Muc5b and Muc5ac are the major oligomeric mucins in equine airway mucus

Karine Rousseau,1 Sara Kirkham,1 Shaun McKane,2 Richard Newton,3 Peter Clegg,1,2 and David J. Thornton1

1Wellcome Trust Centre for Cell-Matrix Research, Faculty of Life Sciences, University of Manchester, Manchester, 2University of Liverpool, Faculty of Veterinary Science, Leahurst, Neston, South Wirral, and 3Animal Health Trust, Lanwades Park, Kentford, Newmarket, United Kingdom

Submitted 10 November 2006; accepted in final form 7 February 2007

Rousseau K, Kirkham S, McKane S, Newton R, Clegg P, Thornton DJ. Muc5b and Muc5ac are the major oligomeric mucins in equine airway mucus. Am J Physiol Lung Cell Mol Physiol 292: L1396–L1404, 2007. First published February 9, 2007; doi:10.1152/ajplung.00444.2006.—Horses frequently suffer from respiratory diseases, which, irrespective of etiology, are often associated with airway mucus accumulation. Studies on human airways have shown that the key structural components of the mucus layer are oligomeric mucins, which can undergo changes of expression and properties in disease. However, there is little information on these gel-forming glycoproteins in horse airways mucus. Therefore, the aims of this study were to isolate equine airway oligomeric mucins, characterize their macromolecular properties, and identify their gene products. To this end, pooled tracheal washes, collected from healthy horses and horses suffering from respiratory diseases, were solubilized with 6 M guanidinium chloride (GdmCl). The oligomeric mucins were purified by density gradient centrifugation followed by size exclusion chromatography. Biochemical and biochemical analyses showed the mucins were stiffened random coils in solution that were polydisperse in size ($M_r = 0.5–10 \mu m$) and size (2–40 MDa). Agarose gel electrophoresis showed that the pooled mucus sample contained at least two populations of oligomeric mucins. Electrospray ionization tandem mass spectrometry of tryptic digests of the unfractionated mucin preparation showed that the oligomeric mucins Muc5b and Muc5ac were present. In summary, we have shown that equine airway mucus is a mixture of Muc5b and Muc5ac mucins that have a similar macromolecular organization to their human counterparts. This study will form the basis for future studies to analyze the contribution of these two mucins to equine airways pathology associated with mucus accumulation.

horses

There are a wide variety of causes of equine respiratory disease, with both infections with viruses and bacteria [inflammatory airway disease (IAD)] as well as allergic airway diseases [recurrent airway obstruction (RAO)] being common (3, 5, 32, 52). Despite differing etiologies, all forms of equine respiratory disease are associated with increased volumes of respiratory secretions and a change in the physical and chemical nature of the mucus that frequently accumulates in the trachea of affected horses (14–16). Whereas the mucus gel in the respiratory tract is vital for the protection and normal function of the airways, in many circumstances, the overproduction and change in composition is a significant pathologic feature of these conditions (14–16). In racehorses, respiratory diseases are the second most common reason why horses fail to train, and accumulation of tracheal mucus has been identified as a risk factor for poor racing performance in such horses (24).

The properties of mucus gels are dependent on a number of factors, but by far the most important are a family of large oligomeric, secreted glycoproteins called mucins. Studies of human airway mucus have shown that it is comprised of a mixture of oligomeric mucins (80% carbohydrate by weight) that are polydisperse in mass (2–40 MDa) and size (0.5–10 μm in length). Component mucin monomers (2–3 MDa) are assembled linearly and held together by disulfide bonds, and they share a generic organization (6, 34, 40, 42, 43, 49). The polypeptides have cysteine-rich motifs in their COOH and NH2 termini that are necessary for oligomerization, a central domain containing repeated serine/threonine/proline-rich regions (STP domains) to which O-glycans are attached, and interspersed among the STP domains are repeated cysteine-rich regions (cys domains) of unknown function (6, 13, 17, 34, 40, 42, 43, 49). The STP domains show little sequence homology between mucins; in contrast, the NH2- and COOH-terminal domains and the cys domains show considerable homologies (4, 12, 13, 17, 19, 50).

