Role of the cystine-knot motif at the C-terminus of rat mucin protein Muc2 in dimer formation and secretion

Sherilyn L. BELL, Gongqiao XU and Janet F. FORSTNER¹

Division of Structural Biology and Biochemistry, Research Institute, The Hospital for Sick Children, 555 University Avenue, Toronto, Ontario, Canada M5G 1X8

DNA constructs based on the 534-amino-acid C-terminus of rat mucin protein Muc2 (RMC), were transfected into COS cells and the resultant ³⁵S-labelled dimers and monomers were detected by SDS/PAGE of immunoprecipitates. The cystine-knot construct, encoding the C-terminal 115 amino acids, appeared in cell lysates as a 45 kDa dimer, but was not secreted. A construct, devoid of the cystine knot, failed to form dimers. Site-specific mutagenesis within the cystine knot was performed on a conserved unpaired cysteine (designated Cys-X), which has been implicated in some cystine-knot-containing growth factors as being important for intermolecular disulphide-bond formation. Dimerization of RMC was effectively abolished. Each cysteine (Cys-1–Cys-6) comprising the three intramolecular disulphide bonds of the cystine knot was then mutated. Dimer formation

was impaired in each case, although much less so for the Cys-3 mutant than the others. Abnormal high-molecular-mass, disulphide-dependent aggregates formed with mutations Cys-1, Cys-2, Cys-4 and Cys-5 and were poorly secreted. It is concluded that the intact cystine-knot domain is essential for dimerization of the C-terminal domain of rat Muc2, and that residue Cys-X in the knot plays a key role. The structural integrity of the cystine knot, maintained by intramolecular bonds Cys-1–Cys-4, Cys-2–Cys-5 and Cys-3–Cys-6, also appears to be important for dimerization, probably by allowing correct positioning of the unpaired Cys-X residue for stable intermolecular cystine-bond formation.

Key words: cysteine mutation, disulphide bond.

INTRODUCTION

The epithelial linings of the gastrointestinal and respiratory tracts, as well as several other organs, are coated with a viscous protective layer of secreted mucins [1]. The gel-like nature of the layer depends critically upon the ability of the highly glycosylated mucin macromolecules to form oligomers and polymers [2]. The monomers of secretory mucins such as human intestinal MUC2, consist of a central polypeptide made up of imperfect tandem repeats rich in serine and/or threonine, and flanking end regions rich in cysteine [3]. Although the precise sequences of the tandem repeats are not conserved amongst different secretory mucins, tandem repeats in all of the mucins share in common the feature of being extensively O-glycosylated. The flanking end regions have unique (non-tandem repeat) sequences, contain some N-linked oligosaccharides, and tend to exhibit considerable interspecies amino acid sequence similarity, as illustrated for example between human MUC2 and rodent Muc2 [4–6]. There is an even more striking similarity of cysteine alignments (implying structural homology), particularly in the C-terminal domains, of secretory mucins of diverse organs and species. The list includes frog integumentary mucin [7], bovine submaxillary mucin [8], porcine submaxillary mucin (PSM) [9], rat Muc2 and human MUC2 [10,11], human MUC5AC, MUC5B and MUC6 [12–15] and porcine gastric mucin [16]. These are the same alignments seen in the oligomer-forming, human blood-clotting protein, von

Willebrand factor (vWF). Thus the mechanisms involved in disulphide-mediated dimer and oligomer formation of mucins are expected to be similar to that of vWF.

To identify subtle post-translational steps and functional roles for C-terminal domains in macromolecules as large as vWF or secretory mucins, it is advantageous to carry out specific domain expression experiments which allow product separation and sizing by gel electrophoresis. In transient expression studies the C-terminal domain of vWF was shown to be critical for disulphide-mediated dimerization [17]. Similar expression systems have shown the C-terminus of PSM [18] and rat intestinal Muc2 (rMuc2) [19,20] to undergo dimer formation. In both vWF and secretory mucins there is a special cysteine-rich motif near the C-terminus, termed the ‘cystine knot’, which is the structural basis for the classification of a growing superfamily of proteins. Members of this superfamily usually undergo dimerization to carry out a specialized function, such as binding to specific receptors, or the formation of functional multimers (reviewed in [21–24]). Detailed structural studies of the cystine-knot motif have not been carried out on any mucin molecules, but X-ray crystallography of transforming growth factor- β 2 (TGF- β 2) and other growth factors has provided a precise picture of the intramolecular disulphide bonds that stabilize the three-dimensional configuration of a cystine knot [25,26]. The knot consists of an internal ring of approximately eight amino acids joined via specific cystine bonds between cysteine residues Cys-2

Abbreviations used: Cys-X, conserved unpaired cysteine within the cystine knot; DMEM, Dulbecco's modified Eagle's medium; DTT, dithiothreitol; ER, endoplasmic reticulum; PDGF, platelet-derived growth factor; PSM, porcine submaxillary mucin; RMC, 534-amino-acid C-terminus of rat Muc2; pCK, construct plasmid containing the cystine knot; pRMC-CK, construct plasmid of RMC lacking the cystine knot; pRMC-CXA, construct plasmid of RMC with Cys⁷⁷⁸ mutated to Ala; pRMC-C1A, pRMC-C2A, pRMC-C3A, pRMC-C4A, pRMC-C5A and pRMC-C6A, construct plasmids of RMC with the Cys residue mutated to Ala at positions 732, 755, 759, 779, 809 and 811, respectively; rMuc2, rat intestinal mucin, Muc2; TGF- β 2, transforming growth factor- β 2; vWF, von Willebrand factor.

