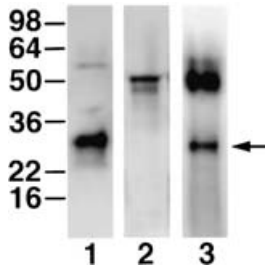

Figure 3 Immunoprecipitation of rMuc3 products

COS-1 cell lysates transfected with p20 were subjected to immunoprecipitation with anti-V5 antibody. The immunoprecipitate was subjected to SDS/PAGE and Western blotting using anti-V5 (lane1), anti-6279 (lane 2) and anti-Myc (lane 3) antibodies. Numbers on the left represent molecular mass markers (kDa). The arrow marks the 30 kDa position.

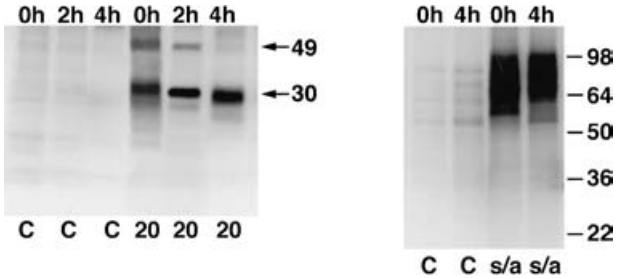

Figure 4 A second cleavage in the C-terminal domain of rMuc3 during biosynthesis

COS-1 cells transfected with p20 (20) or p20s/a (s/a) were pulsed with [^{35}S]Cys/[^{35}S]Met for 15 min and chased for 0, 2 and 4 h. Cell lysates were immunoprecipitated with anti-V5 antibody, separated by SDS/PAGE and detected by autoradiography. C refers to vector controls. Specific bands at 49 and 30 kDa are indicated by arrows. Numbers on the right represent molecular mass markers (kDa).

between 60 and 98 kDa (Figure 2, lane 2). For p20, a faint Myc-positive band was also detected at 30 kDa (lane 1). Since we established previously that the N-terminal cleavage product has a molecular mass of 30 kDa and is intensely stained by anti-V5 antibody [16], anti-Myc reactivity in this location was not expected. This band was initially assumed to be insignificant. However, the band was observed frequently, and therefore we decided to examine it more carefully.

COS-1 cell lysates following p20 transfection were first subjected to affinity purification on a His.Bind resin, and a concentrated sample (two times more concentrated than in previous experiments) was subjected to SDS/PAGE and then Western blotted using anti-Myc antibody. There were two immunoreactive C-terminal products, the expected one at 49–55 kDa and the other at 30 kDa (Figure 2, lane 3, arrow). Neither appeared in cells transfected with p20s/a (Figure 2, lane 2). The presence of the Myc-positive 30 kDa band from p20 expression suggests that once G↓SIVV cleavage occurs, another cleavage in the 49 kDa fragment yields a smaller 30 kDa C-terminal product.

The second proteolytic cleavage occurs at a site beyond the 6279 epitope

A slightly different approach was used to confirm these findings. COS-1 cells were transfected with p20, and cell lysates were subjected to immunoprecipitation using anti-V5 antibody. Anti-V5 was used with the expectation that it would co-precipitate all proteolytic fragments, because they remain associated through non-covalent interactions [16]. The immunoprecipitate was subjected to SDS/PAGE and then Western blotted using anti-V5, anti-6279 and anti-Myc antibodies (Figure 3). An N-terminal V5-positive product appeared at 27–30 kDa (as expected after G↓SIVV cleavage) (Figure 3, lane 1), and the C-terminal 6279- and Myc-positive product appeared at 49–55 kDa (lanes 2 and 3), as expected. There was also a 30 kDa Myc-positive, 6279-negative product (lanes 2 and 3). These findings suggest that the C-terminal 49–55 kDa fragment is cleaved at a site located at or on the C-terminal side of the 6279 epitope (see Figure 1). The 6279 epitope was not detected in any bands in these gels, which probably means that the 6279-positive peptide cleaved from the 49 kDa fragment was too small (<10 kDa) to be retained on the gel. Alternatively, the fragment may not have remained associated with the others, or may have been degraded further and the 6279 epitope destroyed.

