Biosynthesis of a low-molecular-mass rat submandibular gland mucin glycoprotein in COS7 cells

Keith NEHRKE and Lawrence A. TABAK*
Departments of Dental Research and Biochemistry, School of Medicine and Dentistry, University of Rochester, 601 Elmwood Avenue, Box 611, Rochester, NY 14642, U.S.A.

INTRODUCTION

Mucin glycoproteins are a principal component of mucus, the slimy, visco-elastic coat that protects all mucosal surfaces of the body [1]. The unique physicochemical properties of mucins are attributed, in part, to the presence of multiple oligosaccharides that are attached via O-glycosidic linkage between GalNAc and serine and threonine residues of the apomucin. Interactions between clusters of O-linked GalNAc and the core protein stiffen the molecule, leading to a characteristic extended ‘rod-like’ mucin structure [2]. Since O-glycosylation proceeds in a stepwise manner, the addition of the initial GalNAc represents both the initial regulatory step in mucin glycosylation and a key structural determinant of mucin function.

Mucin suprastructure is stabilized by both non-covalent and covalent forces. The marked expansion of mucin shape leads to significant molecular ‘overlap’ and interweaving among carbohydrate side chains, and non-covalent interactions among these oligosaccharides plays a significant role in imparting a high viscosity to mucins in solution [3,4]. The suprastructure of higher-molecular-mass mucins is further stabilized through the formation of disulphide-bond-mediated oligomerization [5,6].

The biosynthesis of several high-molecular-mass mucins has been examined [7–10], and each of these studies underscores the complexity of the process. For example, in some but not all cases, N-linked glycosylation must precede appropriate formation of disulphide-bonded subunits [8]. In contrast, no studies have been reported that have examined the biosynthetic events of a low-molecular-mass mucin. Recently, the protein backbone of rat submandibular gland (RSMG) mucin has been cloned [11]. The RSMG apomucin is a low-molecular-mass species, consisting of only 322 amino acids. In common with other secreted mucins, the RSMG mucin is highly glycosylated, with greater than 60% carbohydrate content by mass. In addition to the multiple sites for the acquisition of O-glycans, one potential N-glycosylation site is present. Although there is no sequence similarity between the RSMG apomucin and human low-molecular-mass salivary mucin (MG-2) [12], architecturally these two mucins are quite similar. Each contains an N-terminal region that has relatively few potential O-glycosylation sites arranged in a non-repeating manner. Two non-conserved cysteine residues are found in this region. The central portion of both the RSMG apomucin and MG-2 is dominated by the presence of tandem repeats that are highly enriched in hydroxyamino acids and proline. Thus the majority of potential O-glycosylation sites are located in the tandem repeat region. The C-terminal region of the RSMG apomucin and MG-2 is serine- and threonine-rich, but is relatively low in proline content. No obvious repeat motif is observed in the C-terminus; rather, long stretches of serines and threonines are observed [13].

Unfortunately, there is no suitably differentiated cell line in which to study the biosynthesis of RSMG mucin. Therefore we have transiently expressed RSMG apomucin in a well-defined heterologous system (COS7 cell) as an initial step in defining the biosynthetic pathway of this class of mucin. We find that the recombinant mucin is heavily glycosylated and is secreted. Neither N-glycosylation nor disulphide-mediated cross-linking events are necessary for secretion to occur. Through the identification of partially glycosylated intermediates we have determined that the addition of GalNAc occurs sequentially rather than simultaneously, as has been described for other O-glycosylated proteins.

MATERIALS AND METHODS

Vectors, primers and reagents

To facilitate identification and purification of recombinantly expressed RSMG mucin, we engineered both a FLAG®-epitope tag and a heart muscle kinase site at the carboxy-terminus of the apomucin by reverse transcriptase PCR amplification of the apomucin coding region from RSMG cDNA using the primers apo1, 5'-CCG CTG CAG ATG AAA AGG GAA ACT TTC ATC TTG GGC-3', and apo3, 5'-AAC ACC TAG TAG GAG CAT ATC TTG GGC-3'.

Abbreviations used: conA, concanavalin A; RSMG, rat submandibular gland; DMEM, Dulbecco’s modified Eagle’s medium.
* To whom correspondence should be addressed.
CAT ACC AAA GTA CAT CGT AAG CCA CTG AAG-3' [14]. The resulting PCR product was cut with PsiI and AarII and cloned into the complementary sites in pKN4 [15], which is based on the expression vector pcDL-SRz296 [16] (DNAX Research Institute of Molecular and Cellular Biology Inc.), to create pKN20, the apomucin expression vector.

Cell culture
All cell-culture reagents were from Life Technologies. COS7 cells were maintained in Dulbecco’s modified Eagle’s medium (DMEM) supplemented with 10% (v/v) fetal-bovine serum and passed twice a week. Cells at approx. 70% confluency were trypsinized with trypsin-EDTA (Promega) and resuspended in DMEM containing 10% fetal-bovine serum and 1% penicillin-streptomycin. Cells were counted with a hemocytometer.