¹ To whom correspondence should be addressed (e-mail jfforst@sickkids.on.ca)

Figure 1 Diagram of the cystine-knot region of rMuc2

Numbers 723–837 refer to amino acids in the C-terminal domain, based on a previously published sequence [10]. Circled cysteine residues (C) and broken line connections refer to the intramolecular cystine bonds that stabilize the cystine knot. C^X is an unpaired cysteine residue adjacent to C⁴. C⁹, C^b, C^c, and C^d are other cysteine residues in the sequence which are conserved in mucins and vWF but absent in many other cystine-knot-containing proteins. Cⁱ and Cⁱⁱ refer to cysteine positions absent in mucins but present in PDGF. Arrowheads mark the position of N-glycan attachment sites.

and Cys-5, and Cys-3 and Cys-6 (numbered from the N-terminus of the knot [22] and shown above in Figure 1). The cystine bond from Cys-1 to Cys-4 passes through the centre of the ring to complete the knot. The peptide regions between the cysteine residues form anti-parallel β -sheets that extend outwards from the central knot and twist into three variable external loops of approx. 13–16 residues each. The loops protect the interior of the cystine knot against protease degradation, and also prevent access of the cystine bonds to solvents. Of special interest in the present context is one cysteine residue (designated C^X in Figure 1) located N-terminally adjacent to Cys-4 of the knot. The Cys-X is conserved in almost all dimer-forming cystine-knot structures reported to date, including the gel-forming secretory mucins. The Cys-X residue of monomeric TGF- β 2 is reported to form a stable intermolecular disulphide bond with the equivalent cysteine in a second molecule of TGF- β 2 [25,26]. The same residue is responsible for dimer formation of glial cell line-derived neurotrophic factor [27]. Absence of this cysteine in the vWF polypeptide prevents dimerization, resulting in type IID von Willebrand disease [28] and mutagenesis of the same cysteine residue in norrin protein disrupts oligomer formation [29]. The equivalent cysteine residue is absent in the cystine-knot structure of the non-dimerizing, asymmetric developmental determinant called 'lefty' [30]. Thus the Cys-X appears to be important for dimerization of various cystine-knot-containing polypeptides. However, not all cystine-knot-containing proteins which form dimers depend to the same extent on the Cys-X. Nerve growth factor and platelet-derived growth factor (PDGF), for example, both of which require a cystine-knot motif for dimerization, do not possess a Cys-X adjacent to Cys-4. Nerve-growth-factor dimerization is mediated by non-covalent interactions [31], and PDGF forms disulphide-dependent dimers by utilizing two other intermolecular cysteine residues [22] (the positions of these two cysteine residues are marked Cⁱ and Cⁱⁱ in Figure 1 above, but cysteines are not actually present in these locations in mucin sequences). In C-terminal domain expression studies of PSM, the Cys-X equivalent was reported to be just one of three essential cysteine residues required for dimer formation [32].

In the present study we have expressed the C-terminal domain (RMC) of rMuc2, as well as several mutant constructs, in COS cells to determine if the cystine knot is needed for dimer formation, and whether the specific Cys-X and the intramolecular cystine bonds of the cystine knot play a role. We have also determined the effect of cysteine mutations on secretory competence of the rMuc2 domain.

EXPERIMENTAL

C-terminal domain constructs of rMuc2

Constructs were generated using the plasmid pRMC as a PCR template. pRMC encodes the signal peptide, a three-residue linker, and the C-terminal 534 amino acids of rMuc2 (containing ten N-glycan sites). This construct has been described in previous studies [19,20]. Near the 5'-end, a sequence designated epitope D4553, is recognized by a specific antibody described previously [10]. Deletion and substitution mutations were confined to sequences within the region encoding the C-terminal 115 residues. The resultant constructs and the primers used for PCR, are listed in Table 1. Primers were synthesized and all mutations confirmed by sequencing in both directions (Biotechnology Service Centre, The Hospital for Sick Children, Toronto, Canada).

The construct plasmid containing the cystine knot (pCK), encodes only the C-terminal 115 amino acids of rMuc2. The cystine knot itself consists of residues 732–816 (Figure 1). One cDNA fragment was generated using four self-annealing primers (1–4 in Table 1) to generate the RMC signal peptide, three residue linker, and D4553-synthetic-peptide epitope. This fragment was annealed to another fragment generated by PCR using primers 5 and 6 to encode the C-terminal 115 amino acids of rMuc2. The final recombinant cDNA was 444 nt, and the glycosylated translation product was expected to have a molecular mass of approx. 20 kDa.

