Published in final edited form as:

Mol Biotechnol. 2013 June ; 54(2): 250–256. doi:10.1007/s12033-012-9562-3.

Increased Understanding of the Biochemistry and Biosynthesis of MUC2 and Other Gel-Forming Mucins Through the Recombinant Expression of Their Protein Domains

Malin Bäckström,

The Mucin Biology Group, Department of Medical Biochemistry, University of Gothenburg, Box 440, 405 30 Göteborg, Sweden

Mammalian Protein Expression Core Facility, University of Gothenburg, Box 440, 405 30 Göteborg, Sweden

Daniel Ambort,

The Mucin Biology Group, Department of Medical Biochemistry, University of Gothenburg, Box 440, 405 30 Göteborg, Sweden

Elisabeth Thomsson,

Mammalian Protein Expression Core Facility, University of Gothenburg, Box 440, 405 30 Göteborg, Sweden

Malin E. V. Johansson, and

The Mucin Biology Group, Department of Medical Biochemistry, University of Gothenburg, Box 440, 405 30 Göteborg, Sweden

Gunnar C. Hansson

The Mucin Biology Group, Department of Medical Biochemistry, University of Gothenburg, Box 440, 405 30 Göteborg, Sweden

Malin Bäckström: malin.backstrom@medkem.gu.se

Abstract

The gel-forming mucins are large and heavily *O*-glycosylated proteins which build up mucus gels. The recombinant production of full-length gel-forming mucins has not been possible to date. In order to study mucin biosynthesis and biochemistry, we and others have taken the alternative approach of constructing different recombinant proteins consisting of one or several domains of these large proteins and expressing them separately in different cell lines. Using this approach, we have determined that MUC2, the intestinal gel-forming mucin, dimerizes via its C-terminal cysteine-knot domain and also trimerizes via one of the N-terminal von Willebrand D domains. Both of these interactions are disulfide bond mediated. Via this assembly, a molecular network is built by which the mucus gel is formed. Here we discuss not only the functional understanding obtained from studies of the recombinant proteins, but also highlight the difficulties encountered when these proteins were produced recombinantly. We often found an accumulation of the

Correspondence to: Malin Bäckström, malin.backstrom@medkem.gu.se.

Parts of the work were performed at the Mammalian Protein Expression Core Facility at the University of Gothenburg.

proteins in the ER and consequently no secretion. This was especially apparent when the cysteine-rich domains of the N- and C-terminal parts of the mucins were expressed. Other proteins that we constructed were either not secreted or not expressed at all. Despite these problems, the knowledge of mucin biosynthesis and assembly has advanced considerably through the studies of these recombinant proteins.

Keywords

Mucin; MUC2; MUC5AC; Recombinant protein; Chinese hamster ovary cells; Disulfide bonds; *O*-glycosylation

Gel-Forming Mucins and Their Protein Domains

Mucins are large glycoproteins that are the main components of the mucus that covers the mucosal surfaces of our bodies. The secreted, gel-forming mucins make up the actual mucus gel and have lubricating and cell-protective functions. The other class of mucins, the transmembrane mucins, is located in the epithelial cell membranes of the mucosal surfaces and form the glycocalyx. The membrane-bound mucins, just as the gel-forming mucins, are involved in cellular protection, but they also take part in cellular signaling via their cytoplasmic tails.

All gel-forming mucins are large proteins with a monomer consisting of about 5,000 amino acids (Fig. 1a) and have at least one serine-, threonine-, and proline-rich sequence (or PTS domain) that serves as attachment sites for *O*-linked glycans [1]. The numerous oligosaccharides attached to these domains give the resulting mucin domains an extended conformation with a “bottle-brush-like” appearance. There are four secreted, gel-forming mucins in the human genome and all are located in a gene cluster on chromosome 11p15 [2]. The four genes *MUC6*, *MUC2*, *MUC5AC*, and *MUC5B* encode mucins with different, but somewhat over-lapping, tissue distributions. *MUC2* is the main intestinal mucin and was recently reviewed [3]. *MUC6* and *MUC5AC* are produced in the stomach with *MUC5AC* also being present in tear fluid and in the airways together with *MUC5B* [2]. This mucin is in addition present in the saliva. *MUC5AC*, *MUC5B*, and *MUC6* are also all found in cervical secretions with *MUC5B* being the most abundant [4]. A fifth human gel-forming mucin gene, *MUC19*, was more recently described [5]. *MUC19* is the human ortholog of the porcine sub-maxillary mucin (PSM), but the expression of the *MUC19* protein has not yet been proven.

