



Proteomic analysis of polymeric salivary mucins: No evidence for MUC19 in human saliva

Karine Rousseau, Sara Kirkham, Lindsay Johnson, Brian Fitzpatrick, Marj Howard, Emily J Adams, Duncan F Rogers, David Knight, Peter Clegg, David J Thornton

► To cite this version:

Karine Rousseau, Sara Kirkham, Lindsay Johnson, Brian Fitzpatrick, Marj Howard, et al.. Proteomic analysis of polymeric salivary mucins: No evidence for MUC19 in human saliva. *Biochemical Journal*, 2008, 413 (3), pp.545-552. 10.1042/BJ20080260 . hal-00478964

HAL Id: hal-00478964

<https://hal.science/hal-00478964>

Submitted on 30 Apr 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Proteomic analysis of polymeric salivary mucins: No evidence for MUC19 in human saliva

Karine Rousseau¹, Sara Kirkham¹, Lindsay Johnson¹, Brian Fitzpatrick¹, Marj Howard¹, Emily J. Adams², Duncan F. Rogers², David Knight¹, Peter Clegg^{1,3} & David J. Thornton¹

¹ Wellcome Trust Centre for Cell-Matrix Research, Faculty of Life Sciences, University of Manchester, Manchester, M13 9PT, UK. ² Airway Disease, National Heart & Lung Institute, Imperial College London, London SW3 6LY, UK, ³ University of Liverpool, Faculty of Veterinary Science, Leahurst, Neston, CH64 7TE, UK.

Running head: Salivary mucins

Corresponding author: David J. Thornton

Wellcome Trust Centre for Cell-Matrix Research
Faculty of Life Sciences
The Michael Smith Building
University of Manchester
Manchester
M13 9PT, UK.

Tel: 0161 275 5647

Fax: 0161 275 1505

E-mail: dave.thornton@manchester.ac.uk

Key words: saliva, mucins, Muc19, Muc5b

Abbreviations: TR: tandem repeat; GdmCl: Guanidinium chloride; PAS: Periodic Acid Schiff's; PSM: Pig submaxillary mucin; BSM: Bovine submaxillary mucin; DTT: dithiothreitol; CsCl: Caesium chloride; MS-MS; Tandem mass spectrometry.

ABSTRACT

MUC5B is the predominant polymeric mucin in human saliva (Thornton, Khan, Mehrotra, Howard, Veerman, Packer, and Sheehan (1999) *Glycobiology* **9**, 293-302) where it contributes to oral cavity hydration and protection. More recently the gene for another putative polymeric mucin, *MUC19*, has been shown to be expressed in human salivary glands (Chen, Zhao, Kalaslavadi, Hamati, Nehrke, Le, Ann, and Wu, (2004) *Am.J.Respir.Cell Mol.Biol.* **30**, 155-165). However, to date, the MUC19 mucin has not been isolated from human saliva. Our aim was therefore to purify and characterize the MUC19 glycoprotein from human saliva. Saliva was solubilised in 4M guanidinium chloride and the high-density mucins were purified by density gradient centrifugation. The presence of MUC19 was investigated using tandem mass spectrometry of tryptic peptides derived from this mucin preparation. Using this approach we found multiple MUC5B-derived tryptic peptides but were unable to detect any putative MUC19 peptides. These data suggest that MUC19 is not a major component in human saliva. In contrast, using the same experimental approach we identified Muc19 and Muc5b glycoproteins in horse saliva. Moreover, we also identified Muc19 from pig, cow and rat saliva; the saliva of cow and rat also contained Muc5b, although, due to the lack of pig *Muc5b* genomic sequence data, we were unable to identify Muc5b in pig saliva. Our data suggest that unlike human saliva which contains MUC5B, cow, horse and rat saliva is a heterogeneous mixture of Muc5b and Muc19. The functional consequence of these species differences remains to be elucidated.

INTRODUCTION

Saliva is a dilute, complex mixture of proteins, glycoproteins, lipids and ions secreted from the major and minor salivary glands [1-3]. Saliva has many diverse functions including; aiding mastication and speech, digestion of food and maintenance of oral health [4]. The major macromolecular components of normal, unstimulated saliva are the mucins. There are two-populations of salivary mucins; the high molecular weight polymeric, gel-forming mucin MUC5B and the lower molecular weight, non-polymeric mucin MUC7 [5-7]. Two distinctive structural characteristics of these glycoproteins have important functional consequences [8]. First, the MUC5B and MUC7 mucin polypeptides, have a large central domain, with tandemly repeated sequences enriched in serine and/or threonine residues (the TR or mucin domain), which are the sites of extensive substitution with O-glycans. Second, upstream and downstream of the MUC5B mucin domain are regions of the polypeptide similar to von Willebrand Factor D domains, which are important for mucin polymer formation [8]. Thus, as a result of their extreme size and abundance of negatively charged O-glycans these glycoproteins play key roles in hydration and lubrication of the oral surfaces. Furthermore, mucins bind to, and sequester bacteria via their glycans and protein domains [9;10].

Until 2003, four gel-forming mucins had been reported, namely MUC6, MUC2, MUC5AC and MUC5B which are encoded by four consecutive genes on chromosome 11p15.5 [11]. These mucins are expressed in a tissue, and cell specific manner. For instance, in humans, MUC6 is expressed in mucous cells of submucosal glands in the stomach, MUC5B is mainly expressed by the mucous cells of the salivary glands and the submucosal glands in the airways, MUC5AC is expressed by the goblet cells of the airways and the stomach, and MUC2 is mainly expressed by the goblet cells in the intestine [12-17]. This tissue-specific distribution has been found to be relatively well conserved between species [18;19]. For example, in the mouse, *Muc5b*, along with *Muc5ac*, have been found to be expressed in the airways, *Muc2* is mainly expressed in the intestine and *Muc5ac* and *Muc6* are both found in the stomach [20-23]. However, differences in expression between mouse and human have also been reported. For example, in contrast to humans, no evidence of *Muc5b* (nor *Muc2*, *Muc5ac* or *Muc6*) expression has been reported in mouse saliva [22;23].

In 2003, MUC19 was reported as a fifth human gel-forming mucin. Although, neither the complete amino acid nor mRNA sequence have been published, MUC19 is predicted to have similar structural features to the other human gel-forming mucins [24]. *MUC19* gene expression has been reported in human airways and, in particular, in human salivary glands [24]. In the mouse, the complete sequence of the *Muc19* gene has been described and it is strongly expressed in sublingual and submandibular salivary glands [24;25]. Furthermore, mRNA sequences from porcine and bovine submaxillary mucins show high sequence similarity to the mouse *Muc19* sequence [26-28]. On the basis of these data, MUC19/*Muc19* would be expected to be present in saliva from humans and other mammals. However, studies on human saliva have identified MUC5B as the predominant polymeric mucin [7;16], and the MUC19 glycoprotein has yet to be identified in human saliva, or for that matter, in the saliva from other mammals. Moreover, in other animals, it

is not known whether Muc5b is a major salivary mucin. Therefore, in this study, our first aim was to purify the polymeric mucins from human saliva in order to identify whether MUC19 was present. Subsequently, we analysed the saliva collected from horse, rat, pig and cow to identify if Muc5b and/or Muc19 were present.

EXPERIMENTAL METHODS

Saliva collection: Unstimulated saliva was collected from 6 healthy human volunteers (between 20 and 35 of age) by spitting into 50 ml tubes that were kept on ice during the collection. Horse, cow and pig saliva were collected from excess secretions drooled while feeding. Necessary ethical approval was obtained for all the above animal sampling. Rats were anaesthetised with ketamine hydrochloride (100mg/kg) and xylazine (10mg/kg) by intraperitoneal injection and were injected intraperitoneally with the cholinergic agonist pilocarpine (25mg/kg) to stimulate salivary secretion. The anaesthetised rats were laid on their side and saliva dripping from the mouth was collected into plastic petri dishes. Experiments were carried out in accordance with the UK Home office guidelines for animal welfare based on the Animals (Scientific Procedures) Act 1986. Equal volumes of 8M guanidinium chloride (GdmCl) were added to each sample at the time of collection and up to 5 volumes of 4M GdmCl were subsequently added to solubilise the mucins. This step was performed by gentle agitation for at least 24 h at 4°C.

Purification and identification of human and horse salivary mucins: Human and equine salivary mucins, solubilised as above, were purified by caesium chloride/4M GdmCl density gradient centrifugation at a starting density of 1.4g/ml. The samples were centrifuged for at least 68 hours at 40 000 rpm and 15°C in a Beckman Ti45 rotor. After centrifugation tubes were emptied from the top into 20 fractions. An aliquot of each fraction was slot blotted onto a nitrocellulose membrane and stained with periodic acid Schiff's (PAS) reagent according to our previously published method [29]. The PAS-rich fractions were pooled, reduced and carboxymethylated, dialysed into water and then lyophilised.