From studies of human airway secretions, it is clear that MUC5AC and MUC5B, and to a much lesser extent, MUC2, are key determinants of mucus gel properties (25). These glycoproteins can be changed in amount, glycoform, and size in respiratory disease, and these alterations can have dramatic clinical consequences (30, 38, 39). However, at present, little is known about the type, structure, and biochemistry of these macromolecules in equine airways mucus. A recent study (21) has shown that Muc5ac, but not Muc2, is expressed in horse airway, but to date there are no reports on the expression of equine Muc5b. In allergic airway disease in the horse (RAO), the altered physical properties of the respiratory mucus (20) have been attributed to both changes in the quality and quantity of its constituent mucins (27), suggesting that mucins are also likely to play a role in other equine respiratory disease such as IAD. Here, we have purified, characterized, and identified the oligomeric mucins from equine airways mucus. We report that Muc5b and Muc5ac are the major oligomeric mucins and that these glycoproteins have a similar macromolecular organization to their counterparts found in human airways mucus.

**EXPERIMENTAL PROCEDURES**

**Mucin purification.** Endoscopic tracheal washes were obtained as part of routine diagnostic procedures in horses suffering from respiratory diseases with tracheal mucus accumulation (including diagnosed cases of IAD or RAO). Similar tracheal washes were also...
obtained from normal horses that had no history or clinical evidence of respiratory disease immediately following euthanasia. The protocol for mucus and tissue sample collection was approved by internal independent review within the Faculty of Veterinary Sciences, The University of Liverpool. All washes were diluted with 3 volumes of 8 M GdmCl, and all the normal (2 horses) and pathological (4 horses) tracheal washes were pooled for analysis of their mucin content. After addition of 6 M GdmCl to bring the sample to 10 × the original volume, the mucins were extracted with gentle stirring at 4°C. Solubilized mucins were purified following a previously published method (47). Briefly, the extract was centrifuged in a cesium chloride (CsCl)/4 M GdmCl density gradient at a starting density of 1.4 g/ml in a Beckman Ti 45 rotor at 40,000 rpm for 65 h at 15°C. After centrifugation, the tubes were emptied from the top. Mucin-containing fractions were pooled, dialyzed against 4 M GdmCl, and further purified by gel filtration chromatography on a Sepharose CL-2B column (1.6 × 32 cm) eluted with 4 M GdmCl at a flow rate of 12 ml/h. Mucin-containing fractions were pooled and stored at 4°C. Preparation of reduced and carboxymethylated mucins and high M₅ mucin glycopeptides. Purified mucins were dialyzed into 6 M urea or 6 M GdmCl (both containing 0.1 M Tris, 5 mM EDTA, pH 8) and reduced by treatment with 10 mM dithiothreitol (DTT) for 1 h at 37°C. Iodoacetamide was added (25 mM final concentration), and the reduced mucins were carboxymethylated for 30 min in the dark at room temperature. High M₅ mucin glycopeptides were prepared by digestion of reduced mucins with trypsin (1 μg) overnight at 37°C in 0.1 M ammonium hydrogen carbonate, pH 8.0.

Rate-zonal centrifugation. Mucins and reduced and carboxymethylated mucins were layered onto 6–8 M GdmCl gradients and centrifuged at 40,000 rpm in a SW 40 Ti Beckman rotor for 2.5 h at 15°C as described previously (37).

Agarose gel electrophoresis. Reduced mucins were electrophoresed in 0.7% agarose gels at 30 V for 15 h in 40 mM Tris-acetate, 1 mM EDTA, pH 8.0, containing 0.1% (wt/vol) SDS. After electrophoresis, the proteins were vacuum-blotted onto nitrocellulose membrane for 2 h at 40 mbar in 0.6 M sodium chloride, 60 mM sodium citrate, as previously described (45).