The four main human gel-forming mucins, *MUC2*, *MUC5AC*, *MUC5B*, and *MUC6*, have highly similar domain organizations (Fig. 1a). Their N- and C-terminal domains, which flank the large, central *O*-glycosylated mucin domains, are involved in the assembly of the polymeric molecules that build up the mucus gels (Fig. 1b). The N- and C-terminal parts contain several conserved von Willebrand D (VWD)-domains which commonly have more than 1 cysteine per 10 amino acid residues. There are three full and one partial VWD-domain (VWD1, VWD2, VWD', and VWD3) in the N-terminal part of *MUC2*, *MUC5AC*, *MUC5B*, and *MUC6*, and an additional VWD-domain (VWD4) in the C-terminal region of all but *MUC6*. All four gel-forming mucins have a cysteine-knot (CK) domain in their far C-

termini (Fig. 1a). One striking difference between the gel-forming mucins is the number of short cysteine-rich domains (CysD) which are interspersed within the large PTS domains. There are two CysD domains in MUC2, but more than nine in MUC5AC (the exact number is unknown, as this gene is not fully sequenced) and seven in MUC5B [2]. MUC6 does not, in contrast, have any CysD domains (Fig. 1a). Gel-forming mucins also exist in other mammals and contain similar domain structures to those described above for the human mucins [1].

Mucin O-glycosylation

An important characteristic of mucins is the presence of numerous *O*-glycans that are attached to the serines and threonines in the PTS domains. The high density of these often branched oligosaccharides gives the mucin domains their extended structure and will bind water molecules to give the mucins their viscous properties. The specific structure of the *O*-glycans of both gel-forming and transmembrane mucins is the result of the action of the specific glycosyltransferases that are expressed in the Golgi apparatus of the mucin-producing cell. The glycosylation of mucins varies between different tissues and has also been shown to be changed during malignant transformation, as is best known for MUC1 in breast cancer [6]. The structure of the *O*-glycans in different cancer types is often determined by LC-MS and LC-MS/MS of released oligosaccharides, and it can sometimes be difficult to obtain enough tumor tissue to allow a full analysis and characterization of the glycans. Therefore, these investigations have, in some cases, been performed using recombinant mucin domains that have been over-expressed in cancer cell lines, in order to get more protein materials from which the glycans could be released for analysis [7–9]. It is also possible to produce mucin domain-containing recombinant proteins with a specific glycosylation pattern, for example, for functional studies, through the use of different producer cell lines. Examples are wild-type Chinese Hamster Ovary (CHO) K1 [8] or its corresponding mutant cells Lec3.2.8.1 or IdID [10] or cells transfected with specific glycosyltransferases, for example, sialyltransferases [11] or fucosyltransferases [12].

Recombinant Production of Gel-Forming Mucins

To understand the biosynthesis and biochemistry of the gel-forming mucins, we and others have worked with the expression of different domains of the mucins as recombinant proteins. So far no complete gel-forming mucin, with both the N- and C-terminal cysteine-rich domains and an *O*-glycosylated mucin domain in between, has been produced recombinantly in any quantity, despite the attempts that have been made in several labs. There are several possible explanations why this has been difficult, the main reasons most likely being the size, complex biosynthesis, and assembly of these molecules into polymers. The mucins are extremely large proteins. For example, the primary sequence of MUC2 consists of approximately 5,200 amino acids (Fig. 1a) leading to protein which has a theoretical mass of 500 kDa. The newly formed polypeptide chain, with its co-translationally added *N*-glycans, is assembled into dimeric complexes in the ER leading to a glycosylated dimer of about 1.2 MDa. This is followed by extensive *O*-glycosylation and further multimerization in the Golgi into molecular complexes larger than 100 MDa (Fig. 1b). Proteins of this size are challenging to work with and analyze, but they are also difficult for

the cells to produce. In the intestine, MUC2 is produced by specialized goblet cells, and it is clear that there are special features of these cells that are important for their mucin production. Unfortunately, there are no appropriate goblet-like cell lines that can be used for an efficient production of recombinant mucins. There are, however, some epithelial cell lines which produce endogenous mucins. The colon carcinoma cell line LS 174T, for example, makes mainly MUC2, and the colonic carcinoma cell line HT29 makes MUC5AC together with some MUC2. These cell lines could potentially be used for the over-expression of recombinant mucins, but they have other drawbacks. The fact that they are difficult to transfect can probably be overcome, but they will also form covalently mixed products of endogenous and recombinant mucins, and this limits their use considerably.