Immunoprecipitation, glycosidase digestion and analysis of glycans
Nuclei were removed from the lysed cells by centrifugation at 16000 g for 15 min and the lysates were immediately immunoprecipitated with anti-FLAG*-agarose (IBI). Following an overnight incubation at 4 °C, the immunoprecipitates were washed twice with lysis buffer and the bound material was eluted from the agarose beads by boiling for 15 min in 100 mM Tris/HCl (pH 7.6)/50 mM NaCl/0.2% SDS/2% (v/v) 2-mercaptoethanol. The resulting eluant was diluted 10-fold into PBS/1% Triton X-100 containing 0.5 mg/ml BSA, and then either reimmunoprecipitated overnight with anti-FLAG*-agarose or precipitated with the lectin VVA (Vicia villosa) conjugated to agarose beads (E-Y Laboratories). The bound material was resuspended in 2× Tricine gel loading buffer, heated to 65 °C for 30 min and run on an 8.25% (v/v)/4% (w/v) Tricine gel as described by Schägger and von Jagow [17].

Alternatively, following the first immunoprecipitation, the antibody-agarose-bound mucin was digested in a volume of 25 μl with α-N-acetylgalactosaminidase (Oxford Glycosystems), according to the recommendations of the manufacturer, before release from the antibody agarose beads by the addition of 25 μl of 2× elution buffer. The digested material was then either reimmunoprecipitated with anti-FLAG*-agarose or lectin-precipitated as described above.

For analysis of the secreted mucin, COS7 cells were either metabolically labelled with Trans-label Express (EY Laboratories), as described for intracellular mucin biogenesis [18], or due to the microheterogeneity of the sample, proteolysis of the sample was ruled out by glycosidase digestion of this material suggested that it is heavily O-glycosylated. Mild-acid hydrolysis or sialidase treatment resulted in a decrease in mobility during SDS/PAGE, presumably due to the removal of charged sialic acid (Figure 1, compare lanes 2 and 4 with lane 1). This suggests that the terminal sialic acid moieties contribute little to the aberrant mobility of the mature mucin compared with the predicted mobility of the protein core. Subsequent treatment of sialidase- or acid-hydrolysed material with O-glycanase (Figure 1, lanes 3 and 6) or with a combination of β-galactosidase and α-N-acetylgalactosaminidase (Figure 1, lane 7) resulted in an increase in the mobility of the mucin. The glycosidase-digested material (lanes 3, 6 and 7, Figure 1) migrated as a diffuse ‘smear’, which may have been due to incomplete enzymic removal of the carbohydrate side chains and/or due to the microheterogeneity of the sample; proteolysis of the sample was ruled out by identical digestions with a control protein containing no oligosaccharide (results not shown). In addition, substituents on a sub-fraction of the sialic acid residues may render them resistant precipitation using agarose-bound lectins concanavalin A (conA), PSA (Pisum sativum), and PHA-E (Phaseolus vulgaris) (EY Laboratories), as described for intracellular mucin biosynthetic intermediates above.

RESULTS AND DISCUSSION

Expression and characterization of recombinant RSMG mucin in COS7 cells
To facilitate the analysis of RSMG apomucin expression in COS7 cells, the recombinant protein was recovered by immunoprecipitation and radiolabelled with 32P using heart muscle kinase. The predominant radiolabelled immunoprecipitated species displayed an apparent molecular mass of over 100 kDa on denaturing Tricine/SDS/PAGE (Figure 1, lane 1). As expected [18], although there are two potential sites for disulphide bond formation, the RSMG mucin, unlike higher-molecular-mass mucins, did not multimerize and migrated as a monomer under non-reducing conditions (Figure 2). Glycosidase digestion of this material suggested that it is heavily O-glycosylated. Mild-acid hydrolysis or sialidase treatment resulted in a decrease in mobility during SDS/PAGE, presumably due to the removal of charged sialic acid (Figure 1, compare lanes 2 and 4 with lane 1). This suggests that the terminal sialic acid moieties contribute little to the aberrant mobility of the mature mucin compared with the predicted mobility of the protein core. Subsequent treatment of sialidase- or acid-hydrolysed material with O-glycanase (Figure 1, lanes 3 and 6) or with a combination of β-galactosidase and α-N-acetylgalactosaminidase (Figure 1, lane 7) resulted in an increase in the mobility of the mucin. The glycosidase-digested material (lanes 3, 6 and 7, Figure 1) migrated as a diffuse ‘smear’, which may have been due to incomplete enzymic removal of the carbohydrate side chains and/or due to the microheterogeneity of the sample; proteolysis of the sample was ruled out by identical digestions with a control protein containing no oligosaccharide (results not shown). In addition, substituents on a sub-fraction of the sialic acid residues may render them resistant precipitation using agarose-bound lectins concanavalin A (conA), PSA (Pisum sativum), and PHA-E (Phaseolus vulgaris) (EY Laboratories), as described for intracellular mucin biosynthetic intermediates above.
Biosynthesis of a low-molecular-mass salivary mucin