Multilayer light scattering. The molecular size distributions of intact, reduced, and carboxymethylated mucins and high M₅ mucin glycopeptides were determined by multilayer light scattering (MALLS) essentially as described previously (35). In brief, samples were chromatographed on a TSK 6000 column (0.78 × 30 cm) eluted at a flow rate of 0.83 ml/min with 0.2 M NaCl/1 mM EDTA. The column effluent was monitored with an inline DAWN EOS laser photometer and an Optilab rEX refractometer (Wyatt Technology, Haverhill, Suffolk, United Kingdom). Light-scattering data were collected at 18 angles between 10 and 151°, and the data were analyzed according to Zimm (53).

Tryptic digestion and mass spectrometry. Reduced and carboxymethylated mucins in 2 M urea/100 mM ammonium bicarbonate, pH 8.0, were digested overnight at 37°C with 1 μg of trypsin. Tryptic peptides were separated from high M₅ mucin glycopeptides by gel chromatography on a Sephacryl S-100 column (1 × 30 cm) eluted with 100 mM ammonium hydrogen carbonate, pH 8.0, and then lyophilized. Tryptic peptides were solubilized in 0.1% formic acid, separated by reverse-phase chromatography, and analyzed inline by positive ion electrospray ionization mass spectrometry/mass spectrometry (ESI-MS/MS) using a Quadrupole-Time of Flight (Q-ToF) Micromass mass spectrometer (Waters, Manchester, United Kingdom). Samples were introduced via a capillary liquid chromatography system fitted with a stream-select module (Waters). Aliquots (10–50 μl) of the samples were separated on a PepMap column (0.075 × 150 mm) with a gradient of 5% acetonitrile in 0.1% formic acid to 25% acetonitrile in 0.1% formic acid over 30 min at a flow rate of 200 nl/min. Data acquired were analyzed using SWISS-PROT and TrEMBL databases using the ProteinLynx global server 1.1 software (Waters). Parameters were set to 100-nDa peptide mass tolerance, methionine oxidation and carboxamidomethyl cysteine modification. The MS/MS spectrum of each peptide matched was examined individually with the acceptance criteria being that the parent and fragment ion masses were within the calibrated tolerance limits and that the spectrum contained a series of at least four consecutive y ions. A custom database (see below) containing putative equine mucin sequences was also used to analyze peptide fragmentation data.

Generation of equine Muc gene genomic sequences. Unprocessed data generated by the Horse Genome Sequencing Project is available through the National Center for Biotechnology Information (NCBI) trace archive website (http://www.ncbi.nlm.nih.gov/Traces/traces.cgi). The database consists of clones of ~1,000 bp. Human MUC5AC, MUC5B, and MUC2 mRNA sequences encoding the NH₂-terminal regions of the three mucins were used to perform a MegaBLAST search of the database. Clones were recovered and assembled into contigs using the program CAP3 (26). The ends of the contigs were then used to repeat the database search to extend the contig sequence. This BLAST and assembly process was repeated several times until the majority of the genomic sequence encoding the putative NH₂ terminus of Muc5ac and Muc5b was obtained, and a smaller fragment of the NH₂ terminus of Muc2 was generated. The approach was repeated using the mRNA sequences encoding the cys domains of the human MUC5AC, MUC5B, and MUC2 genes. The intron/exon boundaries were predicted by comparison of the genomic sequences of the human and mouse mucins, since it is known that these are conserved between mucins as well as between species (12, 18, 19). The predicted exonic sequence was translated, and the resultant amino acid sequences were used to prepare a database of putative equine Muc5ac, Muc5b, and Muc2 mucin sequences.