The cellular machinery of the specialized mucin-producing goblet cells is well adapted for producing these large molecules and assembling them in the correct way. This most likely includes particular ER chaperones and protein disulfide isomerases that assist in protein folding and the formation of the numerous disulfide bonds. One specific example for mucins is the helper protein AGR2 which is necessary for MUC2 to exit the ER, as the lack of AGR2 leads to accumulation of mis-folded MUC2 in the ER [13]. The mucin-producing cells also have a secretory machinery that stores the newly synthesized mucins in regulated secretory vesicles before they are released upon stimulation. The cell lines more commonly used for the production of other recombinant proteins, like CHO and COS cells, may lack specific components or have them in insufficient amounts, for correct folding, assembly, storage, and secretion of full-length mucins. They are therefore not ideal for the purpose of the expression of these molecules.

As it has not been possible to produce full-length mucins, and in order to study the functional properties of individual protein domains, many studies have focused on shorter protein constructs. This has led to an increased knowledge about the roles of both the N-terminal and C-terminal cysteine-rich parts, and also the CysD domains, in mucin biosynthesis and assembly, which are discussed in more detail below.

C-terminal Cysteine-Rich Domains

The C-terminal cysteine-rich domains of MUC2 have been produced recombinantly as a fusion protein with a myc-tag and enhanced green fluorescent protein (GFP). Both tags were located N-terminally to the VWD4 domain, and the protein was termed MG-MUC2C (where M = myc and G = GFP) [14]. The generation of stable CHO K1 clones that secreted this fusion protein was not simple, but was facilitated by the green fluorescence that could be used to identify the transfected cells. During the generation of clones producing MG-MUC2C, we observed that lower levels of intracellular green fluorescence, indicative of lower expression levels, led to a secretion of the MG-MUC2C protein (Fig. 2a). Cells with a high expression of MG-MUC2C showed large green fluorescent aggregates, indicating that the ER was over-loaded with mis-folded mucins (Fig. 2b). Consequently, secretion of MG-MUC2C was undetectable from these cells. In some clones, secretion of MG-MUC2C increased over time, as the expression levels decreased. This phenomenon is probably explained by the very high abundance of cysteine residues (there are more than 80 cysteines in MG-MUC2C) that need to form correct disulfide bonds before they are allowed to exit the

ER. During the screening of clones, we actively selected those with low levels of intracellular green fluorescence. Thereafter we screened these clones for secretion of the MG-MUC2C protein by Western blot. This approach was subsequently also used for CHO K1 cell clone generation protocols for the N-terminal MUC2 and MUC5AC constructs. This inverse correlation between intracellular protein levels and amounts of secreted protein is worthy of consideration when establishing clones for the production of secreted proteins.

CHO K1 clones that secreted the C-terminal cysteine-rich domains of MUC5AC have also been generated. In this case, the protein was produced with an N-terminal myc- and a C-terminal His-tag (M-MUC5AC-CH [15]). Selected clones secreting either MG-MUC2C or M-MUC5AC-CH were then cultured for the collection and purification of both these proteins. It was found that the C-terminal cysteine-rich parts of both MUC2 and MUC5AC formed disulfide-linked covalent dimeric complexes in the ER of CHO K1 cells [14, 15]. These dimers were correctly folded and this allowed ER exit followed by a constitutive secretion. The dimerization occurred via the CK domain at the far C-terminal end. This has also been shown for rat Muc2 using constructs expressed in COS cells, containing either the full C-terminal section or only the last C-terminal 115 amino acids [16, 17]. A construct containing the last 90 residues of PSM, including the CK domain, was also sufficient for the formation of dimers when expressed in COS cells [18]. Dimerization in the C-terminal CK domain was also shown for MUC6 when the whole C-terminal cysteine-rich section of the protein was produced recombinantly in three different cancer cell lines [19]. It can therefore be concluded that the disulfide-linked dimerization in the ER via the C-terminal CK domain is a common theme for all gel-forming mucins and that this process functions in many cell types.