In the course of these experiments, we observed that the anti-Myc and anti-6279 antibodies were very successful for Western blots, but that the anti-V5 antibody was a much superior antibody for immunoprecipitation of rMuc3 products. Therefore subsequent immunopreciptation experiments were carried out using the anti-V5 antibody. The anti-Myc and anti-6279 antibodies were restricted to use in Western blots.

Kinetics of secondary cleavage as judged from metabolic pulse–chase studies

Using a [^{35}S]Met/[^{35}S]Cys pulse–chase metabolic protocol [16], we established previously that the G↓SIVV cleavage occurs at an early stage of biosynthesis (within 15 min), but that the major cleaved fragments (30 and 49 kDa) remain associated via non-covalent interactions. Both fragments were thus co-immunoprecipitated by the anti-V5 antibody. In the present study COS-1 (and Caco-2) cells were transfected with p20 or p20s/a constructs, metabolically labelled (15 min) with ^{35}S-Pro-Mix and chased in non-radioactive medium for 0, 2 and 4 h. Cells were lysed in RIPA buffer, immunoprecipitated with anti-V5 antibody and the products separated on SDS/PAGE. Autoradiographs for COS-1 cell lysates (Figure 4) indicate that at 0 h the labelled 49 kDa C-terminal and 30 kDa products were present. By 4 h of chase, the 49 kDa product had disappeared, whereas the 30 kDa product was much more prominent and showed a slight increase in mobility. Chase periods of 2 h gave intermediate results. Transfected Caco-2 cells gave the same results as COS-1 cells (not shown), indicating that the results are not confined to non-mucin-producing COS-1-cells. The p20s/a product did not yield any proteolytic fragments over the 4 h period (Figure 4, right panel), indicating that neither the first nor the second cleavage occurred in this mutant. There was no evidence of a 49 kDa product (or others) in the secretion media of any of the samples (results not shown), indicating no secretion of the extracellular domain within 4 h. Taken altogether, the data to this point are consistent with the interpretation that the broad 30 kDa band may contain two products: an N-terminal V5-positive fragment resulting from the original G↓SIVV cleavage and present at 0 h chase time, and a second fragment arising from cleavage of the C-terminal 49 kDa fragment. The second cleavage appears to occur slowly over a chase period of 0–4 h, and thus may represent continued post-translational processing at a stage beyond the ER (endoplasmic reticulum).

Resolution of the two 30 kDa cleavage products

It was difficult on gels to differentiate between the 30 kDa N-terminal fragment and putative 30 kDa C-terminal fragment

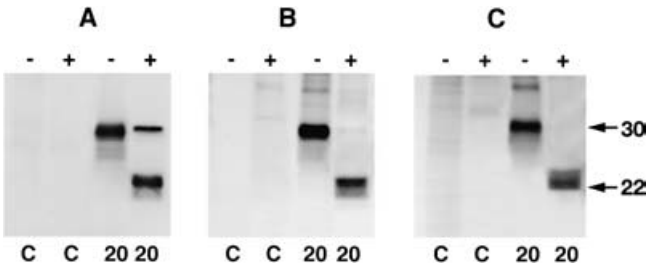

Figure 5 Separation of 30 kDa cleavage products

COS-1 cells transfected with p20 (20) were pulsed with [^{35}S]Cys/[^{35}S]Met for 15 min, chased for 4 h and immunoprecipitated with anti-V5 antibody followed by incubation with (+) or without (−) N-glycosidase F. (**A**) No pretreatment (no SDS, no boiling); (**B**) cell lysates were boiled (95 °C for 5 min) prior to immunoprecipitation; (**C**) cell lysates were incubated with 2% SDS prior to immunoprecipitation. SDS/PAGE and autoradiography were used to detect products in each case. C refers to vector controls. Specific bands at 30 and 22 kDa are indicated by arrows.