The construct pRMC-CK lacks the sequence encoding the 85 amino acids which comprise the cystine knot, two additional amino acids on the N-terminal side, and nine amino acids on the C-terminal side of the cystine knot (i.e. deletion of residues 730–825, Figure 1). The 5'-segment was made by PCR using sense and antisense primers 7 and 8 to generate a product encoding the signal peptide, D4553 epitope and three C-terminally adjacent amino acids. The 3'-segment was made with sense and antisense primers 9 and 10, to generate a product encoding the C-terminal 12 amino acids of rMuc2. Each segment contained complementary overlap sequences (shown in **bold** in Table 1) which were annealed and extended using recombinant PCR conditions described previously [19], except that *Pfu* polymerase was used instead of *Taq* polymerase. The recombinant PCR product was 1390 nt in size, and sequencing confirmed that no frameshifts had been introduced by the deletion. The translation product was expected to contain 440 amino acids, to have nine N-glycan sites, and to have a molecular mass of approx. 70 kDa.

The construct plasmid pRMC-CXA, involves a single cysteine to alanine mutation at position 778 of rMuc2 (marked Cys-X in Figure 1). Site-specific mutagenesis was carried out using the QuikChange mutagenesis kit of Stratagene (La Jolla, CA, U.S.A.), with primers 11 and 12 designed according to the manufacturer's instructions.

The plasmid constructs pRMC-C1A, pRMC-C2A, pRMC-C3A, pRMC-C4A, pRMC-C5A and pRMC-C6A were developed using the QuikChange mutagenesis kit with primers 13–24 (Table 1) to substitute alanine for each cysteine residue involved in the three intramolecular cystine bonds that stabilize the cystine knot (Figure 1; residues 732, 755, 759, 779, 809 and 811, respectively).

Cell culture

COS-1 cells (CRL-1650, A.T.C.C.) were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with penicillin (100 units/ml), streptomycin (100 μ g/ml) and 10% (w/v) fetal-bovine serum (Life Technologies) at 37 °C with 5% CO₂.

Table 1 Mutations of the rMuc2 cystine knot

(s) and (a) refer to sense and antisense primers, respectively. The sites of overlap are in **bold**, and the sites of mutagenesis are underlined.

Construct	Primer number	Primer sequence
pRMC	previously described [19]	
pCK	1(s)	5'-CGTCTAGAATGGGGCTGCCACT AGCTCGCCTGGTGGCT -3'
	2(a)	5'- AAGCCCTTGGCTAAGGCC AGGACTAGGCACAC AGCCACCAGGCGAGCT -3'
	3(s)	5'- GGCCTTAGCCAAGGGCTT GGAAGATGAGTGGTT GGTGAATG -3'
	4(a)	5'- AGGGTTATGGATACAGTGT GGCTTGGACGGGT CATTCAAC -3'
	5(s)	5'- ACACTGTATCCATAACCC TAAACACAC-3'
	6(a)	5'-CGAGCTCCTATCACTTCTTCTAGAAAGC-3'
pRMC-CK	7(s)	5'-CGTCTAGAATGGGGCTGCCACTAGCGCCTGGTGTGTGCTAGTCCTG-3'
	8(a)	5'- AGAACGCCTTGTGTGTTA GGGTTATGG-3'
	9(s)	5'- TAACAACACAAGGCGTTCT AGTCCC-3'
	10(a)	5'-CGAGCTCCTATCACTTCTTCTAGAAAGC-3'
pRMC-CXA	11(s)	5'-ATGGGTGTTCT GCC TGCA-3'
	12(a)	5'-TGCAGGCAGAACCCAT-3'
pRMC-C1A	13(s)	5'-CTAACAACACAGTCCCT GCC TCTGCCATACCCGTC-3'
	14(a)	5'-GACGGGTATGGCAGAG GCA GGGACTGTGTGTTAG-3'
pRMC-C2A	15(s)	5'-CATCTCCATGAATTC GCC GCTGGGTCTGTGG-3'
	16(a)	5'-CCACAGGACCCAGCG GCA AGTTCATGGAGATG-3'
pRMC-C3A	17(s)	5'-CTTCTGCGCTGGGTCC GCT GGGACTTTTGCC-3'
	18(a)	5'-GGCAAAAGTCC AGC GGACCCAGCGCAGAAG-3'
pRMC-C4A	19(s)	5'-ATGGGTGTTCTTGC CC AGGGAGGAGAGGACC-3'
	20(a)	5'-GGTCCTCTCCTCCCTGGCGCAAGAACCCAT-3'
pRMC-C5A	21(s)	5'-ACACACATCGAGAGCGCTCTGTGCCAGGGCACC-3'
	22(a)	5'-GGTGCCCTGGCACAG AGC GCTCTCGATGTGTG-3'
pRMC-C6A	23(s)	5'-ATCGAGAGCTGTCTG CCC AGGGCACCCTG-3'
	24(a)	5'-CAGACGGTGCCCTGG CC CAGACAGCTCTCGAT-3'

Transfection solution (LIPOFECTAMINE[™]; Life Technologies, Rockville, MD, U.S.A.) was added to the COS cells according to the manufacturer's instructions, and the cells incubated for 5 h at 37 °C. Metabolic labelling was performed 48 h after transfection.