cDNA preparation, amplification, and sequencing. Equine tissue samples were obtained from trachea, stomach, and small intestine from a deceased foal at the Leahurst Large Animal Hospital (University of Liverpool) and collected in RNA later (Ambion, Huntington, United Kingdom). The mRNA was extracted using the TRizol reagent and chloroform method (9) and then reverse-transcribed into cDNA using random primers. Oligonucleotide primers (see below) were used to amplify cDNA from equine tracheal mRNA using Abgene buffer IV [750 mM Tris–HCl (pH 8.8 at 25°C), 200 mM (NH₄)₂SO₄, 0.1% (vol/vol) Tween 20, 15 mM MgCl₂] with 0.5 μM each oligonucleotide (MWG, http://www.mwg-biotech.com/html/all/index.php), 0.2 mM each dNTP (Abgene, Epsom, United Kingdom), with 0.125 units of Taq polymerase (Abgene) per 15 μl of reaction. The amplification program consisted of 3-min denaturation at 96°C followed by 32 cycles of 30-s denaturation at 96°C, 30 s of annealing at the appropriate temperature, and 30 s of elongation at 72°C; the reactions were then completed with a 5-min elongation at 70°C. The products were visualized under a UV transilluminator after electrophoresis on 2% (wt/vol) agarose gels containing ethidium bromide. PCR products were extracted from the gel and purified using the GeneClean PCR purification kit (Qiogene, Stretton, United Kingdom) and sequenced using the BigDye Terminator v3.1 chemistry [2.5% (vol/vol) DMSO was also added to the sequencing reaction]. Three pairs of primers were designed based on the predicted expression in different tissues. Fragment 1 (FR1) was amplified with 5′ primer CTGCTCCAAACCTTGAC-GCTAC and 3′ primer CAGCTTGTGGTGAAAGGAGG for Muc5b; FR2 using 5′ primer TCAAATCTGGTACCGGATG and 3′ primer AGTTGTTGACGTGTTGGGATA for Muc5ac, and FR3 using 5′ primer AGCGGAGCCTCAGGCTGTC for Muc5b and 3′ primer CGTGCGTCGGTCACCCCTCA for Muc2.

Analytical methods. Molecules transferred to nitrocellulose membranes by slot blotting or Western blotting were stained for total carbohydrate with periodic acid-Schiff (PAS) reagent (44) and protein with the Coomassie blue reagent. Immunodetection was performed using a polyclonal antiserum that recognizes reduced and carboxymethylated mucin subunits (34). Immuno- and Western blots were visualized using horseradish peroxidase-labeled secondary anti-
bodies in conjunction with an ECL Western detection kit. Band intensities were measured using a Bio-Rad model GS-800-calibrated densitometer.

RESULTS

Purification of oligomeric mucins from equine airways mucus. Pooled tracheal washes from healthy horses, and horses suffering from respiratory disease, were solubilized in 6 M guanidinium chloride (GdmCl) and subjected to CsCl density gradient centrifugation (Fig. 1A). This resulted in two carbohydrate-rich peaks (monitored by PAS staining) in the high-density fractions (1.33–1.56 g/ml density) that were separated from the majority of the proteins (monitored by A280 measurements), which were concentrated in the lower density fractions (1.24–1.32 g/ml density). Agarose gel electrophoresis of fractions from across the density distribution showed bands in fractions 10–20 that were reactive with an antiserum that recognizes human reduced mucin monomers (Fig. 1A, inset); these were the same fractions that contained the carbohydrate-rich material. The agarose gel analysis also showed evidence of different populations of molecules with different electrophoretic mobilities. Taken together, these data suggested that at least two populations of oligomeric mucins were present in the high-density fractions (Fig. 1A). Therefore, these fractions were pooled and dialyzed against 4 M GdmCl, and the mucins were further purified by gel filtration chromatography on a Sepharose CL-2B column (Fig. 1B). The high-molecular weight mucins (PAS staining material in fractions 13–25) were separated from lower M_r glycoproteins (PAS staining material in fractions 26–40) and the remaining smaller proteins (Coomassie blue staining material in fractions 30–52). Mucin-containing fractions (fractions 13–25) were pooled (Fig. 1B) and used for further characterization.

Mucin characterization. The molecular size distribution of the mucins was monitored by rate-zonal centrifugation on 6–8 M GdmCl gradients (Fig. 2). The intact molecules exhibited a broad range of sedimentation rates consistent with a polydisperse population of high-molecular weight mucins (Fig. 2). However, treatment with the reducing agent DTT caused a marked decrease in sedimentation rate and polydispersity (Fig. 2) and suggested the mucins were oligomeric assemblies stabilized by disulfide bonds. To gain more insight into the size and molecular architecture of the mucins, MALLS was performed on the intact mucins and the major fragments produced after reduction (reduced monomers) and reduction followed by trypsin digestion (high M_r glycopeptides). Each of the three samples was subjected to gel filtration chromatography on a TSK 6000 column (Fig. 3), and the absolute size (radius of gyration, R_g) and M_r distributions were determined across the chromatograms (Table 1). The data showed that the mucin oligomers...
MUCINS FROM EQUINE AIRWAYS MUCUS