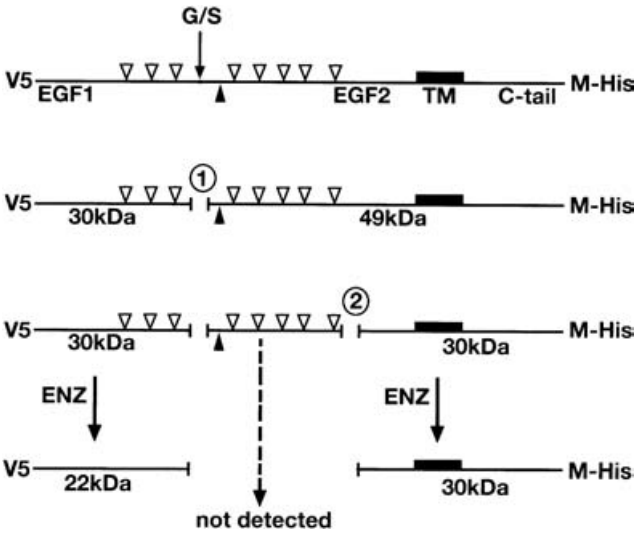

Figure 6 Post-translational cleavages in the C-terminal domain of rMuc3

The schematic shows the rMuc3 product of p20 (top), the G↓S cleavage (1) in the SEA module and the proposed second cleavage region (2) between the 6279 epitope (solid arrowhead) and EGF2. ENZ refers to digestion with N-glycosidase F. EGF1, EGF2, SEA and TM represent two EGF-like domains, the SEA module and the transmembrane region respectively. V5, M and His refer to V5 epitope, Myc and polyhistidine tags.

because of their size similarity, the tendency of cleavage fragments to remain associated via non-covalent interactions, and the inefficiency of the anti-Myc (and anti-6279) antibodies for immunoprecipitation. To resolve the 30 kDa products, two approaches were used: irreversible separation of cleavage products by boiling or SDS treatment [16] prior to immunoprecipitation with anti-V5, and deglycosylation of cleavage products by N-glycosidase F. We reasoned that the N-terminal V5 fragment would be more heavily glycosylated (three consensus sites for N-glycans) (Figure 1) than the C-terminal product of 30 kDa after a second cleavage reaction in the 49 kDa fragment (one or no N-glycan sites). Differential mobility after enzyme treatment was expected to separate the two products.

COS-1 cells transfected with p20 were metabolically labelled with ^{35}S-Pro-Mix for 15 min and chased in non-radioactive medium for 4 h. Prior to immunoprecipitation with anti-V5, lysates were either boiled for 5 min or subjected to 2% (w/v) SDS to dissociate the cleavage fragments irreversibly. The immunoprecipitates were then incubated with N-glycosidase F followed by SDS/PAGE. In each case, appropriate controls (no boiling, no SDS, no enzyme, vector alone) were run to assess the individual effects of each treatment on the immunoprecipitates.

As shown in Figure 5, anti-V5 immunoprecipitates of non-boiled lysates produced a broad band of 27–30 kDa before, and two distinct bands after, incubation with N-glycosidase F (Figure 5A). Most of the 30 kDa band shifted to 22 kDa after N-glycosidase F treatment. On the basis of results published previously [20], this behaviour was expected for the N-terminal V5-positive cleavage fragment, and is consistent with removal of three N-linked oligosaccharides. The remainder of the 30 kDa band was insensitive to the enzyme, and thus its mobility did not change. The enzyme-insensitive product was not present (or gave only a very faint band) in samples that were preboiled (Figure 5B) or treated with SDS (Figure 5C) prior to immunoprecipitation with anti-V5. This means that anti-V5 could not precipitate the 30 kDa enzyme-insensitive product from cell lysates once the fragments were first dissociated by boiling or SDS. This finding is consistent with the interpretation that the enzyme-insensitive 30 kDa product represents a poorly (or non-) glycosylated (Myc-containing) C-terminal cleavage product, while the rest of the 30 kDa band is the V5-positive enzyme-sensitive N-terminal fragment.

The data thus support the conclusion that the C-terminal domain of the membrane rMuc3 undergoes at least two post-translational cleavage reactions, one at the G↓SIVV site in the SEA module and one at a second site that, on the basis of size and immunoreactivity considerations, is judged to be located between the 6279 sequence and the EGF2 region, as shown in the schematic of Figure 6.