During the purification of the MG-MUC2C protein, it was found that prolonged incubation at pH < 6 resulted in a specific cleavage of the protein into two smaller fragments. The cleavage site was located between the aspartic acid and proline in a GDPH sequence in the VWD4 domain. The cleavage occurred via a pH-dependent autocatalytic mechanism [20]. This cleavage generates a new reactive C-terminus that is potentially involved in cross-linking MUC2 to itself or to other molecules. This is likely of biological importance as the pH in the late secretory pathway is below six, and the cleavage and chemical cross-linking most likely occurs also in vivo. The same GDPH sequence is present in the VWD4 domain of MUC5AC. When the M-MUC5AC-CH protein was expressed in CHO K1 cells, this protein was also found to be auto-catalytically cleaved at this site. In MUC5AC, however, the cleavage was not dependent on low pH as it occurred at a neutral pH, as found in the ER, although it was accelerated by a reduction of the pH [15].

An additional observation from the studies of MG-MUC2C is that when expressed in the colonic cell line LS 174T, the molecule formed heterodimers with the endogenous, full-length MUC2 that was stored in the regulated secretory vesicles of these cells. It was, however, also constitutively secreted, but then as homodimers [14]. As MG-MUC2C, but not the endogenous MUC2 was secreted from LS 174T cells, one can conclude that the information for targeting MUC2 to the regulated secretory vesicles does not reside within the C-terminal domains that are present in MG-MUC2C.

N-terminal Cysteine-Rich Domains

The MUC2 N-terminal VWD1, VWD2, VWD', and VWD3 domains have also been produced as a fusion protein with myc and GFP (MUC2N-MG). Transfection of CHO K1 cells gave rise to a secreted protein that migrated at 250 kDa in SDS gel electrophoresis after reduction of its disulfide bonds [21] (Fig. 3). During the process of selecting cell clones producing MUC2N-MG, we observed several clones that showed large intracellular aggregates of the GFP fusion protein, just as was found in some CHO K1 clones expressing MG-MUC2C (Fig. 2). We finally managed to establish CHO K1 clones that secreted some MUC2N-MG that could be collected and purified for further studies. When analyzed under reducing and nonreducing conditions, it was found that this protein formed disulfide bond stabilized oligomers [21] and that this oligomerization occurred in the late Golgi. Due to their large size, it was difficult to determine their exact size and the number of subunits that formed the oligomer. After trypsin-cleavage of the oligomer, a trypsin-resistant core fragment was found, which contained parts of the VWD3 domain. Analysis of this smaller oligomer led to the conclusion that the oligomers had a trimeric structure [21]. This is in contrast to the VWF, whose VWD3 domain forms dimeric complexes, also in the late Golgi. An intramolecular disulfide bond between VWD1D2 and VWD3 prevents the formation of these disulfide-bonded dimers earlier during its biosynthesis in the ER [22]. A similar mechanism probably also exists in MUC2, as the MUC2 VWD3 domain alone, when produced as a shorter recombinant protein without the VWD1D2 domains, trimerizes in the ER (unpublished data). The difference in oligomerization between VWF and MUC2, in that they dimerize and trimerize, respectively, is important for the final oligomeric structures of these proteins. VWF is a linear thread-like molecule, and MUC2, on the contrary, builds up nets or flat sheets [3].

The N-terminal cysteine-rich sections of PSM have been expressed in COS cells and were also suggested to form trimers [23]. Recently, a recombinant MUC6 N-terminal protein was produced in different cancer cell lines [19]. The oligomeric form of this protein may be interpreted as larger than a dimeric complex, but again, due to its large size, it was difficult to conclude as to the number of subunits in the complex.

We have for some time attempted to produce the N-terminal cysteine-rich domains of MUC5AC, either alone or as a fusion protein with both a myc-tag and GFP in order to elucidate its oligomeric nature. When expressed in CHO K1 cells, the MUC5AC N-terminal protein MUC5ACN-MG was not secreted and was most likely retained within the ER (Fig. 3). The reason as to why MUC5ACN-MG is not secreted from the cell in contrast to the corresponding proteins from MUC2 and MUC6 is not obvious. This illustrates the difficulties in designing recombinant proteins when the domain borders are not fully known. It is possible that some parts of the MUC5AC protein that are important for folding or disulfide bond formation have been removed in the recombinant protein. It is also possible, and even very likely, that the CHO K1 cells do not express all the required co-factors or chaperones for the folding of the MUC5AC N-terminal domains as discussed above.