were polydisperse in both size and $M_r$, and, as expected from the rate-zonal centrifugation analysis, reduction of disulfide bonds caused a decrease in $M_r$ size, and polydispersity of the mucins (Table 1). Trypsin digestion of the reduced mucins caused a further reduction in $M_r$ and size of the majority of the preparation (Table 1). These fragmentation patterns are consistent with the behavior of oligomeric mucins isolated from human airways mucus (38–40, 48, 49). It should be noted that a minor component was also produced after trypsin digestion of the reduced mucins with average $M_r$ of 6.2 MDa and $R_G$ of 220 nm. This might represent undigested oligomeric mucins or another trypsin-resistant glycoprotein present in the preparation.

The relationship between molecular shape ($R_G$) and $M_r$ of molecules in solution can be expressed by the value of the Mark-Houwink parameter ($\alpha$), which can be obtained from the equation $R_G = A M_r^\alpha$ (A is a constant related to the mass per unit length of the molecule). A plot of log $R_G$ against log $M_r$ across the chromatographic distribution of the intact oligomers is shown in Fig. 4. The slope of the line yielded a value of $\alpha = 0.64$, which suggests that the mucins have a stiffened random coil conformation in solution.

Table 1. Molecular size data for intact mucins and mucin fragments

<table>
<thead>
<tr>
<th></th>
<th>Range of $M_r$ (MDa)</th>
<th>Average $M_r$ (MDa)</th>
<th>Range of $R_G$ (nm)</th>
<th>Average $R_G$ (nm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intact mucins</td>
<td>6–20</td>
<td>14</td>
<td>140–270</td>
<td>225</td>
</tr>
<tr>
<td>Reduced mucin subunits</td>
<td>4.5–12.5</td>
<td>7.5</td>
<td>98–180</td>
<td>130</td>
</tr>
<tr>
<td>Mucin glycopeptides</td>
<td>2–3.5</td>
<td>2.6</td>
<td>62–77</td>
<td>65</td>
</tr>
</tbody>
</table>

$M_r$ and radius of gyration ($R_G$) were determined from the multilangle laser light scattering (MALLS) data measured across the chromatographic distributions of each component (see Fig. 3). The data used in the analyses are for the intact mucins eluting between 5.5–7.5 ml, the reduced mucins eluting between 6.0–7.0 ml, and the glycopeptides eluting between 6.8–7.5 ml.

Fig. 3. Determination of molecular size distributions of equine respiratory mucins before and after fragmentation. The intact mucins (solid line), the reduced and carboxymethylated mucins (dotted line), and the reduced, carboxymethylated, and trypsin-treated mucins (dashed line) were chromatographed on a TSK 6000 column, and the effluent was monitored for light scattering and refractive index (see EXPERIMENTAL PROCEDURES). The calculated values for $M_r$ and radius of gyration ($R_G$) across the distributions and the average values for $M_r$ and $R_G$ are presented in Table 1.

Fig. 4. Relationship between the size and shape of the oligomeric mucins. The unreduced mucins were chromatographed on TSK 6000, and the column eluent was monitored for light scattering and refractive index as described in EXPERIMENTAL PROCEDURES. The data used in this analysis are for the molecules eluting between 5.5 and 7.5 ml (see Fig. 3). The slope ($\alpha$) of the log-log plot of the dependence of $M_r$ and $R_G$ yields a value of 0.64.