The mucins present in the sample were identified by using tandem mass spectrometry as described previously [19]. In brief, samples were re-suspended into 0.1M ammonium bicarbonate prior to digestion with trypsin. The low molecular weight tryptic peptides were isolated by Sephacryl S-100 gel filtration chromatography and then separated by reverse-phase chromatography and analysed in-line by positive ion electrospray ionisation-MS-MS using a Q-ToF micro mass spectrometer (Waters, Manchester, UK). MS-MS data were analysed using a custom database containing putative mucin sequences (see below).

Mucin sequence prediction: Published *Muc19* and *MUC5B* sequences (accession numbers [NM_207243](#), [AF005273](#) and [AJ004862](#) [25;27;30]) obtained from mRNA sequencing data were used to perform a BLAST search of the genome sequence database of human, horse, bovine, pig and rat genomes. The *Muc5b* and *Muc19* genes in each genome were determined based on gene synteny and sequence similarity. In the human (chromosome 12), horse (chromosome number not identified) and rat (chromosome 7) genomes, the *LRRK2* (Leucine Rich Repeat Kinase 2) gene is located upstream of *Muc19* while the *CNTN1* (Contactin 1) gene is located downstream of *Muc19*. In the current version of the cow genome, the *Muc19* gene is located downstream of the *LRRK2* gene on chromosome 5 but the *CNTN1* gene is not in the same loci. However, the genome contains sequence gaps and the assembly might not be complete. In the human genome, *MUC5B* is part of a MUC gene complex located between the *APA2* gene (adaptor-related protein complex 2, alpha 2 subunit) and the *Tollip* gene (Toll-interacting protein). This was also the case in the horse

genome, while, in the rat and cow genomes, it is not yet clear if synteny is conserved because of the gaps in the genomic sequence and the incomplete assembly of these genomes. The structure of the genes was predicted based on the conservation of the exon/intron structure and amino acid sequence, which has been shown to be very tightly conserved between oligomeric mucins across species [22;31]. The genome databases were accessed through the UCSC browser (<http://genome.cse.ucsc.edu/>) and the ensemble browser (<http://www.ensembl.org/index.html>). It should be noted that the predicted N-terminal sequences do not represent the entire amino terminal sequence, but start from exon 4 and 2 for Muc19 and Muc5b respectively. The predicted exonic sequences were translated and the resultant amino acid sequences of putative Muc5b and Muc19 mucin sequences were added to an in-house custom mucin database.

Database searching-- Tandem mass spectra were extracted, charge state deconvoluted and deisotoped by Masslynx version 4.0 (Waters, Manchester, UK). All MS-MS samples were analyzed using Mascot (Matrix Science, London, UK; version 2.2.03) and X! Tandem (www.thegpm.org; version 2007.01.01.1). X! Tandem was set up to search a subset of our in-house custom mucin database. Mascot was set up to search the MSDB_2006-Sep-08 database (selected for Mammalian) and our in-house custom mucin database. Mascot and X! Tandem were searched with a fragmentation mass tolerance of 0.4 Da and a parent ion tolerance of 1.3 Da. Iodoacetamide derivative of cysteine was specified in Mascot and X! Tandem as a fixed modification. Oxidation of methionine was specified in Mascot and X! Tandem as a variable modification.

Criteria for protein identification-- Scaffold (version Scaffold-01_07_00, Proteome Software Inc., Portland, OR) was used to validate MS-MS based peptide and protein identifications. Peptide identifications were accepted if they could be established at greater than 90% probability as specified by the PeptideProphet algorithm [32]. Protein identifications were accepted if they could be established at greater than 90% probability and contained at least 1 identified peptide. Protein probabilities were assigned by the ProteinProphet algorithm [33]. Proteins that contained similar peptides and could not be differentiated based on MS-MS analysis alone were grouped to satisfy the principles of parsimony.

Identification of mucins in the saliva of cow, rat and pig: To determine if Muc19 and Muc5b were present in cow, rat and pig saliva, we employed a less lengthy procedure to isolate the mucins. This was because we only wished to determine if the mucin was present as opposed to obtaining maximum coverage of the polypeptide. For this approach, our aim was to enrich rather than purify the mucins. In brief, saliva from cow, rat and pig was dissolved with 8M GdmCl and then dialysed into 6M urea. An aliquot of this non-purified saliva was reduced and subjected to SDS-PAGE using NuPAGE 4 to12% bis-acrylamide gels (Invitrogen, Paisley, UK). Gels were run for 2 hours at 150V in NuPAGE MES-SDS running buffer and NuPAGE LDS sample buffer. After electrophoresis the gels were stained with PAS [29] and the high molecular weight smears at the top of the gel were excised, in-gel digested with trypsin [34] and analysed by tandem mass spectrometry as described above.

Rate-zonal centrifugation: To determine if Muc19 is a polymeric mucin stabilised by disulphide bonds, untreated mucins and reduced, and carboxymethylated mucins from horse and rat, were layered onto 6-8M GdmCl gradients and centrifuged at 40000 rpm in a SW40 Ti Beckman rotor for 2 ½ h at 15°C as described previously [19]. The tubes were emptied into 24 fractions from the top and the mucin distribution analysed by PAS-staining.

MUC5B and MUC19 mRNA relative expression in human salivary glands: Normal human salivary gland total RNA (Clontech 636552, 24 male/female Caucasians, age 16 to 60) was reverse transcribed into cDNA using random primers. The SYBR green system (Eurogentec) was then used to amplify *MUC5B* and *MUC19* cDNA by Real-Time quantitative PCR using the protocol supplied by the manufacturer. The oligonucleotide primers used to amplify *MUC5B* were; *MUC5B*-forward primer 5' ACCCAACGGTGCAATGTCA 3' and *MUC5B*-reverse primer 5' CGGCCACTCTCTTGTACTCAAAG 3' and the oligonucleotides primers used to amplify *MUC19* were; *MUC19*-forward primer 5' GAGTTCAGATGGCAAATGCACA 3' and *MUC19*-reverse primer 5' TGCCATCAGGACAGTCAAGTACA 3' [24]. The efficiency of each pair of primers was determined by using 5 serial dilutions of the cDNA and was $\geq 99.8\%$. The identity of the amplified products was confirmed by sequencing.

RESULTS

The first aim of the study was to determine whether MUC19 was present in human saliva. Our approach, which we have employed previously to identify human and horse polymeric mucins [19;35], was to purify salivary mucins, and then use tandem mass spectrometry of tryptic peptides to identify MUC19. However, only a partial MUC19 sequence, corresponding to the C-terminus of the mucin, has been reported [24]. Therefore, to maximise the chance of identifying MUC19, most of the polypeptide sequence of the N-terminus of MUC19 was predicted (for details see the experimental section). This resulted in a 1178 amino acid sequence, which was added to an in-house mucin protein sequence database for subsequent analysis of tandem MS data.

Identification of polymeric mucins in human saliva: Human saliva, pooled from 6 individuals, was solubilised in 4M GdmCl and fractionated by density gradient centrifugation. PAS-staining of fractions across the density distribution showed that the high-density fractions contained the bulk of the mucins (Fig. 1). The major PAS-rich peak was pooled (fractions 11-18), reduced and carboxymethylated, dialysed against water and lyophilised in order to concentrate the sample. The lyophilised material was digested with trypsin and the low molecular weight tryptic peptides were analysed by tandem MS [19]. This revealed 84 unique peptides which matched MUC5B (Fig. 2), but no peptides that matched MUC19 (Fig 3). Since MUC19 was not found in the high-density, PAS-rich fractions, the lower density material (fractions 1 to 10) from the density gradient was pooled and analysed in the same manner as above. The tandem MS analysis showed the presence of 5 peptides that matched MUC5B but no peptides that matched MUC19.

mRNA expression in human salivary glands: The apparent absence of MUC19 in human saliva extracts is at odds with published mRNA expression data [24]. Therefore, to confirm previous expression data and, to determine the relative expression of *MUC5B* to *MUC19* mRNA we performed quantitative Real-Time PCR on pooled human salivary gland cDNA. This analysis showed *MUC19* mRNA was expressed in human salivary glands and that there was approximately 10-fold more *MUC5B* mRNA than *MUC19* mRNA.

The presence of *MUC19* mRNA in salivary glands suggests that our mass spectrometry-based approach might not have been able to detect this mucin. Therefore, to investigate whether our approach could identify this mucin, we collected saliva from other mammals (cow, horse, rat, and pig), isolated the mucins and tested for Muc19, and the homologue of the major human salivary mucin, Muc5b, using tandem MS of tryptic peptides. However, for this to be successful we first had to predict the polypeptide sequences for horse, cow, pig and rat Muc19 and Muc5b.