DISCUSSION

The second cleavage of rMuc3

The first proteolytic cleavage of rMuc3 within the SEA module occurs in the ER and may represent the first step in the eventual solubilization of rMuc3 after it is transported to the apical membrane of intestinal epithelial cells. There is a growing appreciation that cleaved SEA modules in large O-glycosylated membrane proteins, including mucins, generate new ligand–receptor 'alliance' systems [11,22]. The newly active receptor (cytoplasmic membrane fragment) is recruited for phosphorylation-dependent signal transduction and hence downstream regulation of cell processes including proliferation, differentiation, cell–cell adhesion and metastatic potential of tumour cells.

The present study provides evidence for a second cleavage occurring within or near the SEA module of rMuc3 during post-translational processing. Because the N- and C-terminal 30 kDa peptides remain associated, the two were difficult to distinguish on autoradiographs after immunoprecipitation with anti-V5 antibody. However, treatment of immunoprecipitates with N-glycosidase F separated the broad 30 kDa band into two bands. The 22 kDa band represents the deglycosylated V5-positive fragment, since earlier studies indicated a similar increase in mobility after N-glycosidase F treatment [20]. The enzyme-insensitive 30 kDa band represents a Myc-containing non-glycosylated C-terminal peptide derived from the original 49 kDa fragment. These conclusions were supported by experiments in which the cell lysates were first treated by boiling or SDS to dissociate the fragments. On subsequent immunoprecipitation using anti-V5 antibody, and N-glycosidase F treatment of the immunoprecipitates, the dominant band on autoradiographs was at 22 kDa. The

30 kDa peptide was absent or gave only a very faint band even after long exposure times. This would be expected, since, once dissociated from the other fragment, the 30 kDa Myc-positive product would not be immunoprecipitated by anti-V5.

The central excised peptide containing epitope 6279 (Figure 6) was not recovered on 4–20 % (w/v) polyacrylamide gels, irrespective of the various pretreatments, from which we conclude that its small size may have contributed to rapid degradation during the chase period, or else that the epitope was 'masked' by conformational changes attributed to the presence of N-glycans [23,24]. Another possibility is that it may not be associated with the other fragments even in the absence of SDS or boiling, and therefore was not co-precipitated with anti-V5 antibody. Unfortunately the anti-6279 antibody was not an effective immunoprecipating antibody and thus could not be used to purify or enrich the small central fragment directly for detection. We also failed to detect this peptide in Western blots of COS-1 transfectants of p20 using the anti-6279 antiserum, which favours degradation or epitope loss as the most likely explanations. A schematic illustrating the interpretation of our data is provided in Figure 6.

The specific site of the second cleavage has not been identified. On the basis of observed size and electrophoretic mobility shifts of proteolytic fragments after enzymic deglycosylation, the second site is likely to be near the beginning of or just within the EGF2 motif. The resulting C-terminal peptide would contain approx. 250 amino acids, including the EGF2 motif, the transmembrane region, the C-tail, and the Myc and poly(His) tags. This would account for a product of approx. 27.5 kDa, based on an average mass of 110 Da per amino acid. The proteolytic enzyme responsible and the specific site of cleavage are not known at this time, although a putative serine protease site (RS) is located 20 amino acids inside the N-terminal border of EGF2 beyond the last N-glycan site, and may be relevant.

A second cleavage in MUC4 and Ig-Hepta

Rat Muc4 shares some structural features in common with rMuc3, and has been extensively investigated (both in the rodent as well as its human homologue MUC4) with regard to its role in signal transduction and influences on cell–cell adhesion, cell growth and metastases [5]. Muc4 does not carry a SEA module but is cleaved in the ER at an Asp–Pro sequence, yielding a large extracellular subunit [ASGP1 (ascites sialoglycoprotein 1)] and a C-terminal 120 kDa glycopeptide, ASGP2. The latter, like rMuc3, contains two similarly spaced EGF-like motifs, a transmembrane domain and a cytoplasmic tail. The two subunits of rat Muc4 remain non-covalently associated during cell residence unless treated by boiling, SDS or other denaturants such as urea or guanidine hydrochloride. Komatsu et al. [25] showed recently that a second cleavage takes place near the EGF2 domain to yield a C-terminal peptide of approx. 25 kDa. The investigators propose that the (unidentified) protease responsible for the second cleavage may be extremely important *in vivo*, since down-regulation of the protease would leave the ectodomain of rat Muc4 attached at tumour cell surfaces, where it would maintain a barrier to immunosurveillance mechanisms, and permit continued tumour growth, stimulated by the continued presence of EGF1 in the ASGP2 subunit. Since rMuc3 also carries a similarly placed EGF1 sequence, cleavage of the ectodomain of rMuc3 by a second cleavage may have a similar significance *in vivo*.