CysD Domains

The CysD domain is a small domain consisting of only 110 amino acid residues. It contains 10 cysteines that are involved in intramolecular disulfide bonds. CysD domains are interspersed within the *O*-glycosylated mucin domains in the MUC2, MUC5AC, and MUC5B mucins. The function of the CysD domains is unknown. Their biochemical properties have made recombinant production difficult. However, a fusion protein between the second MUC2 CysD domain and the Fc section of a mouse immunoglobulin (Ig Fc) led to the successful production of the CysD domain in CHO K1 cells. After cleaving off the Ig Fc part using a built-in enterokinase cleavage site, the CysD domains were analyzed using both the gel electrophoresis under denaturing or native conditions and by analytical gel filtration. It was found that this CysD domain of MUC2 forms non-covalent dimers, probably by hydrophobic interactions. This suggests that the CysD domains may be involved in cross-linking of the mucus gel after secretion [24]. The different frequencies of, and thus the distances between, the CysD domains in MUC2, MUC5AC, and MUC5B would then give different densities of these crosslinks in the respective mucus gels. This may regulate the porosity and permeability of the mucus.

CysD domains from MUC5AC and MUC5B have previously been produced recombinantly in COS and CHO cells [25]. Using denaturing SDS gel electrophoresis, these CysD domains have been suggested to migrate as monomers under both reducing and non-reducing conditions, but some higher molecular mass aggregates were also observed. We expressed the MUC5AC CysD domains in CHO K1 cells as Ig Fc fusion proteins and removed the Ig Fc tail by enterokinase cleavage. When analyzed by gel filtration under native conditions, these CysD domains also formed pH-independent, non-covalent dimers just as the MUC2 CysD (unpublished data).

In COS cells, it was suggested that the MUC5AC and MUC5B CysD domains carry C-mannosylation at a WXXW motif, as the proteins were found to bind mannose-binding lectins [25]. The same proteins with point-mutations removing their motifs were retained in the ER, and it has been suggested that C-mannosylation was required for exit from the ER [25]. We could not find any C-mannosylation in the same motif in the recombinant MUC2 CysD that had been produced in CHO K1 cells when analyzed by mass spectrometry [24], even though C-mannose has been identified in other proteins from CHO K1 cells [26]. It seems therefore that MUC2 CysD is not C-mannosylated, at least not when produced as an Ig Fc fusion protein.

Conclusions

The gel-forming mucins are complex molecules with characteristic large heavily *O*-glycosylated mucin domains. Several cysteine-rich VWD domains flank the mucin domains and are involved in the formation of disulfidelinked oligomers. So far it has not been possible to produce recombinant full-length mucins. In contrast, both the *O*-glycosylated mucin domains and the different cysteine-rich VWD and CysD domains have been produced separately in different cell lines. Using these proteins, the knowledge of both the glycosylation and the oligomerization of mucins has been increased considerably. The

production of these recombinant proteins has, however, not always been easy, especially for those containing cysteine-rich VWD domains. This is mainly due to their large sizes and the many disulfide bonds, but also because it is not always clear where the domain boundaries are located. In some cases, as with the small CysD domains and also with the extracellular region of some transmembrane mucins, it has been helpful to make fusion proteins with an Ig Fc domain to facilitate the passage of the protein through the secretory pathway. In other cases, it has not been possible to produce a secreted protein despite many different approaches. Many different variants have often been investigated and many of these have led to proteins that never passed the ER exit control and were not secreted. Further investigations on the different domains of the gel-forming mucins is thus required, warranting more research into the area of complex, multimeric protein biosynthesis, folding, and assembly.

Acknowledgments

This work was supported by the Swedish Research Council (no. 7461, 21027, and 342-2004-4434), The Swedish Cancer Foundation, The Knut and Alice Wallenberg Foundation (KAW2007.0118), IngaBritt and Arne Lundberg Foundation, Sahlgren's University Hospital (LUA-ALF), Wilhelm and Martina Lundgren's Foundation, Torsten och Ragnar Söderbergs Stiftelser, The Sahlgrenska Academy and The Swedish Foundation for Strategic Research—The Mucus–Bacteria–Colitis Center (MBC) of the Innate Immunity Program.