Predicted N-terminal sequences of Muc5b and Muc19 from cow, horse, rat and pig:

Mucin polypeptide sequences for cow, horse, rat, and pig were predicted as described in the experimental section. The Muc19 and Muc5b N-terminal deduced amino acid sequence alignments are shown in Figures 3 and 4 respectively. The Muc19 sequences show approximately 68% identity and that 99% of the cysteine residues (100/101) are conserved.

The alignment of the Muc5b sequences shows they share 67% identity and all of the 90 cysteine residues are conserved.

Identification of Muc19 and Muc5b mucins from horse saliva: The purification and identification strategy used for analysis of human salivary mucins was employed for the analysis of horse saliva. The CsCl/4M GdmCl density gradient fractionation of horse saliva is shown in Figure 5. The high-density mucins (fractions 10-16) were reduced and carboxymethylated, digested with trypsin and subjected to tandem MS analysis. The results showed that 13 peptides matched the predicted N-terminal polypeptide of Muc5b and 25 peptides matched the predicted N-terminal polypeptide sequence of Muc19 (Table 1, Figs. 3 & 4). This validated our approach and showed the presence of Muc5b and Muc19 mucins in horse saliva. We then investigated whether one or both of these two mucins were present in the saliva of the cow, pig and rat.

Identification of Muc19 and Muc5b mucins from cow, pig and rat saliva: To more rapidly identify the mucins present in the saliva of these 3 animals we exploited the low electrophoretic migration of polymeric mucins in SDS-PAGE [36]. Solubilised saliva collected from rat, cow and pig was subjected to SDS-PAGE on a 4 to 12% gradient gel. After electrophoresis, gels were stained with the PAS reagent and the carbohydrate-rich smears located at the top of the gels (data not shown), were excised and digested with trypsin. Tandem MS analysis showed the presence of both Muc5b and Muc19 in rat and cow saliva (Table 1, Fig 3 & 4). Muc19 was also present in pig saliva with 8 peptide matches (Table 1, Fig. 3). However, the presence of Muc5b in pig saliva could not be tested because no *Muc5b* sequence is currently available in the pig genome.

Is Muc19 a polymeric mucin?: Muc19 is reported to be a polymeric mucin based on the presence of the von Willebrand factor-like D domains which are necessary for the disulphide-bond mediated oligomerisation of gel-forming mucins [24]. In order to test this, we examined if there was a change in size distribution of horse and rat salivary mucins (both mixtures of Muc5b and Muc19), before and after treatment with a reducing agent (10mM DTT), by rate zonal centrifugation (Fig. 6). The unreduced mucins were characterized by a broad range of sedimentation rates, characteristic of a polydisperse distribution of mucins [37;38], which, after reduction, exhibited a lower sedimentation rate. These results provide evidence that Muc19 (and Muc5b) is a polymeric mucin stabilised by disulphide bonds.

DISCUSSION

The finding that the *MUC19* gene is expressed in human salivary glands [24] raised the issue that a major component of the protective barrier in the mouth has been overlooked in previous studies on saliva. For instance, we have reported that MUC5B was the polymeric mucin responsible for the properties of saliva [7;35;39]. However, using standard methods for extraction and purification of polymeric mucins [40], we were unable in the present study to detect MUC19 either in the main mucin preparation after density gradient purification, or in any of the other fractions from the density gradient. Thus, the data presented here confirms these earlier studies and suggests that MUC5B is the predominant polymeric mucin in adult human saliva (pooled from 6 healthy volunteers); and that MUC19 is apparently absent. There are a number of reasons why MUC19 might not have been detected, (i) it was not solubilised in 4M GdmCl; (ii) peptides may not have been produced by trypsin digestion; thus the mucin would not have been identified using the MS-based approach; (iii) it is not present in saliva, or (iv) MUC19 is present but at levels below the limit of detection of the MS analysis. It seems unlikely that the first two reasons explain the lack of MUC19 since, using the same methodology, Muc19 was found in horse saliva. In addition, using the MS-based identification strategy, Muc19 was also present in rat, pig and cow saliva. We therefore propose that MUC19 is either not present in adult human saliva or is present at comparatively low levels. The latter proposal would seem to be favoured by the mRNA expression data which showed, in pooled human salivary gland mRNA, that *MUC19* mRNA is only around 10% of the level of *MUC5B* mRNA. Thus, if the two mRNAs are translated with similar efficiency this would suggest MUC19 is a more minor component of saliva than MUC5B. However, on the basis on the large number of peptides (84 in total) matched to MUC5B from the mass spectrometry data, we might reasonably have expected to detect peptides from MUC19 if it were present at 10% of the level of MUC5B in saliva. Therefore, we cannot discount other possibilities to explain the apparent absence of MUC19 from human saliva. It is possible that the mRNA is not, or is only poorly translated into protein, the mucin is only present during development or childhood, or the protein is made but not secreted under normal circumstances. Interestingly, MUC19 levels (mRNA and protein) have been shown to be altered in the conjunctival tissue from a patient with Sjögrens syndrome [41], a condition that also affects saliva production. However, these authors did not investigate changes in MUC19 production in salivary glands.

Compositional differences of saliva have been reported between different mammals; for example, in amylase, histatin and proline-rich proteins [42;43]. Thus, the difference in mucin composition between the different animals is not surprising. The functional significance of having or not having Muc19 in saliva is at present obscure. One might speculate that it impacts on the physical protective properties of saliva, in terms of hydration and lubrication, and/or it influences the innate defence properties of saliva, i.e. pathogen binding and sequestration of toxins. It is possible that the absence or low level of MUC19 in human saliva is an evolutionary adaptation, relating to diet or oral flora.

The polymeric properties of MUC5B/Muc5b are an established feature of this mucin [16;19;44-46], however, the oligomerisation potential of Muc19 was based on its

sequence similarity to the other gel-forming mucins [24]. Here, we have provided experimental evidence that Muc19 from horse and rat forms polymers held together by disulphide bounds. While this is the first study that specifically demonstrates Muc19 is a polymeric glycoprotein, earlier studies on recombinant N- and C-terminal fragments of PSM (now known to be Muc19) showed they were able to assemble in a disulphide-bond dependent manner [47-49], further confirming the polymeric nature of Muc19/PSM.

Horse, rat, and cow saliva are a mixture of two polymeric mucins; Muc5b and Muc19. Mucus gels from other tissues have also been shown to be mixtures of two mucins; in the airways MUC5AC/Muc5ac and MUC5B/Muc5b [19;50] and in the stomach MUC5AC and MUC6 [13]. In all of these tissues the specific role of each mucin is not well understood. However, potential roles for the glycan structures on the mucins in the stomach have been proposed in defence of the epithelium against the bacterium *H. pylori*. Le^b and sialyl Le^x glycans on MUC5AC can act as receptors for *H. pylori* [51] and α 1-4 linked N-acetylglucosamine termini of glycans on MUC6 act as an antibiotic [52]. Maybe Muc5b and Muc19 have similar roles in the oral cavity. Interestingly, in the rat we do have preliminary evidence that Muc19 has higher charge density than Muc5b (K. Rousseau and D.J. Thornton, unpublished observation).

In summary, our results confirm the earlier findings of Chen and co-workers [24] that *MUC19* mRNA is expressed in human salivary glands; however, we have demonstrated that the MUC19 glycoprotein is not a major component of human adult saliva. Furthermore, we have shown for the first time that horse, cow and rat saliva do contain Muc19 and these secretions are mixtures of Muc5b and Muc19. Future studies are required to investigate the functional relevance of differences in mucin composition of saliva in the physiology of the oral cavity.

ACKNOWLEDGMENTS

We would like to thank Emma Keevill and Julian Selley for their technical assistance with mass spectrometry analyses. This work was supported by the Wellcome Trust, the Horserace Betting Levy Board and the BBSRC.

Table 1: Numbers of peptides identified by tandem mass spectrometry from saliva extracts.

species	MUC5B*	Muc5b	MUC19	Muc19	
Human	84	-	none	-	Purified mucin
Horse	-	13	-	25	
Rat	-	5	-	20	Non-purified saliva
Cow	-	7	-	3	
Pig	-	NT	-	8	

NT: not tested, the pig Muc5b sequence was not available in the pig genome. *Peptides matched to full length MUC5B sequence, for the others we only searched for matches against N-terminal polypeptide sequences.

Dashes mean that samples were not analysed against these sequences.

FIGURE LEGENDS

Fig.1. Purification of human salivary mucins.

Human salivary mucins were extracted from saliva with 4M GdmCl and subjected to CsCl/4M GdmCl density gradient centrifugation. After centrifugation the tube was emptied from the top and fractions were analysed for carbohydrate by PAS assay [29] (filled circles) and their density was determined by weighing (open squares). Fractions 11 to 18, which contain the mucins, were pooled for subsequent tandem MS analysis.