A second cleavage in the C-terminal domain has also been shown for Ig-Hepta [26], a membrane protein bearing extracellular immunoglobulin repeats and a C-terminus containing two EGF motifs, a SEA module with a protease-cleavable G↓SVVV sequence, seven transmembrane segments and a C-tail. Ig-Hepta belongs to the family of type II G-protein-coupled receptors, which play numerous roles in cell adhesion and signalling. The second cleavage in Ig-Hepta occurs in the membrane-proximal stalk region just beyond a cystine box motif close to the extracellular face of the cell membrane. The Cys-box and consensus second cleavage site (CKCNHL↓TSFS) are present in a number of SEA module-containing membrane proteins, including some with only one transmembrane domain [27,28]. Although these specific sites are absent from rMuc3, the presence of a SEA module, EGF domains and second juxtamembrane cleavage in Muc3 is sufficiently similar to Ig-Hepta to invite consideration of a similar cleavage mechanism.

Relationship of mucins to signal transduction processes

One of the most intriguing aspects of membrane mucin peptide cleavages is that the cleaved fragments remain associated via non-covalent interactions throughout their passage from the ER to the cell membrane. This is also characteristic of the large family of protease-activated receptors, of which the G-protein-coupled receptors constitute one family [29]. In each case, including Ig-Hepta, a single gene encodes a large peptide which is cleaved by a protease such as thrombin or another serine protease. Cleavage yields an extracellular N-terminal subunit, which serves as a new membrane-associated ligand, and a membrane-tethered subunit which can be phosphorylated and becomes a receptor, and the two remain physically coupled. The protease can thus be viewed as an 'activator' as opposed to a degradative enzyme for the target protein. Once activated, the ligand and receptor moieties remain physically close, which presumably enhances G-protein-mediated signal transduction phosphorylation cascades, and their end results on processes such as cell growth, development, mitogenesis and/or cytokine-mediated inflammatory responses.

Carraway et al. [30] have referred to this type of ligand–receptor coupling as a 'sensor' system for epithelial cell function. In the case of rat Muc4, the EGF1 sequence of protease-released ASGP2 acts as an intrinsic ligand for the Erb B family of receptor tyrosine kinases, forming a stable complex initially with Erb B2. Recruitment of neuregulin and Erb B3 tyrosine kinases, and eventually Erb B1, initiates downstream signalling to mitogen-activated protein kinases, and hence cell transformation and/or proliferation. In the case of MUC1, the mechanism of signal transduction is different, since there are no EGF motifs. Wreschner et al. [22] proposed that protease-stimulated SEA module cleavage causes sequential on–off (association–dissociation) reactions of the partners. This signals sequential phosphorylation–dephosphorylation of multiple 'docking' site tyrosine residues in the cytoplasmic tail receptor [11]. Another signalling mechanism recognized for MUC1 involves the binding of the cytoplasmic subunit to β-catenin, which is enhanced by phosphorylation of a specific tyrosine phosphorylation site (YEKV) and a protein kinase C-δ phosphorylation reaction [31]. It is unlikely that Muc3 shares these mechanisms, since rMuc3 only contains a single putative tyrosine phosphorylation site (YSNF) in the C-tail, and it lacks a protein kinase C site and a catenin binding sequence.

To date, the roles of Muc3 in signal transduction and intercellular adhesion have not been investigated. In view of the presence of EGF domains, a functional SEA module, the generation of soluble forms and a C-tail tyrosine phosphorylation site, however, this appears to be a very fruitful area for future research.

Financial support was obtained from the Canadian Institutes of Health Research and the Canadian Cystic Fibrosis Foundation. R. W. was a postdoctoral fellow supported by the Hospital for Sick Children, Toronto, from 2000 to 2002.