Metabolic labelling and immunoprecipitation of expressed proteins

Pulse-chase studies were performed with ³⁵S-Pro-mix *in vitro* cell-labelling mixture (Amersham, Oakville, Ontario, Canada) essentially as previously described [20], except that the chase medium contained 0.24 mg/ml cysteine and 0.15 mg/ml methionine. Cells were pulsed for 15 min, chased in non-radioactive medium for 0–2 h, and lysed in 200 µl of lysis buffer [50 mM Tris/HCl, pH 8.0, 150 mM NaCl, 1 mM EDTA, 1 % Nonidet P40, 10 mM iodoacetamide and 1 × Complete[™] protease inhibitor (Roche Molecular Biochemicals, Laval, Quebec, Canada)] for 1 h in darkness at 4 °C. Immunoprecipitations were also carried out as previously described [20], except that 4 µl of pre-immune and immune sera were used in each reaction. The polyclonal antibody used was generated earlier against the synthetic peptide D4553, corresponding to the N-terminal residues of the 534 C-terminal residues of rMuc2 core polypeptide [10]. Autoradiograms were digitally scanned with Grab-It[™] version 1.57b (UVP, Upland, CA, U.S.A.), and printed using a Kodak 8600 dye-sublimation printer (Rochester, NY, U.S.A.).

RESULTS

Products of pRMC and pCK in cell lysates and medium

Figure 2(A) shows that, after a 15 min pulse-label, the major product of pRMC was an immunospecific ³⁵S-labelled dimer

Figure 2 Products of pRMC and pCK after 0 and 2 h pulse-chase experiments

COS-1 cells were transfected with pSVL (vector alone), pRMC (wild-type) and pCK (cystine knot), and labelled with ³⁵S-Pro-mix for 15 min. Samples were collected after 0 or 2 h chase with non-radioactive medium and were immunoprecipitated with anti-D4553 antibody. Lysates (A and B) and medium (C and D) were subjected to SDS/PAGE under non-reducing conditions (A and C, 8 % gels) and after treatment with 10 mM DTT (B and D, 8 % and 16 % gels). Molecular-mass markers are shown in kDa.

Figure 3 Translation and secretion of pRMC-CK

COS-1 cells were transfected with pSVL (vector control), pRMC (wild-type) and pRMC-CK (without the cystine knot), labelled with ^{35}S -Pro-mix and processed after a 2 h pulse-chase. Lysates (**A** and **B**) and medium (**C** and **D**) were immunoprecipitated and analysed by SDS/PAGE under non-reducing conditions (**A** and **C**, 4–12% gels) and after reduction (**B** and **D**, 8% gels).

present at 165 kDa. There was also a minor product at approximately 100 kDa, which faded dramatically by 2 h. The latter band has been noted in previous studies [20], and is assumed to represent a small amount of precursor monomer. Occasionally there were higher molecular mass oligomers or aggregates seen near the top of the gels. After thiol reduction, however, all of the immunospecific bands were reduced to a 90 kDa monomer (Figure 2B). Thus the C-terminal domain of rMuc2 is able to form disulphide-dependent dimers at an early co- or post-translational period of biosynthesis, presumably in the endoplasmic reticulum (ER).

The construct pCK gave rise in cell lysates (non-reduced) to an immunoreactive product at approximately 45 kDa and a minor band near the bottom of the gel (approximately 30 kDa on 8% gels) (Figure 2A). By 2 h the 30 kDa band faded and the 45 kDa band was more intense, suggesting a precursor-product relationship. After reduction the only immunospecific band seen at 2 h was the 30 kDa band, which was more accurately sized on 16% gels as 22 kDa (Figure 2B, right lane). The 45 and 22 kDa bands are close to the expected molecular mass for glycosylated dimer and monomer products of the cystine knot (40 and 20 kDa respectively, based on the sequence shown in Figure 1). We

conclude therefore that the cystine knot by itself is capable of forming a disulphide-bonded dimer.

Secretion of wild-type RMC was apparent by 2 h (Figure 2C), with the dimer positioned at about 200 kDa and the reduced monomer at 100 kDa (Figure 2D). The slightly lower mobility of both products in the medium relative to cell lysates has been observed previously, and was shown to be due to a late pre-secretory step of sialylation [19]. Unlike RMC, however, the dimer product of the cystine-knot-construct pCK was not secreted (Figures 2C and 2D), which may mean that it fails to pass the quality control system of the ER [33] to exit into the secretory pathway.

Deletion of the cystine knot from RMC prevents dimer formation

The product of construct pRMC-CK lacks residues 730–825, which contain the cystine knot of rMuc2 (Figure 1). The translated immunospecific product (monomer) was expected to be about 70 kDa, which is where it appeared on SDS/PAGE (Figure 3A). Unlike wild-type RMC products compared within the same experiment, almost no dimer was detected, although a faint band (smear) was seen at about 150 kDa. After reduction, the only immunoreactive product of RMC-CK was centred at about 66 kDa (Figure 3B). The monomer was detected in medium samples too, with or without reducing conditions (Figures 3C and 3D). Thus removal of the cystine knot from the C-terminal RMC domain essentially prevented normal dimer formation, but permitted release of some of the mutant monomer into the medium. Together these results suggest that the cystine-knot region of the Muc2 C-terminus is essential for dimer formation, but some other sequence information present in RMC (and also in RMC-CK) may be needed to ensure release into the constitutive secretion pathway.