Fig.2. MUC5B peptide matches from tandem mass spectrometry analyses of purified human salivary mucins.

Tryptic peptides generated from the mucins purified from human saliva by density gradient centrifugation were analysed by tandem MS. The 84 peptides that matched the human MUC5B sequence (**Q9HC84**) are highlighted in the black boxes.

Fig.3. Alignment of human MUC19, with horse, cow, pig (**AAC62527**) and rat Muc19 partial N-terminal sequences.

The amino acids which are identical between all species are shown with an asterisk (*) under the alignment. The cysteine residues are highlighted in grey. The translations shown here are from the second exon of each *Muc19* gene. The peptide matches found by tandem mass spectrometry analysis of mucins from the saliva from each animal are highlighted in the black boxes. No peptides were found that matched MUC19 in the human salivary mucin preparation. In contrast, peptides that matched to the putative Muc19 sequence were found in horse (25 peptides), cow (3 peptides), pig (8 peptides) and rat (20 peptides).

Fig.4. Alignment of human MUC5B (**AAG33673**) with horse, cow and rat Muc5b partial N-terminal sequence.

The amino acids which are identical between all species are shown with an asterisk (*) under the alignment. The cysteine residues are highlighted in grey. The translations shown here are from the fourth exon of each *Muc5b* gene. The cow Muc5b sequences contains two gaps (due to the incomplete genome data) which are indicated with an x. The peptide matches found by tandem mass spectrometry analysis of mucins from the saliva from each animal are highlighted in the black boxes. Peptides that matched to the putative Muc5b sequence were found in human (42 peptides) horse (13 peptides), cow (7 peptides), rat (5 peptides). The pig Muc5b sequence was not found in the pig genome and therefore no peptide matches could be identified.

Fig.5. Purification of horse salivary mucins

Horse saliva was fractionated by CsCl/4M GdmCl density gradient centrifugation. After centrifugation the tube was emptied from the top and fractions were analysed for carbohydrate by PAS assay [29] (filled circles) and their density was determined by weighing (open squares). Fractions 10 to 16, which contain the mucins, were pooled for subsequent tandem MS analysis.

Fig.6. Size distribution of horse and rat salivary mucins.

Horse (top panel) and rat (bottom panel) saliva extracts (4M GdmCl), before (filled circles) and after reduction (open circles), were centrifuged on 6-8M GdmCl gradients. After centrifugation tubes were emptied from the top and fractions were analyzed for carbohydrate using a PAS assay [29]. Fractions 1 (6M GdmCl) and 24 (8M GdmCl) are the top and the bottom of the tube respectively.

Reference List

- 1 Helmerhorst, E. J. and Oppenheim, F. G. (2007) Saliva: a dynamic proteome. *J. Dent. Res.* **86**, 680-93
- 2 Hu, S., Xie, Y., Ramachandran, P., Ogorzalek Loo, R. R., Li, Y., Loo, J. A., and Wong, D. T. (2005) Large-scale identification of proteins in human salivary proteome by liquid chromatography/mass spectrometry and two-dimensional gel electrophoresis-mass spectrometry. *Proteomics* **5**, 1714-28
- 3 Whelton, H. (1996) in *Saliva in oral health* (Edgar, W. M. and O'Mullane, D. M., eds.), pp. 1-8, British Dental Journal, London
- 4 Nieuw Amerongen, A. V., Ligtenberg, A. J., and Veerman, E. C. (2007) Implications for diagnostics in the biochemistry and physiology of saliva. *Ann. N. Y. Acad. Sci.* **1098**, 1-6
- 5 Bobek, L. A., Tsai, H., Biesbrock, A. R., and Levine, M. J. (1993) Molecular cloning, sequence, and specificity of expression of the gene encoding the low molecular weight human salivary mucin (MUC7). *J. Biol. Chem.* **268**, 20563-69
- 6 Mehrotra, R., Thornton, D. J., and Sheehan, J. K. (1998) Isolation and physical characterization of the MUC7 (MG2) mucin from saliva: evidence for self-association. *Biochem. J.* **334**, 415-22
- 7 Thornton, D. J., Khan, N., Mehrotra, R., Howard, M., Veerman, E., Packer, N. H., and Sheehan, J. K. (1999) Salivary mucin MG1 is comprised almost entirely of different glycosylated forms of the MUC5B gene product. *Glycobiology* **9**, 293-302
- 8 Thornton, D. J., Rousseau, K., and McGuckin, M. A. (2008) Structure and Function of the Polymeric Mucins in Airways Mucus. *Annu. Rev. Physiol* **70**, 459-86
- 9 Veerman, E. C., Ligtenberg, A. J., Schenkels, L. C., Walgreen-Weterings, E., and Nieuw Amerongen, A. V. (1995) Binding of human high-molecular-weight salivary mucins (MG1) to *Hemophilus parainfluenzae*. *J. Dent. Res.* **74**, 351-57
- 10 Veerman, E. C., Bank, C. M., Namavar, F., Appelmelk, B. J., Bolscher, J. G., and Nieuw Amerongen, A. V. (1997) Sulfated glycans on oral mucin as receptors for *Helicobacter pylori*. *Glycobiology* **7**, 737-43
- 11 Pigny, P., Guyonnet-Duperat, V., Hill, A. S., Pratt, W. S., Galiegue-Zouitina, S., d'Hooze, M. C., Laine, A., Van-Seuning, I., Degand, P., Gum, J. R., Kim, Y. S., Swallow, D. M., Aubert, J. P., and Porchet, N. (1996) Human mucin genes assigned to 11p15.5: identification and organization of a cluster of genes. *Genomics* **38**, 340-52
- 12 Davies, J. R., Svitacheva, N., Lannefors, L., Kornfalt, R., and Carlstedt, I. (1999) Identification of MUC5B, MUC5AC and small amounts of MUC2 mucins in cystic fibrosis airway secretions. *Biochem. J.* **344**, 321-30
- 13 Nordman, H., Davies, J. R., Lindell, G., De Bolos, C., Real, F., and Carlstedt, I. (2002) Gastric MUC5AC and MUC6 are large oligomeric mucins that differ in size, glycosylation and tissue distribution. *Biochem. J.* **364**, 191-200
- 14 Toribara, N. W., Gum, J. R., Jr., Culhane, P. J., Lagace, R. E., Hicks, J. W., Petersen, G. M., and Kim, Y. S. (1991) MUC-2 human small intestinal mucin gene structure. Repeated arrays and polymorphism. *J. Clin. Invest.* **88**, 1005-13
- 15 Vinall, L. E., Fowler, J. C., Jones, A. L., Kirkbride, H. J., De Bolos, C., Laine, A., Porchet, N., Gum, J. R., Kim, Y. S., Moss, F. M., Mitchell, D. M., and Swallow, D.

- M. (2000) Polymorphism of human mucin genes in chest disease: possible significance of MUC2. *Am. J. Respir. Cell Mol. Biol.* **23**, 678-86
- 16 Wickstrom, C., Davies, J. R., Eriksen, G. V., Veerman, E. C., and Carlstedt, I. (1998) MUC5B is a major gel-forming, oligomeric mucin from human salivary gland, respiratory tract and endocervix: identification of glycoforms and C-terminal cleavage. *Biochem. J.* **334**, 685-93
 - 17 Ho, S. B., Takamura, K., Anway, R., Shekels, L. L., Toribara, N. W., and Ota, H. (2004) The adherent gastric mucous layer is composed of alternating layers of MUC5AC and MUC6 mucin proteins. *Dig. Dis. Sci.* **49**, 1598-1606
 - 18 Inatomi, T., Tisdale, A. S., Zhan, Q., Spurr-Michaud, S., and Gipson, I. K. (1997) Cloning of rat Muc5AC mucin gene: comparison of its structure and tissue distribution to that of human and mouse homologues. *Biochem. Biophys. Res. Commun.* **236**, 789-97
 - 19 Rousseau, K., Kirkham, S., McKane, S., Newton, R., Clegg, P., and Thornton, D. (2007) Muc5b and Muc5ac are the major oligomeric mucins in equine airway mucus. *Am. J. Physiol. Lung Cell Mol. Physiol.* **292**, L1396-L1404
 - 20 Borchers, M. T., Wesselkamper, S., Wert, S. E., Shapiro, S. D., and Leikauf, G. D. (1999) Monocyte inflammation augments acrolein-induced Muc5ac expression in mouse lung. *Am. J. Physiol.* **277**, 489-97
 - 21 Chen, Y., Zhao, Y. H., and Wu, R. (2001) In silico cloning of mouse Muc5b gene and upregulation of its expression in mouse asthma model. *Am. J. Respir. Crit. Care Med.* **164**, 1059-66
 - 22 Escande, F., Porchet, N., Aubert, J. P., and Buisine, M. P. (2002) The mouse Muc5b mucin gene: cDNA and genomic structures, chromosomal localization and expression. *Biochem. J.* **363**, 589-98
 - 23 Escande, F., Porchet, N., Bernigaud, A., Petitprez, D., Aubert, J. P., and Buisine, M. P. (2004) The mouse secreted gel-forming mucin gene cluster. *Biochim. Biophys. Acta* **1676**, 240-50
 - 24 Chen, Y., Zhao, Y. H., Kalaslavadi, T. B., Hamati, E., Nehrke, K., Le, A. D., Ann, D. K., and Wu, R. (2004) Genome-Wide Search and Identification of a Novel Gel-Forming Mucin MUC19/Muc19 in Glandular Tissues. *Am. J. Respir. Cell Mol. Biol.* **30**, 155-65
 - 25 Culp, D. J., Latchney, L. R., Fallon, M. A., Denny, P. A., Denny, P. C., Couwenhoven, R. I., and Chuang, S. (2004) The gene encoding mouse Muc19: cDNA, genomic organization and relationship to Smgc. *Physiol. Genomics* **19**, 303-18
 - 26 Bhargava, A. K., Woitach, 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
 - 27 Eckhardt, A. E., Timpte, C. S., DeLuca, A. W., and Hill, R. L. (1997) The complete cDNA sequence and structural polymorphism of the polypeptide chain of porcine submaxillary mucin. *J. Biol. Chem.* **272**, 33204-210
 - 28 Jiang, W., Woitach, J. T., Gupta, D., and Bhavanandan, V. P. (1998) Sequence of a second gene encoding bovine submaxillary mucin: implication for mucin heterogeneity and cloning. *Biochem. Biophys. Res. Commun.* **251**, 550-56