Figure 2  Tricine gel analysis of labelled mucin under reducing (lane 1) or non-reducing (lane 2) conditions

Dithiothreitol was omitted from the heart muscle kinase reconstitution buffer used to label the immunoprecipitated mucin, and 2-mercaptoethanol was omitted from the Tricine gel loading dye mix of the sample in lane 2. Molecular-mass markers are given in kDa.

Figure 3  Lectin precipitations of RSMG low-molecular-mass mucin produced in COS7 cells

Mucin was immunoprecipitated with anti-FLAG<sup>4</sup>-antibody and then radiolabelled with <sup>32</sup>P using heart muscle kinase. The mucin was then eluted from the antibody and either re-immunoprecipitated (lane 1), or precipitated with concanavalin A (conA) agarose (lanes 2 and 3), PSA agarose (lanes 4 and 5) or PHA-E agarose (lanes 6 and 7). The saccharides α-methylmannoside (lanes 3 and 5) or galactose (lane 7) were used at 100 mM to inhibit lectin binding and demonstrate specificity. All precipitated materials were then analysed by Tricine/SDS/PAGE. Molecular-mass markers are given in kDa.

Figure 4  Pulse–chase maturation of RSMG low-molecular-mass mucin in normal and in tunicamycin-treated COS7 cells

Mucin in cell medium (a) and lysates (b), from untreated COS7 cells or cells treated before and during the pulse–chase with tunicamycin, was immunoprecipitated using anti-FLAG<sup>4</sup>-antibody and then separated by Tricine/SDS/PAGE. Molecular-mass markers are given in kDa.

Time course of RSMG mucin glycosylation and secretion

Glycosylation and secretion of the RSMG mucin was examined using a pulse–chase protocol. Cell lysate derived from transfected (with apomucin reporter) metabolically labelled COS7 cells harvested before the chase period contained a predominant anti-FLAG<sup>4</sup>-reactive species of approx. 63 kDa (Figure 4b, lane 1). Over the course of several hours, the 63 kDa species disappeared and a new species of greater than 100 kDa appeared (Figure 4b, lanes 2–6). At approx. 80 min the protein began to appear in the cell culture media (Figure 4, lane 2) and by 320 min release of the mature mucin was nearly complete (Figure 4a, lane 4). At intermediate times a ‘smear’ of material that could represent partially glycosylated mucin was detected (Figure 4b, lanes 2–5).

It should be noted that we observed an approx. 5-fold less recovery of the fully glycosylated mucin compared with the starting unglycosylated apomucin. One possible reason for this is that only a portion (approx. 20%) of the apomucin is processed to yield secreted mucin. Alternatively, glycosylation may lead to a partial sequestration of the antibody recognition site, thereby diminishing the detectability of the FLAG<sup>4</sup>-epitope tagged mucin.
Figure 5  Lectin precipitation of RSMG low-molecular-mass mucin from COS7 cells treated with benzyl-N-acetyl-α-galactosaminide, an inhibitor of O-glycan chain extension

Metabolically labelled mucin was immunopurified with anti-FLAG™-agarose from the culture media of untreated (lanes 1 and 3) or benzyl-N-acetyl-α-galactosaminide-treated COS7 cells (lanes 2 and 4). Following desorption, the immunopure materials were either re-precipitated with anti-FLAG™-agarose (lanes 1 and 2) or precipitated with VVA, a lectin that recognizes unextended GalNAc in α-linkage to Ser/Thr in the protein core. All precipitated materials were then analysed by Tricine/SDS/PAGE. Molecular-mass markers are given in kDa.

Role of N-linked glycosylation in RSMG mucin processing and secretion

To determine the extent to which N-glycosylation influences the addition of O-glycans, oligosaccharide processing, or secretion, we treated the COS7 cells with tunicamycin, an inhibitor of N-glycosylation, before and during the pulse–chase period. The greater mobility of the pre-chase material from tunicamycin-treated cells suggested that N-glycosylation occurs within the 20 min labelling period before the chase (Figure 4b, compare lanes 1 and 7). However, RSMG mucin processing appeared normal in the absence of N-glycosylation. The rate of sugar addition to the core protein (Figure 4b, lanes 7–12) and the appearance of the mature form of the protein in the media (Figure 4a, lanes 6–10) were essentially unchanged in the presence of tunicamycin.