References

1. Lang T, Hansson GC, Samuelsson T. Gel-forming mucins appeared early in metazoan evolution. *Proceedings of the National Academy of Sciences of the United States of America*. 2007; 104:16209–16214. [PubMed: 17911254]
2. Thornton DJ, Rousseau K, McGuckin MA. Structure and function of the polymeric mucins in airways mucus. *Annual Review of Physiology*. 2008; 70:459–486.
3. Johansson ME, Ambort D, Pelaseyed T, Schutte A, Gustafsson JK, et al. Composition and functional role of the mucus layers in the intestine. *Cellular and Molecular Life Sciences*. 2011; 68:3635–3641. [PubMed: 21947475]
4. Andersch-Björkman Y, Thomsson KA, Holmen Larsson JM, Ekerhovd E, Hansson GC. Large scale identification of proteins, mucins, and their O-glycosylation in the endocervical mucus during the menstrual cycle. *Molecular and Cellular Proteomics*. 2007; 6:708–716. [PubMed: 17220477]
5. Zhu L, Lee P, Yu D, Tao S, Chen Y. Cloning and characterization of human MUC19 gene. *American Journal of Respiratory Cell and Molecular Biology*. 2011; 45:348–358. [PubMed: 21075863]
6. Burchell JM, Mungul A, Taylor-Papadimitriou J. O-linked glycosylation in the mammary gland: changes that occur during malignancy. *Journal Mammary Gland Biol Neoplasia*. 2001; 6:355–364. [PubMed: 11547903]
7. Müller S, Hanisch FG. Recombinant MUC1 probe authentically reflects cell-specific O-glycosylation profiles of endogenous breast cancer mucin. High density and prevalent core 2-based glycosylation. *Journal of Biological Chemistry*. 2002; 277:26103–26112. [PubMed: 12000758]
8. Bäckström M, Link T, Olson FJ, Karlsson H, Graham R, et al. Recombinant MUC1 mucin with a breast cancer-like O-glycosylation produced in large amounts in Chinese-hamster ovary cells. *The Biochemical Journal*. 2003; 376:677–686. [PubMed: 12950230]
9. Premaratne P, Welen K, Damber JE, Hansson GC, Bäckström M. O-glycosylation of MUC1 mucin in prostate cancer and the effects of its expression on tumor growth in a prostate cancer xenograft model. *Tumour Biology*. 2011; 32:203–213. [PubMed: 20872286]
10. Sihlbom C, van Dijk-Härd I, Lidell ME, Noll T, Hansson GC, et al. Localization of O-glycans in MUC1 glycoproteins using electron-capture dissociation fragmentation mass spectrometry. *Glycobiology*. 2009; 19:375–381. [PubMed: 19095697]