**Mucin identification.** The mucin preparation was subjected to reduction and carboxymethylation followed by digestion with trypsin to generate peptides for subsequent analysis by ESI-MS/MS. The peptide fragmentation data produced were searched against the SWISS-PROT and TrEMBL databases to identify matches with known mucin sequences. No equine mucin sequences were identified that matched the peptide fragmentation data. However, four peptide matches with sequences from human MUC5B (Fig. 5A), one peptide from mouse Muc5b (Fig. 5B), two peptide matches from mouse Muc5ac (Fig. 5C), and one peptide that matched a sequence present in both mouse Muc5ac and Muc5b were identified (Fig. 5D). No perfect matches were found to human MUC2 or mouse Muc2. These data suggest that homologs of the two major human oligomeric mucins, MUC5AC and MUC5B, are present in equine airways mucus.

**Predicted sequences of equine Muc5ac, Muc5b, and Muc2.** The lack of peptide matches to equine mucins is most likely because of the scarcity of equine mucin sequences in the SWISS-PROT and TrEMBL databases. Therefore, to gain a more comprehensive analysis of the peptide fragmentation data, we assembled sequences corresponding to these three genes from unprocessed clones from the Horse Genome Sequencing Project (http://www.ncbi.nlm.nih.gov/Traces/tracerep.cgi). To achieve this, we performed a BLAST search of the database using mRNA sequences encoding the NH2-terminal region of human MUC5B, MUC5AC, and MUC2. One hundred forty-six clones were isolated and assembled to generate nine different contigs. On the basis of their similarity to the human mucins at the level of both nucleotide and the translated amino acid sequences, four contigs (assembled from 70 clones) were assigned to Muc5b, three contigs (assembled from 69 clones) to Muc5ac, and two contigs (assembled from 7 clones) to Muc2. The translation and alignment of these sequences with human MUC5B (Fig. 6A), MUC5AC (Fig. 6B), and MUC2 (Fig. 6C) showed the equine Muc5b, Muc5ac, and Muc2 sequences to have 80.93%, 78.95%, and 82.85% identity to human MUC5B, MUC5AC, and MUC2, respectively. In ad-
The database was also searched with sequences encoding the cysteine-rich domains of the three human genes. Using human MUC5B cysteine-rich domain sequences, four clones were isolated that were assembled into two contigs, which, after translation, resulted in a polypeptide sequence similar to MUC5B cysteine-rich domain sequences. For Muc5ac, only peptides corresponding to the cysteine-rich domains were found (Fig. 7A). For Muc5b, only peptides from Muc5ac (Figs. 6 and 7). For Muc5b, 17 peptides were in the NH2-terminal region (Fig. 6A), and 16 were from the putative cys domains. These data suggest that the mucin monomers contain glycosylated proteinase-resistant domains flanked by cysteine-containing regions of the protein core less substituted with glycan chains. This organization is consistent with the identification of peptides in the ESI-MS/MS analysis that likely arise from trachea, whereas putative Muc5ac fragment was only expressed in the small intestine (Fig. 8). Primer pairs FR1 (Muc5b) and FR2 (Muc5ac) were also amplified from equine genomic DNA, and all four amplified products (2 from cDNA and 2 from genomic DNA) were of the expected size, and their sequences were identical to the predicted sequences (data not shown). Furthermore, the sequence of the PCR fragment amplified with the primer pair FR3 (Muc2) was identical to the predicted sequence. These results indicate that the assembly and assignment was correct for these three mucin fragments.

**DISCUSSION**

Using a two-step purification scheme (density gradient centrifugation followed by gel filtration), we have successfully isolated a heterogeneous population of oligomeric mucins from equine airways mucus. Density gradient centrifugation and agarose gel electrophoresis indicated that at least two populations of oligomeric mucins were present in the preparation. Rate-zonal centrifugation and MALLS analyses of the purified mucins, before and after reduction, indicated that the equine mucins are polydisperse oligomers (average Mr = 14 MDa) comprised of monomeric units (average Mr = 7 MDa) held together by disulfide bridges. Trypsin digestion of the reduced monomers gave rise to high Mr glycopeptides (average Mr = 2.6 MDa) that correspond to the glycosylated regions of the molecule. These data suggest that the mucin monomers contain glycosylated proteinase-resistant domains flanked by cysteine-containing regions of the protein core less substituted with glycan chains. This organization is consistent with the identification of peptides in the ESI-MS/MS analysis that likely arise from trachea, stomach, and small intestine. The putative Muc5b fragment was expressed only in trachea, whereas putative Muc5ac was expressed in trachea as well as stomach, and the Muc2 fragment was only expressed in the small intestine (Fig. 8).
from cys domains that, in human mucins, are situated between the STP domains (13, 17, 46). Taken together, these data demonstrate that horse airways oligomeric mucins are very similar in molecular architecture to their human counterparts.