A mucin that contains underextended oligosaccharides is secreted normally

Benzyl-N-acetyl-α-galactosaminide acts as a competitive inhibitor of UDP Gal:GalNAc-β1,3-galactosyltransferase and can prevent the extension of the initial sugar residue added during O-glycosylation [19–21]. Treatment of apomucin-transfected COS7 cells with aryl-N-acetyl-α-galactosaminide resulted in the production of a mucin that migrated with decreased mobility (Figure 5, lanes 2 and 4), resembling mucin from which the sialic acid had been removed (Figure 1, lane 4). To ensure the absence of carbohydrate chains that had been extended beyond the initial GalNAc residue, immunopurified recombinant RSMG mucin from untreated COS7 cells and from cells treated with aryl-N-acetyl-α-galactosaminide were precipitated a second time with the lectin VVA, which recognizes predominantly unsubstituted GalNAc α-linked to threonine or serine. The mucin purified from aryl-N-acetyl-α-galactosaminide-treated cells was recognized by the lectin VVA, whereas mucin from untreated cells was not (Figure 5, lanes 3 and 4). Recent evidence indicates that aryl-N-acetyl-α-galactosaminides may not fully inhibit carbohydrate extension under certain circumstances. Our data would fit a model where either extension or sialic acid addition was inhibited, given that a fraction of the carbohydrate chains were terminated at the initial GalNAc residue.

Although it has not been shown whether aryl-N-acetyl-α-galactosaminide-induced intermediate products are transported normally or broken down during secretion, we show here, using pulse-chase analysis, that the aryl-N-acetyl-α-galactosaminide-induced RSMG mucin is processed and secreted in the same manner in which mucin from untreated COS7 cells was. Thus mucin appears in the media by 80 min, continues to be produced for at least 5 h, and displays a reduced mobility upon maturation, consistent with a lack of terminal sialic acid (Figure 6). These data suggest that the observed decrease in mobility during RSMG mucin biosynthesis is due to the addition of the core GalNAc moiety and that extension (or sialation) of this core is not required for normal transport and secretion.

Addition of core GalNAc does not occur simultaneously

During pulse-chase analysis of RSMG mucin in untreated COS7 cells, we noted the appearance of intermediate-sized species that migrated between the unmodified and fully glycosylated material...
GalNAc, as was observed during previous glycosidase digestions (Figures 1 and 2). This fact that VVA also recognizes a small fraction of the more mature, glycosidase-resistant mucin implies that some portion of the carbohydrate side chains on these molecules have not yet acquired substituents that would mask their affinity for VVA.

These results provide evidence for O-glycosylation intermediates that contribute to mucin microheterogeneity. In contrast with other systems, where the addition of the initial O-linked GalNAc has been reported to be simultaneous [23], we find evidence of initiation that is spaced in time. While this raises the interesting possibility that changes in transport time through the Golgi may alter the patterns of glycosylation, it remains to be determined if addition is also spaced in location and occurs in a hierarchical or random manner. The recent demonstration of multiple UDP-N-acetylgalactosamine:polypeptide galactosaminyltransferases [24,25] would suggest that the former is a possibility; i.e. each isoform acts on specific hydroxyamino acids of the protein backbone.

**Prospects for recombinant mucins as therapeutics**

Mucins produced by salivary glands play a significant role in protecting both the hard and soft tissues of the mouth [13]. The sensation of oral dryness (xerostomia) is a prevalent problem, particularly among the elderly, with estimates ranging from 12 to 29% in different populations [26,27]. Individuals who experience a complete loss of salivary function are at great risk of developing xerostomia and other pathologies [26,28] due, in part, to the loss of the protective qualities provided by the mucins [29,30]. Native animal gastro-intestinal mucins have been used in various formulations of artificial saliva in an attempt to deliver adequate levels of mucins to compromised individuals with varied success [29,30].

While a number of high-molecular-mass apomucins have been cloned, the expression of such recombinant glycoproteins would be expected to be complicated by the diverse range of co- and post-translational processing steps that must occur. In contrast, our studies with the low-molecular-mass RSMG mucin suggest that its biosynthetic pathway is relatively simple and would thus represent a reasonable candidate for expressing in a heterologous system. The cost inherent in the use of mammalian cell host systems makes it unlikely that COS7 cells could be used for such a purpose. However, the successful overexpression of the UDP-N-acetylgalactosamine:polypeptide galactosaminyltransferase in SF9 cells [31] suggests that such cells could be re-engineered to produce the desired recombinant glycoprotein at lower cost.

We thank Dr. Fred Hagen and Dr. Kelly Ten Hagen for reviewing the manuscript and helpful comments, and Patricia Noonan for her help in preparing the manuscript. Supported in part by National Institute of Health Grant DE-08108 (to L.A.T.). K.N. was supported by training grant T32 DE-07202.

**REFERENCES**

Received 17 May 1996/26 November 1996; accepted 12 December 1996