Importance of Cys-X in dimer formation

The construct pRMC-CXA encodes the entire pRMC product, but the unpaired cysteine residue Cys-X of the cystine knot is mutated to alanine. There was very low dimer formation of RMC-CXA, as shown for the 2 h pulse-chase period (Figure 4A). Under reducing conditions (Figure 4B), both wild-type and mutant constructs gave identically sized monomers, as expected. In medium (Figures 4C and 4D), the monomer of RMC-CXA appeared even under non-reducing conditions, but no dimer was detected. These findings suggest that the unpaired cysteine residue, Cys-X, of the cystine knot is an absolute necessity for dimer formation of the RMC domain, but that the Cys-X plays little or no role in the regulation of secretion. They also show that when overexpressed in COS cells, the mutant monomer of RMC-CXA was secreted.

Structural integrity of the cystine knot is required for correct dimer formation

On the basis of published X-ray crystallographic results of four cystine-knot structures [25,26,31,34,35], it is likely that cysteine residues other than the Cys-X residue play important auxiliary roles in correct mucin-domain-dimer formation and stability. To test this hypothesis, mutations were carried out on pRMC to generate products in which the stabilizing intramolecular cystine bonds (Cys-1–Cys-4, Cys-2–Cys-5, Cys-3–Cys-6) would be disrupted. Although all six mutants under non-reducing conditions

Figure 4 Mutation of Cys-X in the rMuc2 C-terminal-cystine-knot domain

COS-1 cells were transfected with pSVL, pRMC and pRMC-CXA, labelled with ^{35}S -Pro-mix and processed after a 2 h pulse-chase. Lysates (**A** and **B**) and medium (**C** and **D**) were immunoprecipitated and analysed by SDS/PAGE under non-reducing conditions (**A** and **C**, 4–12% gels) and after reduction (**B** and **D**, 8% gels).

produced monomers and even low amounts (faint bands) of dimers in cell lysates (Figure 5A), dimer formation was dramatically impaired relative to the unmutated RMC construct. In the case of Cys-1 and Cys-2 mutations, four or five new high-molecular-mass aggregates ('laddered' over 250 kDa) were also prominent. The mutation at Cys-3 did not appear to cause as much structural damage, since normal sized dimers were more visible and aggregate formation was minimal (Figure 5A). The Cys-4, Cys-5 and Cys-6 mutations gave results that were similar to those of Cys-1, Cys-2 and Cys-3 respectively (results not shown). After reduction the only product of RMC and each mutated species was the monomer (shown for Cys-1, Cys-2 and Cys-3 in Figure 5B). We interpret these results to indicate that disruption of the intramolecular cystine bonds of the cystine knot, particularly those involving Cys-1–Cys-4 and Cys-2–Cys-5, destabilizes the monomer, thereby impairing normal dimer formation and leading to abnormal or scrambled disulphide bonds, and higher-molecular-mass aggregates. Very little secretion of the monomers, dimers or aggregates was observed with any of these mutations, although the product of mutants Cys-3 (dimers and monomers) and Cys-6 (monomers only), were secreted more than the others (shown for Cys-3 in Figures 5C and 5D). These findings are consistent with the interpretation

Figure 5 Mutation of cysteine residues Cys-1, Cys-2 and Cys-3 of the cystine knot

COS-1 cells were transfected with vector alone, pRMC, pRMC-C1A, pRMC-C2A, and pRMC-C3A, labelled with ^{35}S -Pro-mix and processed after a 2 h pulse-chase. Cell lysates (**A** and **B**) and medium (**C** and **D**) were analysed by SDS/PAGE under non-reducing conditions (**A** and **C**, 4–12% gels) and after reduction with 10 mM DTT (**B** and **D**, 8% gels). Abbreviations: C₁, Cys-1; C₂, Cys-2; C₃, Cys-3.

that loss of the intramolecular Cys-3–Cys-6 bond causes less disruption of the tertiary structure of the cystine knot, and less interference with normal Cys-X-mediated-dimer formation, than loss of the Cys-1–Cys-4 and Cys-2–Cys-5 bonds.

DISCUSSION

Mucins have traditionally presented unique difficulties for analysis due to their large size, extensive oligomerization and O-glycosylation. Full-length and polymerized mucins enter even highly porous gels with extreme difficulty, and there is only minimal separation during electrophoresis of biosynthetic intermediates or oligomers, even though these species may differ greatly in molecular size [36–39]. This phenomenon has hampered attempts to characterize mucin dimers and oligomers, and to elucidate the protein features responsible for oligomerization. Although it is an approach not yet used widely, expression of smaller, individual mucin domains using recombinant DNA methodology permits easier detection of post-translational modifications [18,19,32]. The results presented here support the conclusion that the structurally intact cystine knot of rMuc2 is necessary to mediate correct disulphide-dependent-dimer formation of the C-terminal domain. The cystine knot alone (construct pCK) formed a dimer, and the RMC construct lacking the cystine knot (pRMC-CK) formed virtually no dimers. The cysteine adjacent to Cys-4 of the knot (Cys-X) played a pivotal role, since its mutation to alanine (construct pRMC-CXA) essentially prevented dimer formation. The Muc2-domain-dimerization mechanism thus resembles that of TGF- β 2 [25,26],

vWF [28] and glial cell-line derived neurotrophic factor [27] in requiring this specific cysteine residue for dimer formation.