- 29 Thornton, D. J., Carlstedt, I., and Sheehan, J. K. (1996) Identification of glycoproteins on nitrocellulose membranes and gels. *Mol. Biotechnol.* **5**, 171-76
- 30 Desseyn, J. L., Buisine, M. P., Porchet, N., Aubert, J. P., and Laine, A. (1998) Genomic organization of the human mucin gene MUC5B. cDNA and genomic sequences upstream of the large central exon. *J. Biol. Chem.* **273**, 30157-164
- 31 Buisine, M. P., Desseyn, J. L., Porchet, N., Degand, P., Laine, A., and Aubert, J. P. (1998) Genomic organization of the 3'-region of the human MUC5AC mucin gene: additional evidence for a common ancestral gene for the 11p15.5 mucin gene family. *Biochem. J.* **332**, 729-38
- 32 Keller, A., Nesvizhskii, A. I., Kolker, E., and Aebersold, R. (2002) Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal. Chem.* **74**, 5383-92
- 33 Nesvizhskii, A. I., Keller, A., Kolker, E., and Aebersold, R. (2003) A statistical model for identifying proteins by tandem mass spectrometry. *Anal. Chem.* **75**, 4646-58
- 34 Shevchenko, A., Wilm, M., Vorm, O., and Mann, M. (1996) Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Anal. Chem.* **68**, 850-58
- 35 Raynal, B. D., Hardingham, T. E., Sheehan, J. K., and Thornton, D. J. (2003) Calcium-dependent protein interactions in MUC5B provide reversible cross-links in salivary mucus. *J. Biol. Chem.* **278**, 28703-10
- 36 Swallow, D. M., Gendler, S., Griffiths, B., Corney, G., Taylor-Papadimitriou, J., and Bramwell, M. E. (1987) The human tumour-associated epithelial mucins are coded by an expressed hypervariable gene locus PUM. *Nature* **328**, 82-84
- 37 Sheehan, J. K. and Thornton, D. J. (2000) Heterogeneity and size distribution of gel-forming mucins. *Methods Mol. Biol.* **125**, 87-96
- 38 Thornton, D. J., Davies, J. R., Kraayenbrink, M., Richardson, P. S., Sheehan, J. K., and Carlstedt, I. (1990) Mucus glycoproteins from 'normal' human tracheobronchial secretion. *Biochem. J.* **265**, 179-86
- 39 Raynal, B. D., Hardingham, T. E., Thornton, D. J., and Sheehan, J. K. (2002) Concentrated solutions of salivary MUC5B mucin do not replicate the gel-forming properties of saliva. *Biochem. J.* **362**, 289-96
- 40 Davies, J. R. and Carlstedt, I. (2000) in *Glycoprotein Methods and protocols, The mucins* (Corfield, A. P., ed.), pp. 3-13, Humana Press,
- 41 Yu, D. F., Chen, Y., Han, J. M., Zhang, H., Chen, X. P., Zou, W. J., Liang, L. Y., Xu, C. C., and Liu, Z. G. (2008) MUC19 expression in human ocular surface and lacrimal gland and its alteration in Sjogren syndrome patients. *Exp. Eye Res.* **86**, 403-11
- 42 Jacobsen, N., Lyche Melvaer, K., and Hensten-Pettersen, A. (1972) Some Properties of Salivary Amylase: A Survey of the Literature and Some Observations. *J. Dent. Res. Supplement to No 2* **51**, 381-88
- 43 Shimada, T. (2006) Salivary proteins as a defense against dietary tannins. *J. Chem. Ecol.* **32**, 1149-63
- 44 Sheehan, J. K., Richardson, P. S., Fung, D. C., Howard, M., and Thornton, D. J. (1995) Analysis of respiratory mucus glycoproteins in asthma: a detailed study from a patient who died in status asthmaticus. *Am. J. Respir. Cell Mol. Biol.* **13**, 748-56

- 45 Thornton, D. J., Sheehan, J. K., Lindgren, H., and Carlstedt, I. (1991) Mucus glycoproteins from cystic fibrotic sputum. Macromolecular properties and structural 'architecture'. *Biochem. J.* **276**, 667-75
- 46 Wickstrom, C., Christersson, C., Davies, J. R., and Carlstedt, I. (2000) Macromolecular organization of saliva: identification of 'insoluble' MUC5B assemblies and non-mucin proteins in the gel phase. *Biochem. J.* **351**, 421-28
- 47 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**, 9845-50
- 48 Perez-Vilar, J. and Hill, R. L. (1998) Identification of the half-cystine residues in porcine submaxillary mucin critical for multimerization through the D-domains. Roles of the CGLCG motif in the D1- and D3-domains. *J. Biol. Chem.* **273**, 34527-34
- 49 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**, 6982-88
- 50 Kirkham, S., Sheehan, J. K., Knight, D., Richardson, P. S., and Thornton, D. J. (2002) Heterogeneity of airways mucus: variations in the amounts and glycoforms of the major oligomeric mucins MUC5AC and MUC5B. *Biochem. J.* **361**, 537-46
- 51 Linden, S., Nordman, H., Hedenbro, J., Hurtig, M., Boren, T., and Carlstedt, I. (2002) Strain- and blood group-dependent binding of *Helicobacter pylori* to human gastric MUC5AC glycoforms. *Gastroenterology* **123**, 1923-30
- 52 Kawakubo, M., Ito, Y., Okimura, Y., Kobayashi, M., Sakura, K., Kasama, S., Fukuda, M. N., Fukuda, M., Katsuyama, T., and Nakayama, J. (2004) Natural antibiotic function of a human gastric mucin against *Helicobacter pylori* infection. *Science* **305**, 1003-06