11. Sewell R, Bäckström M, Dalziel M, Gschmeissner S, Karlsson H, et al. The ST6GalNAc-I sialyltransferase localizes throughout the Golgi and is responsible for the synthesis of the tumor-associated sialyl-Tn O-glycan in human breast cancer. *Journal of Biological Chemistry*. 2006; 281:3586–3594. [PubMed: 16319059]
12. Domino SE, Hurd EA, Thomsson KA, Karnak DM, Holmen Larsson JM, et al. Cervical mucins carry alpha(1,2)fucosylated glycans that partly protect from experimental vaginal candidiasis. *Glycoconjugate Journal*. 2009; 26:1125–1134. [PubMed: 19326211]
13. Park SW, Zhen G, Verhaeghe C, Nakagami Y, Nguyenvu LT, et al. The protein disulfide isomerase AGR2 is essential for production of intestinal mucus. *Proceedings of the National Academy of Sciences of the United States of America*. 2009; 106:6950–6955. [PubMed: 19359471]
14. Lidell ME, Johansson ME, Mörgelin M, Asker N, Gum JR Jr, et al. 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. *The Biochemical Journal*. 2003; 372:335–345. [PubMed: 12597771]
15. Lidell ME, Hansson GC. Cleavage in the GDPH sequence of the C-terminal cysteine-rich part of the human MUC5AC mucin. *The Biochemical Journal*. 2006; 399:121–129. [PubMed: 16787389]
16. Bell SL, Khatri IA, Xu G, Forstner JF. Evidence that a peptide corresponding to the rat Muc2 C-terminus undergoes disulphide-mediated dimerization. *European Journal of Biochemistry*. 1998; 253:123–131. [PubMed: 9578469]
17. Bell SL, Xu G, Forstner JF. Role of the cysteine-rich motif at the C-terminus of rat mucin protein Muc2 in dimer formation and secretion. *The Biochemical Journal*. 2001; 357:203–209. [PubMed: 11415450]
18. Perez-Vilar J, Hill RL. The carboxyl-terminal 90 residues of porcine submaxillary mucin are sufficient for forming disulfide-bonded dimers. *Journal of Biological Chemistry*. 1998; 273:6982–6988. [PubMed: 9507005]
19. Leir SH, Harris A. MUC6 mucin expression inhibits tumor cell invasion. *Experimental cell research*. 2011; 317:2408–2419. [PubMed: 21851820]
20. Lidell ME, Johansson ME, Hansson GC. An autocatalytic cleavage in the C terminus of the human MUC2 mucin occurs at the low pH of the late secretory pathway. *Journal of Biological Chemistry*. 2003; 278:13944–13951. [PubMed: 12582180]
21. Godt K, Johansson ME, Lidell ME, Mörgelin M, Karlsson H, et al. The N terminus of the MUC2 mucin forms trimers that are held together within a trypsin-resistant core fragment. *Journal of Biological Chemistry*. 2002; 277:47248–47256. [PubMed: 12374796]
22. Purvis AR, Gross J, Dang LT, Huang RH, Kapadia M, et al. Two Cys residues essential for von Willebrand factor multimer assembly in the Golgi. *Proceedings of the National Academy of Sciences of the United States of America*. 2007; 104:15647–15652. [PubMed: 17895385]
23. Perez-Vilar J, Eckhardt AE, DeLuca A, Hill RL. Porcine submaxillary mucin forms disulfide-linked multimers through its amino-terminal D-domains. *Journal of Biological Chemistry*. 1998; 273:14442–14449. [PubMed: 9603957]
24. Ambort D, van der Post S, Johansson ME, Mackenzie J, Thomsson E, et al. Function of the CysD domain of the gel-forming MUC2 mucin. *The Biochemical Journal*. 2011; 436:61–70. [PubMed: 21338337]
25. Perez-Vilar J, Randell SH, Boucher RC. C-Mannosylation of MUC5AC and MUC5B Cys subdomains. *Glycobiology*. 2004; 14:325–337. [PubMed: 14718370]
26. Krieg J, Glasner W, Vicentini A, Doucey MA, Löffler A, et al. C-Mannosylation of human RNase 2 is an intracellular process performed by a variety of cultured cells. *Journal of Biological Chemistry*. 1997; 272:26687–26692. [PubMed: 9334252]

Fig. 1.

A schematic representation of the human main gel-forming mucins. **a** Domain structure of MUC2, MUC5AC, MUC5B, and MUC6 with the number of amino acids in brackets. **b** The model for the MUC2 net-like polymerization that we have proposed based on our work with recombinant domains from MUC2, as described in this review. The C-termini form disulfide-linked dimeric complexes via their cysteine-knot (CK) domains and the N-termini form disulfide-linked trimeric complexes via their VWD3 domains. The CysD domains cross

link the mucin network by non-covalent dimerization. Green dashes represent the *O*-glycans in the mucin domains (Color figure online)

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Fig. 2.
Confocal images of CHO K1 clones expressing MG-MUC2C. GFP fluorescence in *green* and propidium iodide staining of nuclei in *red*. **a** A clone with lower expression levels which secreted the MG-MUC2C protein. **b** A clone with higher expression levels which showed the presence of large aggregates of non-correctly folded protein in the ER. This clone showed no secretion of MG-MUC2C (Color figure online)

Fig. 3. Western blot of CHO K1 transiently transfected with MUC2N-MG or MUC5ACN-MG. Proteins in the lysates and immunoprecipitated supernatants (with anti-myc) were separated in 4–12 % polyacrylamide gels, and the blot was stained with anti-GFP. MUC2N-MG can be detected in both lysate and supernatant, whereas MUC5ACN-MG can only be found in the lysate