The results in the present study differ somewhat from those derived from transient expressions of the C-terminal domain of PSM [32], particularly with respect to the importance of the Cys-X residue. In PSM the C-terminal 90 amino acids were required for dimer formation, but mutation of the corresponding Cys-X (numbered 13246) to alanine or serine gave products that were still secreted as dimers. Cysteine residues equivalent to Cys-X, Cys-b and Cys-2 in the cystine knot (Figure 1) all appeared to be required for proper dimer formation of the PSM domain, with mutation of Cys-2 having the most deleterious effect. Since the present results place more emphasis on the Cys-X residue for dimer formation of rMuc2, it is predicted that as more research is reported, it will become clear that there is variability amongst secretory mucins in the specific intermolecular disulphide bonds that form and the overall stability of their dimer forms. The key requirement in rMuc2 for the Cys-X does not, of course, rule out auxiliary roles for other amino acids, including other cysteine residues, in dimer formation. Other half-cystines are involved, as shown by mutagenesis of cysteines which form the intramolecular disulphide bonds that stabilize the tertiary structure of cystine knots [25,26,31,34]. Substitution of each of these cysteine residues by alanine severely impaired dimer formation. In the case of Cys-1, Cys-4, Cys-2 and Cys-5, mutation caused the appearance, in cell lysates, of abnormally 'laddered' aggregates at the top of the gel. The aggregates were not (or were poorly) secreted, and they collapsed into monomers under reducing conditions. We conclude, therefore, that the cysteine mutations liberated unpaired cysteine residues which engaged in scrambled or random intermolecular disulphide bonds with similar free cysteine residues of other monomers, resulting in unstable oligomers. It is presumed that the proper folding of the cystine-knot motif in the mucin monomer is needed to position the Cys-X residue for correct intermolecular cystine-bond formation with an equivalent Cys-X residue in other monomers. Substitutions of Cys-3 and Cys-6 were better tolerated (gave more dimer formation and fewer aggregates), thus the Cys-3-Cys-6 bond may not be as important as the other two in stabilizing the tertiary structure and topology of the knot. Intramolecular disulphide bridges are also important in other cystine-knot-containing proteins, with mutations in Cys-2 being especially deleterious in PSM [32] and in norrin protein function in Norrie disease [40]. A naturally occurring mutation of Cys-6 in vWF, described in a patient with von Willebrand disease, reduced the ability of vWF to undergo C-terminal dimerization [41], implying that the Cys-3-Cys-6 bond is very important in stabilizing the cystine-knot structure of this protein.

In full-length secretory mucins, including human MUC2, it has been reported from our laboratory and others [36,37,42-44] that mucin monomers form dimers in the ER before constitutive secretion is possible. The results in this study, in which the C-terminal domain and mutants of rMuc2 were overexpressed in COS cells, showed that dimerization was not always a prerequisite for secretion. For example, the products of pCK formed dimers, but these were not secreted, yet the products of pRMC-CK and pRMC-CXA did not form dimers, but were secreted as monomers. These findings lead to two considerations: (a) that some peptide sequence on the N-terminal side of the cystine knot may be needed as an additional folding 'determinant' to ensure that dimers are not retained in the ER [33]. Relevant to this suggestion is a report [17] that removal of the C domain in vWF (which lies upstream of the cystine knot), allows dimers to form, but prevents their secretion; (b) that mutants either lacking the cystine knot or containing a non-functional CXA mutation,

when overexpressed in COS cells, are released from cells as monomers by some form of 'overflow' pathway. Even in wild-type pRMC constructs some monomer was often noted in medium samples, indicating that the 'overflow' process was not entirely restricted to mutant forms. It is possible that these effects are an artifact of COS-cell-amplification conditions and would not normally occur *in vivo* [45]. Relevant to these interpretations, it was noted in expression studies of PSM in COS-7 cells that some C-terminal domain mutants were secreted as monomers [32]. In addition, expression of the N-terminal domain of human MUC2 in Chinese-hamster ovary cells under neutralized pH conditions also led unexpectedly to the secretion of monomers [46]. At this early stage of knowledge of mucin-secretory-pathway events, we are not able to attribute much significance to domain-monomer secretion in transfected cells.

The major observations from the present study include the importance of the cystine knot in dimer formation of the rMuc2 domain, the prominent role of the unpaired cysteine, Cys-X, in mediating dimerization, and the need for the three internal cystine bonds of the knot to stabilize its structure for the Cys-X residue to function. Both the pivotal role of the Cys-X residue and the relative importance of each cysteine in the knot differ from roles ascribed to cysteine residues of the C-terminus of PSM [32], indicating that there is more than one molecular blueprint for mucin dimerization mechanisms. This is not unlike the situation for the much smaller cystine-knot-bearing growth factors, in which at least three different dimerization mechanisms have been described [25,26,31,34]. Our findings thus offer new insights into early co- and post-translational events that may govern mucin biosynthetic processing and secretion *in vivo*.