Figure 1

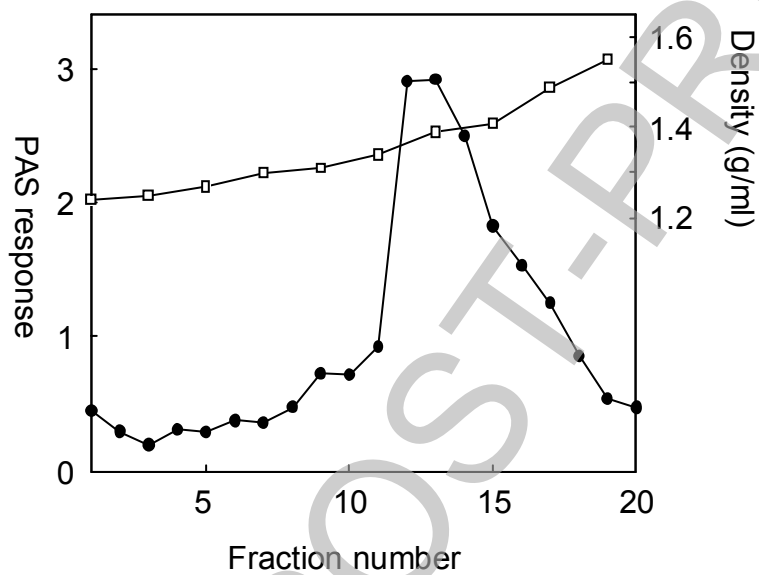


FIGURE 2

MGAPSACRTLVLALAAMLVFPQAE TQGPVE PSWGNAGHTMDGGAPTSSPTRRVS FVPVPTV F PSL SPLNPAHNGR **VCSTWCD FHYKTFDGD**
VFRFPGLCNVYF SEHCRAAYED FNVQLRRGLVGSRPVVTR VV IKAQGLVLEASNGSVLINGQR **EEELPYSRGTGLLVEQSGDYIKV** S IRLVLT
 FLWNGEDSALLELDPKYANQTCGLCGDFNLPAFNEFYAHNARLT **PLQFQNLQKLDGPTEQCPDPLP** PAGNCTDEEGICHR **TLLGPAFAE**
CHALVDS TAYLAACAQDLRCPTCPCATFVEYSRQCAHAGGQPRNWR **CPELCPR** T C P LNMQH QECGSPCTDTCSNPQRAQ **LCEDHCV DGC** F
CEPGS TVLDDTTHSGCLPLGQCPC THGGRTYSPGTSFNNTCS SQTCSGGGLVQCQDLPCPGTCSVQGGAHISTYDEK **LYDL LBGDCSVVLSKK**
CADSSFTVLAELRKCGLTNDNENCLKAVTSLDGGDTAIR VQADGGVFLNSIYTQLPLSAANITLFTPSSFFIVVQ **TGLGLQLLVQLVPMQ**
 VEVRLDPAHQOMCGLCGNFNQNAADDFALSGVVEATGAFAANTWK **AQAACANARN SFEDPCSLSVENENYARH** WCS **SLTDPNSAFSRCH**
SIINPKPFHNSCMFDTNCERSEDCLCAALSSYHACAAGVQLSDWR DGVCTKYMQNCPR **SQRYAYVVDACOPTCR** GLSEADVTCSVSFV
 PVDGCTCPAGTFLNDAGACVPAQECPCYAHGTVLAPGEVHVHDEGAVCSCTGGK **LSCLGASLQK** STGCAAPMVYLDCSNSSAGT **PGAEC LRS**
 CHTLDVGCFSTHCVSGCVCPPGLVSDSGGGCIAEEDPCVHNEATYKPGETIRVDCNTCTCRNRRWEC **SHRLCLGTCVAYCDGHFITFDGD**
RYSFEGSCEYILAQDYCDGNTHGTFRIVTENIPCGTGTGTTCSKA IKLFEVSEYELILQEGTFKAVAR **GGGDDPPYK** IRYMGI FLVIETHGM
 AVSWDRKTSVFIRLHQDYKGF **VCGLCGNFDDNAINDFATR** SSRVVGDALEFGNSWKLSPSCPDALAPKDPCTANPFRKSWAQ **QCSILHGF**
TFAACRSQVDDSTKYEACVNDACADSGDCECFCTA VAAQAACHDAGLCVSWR **TPDTCPLFCDFYNPHGGCEWHYQPCGAPCLKT** CRNP
SGHCLVDLPGLEGCYKCPSPQPF FNEDQMKCVAQCGCYDKDGNYYDVGARVPTAENCQSCNCTPSGIQCAHSLEACTCTYEDRTYSYQDV
 IYNTTDLGLAGCLIAICGSNGTIIRKAVACPGTPTTPTFTTAWVPHSTTSPALPVSTVCVREVCR **WSWYNGHREPEPGLGGGDFETFENL**
RQGYQVCPVLADIECRAAQQLPDMPLIEELGQVQVDCDRMR GLMCANSQOSPPLCHDYELRLCCEYVCGPSFAPGTSPQPSLSASTEPVAF
 TPTQTATEKTTLWVTPSIRSTAALTSQTGSSSGPVTVTPSAPGTTTQCP **PCQWTEWEDDYPKSEQOLGGVSEYDKIRAGGHLCQ** PKD
IECQAESFPNWTLAQVQKQVHCDVHFLVCRNWEQEGVFKM CNYRIRV **LCCSDDHCF** ERATTPPETTELETTATTTTQALFSTPQPTSSP
 GLTRAPPASTTAVPTLSEGLTS **PRYTSLTGTATTGGPRQ** SAGSTEPTVPGVATSTLPTRSALPGT **TGSLGTWRPSQPPTLAPTMTATSRR**
 GTGATASKEPLTTSLAPLTSELSTQAETSTPRTEETMSPLNTNTTATGSTTNPSSTPGTTP **IPVLT** TATTATPAATSSVTVPSSALGT
 KMCWAPKSI **ECRAENYPEVS IDQVQVLTCSLETGLTCK** NEDQGTGR **FMCMFNYN** VRV **LCCDDY** SHCPSTLATSSTATPSSPTGTTWILTKP
 TTTATTTASTGSTATASSTQATAGTPHVSTTATTPTVTSKATPFS **SSPGTATALPALRSTATTPTAT** SFTAI **PSSSLGTTWRLSQT** TTPM
 ATMSTATPSSPTETVHTSTVLTTTATTTGATGSVATPSSTPGTAHTTKVLT **TTTTGFTATPSSSPGRARTLPVWI** STTTPTTRGSTVTPS
 SLPGTHTPTVLTTTPTTTTATGSMATPSSSTQTSCTGTPPSLTTTATTTATGSTTNPSSTPGTTP **IPVLT** TATTATPAATSSVTVPSSALGT
 THTTPVPNTTATTHGRSLSPSSPHTVCTAWTSATSGILGTHITEPSTGTSHTPAATGT **TQHSTPALSSPHSSRTTESPPSPGTTT** PGH
 TTATSRITATATPSKTRTS **LLPSQPT** SAPITTVVTMGCEPQCAWSEWLDYSYMPGPGSGGDFDTYSNIR **AAGGAVCEQPLGLECRAQAQ**
GVPLRELQGVVECSLDFGLVCRNREQVQKFKMCFN YEIRVFCNNGHCPSTPATSSATPSSTPGTTWILTELT **TTTATTTTESTGSTATPTS**
 TLRTAPPKVLTTTATPTPTVTSKATPSSSPGTATALPALRS **TATPTAT** SVTPIPS **SSLGTTWRLSQT** TTPATMSTATPSSPETAHT
 STVLTATATTTGATGSVATPSSTPGTAHTTKVPTTTTGTGATPSSSPGTALTPPVWISTTTTPTTRGSTVTPSSIPGTTHTATVLT **TTTT**
 TVATGSMATPSSSTQTSCTGTPPSLTTTATTITATGSTTNPSSTPGTTRIP **PVLT** TATTATPAATSSVTVPSSALGTHTTPPVNTTATTHGRS
 LSPSSPHTVRTAWTSATSGTLGTHITEPSGTSTHTPAATGT **TQHSTPALSSPHSSRTTESPPSPGTTT** PGHTTATSRITATATP **SKTR**
 TSTLLPSPPTAPITTVVTMGCEPQCAWSEWLDYSYMPGPGSGGDFDTYSNIR **AAGGAVCEQPLGLECRAQAQGVPLRELQGVVECSLDF**
GLVCRNREQVQKFKMCFN YEIRVFCNNGHCPSTPATSSATPSSTPGTTWILTELT **TTTATTTTESTGSTATPTS**
 PTATSSKATSSSPRTATTLPLVLTSTATKSTATSF **TPSSST** LGTTGTSQNRPPHPMATMSTIHPSTPETHTSTVLTTKATTTTRATSSM
 STPSSPTGTTWILTELTAAATTAALPHGTSSPTGTTWILTELTSTATTATVPTGSTATASSTRATAGTLKVLTS **TATTPV** ISSRATPSS
 SPGTATALPALRSATTTPTATSTVTAIPSSSLGTAWTRLS **QTTPTAT** MSTATPSSPETVHTSTVLTTTATTTTRTGSVATPSSPGTAHT
 KVPTTTTGTGATPSSSPGTALTPPVWISTTTTPTTRGSTVTPSSIPGTTHTATVLT **TTTTVATGSMATPSSSTQTSCTGTPPSLTT** TATTI
 TATGSTTNPSSTPGTTP **IPVLT** TATTATPAATSSVTVPSSALGTHTTPPVNTTATTHGRSLSPSSPHTVPTAWTSATSGILGTHITEPS
 TGHSTHTPAATGTGTTQSTPALSSPHSSRTTESPPSPGTTT **PGHTTATSRITATATP** SKTR
 PGTARTPPVWISTTTTPTTSGSTVTPSSIPGTTHTPTVLTTTQPVATGSMATPSSSTQTSCTGTPPSLIT **TATTITATGSTTNPSSTPGTTP**
 IPPELTATTPAATSSVTVPSSALGTHTTPPVNTTATTHGRSLSPSSPHTVPTAWTSATSGTLGTHITEPSTGTSHTPAATGT **TTTT**
 TPALSSPHSSRTTESPPSPGTTT **PGHTTATSRITATATP** SKTR **STLLPSQPT** SAPITTVVTMGCEPQCAWSEWLDYSYMPGPGSGGDFDTYSNIR
AAGGAVCEQPLGLECRAQAQGVPLRELQGVVECSLDFGLVCRNREQVQKFKMCFN YEIRVFCNNGHCPSTPATSSSTAMPSSPT
 GTTWILTELTATTTASTGSTATPSSTPGTAPPEKVLTSPTATPTATSSKATSSSPRTATTLPLVLTSTATKSTATSV **TPIPSSLTGTTG**
 TLPEQTTTPVATMSTIHPSTPETHTSTVLTTKATTRATSSSTPSSPTGTTWILTELTAAATTTAGTGPTATPSSTPGTTWILTELT **TTT**
 ATTTASTGSTATLSSTPGTTWILTELTATVTVPTGSTATASSTQATAGTPHVSTTATTPTVTSKATPSSSPGTATALPALRSTATTPT
 ATSFTAI **PSSSLGTTWRLSQT** TTPATMSTATPSSPETVHTSTVLTTTATTTGATGSVATPSSTPGTAHTTKVPTTTTGTGFTATPSSSP
 GTALTTPVWISTTTTPTTTPTTSGSTVTPSSIPGTTHTARVLT **TTTTVATGSMATPSSSTQTSCTGTPPSLTT** TATTITATGSTTNPSSTP
 GTTPI **PPVLT** SMATTPAATSSKATSSSPRTATTLPLVLTSTATKSTATSF **TPSSST** LWTWTVPAQTTPMSTMTSTIHTSTPETHTST
 VLTTTATMTRATNSTATPSS **SLGTT** TRI **LTELT** TTTATTTAATGSTATLSSTPGTTWILTELTSTIATVMVPTGSTATSS **SLGTAHTPKVVA**
 MATMPTATASTVPSSSTVGTTTRTPAVLPSSLPFTSVSVSSVLT **TLRPTGFPSSSHFTPCPCRAFGQFFSPGEVIYNKDRAGCHFYAVC**
NOHCDIDR FQACPTSPPPVSSAPLSSPSPAPGCDNAIPLRQVNETWLENTVARCVGDNRVLLDPKPVANVTCVNKHLPIKVSDPSQF
 CDFHYECCEICSMWGGSHYSTFDGTSYTFRGNCTYVLMREIHARFGNLS **LYLDNHYCTASATAAAARCP** RALS **IHYK** **SMDIVLTVTMVHGK**
EEGLILFDQIPVSSGFSKNGVLVSLGTTTMR VDI **PALGVT** VTFNGQVQARLPYSLFHNHTEGQCGTCTNNQRDCLQRDGTTAASCKDM
AKTWLVPDSR KDGCAWPTGPTTASPAAVSSSTPTPTCPQPLCDLMSQVFAECHNLVPPGPFNACISDHCR **GRLEVEPCQSL** EAYAEI
CRARGVCS DWRGATGGLCD **LTCPPTKVYKPCGPIQ** PATCNSR **NQSPQLE** GMAEGCFCPEDQILFNAHMGICVQACPCVGPDGFPKFPGERW
 VSNQCSCVCDGGSVSVQCKPLPCDAQGQPPPCNRPFGVTVTRPRAENPCCPPTVCVNTTTC **QPSLPVCP** PGQES **ICTQ** EEGDCPTFRCR
 PQLCSYNGTIFYGVGATFGALPCHMCTCLSGDTQDPTVQCCEADCNNTTC **PQGF** EYKRVAGQCCGECVQTACLTPDGQPVQLNETWVNSHV
 DNCTVYLCEAEGGVHLLTPQASCPDVSSCRGSLRK **KTGCYSCEEDSCQVR** ENT **TILWHQ** GCE **TEVNI** TFCGEGSPGAS **YSAEQAQAMQH**
CTCCQER RVHETVPLHCPNGSAI **LHTY** THVDECGCTPFCVPAPMAPPHTR **GFFPAQEATAV**