Interestingly, the reduced monomers and the high M_r glycopeptides from the horse mucins appear on average two to three times larger than those from human mucins (35, 46, 47). We cannot rule out that these differences in size are due to

Fig. 6. Alignment of human MUC5B (Q9HC84), MUC5AC (Q8WWQ5), and MUC2 (Q02817) NH2-terminal polypeptide sequences with the predicted equine Muc5b, Muc5ac, and Muc2 sequences. Human MUC5B sequence Q9HC84 (from amino acid 68), MUC5AC sequence Q8WWQ5 (from amino acid 72), and MUC2 Q02817 (from amino acid 117) were used for this alignment (A, B, and C, respectively). The asterisks indicate differences between the human and predicted equine (eqMuc) sequences, and the gaps necessary for the alignment are shown with dashes. The gaps between the assembled contigs are shown with dots. The peptide matches found by ESI-MS/MS analysis are highlighted in the black boxes.
association between molecules in the nonchaotropic solvent (0.2 M NaCl) employed for the column chromatography; however, a similar solvent was also employed for studies on human mucins (35, 46, 47). Analysis of the relationship between size and shape of the mucins provides further details on their macromolecular architecture. The value for the shape-dependent Mark-Houwink parameter ($\alpha = 0.64$) suggests that the mucins adopt a stiffened random coil conformation in solution. This value is higher than that determined for other mixtures of oligomeric mucins (10, 36, 40) but very similar to that determined for MUC5AC mucins ($\alpha = 0.69$) produced by cultured HT29 human intestinal cells in culture (35). Overall, the data presented are consistent with the equine oligomeric mucins being linear, flexible macromolecules in solution.

ESI-MS/MS fragmentation analysis of tryptic peptides derived from the purified mucin preparation identified a small number of peptides that matched with sequences from human (and mouse) oligomeric mucins MUC5AC and MUC5B. This suggested that equine airways mucus contained homologs of these two major human airways mucins. No equine mucin sequences were found by this analysis, and this can be explained because of the lack of published equine mucin sequences in the databases used for the analysis of the peptide fragmentation data (SWISS-PROT and TrEMBL). To gain more certainty in the identification of these two mucins, we derived predicted amino acid sequences for putative equine Muc5ac and Muc5b using unprocessed genomic sequence data (http://www.ncbi.nlm.nih.gov/Traces/trace.cgi). Reanalyzing the MS/MS data against these novel equine sequences identified peptides derived from Muc5b (33 peptides) and Muc5ac (5 peptides). Muc5b peptides were from both the NH2-terminal (17 peptides) and the cys domains (16 peptides), whereas the five peptides that matched putative Muc5ac were only from the cys domains. Taken together, the ESI-MS/MS data suggest that Muc5b and Muc5ac are major oligomeric mucins in equine airways mucus, and furthermore, we conclude

Fig. 7. Alignment of human MUC5B (Q9HC84), MUC5AC (Q8WWQ5), and MUC2 (Q02817) cys domain polypeptide sequences with the predicted equine Muc5b (A), Muc5ac (B), and Muc2 (C) sequences. The asterisks indicate differences between the human and predicted equine (eqMuc) sequences. The peptide matches found by ESI-MS/MS analysis are highlighted in the black boxes. Human MUC5AC cys domains 2 and 4, and 3 and 5 differ by a few amino acids; however, only 1 sequence is shown, and the alternative amino acids are indicated on the line above the sequences.