Financial support was provided by the Canadian Cystic Fibrosis Foundation. We thank laboratory members Dr Ismat Khatri and Dr Dorothy McCool for their valuable advice.

REFERENCES

- Forstner, J. F. and Forstner, G. G. (1994) Gastrointestinal mucus. In *Physiology of the Gastrointestinal Tract* (Johnson, L. R., ed.), pp. 1255-1283, Raven Press, New York
- Allen, A. (1989) Gastrointestinal mucus. In *Handbook of Physiology: The Gastrointestinal System, III* (Forte, J. G., ed.), pp. 359-382, American Physiology Society, Bethesda
- Gendler, S. J. and Spicer, A. P. (1995) Epithelial mucin genes. *Annu. Rev. Physiol.* **57**, 607-634
- 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. *J. Biol. Chem.* **269**, 2440-2446
- Ohmori, H., Dohrman, A. F., Gallup, M., Tsuda, T., Kai, H., Gum Jr, J. R., Kim, Y. S. and Basbaum, C. B. (1994) Molecular cloning of the amino-terminal region of a rat MUC2 mucin gene homologue. *J. Biol. Chem.* **269**, 17833-17840
- Xu, G., Huan, L., Khatri, I., Sajjan, U. S., McCool, D., Wang, D., Jones, C., Forstner, G. and Forstner, J. (1992) Human intestinal mucin-like protein (MLP) is homologous with rat MLP in the C-terminal region, and is encoded by a gene on chromosome 11p15.5. *Biochem. Biophys. Res. Commun.* **183**, 821-828
- Probst, J. C., Gertzen, E.-M. and Hoffmann, W. (1990) An integumentary mucin (FIM-B.1) from *Xenopus laevis* homologous with von Willebrand factor. *Biochemistry* **29**, 6240-6244
- Bhargava, A. K., Weitach, 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
- Eckhardt, A. E., Timpl, C. S., Abernethy, J. L., Zhao, Y. and Hill, R. L. (1991) Porcine submaxillary mucin contains a cystine-rich, carboxyl-terminal domain in addition to a highly repetitive, glycosylated domain. *J. Biol. Chem.* **266**, 9678-9686
- Xu, G., Huan, L.-J., Khatri, I. A., Wang, D., Bennick, A., Fahim, R. E., Forstner, G. G. and Forstner, J. F. (1992) cDNA for the carboxyl-terminal region of a rat intestinal mucin-like peptide. *J. Biol. Chem.* **267**, 5401-5407
- Gum Jr, J. R., Hicks, J. W., Toribara, N. W., Rothe, E.-M., Legace, 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