THIS IS NOT THE FINAL VERSION - see doi:10.1042/BJ20080260

Human KGSHPPEATPKYSETNAIIGEISTWSKGAYKSFNGRIFFESSCYTTFCHRCIESGGDFNIEIKRNDSEIEKITVLIDN
Horse KGSHPPEATPKYSETDAIIGEASTWKGAYKAFNGRIFSFESSCYTTFCHRCVDSGGDFNVEIKRNDNEIERITVIDS
Cow -----EASTWKGAYKAFNGRIFSFESSCYTTFCHRCVDSGGDFNIEIKRNDSEIEKITVLIDN
Pig KGSHPPEATPKYSETDAIIGEASTWKGAYKAFNGRIFSFESSCYTTFCHRCVDSGGDFNVEIKRNDSEIEKITVLIDN
Rat KASHVPEATPTTYSEANEVAGEAATWKGTYKALNGRIFSFESSCTFTTFCRDCAESGEDFNVEIKRHENGDIIEIKALIDD
* * * * *

Human NDVSIFGDTILVNGESVQIPYNNKLIHIKKYGEHNVLSRRGILTLMDKNNKLSLTLHKQYPTGGLGNFNSTPGQDIN
Horse NDVSILGDGILVNGESVQIPYNNKLIHIKKYGEHNVLSRRGILTLMDKNNKLSLTLHKQYSTGGLGNFNSTPGDIN
Cow NDVSIFGDILVNGESVQIPYNNKLIHIKKYGEHNVLSRRGILSLMDKNNKLSLTLHKQYPTGGLGNFNSTPGDIN
Pig NDISISGDVILVNGESVHPIYNNKLIHIKKYGEHNVLSRRGILSLMDKNNKLSLTLHKQYPTGGLGNFNSTPGDID
Rat VGLLVDRDITISVNEERVQMPFSNKMIIHIKKQGDHYVLKTRRKILSLSWGK-NKLSLTLTYRQYITGGLGDFNSVPGEDIN
* * * * *

Human EHIANSKIPGDCPNAVGKSYEVCEGDIQHCNKIIGTYFEKCKVAAALSNDYKMICIDEYCQOT-RDKTSTCDTYSLSRLC
Horse EHIANSKIPGDCPNAVSKSYEVCEGDIQHCNKIIGTYFEKCKVTAALSNDYKMICIDEYCQON-GGKSTCDTYSLSRLC
Cow EHIANSKIPGDCPNAVSKSYEVCEGDIQHCNKIIGTYFEKCKVSTALSNDYKMICIDEYCQOS-RDKTSTCDTYSLSRLC
Pig EHIANSKIPGDCPNAVSKSYELCEGDIQHCNKIIGTYFEKCKVSTALSNDYKMICIDEYCQON-RDAKSTCDTYSLSRLC
Rat EHIANSKISNDPSTLSRNNEVCEGDIQHCNKIIGTYFEKCKVSTALSSEYKNCVDEYCHTGGNKQTTCTNTYSLSRLC
* * * * *

Human ASDGPGTFESNRSDSDVVGTCPECHYIYKCGPSNPATCSNVAPFDQSECVSGCTCEGYLLDDIGEKGRVCLKAECF
Horse ASDGPGTYEYWRDDESEVCEKPICFEKHIYKCGPSNPATCSNVAPFDQSECVSGCTCEGYLLDDIGEKGRVCLKTECF
Cow ASDGPGTFESNRDSDPDVCEKPICFEKHIYKCGPSNPATCSNVAPFDQSECVSGCTCEGYLLDDIGEKGRVCLKSDCF
Pig ASDGPGTYEYWRDDESEVCEKPICFEKHIYKCGPSNPATCSNVAPFDQSECVSGCTCEGYLLDDIGEKGRVCLKAYCF
Rat AYDGPVGFEDHVRDDESEVCEKPICFEKHIYKCGPSNPATCSNVAPFDQSECVSGCTCEGYLLDDIGEKGRVCLKAECF
* * * * *

Human CESSGTVYOPGEVREGPGCSOCTCODAKWSCTEALCPGRCKVEGSSLTTFDGVKYNFPNGNCHFLAVHNEDWSISVELRCP
Horse CESSGTVYOPGEVREGPGCSOCTCODAKWSCTEALCPGRCKVEGSSITTFDGVKYNHPGNCHEFLAHDWDWSISVELRCP
Cow CESSGTVYOPGEVREGPGCSOCTCODAKWSCTEALCPGRCKVEGSSLTTFDGVKYNHPGNCHEFLAHDWDWSISVELRCP
Pig CESSGTVYOPGEVREGPGCSOCTCODAKWSCTEALCPGRCKVEGSSLTTFDGVKYNHPGNCHEFLAHDWDWSISVELRCP
Rat CESSGTVYOPGEVREGPGCSOCTCODAKWSCTEALCPGRCKVEGSSITTFDGVKYNHPGNCHEFLAHDWDWSISVELRCP
* * * * *