REFERENCES

1 Perez-Vilar, J. and Hill, R. L. (1999) The structure and assembly of secreted mucins. J. Biol. Chem. **274**, 31751–31754
2 Longman, R. J., Douthwaite, J., Sylvester, P. A., Poulsom, R., Corfield, A. P., Thomas, M. G. and Wright, N. A. (2000) Coordinated localisation of mucins and trefoil peptides in the ulcer associated cell lineage and the gastrointestinal mucosa. Gut **47**, 792–800
3 Thim, L., Madsen, F. and Poulsen, S. S. (2002) Effect of trefoil factors on the viscoelastic properties of mucus gels. Eur. J. Clin. Invest. **32**, 519–527
4 Carraway, K. L., Fregien, N., Carraway, K. L. R. and Carraway, C. A. (1992) Tumor sialomucin complexes as tumor antigens and modulators of cellular interaction and proliferation. J. Cell Sci. **103**, 299–307
5 Carraway, K. L., Price-Schiavi, S. A., Komatsu, M., Idris, N., Perez, A., Li, P., Jepson, S., Zhu, X., Carvajal, M. E. and Carraway, C. A. C. (2000) Multiple facets of sialomucin complex/MUC4, a membrane mucin and erb B2 ligand, in tumors and tissues. Front. Biosci. **5**, d95–d107
6 Komatsu, M., Carraway, C. A. C., Fregien, N. L. and Carraway, K. L. (1997) Reversible disruption of cell-matrix and cell-cell interaction by overexpression of sialomucin complex. J. Biol. Chem. **272**, 33245–33254
7 McNeer, R. R., Huang, D., Fregien, N. L. and Carraway, K. L. (1998) Sialomucin complex in the rat respiratory tract: a model for its role in epithelial protection. Biochem. J. **330**, 737–744
8 Moniaux, N., Nollet, S., Porchet, N., Degand, P., Laine, A. and Aubert, J. P. (1999) Complete sequence of the human mucin MUC4: a putative cell membrane-associated mucin. Biochem. J. **338**, 325–333
9 Sheng, Z., Wu, K., Carraway, K. L. and Fregien, N. (1992) Molecular cloning of the transmembrane component of the 13762 mammary adenocarcinoma sialocarcinoma complex: a new member of the epidermal growth factor superfamily. J. Biol. Chem. **267**, 16341–16346
10 Sherblom, A. P. and Moody, C. E. (1986) Cell surface sialomucin and resistance to natural cell-mediated cytotoxity of rat mammary tumor ascites cells. Cancer Res. **46**, 4543–4546
11 Baruch, A., Hartmann, M., Yoeli, M., Adereth, Y., Greenstein, S., Stadler, Y., Skornik, Y., Zaretsky, J., Smorodinsky, N. I., Keydar, I. and Wreschner, D. H. (1999) The breast cancer-associated MUC1 gene generates both a receptor and its cognate binding protein. Cancer Res. **59**, 1552–1561
12 Choudhury, A., Moniaux, N., Ringel, J., King, J., Moore, E., Aubert, J. P. and Batra, S. (2001) Alternative splicing at the 3′-end of the human pancreatic tumor-associated mucin MUC4 cDNA. Teratogen. Carcinogen. Mutagen. **21**, 83–96
13 Choudhury, A., Moniaux, N., Winpenny, J. P., Hollingworth, M. A., Aubert, J. P. and Batra, S. K. (2000) Human MUC4 mucin cDNA and its variants in pancreatic carcinoma. J. Biochem. (Tokyo) **128**, 233–243
14 Moniaux, N., Escande, F., Bartra, S. K., Porchet, N., Laine, A. and Aubert, J. P. (2000) Alternative splicing generates a family of putative secreted and membrane-associated MUC4 mucins. Eur. J. Biochem. **267**, 4536–4544
15 Smorodinsky, N., Weiss, M., Hartmann, M. L., Baruch, A., Harness, E., Yaakobovitz, M., Keydar, I. and Wreschner, D. H. (1996) Detection of a secreted MUC1/SEC protein by MUC1 isoform specific monoclonal antibodies. Biochem. Biophys. Res. Commun. **228**, 115–121