Fig. 8. Tissue expression of equine Muc5ac, Muc5b, and Muc2. Oligonucleotide primers, designed based on the predicted sequences, amplified products of the expected size (179 bp for Muc5b, 230 bp for Muc5ac, and 239 bp for Muc2) and sequence. The amplified products of Muc5b, Muc5ac, and Muc2 genes from trachea (lane 1), stomach (lane 2), and small intestine (lane 3) are shown.
that Muc5b is more abundant than Muc5ac. We draw this conclusion for two reasons. First, based on the fact that more peptides were matched with sequences from the putative Muc5b, and second, three of the five peptides from Muc5ac are from repeated cysteine domains, which should be at higher relative abundance compared with the unique sequence in the NH₂ terminus (46) where none were found.

The identification of Muc5ac and Muc5b in equine airways mucus is in good agreement with studies on human airways mucus that have demonstrated that MUC5AC and MUC5B are the major oligomeric mucins (46) and preliminary data that showed cross-reactivity of glycoproteins from equine airways mucus with anti-human MUC5AC and MUC5B polyclonal antisera (51). However, whereas Muc5ac expression has been demonstrated previously in equine airways tissue (21), this is the first study to validate the presence of this mucin in equine respiratory mucus. Moreover, this is the first study to demonstrate the expression of Muc5b in equine airways tissue and to isolate and characterize the glycoprotein product. The horse is the third species (in addition to human and mouse) in which both Muc5b and Muc5ac have been identified in the airways (8, 18, 29, 41, 46, 54). This provides evidence that the production of these mucins is important for the protective function of the mucus gel in the airways; however, the role of each individual mucin is still unknown.

Our data on mRNA expression and protein identification indicate that Muc2 is not likely to be a major component of equine airways mucus. This is in agreement with Gerber and colleagues (21), who have shown that Muc2 is not expressed in the airways of horses suffering from RAO. It is important to note that our data does not rule out the presence of other oligomeric mucins in the secretion.

Because of the small number of published equine mucin sequences, it was necessary to assemble putative Muc5ac and Muc5b equine mucin sequences using the unprocessed data available from the Horse Genome Sequencing Project to effectively analyze the ESI-MS/MS data. The assumption on which we based our approach was that equine mucin genes from STP-rich regions interspersed with cysteine domains. This is a reasonable assumption considering that the mucin genes from other species have been found to have the same organization as the human (12, 17, 19, 50), and furthermore, the biochemical and biophysical characterization data presented here showed the mucins to have the same molecular architecture as the human counterparts. The successful identification of peptides using this data verifies that our approach was valid but it is possible that we might have assembled hybrid sequences. However, amplification and sequencing of cDNA fragments from the corresponding region of each putative gene showed that the predicted sequences were correct. Moreover, the two putative mucins showed differential tissue expression: the fragment from Muc5b was expressed in trachea, whereas Muc5ac was expressed in trachea and stomach, and Muc2 in small intestine. This is in agreement with expression pattern of these mucins in human tissues (1, 22, 23, 32, 42, 47). Thus both the sequencing data and tissue expression data indicate that the assembly and assignment was at least correct for these two fragments.

The respiratory syndromes commonly referred to as IAD (5, 48) and RAO (formerly known as chronic obstructive pulmonary disease or COPD; Refs. 19, 25, 29) are both characterized by excess airway mucus and associated neutrophilic inflammation. Therefore, despite probably different underlying etiologies (infectious for IAD and environmental for RAO) and opposing trends in population prevalence with age (decreasing for IAD and increasing for RAO), it remains difficult to readily clinically distinguish these conditions in individual animals. Therefore, the ability to demonstrate detectable differences in the mucins involved in IAD and RAO will theoretically permit development of differential diagnostic tests for these important and prevalent equine respiratory syndromes, thereby informing the most appropriate management of cases.

In summary, we have shown that airways equine mucus is a mixture of two oligomeric mucins, Muc5b and Muc5ac, with Muc5b being more abundant. These findings represent the groundwork for future studies on the role of these gel-forming glycoproteins in equine airways mucus in health and disease.

ACKNOWLEDGMENTS

We acknowledge Drs. David Knight and Thomas A. Jowitt for advice with mass spectrometry and light-scattering analyses, respectively. We also thank Marj Howard and Emma Keevill for technical assistance.

GRANTS

This work was funded by the Horserace Betting Levy Board.

REFERENCES