- 12 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
- 13 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. *J. Biol. Chem.* **270**, 13665–13673
- 14 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
- 15 Toribara, N. W., Ho, S. B., Gum, E., Gum Jr, J. R., Lau, P. and Kim, Y. S. (1997) The carboxyl-terminal sequence of the human secretory mucin, MUC6. *J. Biol. Chem.* **272**, 16398–16403
- 16 Turner, B. S., Bhaskar, K. R., Hadzopoulou-Cladaras, M. and LaMont, J. T. (1999) Cysteine-rich regions of pig gastric mucin contain von Willebrand factor and cystine knot domains at the carboxyl terminal. *Biochim. Biophys. Acta* **1447**, 77–92
- 17 Voorberg, J., Fontijn, R., Calafat, J., Janssen, H., van Mourik, J. A. and Pannekoek, H. (1991) Assembly and routing of von Willebrand factor variants: the requirements for disulfide-linked dimerization reside within the carboxy-terminal 151 amino acids. *J. Cell Biol.* **113**, 195–205
- 18 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
- 19 Bell, S. L., Khatri, I. A., Xu, G. 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
- 20 Xu, G., Bell, S. L., McCool, D. and Forstner, J. F. (2000) The cationic C-terminus of rat Muc2 facilitates dimer formation post translationally and is subsequently removed by furin. *Eur. J. Biochem.* **267**, 2998–3004
- 21 McDonald, N. Q. and Hendrickson, W. A. (1993) A structural superfamily of growth factors containing a cystine knot motif. *Cell* **73**, 421–424
- 22 Murray-Rust, J., McDonald, N. Q., Blundell, T. L., Hosang, M., Oefner, C., Winkler, F. and Bradshaw, R. A. (1993) Topological similarities in TGF- β 2, PDGF-BB and NGF define a superfamily of polypeptide growth factors. *Structure* **1**, 153–159
- 23 Sun, P. D. and Davies, D. R. (1995) The cystine-knot growth-factor superfamily. *Annu. Rev. Biophys. Biomol. Struct.* **24**, 269–291
- 24 Isaacs, N. W. (1995) Cystine knots. *Curr. Opin. Struct. Biol.* **5**, 391–395
- 25 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
- 26 Schlunegger, M. P. and Grutter, M. G. (1992) An unusual feature revealed by the crystal structure at 2.2 Å resolution of human transforming growth factor- β 2. *Nature (London)* **358**, 430–434
- 27 Haniu, M., Hui, J., Young, Y., Le, J., Katta, V., Lee, R., Shimamoto, G. and Rohde, M. F. (1996) Glial cell line-derived neurotrophic factor: selective reduction of the intermolecular disulfide linkage and characterization of its disulfide structure. *Biochemistry* **35**, 16799–16805
- 28 Schneppenheim, R., Brassard, J., Krey, S., Budde, U., Kunicki, T. J., Holmberg, L., Ware, J. and Ruggeri, Z. M. (1996) Defective dimerization of von Willebrand factor subunits due to a Cys $\rightarrow$ Arg mutation in type IID von Willebrand disease. *Proc. Natl. Acad. Sci. U.S.A.* **93**, 3581–3586
- 29 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
- 30 Meno, C., Saijoh, Y., Fujii, H., Ikeda, M., Yokoyama, T., Yokoyama, M., Toyoda, Y. and Hamada, H. (1996) Left-right asymmetric expression of the TGF β -family member *lefty* in mouse embryos. *Nature (London)* **381**, 151–155
- 31 McDonald, N. Q., Lapatto, R., Murray-Rust, J., Gunning, J., Wlodawer, A. and Blundell, T. L. (1991) New protein fold revealed by a 2.3 Å resolution crystal structure of nerve growth factor. *Nature (London)* **354**, 411–414
- 32 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
- 33 Ellgaard, L., Molinari, M. and Helenius, A. (1999) Setting the standards: quality control in the secretory pathway. *Science* **286**, 1882–1888
- 34 Oefner, C., D'Arcy, A., Winkler, F. K., Eggmann, B. and Hosang, M. (1992) Crystal structure of human platelet-derived growth factor BB. *EMBO J.* **11**, 3921–3926
- 35 Muller, Y. A., Li, B., Christinger, H. W., Wells, J. A., Cunningham, B. C. and De Vos, A. M. (1997) Vascular endothelial growth factor: crystal structure and functional mapping of the kinase domain receptor binding site. *Proc. Natl. Acad. Sci. U.S.A.* **94**, 7192–7197
- 36 McCool, D. J., Forstner, J. F. and Forstner, G. G. (1994) Synthesis and secretion of mucin by the human colonic tumour cell line LS180. *Biochem. J.* **302**, 111–118
- 37 Tytgat, K. M. A. J., Boveland, F.-J., Opdam, F. J. M., Einerhand, A. W. C., Büller, H. A. and Dekker, J. (1995) Biosynthesis of rat MUC2 in colon and its analogy with human MUC2. *Biochem. J.* **309**, 221–229
- 38 Tytgat, K. M. A. J., Swallow, D. M., Van Klinken, B. J.-W., Büller, H. A., Einerhand, A. W. C. and Dekker, J. (1995) Unpredictable behaviour of mucins in SDS/polyacrylamide-gel electrophoresis. *Biochem. J.* **310**, 1053–1054
- 39 Van Klinken, B. J.-W., Einerhand, A. W. C., Büller, H. A. and Dekker, J. (1998) Strategic biochemical analysis of mucins. *Anal. Biochem.* **265**, 103–116
- 40 Strasberg, P., Liede, H. A., Stein, T., Warren, I., Sutherland, J. and Ray, P. N. (1995) A novel mutation in the Norrie disease gene predicted to disrupt the cystine knot growth factor motif. *Hum. Mol. Genet.* **4**, 2179–2180
- 41 Montgomery, R. R., Jozwiak, M. A., Hutter, J. J., Endres, J. L., Foster, P. A. and Friedman, K. D. (1999) A homozygous variant of von Willebrand factor (vWF) that fails to C-terminal dimerize resulting in loss of vWF multimers larger than dimer. *Blood* **94**, (suppl 1), 443a
- 42 Dekker, J. and Strous, G. J. (1990) Covalent oligomerization of rat gastric mucin occurs in the rough endoplasmic reticulum, is N-glycosylation-dependent, and precedes initial O-glycosylation. *J. Biol. Chem.* **265**, 18116–18122
- 43 Klomp, L. W. J., van Rens, L. and Strous, G. J. (1994) Identification of a human gastric mucin precursor: N-linked glycosylation and oligomerization. *Biochem. J.* **304**, 693–698
- 44 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
- 45 Brown, C. A. and Mahuran, D. J. (1993) β -hexosaminidase isozymes from cells cotransfected with α and β cDNA constructs: analysis of the α -subunit missense mutation associated with the adult form of Tay–Sachs disease. *Am. J. Hum. Genet.* **53**, 497–508
- 46 Godl, K., Johansson, M. E. V., Axelsson, K. and Hansson, G. C. (2000) Assembly of the N-terminal part of the MUC2 mucin. In *Sixth International Workshop on Carcinoma-associated Mucins: Mucins in Health and Disease*, p. 94, (Imperial Cancer Research Fund) Cambridge, UK