16 Wang, R., Khatri, I. A. and Forstner, J. F. (2002) C-terminal domain of rodent intestinal mucin Muc3 is proteolytically cleaved in the endoplasmic reticulum to generate extracellular and membrane components. Biochem. J. **366**, 623–631
17 Ligtenberg, M. J. L., Kruijshaar, L., Buijs, F., Meijer, M. v. , Litvinov, S. V. and Hilkens, J. (1992) Cell-associated episialin is a complex containing two proteins derived from a common precursor. J. Biol. Chem. **267**, 6171–6177
18 Parry, S., Silverman, H. S., McDermott, K., Willis, A., Hollingsworth, M. A. and Harris, A. (2001) Identification of MUC1 proteolytic cleavages sites in vivo. Biochem. Biophys. Res. Commun. **283**, 715–720
19 Stacey, M., Lin, H., Gordon, S. and McKnight, A. J. (2000) LNB-TM7, a group of seven-transmembrane proteins related to family-B G-protein-coupled receptors. Trends Biochem. Sci. **25**, 284–289
20 Khatri, I. A., Wang, R. and Forstner, J. F. (2003) SEA (sea-urchin sperm protein, enterokinase and agrin) - module cleavage, association of fragments and membrane targeting of rat intestinal mucin. Biochem. J. **372**, 263–270
21 Khatri, I. A., Ho, C., Specian, R. D. and Forstner, J. F. (2001) Characteristics of rodent intestinal mucin Muc3 and alterations in a mouse model of human cystic fibrosis. Am. J. Physiol. Gastrointest. Liver Physiol. **280**, G1321–G1330
22 Wreschner, D. H., McGuckin, M. A., Williams, S. J., Baruch, A., Yoeli, M., Ziv, R., Okun, L., Zaretsky, J., Smorodinsky, N., Keydar, I. et al. (2002) Generation of ligand-receptor alliances by "SEA" module-mediated cleavage of membrane-associated mucin proteins. Protein Sci. **11**, 698–706
23 Helenius, A. (1994) How N-linked oligosaccharides affect glycoprotein folding in the endoplasmic reticulum. Mol. Biol. Cell **5**, 253–265
24 Wormald, M. R., Petrescu, A. J., Pao, Y. L., Glithero, A., Elliot, T. and Dwek, R. A. (2002) Conformational studies of oligosaccharides and glycopeptides: Complementarity of NMR, X-ray crystallography, and molecular modelling. Chem. Rev. **102**, 371–386
25 Komatsu, M., Arango, M. E. and Carraway, K. L. (2002) Synthesis and secretion of Muc4/sialomucin complex: implication of intracellular proteolysis. Biochem. J. **368**, 41–48
26 Abe, J., Fukuzawa, T. and Hirose, S. (2002) Processing and subunit structure of Ig-Hepta, a seven-transmembrane receptor with a long N-terminal extracellular domain. J. Biol. Chem. **277**, 23391–23398
27 Moy, G. W., Mendoza, L. M., Schulz, J. R., Swanson, W. J., Glabe, C. G. and Vacquier, V. D. (1996) The sea urchin sperm receptor for egg jelly is a modulator protein with extensive homology to the human polycystic kidney disease protein, PKD1. J. Cell Biol. **133**, 809–817
28 Vacquier, V. D. and Moy, G. W. (1997) The fucose sulfate polymer of egg jelly binds to sperm REJ and is the inducer of the sea urchin sperm acrosome reaction. Dev. Biol. **192**, 125–135
29 Déry, O., Corvera, C. U., Steinhoff, M. and Bunnett, N. W. (1998) Am. J. Physiol. **274**, C1429–C1452
30 Carraway, K. L., Ramsauer, V. P., Haq, B. and Carraway, C. A. C. (2002) Cell signalling through membrane mucins. BioEssays **25**, 66–71
31 Yamamoto, M., Bharti, A., Li, Y. and Kufe, D. (1997) Interaction of the DF3/MUC1 breast carcinoma-associated antigen and β-catenin in cell adhesion. J. Biol. Chem. **272**, 12492–12494

Received 25 June 2003/9 October 2003; accepted 23 October 2003
Published as BJ Immediate Publication 23 October 2003, DOI 10.1042/BJ20030944