Human PSCQGTGLNSVTLNLSVSDVKYVFNSDGTVTNDRKIRNQGYYSDDKIQIFNASSYLOQVETTFHVKLQIQIVPVMQLY
Horse PSCQGTGLNSVTLNLSVSDVKYVFNDRDGTVTNDRKIRNESYYSDDKIQIFKASSYLOQVETTFHVKLQIQIVPVMQLY
Cow PSCQGTGLNSVTLNLSVSDVKYVFNDRDGTVTNDRKIRNQGYYSDDKIQIFNASSYLOQVETTFHVKLQIQIVPVMQLY
Pig PSCQGTGLNSVTLNLSVSDVKYVFNDRDGTVTNDRKIRNQGYYSDDKIQIFKASSYLOQVETTFHVKLQIQIVPVMQLY
Rat GDCQGTGLNSVTLNLSVSDGNRFVFNDRDGTVTNDRKIRNQGYYSDDKIQIFNASSYLYMAEILSHIRLQIQVMPKQLY
* * * * *

Human VSMFPNGFTDTVGLGCSYNNKAEDDFMSSONILEKTSQAFANSWEMMSCPKGNPSSCISIEKEKFAERHCGLILDSSGPF
Horse VSMFPNGFTDTVGLGCSYNNKAEDDFMSSONILEKTSQAFANSWEMMSCPKGNPSSCISIEKEKFAERHCGLILDSSGPF
Cow VSMFPNGFTDTVGLGCSYNNKAEDDFMSSONILEKTSQAFANSWEMMSCPKGNPSSCISIEKEKFAERHCGLILDSSGPF
Pig VSMFPNGFTDTVGLGCSYNNKAEDDFMSSONILEKTSQAFANSWEMMSCPKGNPSSCISIEKEKFAERHCGLILDSSGPF
Rat VSLAPDSTDTVGLGCSYNNKAEDDFMSSONILEKTSQAFANSWEMMSCPKGNPSSCISIEKEKFAERHCGLILDSSGPF
* * * * *

Human ASCHPIVNPCKPYHECKKRYTCTCENSODCLCTILGNVYKACAEKETIYVWGRAGLCEHSCPSGLVFKNYVKAENSSQSL
Horse ASCHAIYNPKPYHECKKRYTCTCENSODCLCTILGNVYKACAEKETIYVWGRAGLCEHSCPSGLVFKNYVKAENSSQSL
Cow ASCHPIVNPCKPYHECKKRYTCTCENSODCLCTILGNVYKACAEKETIYVWGRAGLCEHSCPSGLVFKNYVKAENSSQSL
Pig ASCHPIVNPCKPYHECKKRYTCTCENSODCLCTILGNVYKACAEKETIYVWGRAGLCEHSCPSGLVFKNYVKAENSSQSL
Rat AACHOTVDPKPYHECKKRYTCTCENSODCLCTILGNVYKACAEKETIYVWGRAGLCEHSCPSGLVFKNYVKAENSSQSL
* * * * *

Human SERDRSDDEVDVVDGCTCPDAMYQNNREGNCVKSQCDYINDEVMQPKLIHIDNKCVRDGLLQCTPIDLTLLQNC
Horse SERDRSDDEVDVVDGCTCPDAMYQNNREGNCVKSQCDYINDEVMQPKLIHIDNKCVRDGLLQCTPIDLTLLQNC
Cow SERDRSDDEVDVVDGCTCPDAMYQNNREGNCVKSQCDYINDEVMQPKLIHIDNKCVRDGLLQCTPIDLTLLQNC
Pig SERDRSDDEVDVVDGCTCPDAMYQNNREGNCVKSQCDYINDEVMQPKLIHIDNKCVRDGLLQCTPIDLTLLQNC
Rat SARDRSDDEVDVVDGCTCPDAMYQNNREGNCVKSQCDYINDEVMQPKLIHIDNKCVRDGLLQCTPIDLTLLQNC
* * * * *

Human GGAAYVDCSDPKAQRRTNRTCTSTRNIPFEDNLPCRRCGCFPEGMVRNSKGCNVFPDDCPGSGGREGYDQGSVTSVGCNE
Horse GGAAYVDCSDPKAQRRTNRTCTSTRNIPFEDNLPCRRCGCFPEGMVRNSKGCNVFPDDCPGSGGREGYDQGSVTSVGCNE
Cow GGAAYVDCSDPKAQRRTNRTCTSTRNIPFEDNLPCRRCGCFPEGMVRNSKGCNVFPDDCPGSGGREGYDQGSVTSVGCNE
Pig GGAAYVDCSDPKAQRRTNRTCTSTRNIPFEDNLPCRRCGCFPEGMVRNSKGCNVFPDDCPGSGGREGYDQGSVTSVGCNE
Rat RGAAYVDCSDPKAQRRTNRTCTSTRNIPFEDNLPCRRCGCFPEGMVRNSKGCNVFPDDCPGSGGREGYDQGSVTSVGCNE
* * * * *

Human CTGKGSWSCTQNECQTTCHYIYEGGHVTFDGSYSFDDGLQYYSFLEDYCGHENGTFRIILTESVPCEDGLTCSRRIIVA
Horse CTGKGSWSCTQNECQTTCHYIYEGGHVTFDGSYSFDDGLQYYSFLEDYCGHENGTFRIILTESVPCEDGLTCSRRIIVA
Cow CTGKGSWSCTQNECQTTCHYIYEGGHVTFDGSYSFDDGLQYYSFLEDYCGHENGTFRIILTESVPCEDGLTCSRRIIVA
Pig CTGKGSWSCTQNECQTTCHYIYEGGHVTFDGSYSFDDGLQYYSFLEDYCGHENGTFRIILTESVPCEDGLTCSRRIIVA
Rat CTGKGSWSCTQNECQTTCHYIYEGGHVTFDGSYSFDDGLQYYSFLEDYCGHENGTFRIILTESVPCEDGLTCSRRIIVA
* * * * *

Human FQDQNIIVLDGKVTAVKSTESKKEELNANAYSIHTVGLYLILKFLQNGIIVWDKNTRLSVILDPNWNQKVCGLCGNNNGD
Horse FQDQNIIVLDGKVTAAKTSTESKKEEPNGNTYSVHTVGLYLILKFLNGITIIWDKNTRMSVILDPNWNQKVCGLCGNNNGD
Cow FQDQNIIVLDGKVTAVKSTESKKEELNANAYSIHTVGLYLILKFLNGITIIWDKNTRLSVILDPNWNQKVCGLCGNNNGD
Pig FQDQNIIVLDGKVTAVKSTESKKEELNANAYSIHTVGLYLILKFLNGITIIWDKNTRLSVILDPNWNQKVCGLCGNNNGD
Rat FQDQNIIVLDGKVTAVKSTESKKEELNANAYSIHTVGLYLILKFLNGITIIWDKNTRLSVILDPNWNQKVCGLCGNNNGD
* * * * *

Human LKDDFTTRYSSVAGLEFGNSWKTSECSQSDTVTAQTFPCDSNPYCKAWAVRKEIIRDSTFRDCHNKVDPSAYHDVIEE
Horse LKDDFTTRYSSVAGLEFGNSWKTSECSQSDTVTAQTFPCDSNPYCKAWAVRKEIIRDSTFRDCHNKVDPSAYHDVIEE
Cow LKDDFTTRYSSVAGLEFGNSWKTSECSQSDTVTAQTFPCDSNPYCKAWAVRKEIIRDSTFRDCHNKVDPSAYHDVIEE
Pig LKDDFTTRYSSVAGLEFGNSWKTSECSQSDTVTAQTFPCDSNPYCKAWAVRKEIIRDSTFRDCHNKVDPSAYHDVIEE
Rat LKDDFTTRYSSVAGLEFGNSWKTSECSQSDTVTAQTFPCDSNPYCKAWAVRKEIIRDSTFRDCHNKVDPSAYHDVIEE
* * * * *

Human ACACDMEGKYLGFCTAVAMYAEACSAVGVCVSWRKPNLCVYCDYNAFGECSWHYEPCTGTVTAKTKDQLVQKQESALL
Horse ACACDMEGKYLGFCTAVAMYAEACSAVGVCVSWRKPNLCVYCDYNAFGECSWHYEPCTGTVTAKTKDQLVQKQESALL
Cow ACACDMEGKYLGFCTAVAMYAEACSAVGVCVSWRKPNLCVYCDYNAFGECSWHYEPCTGTVTAKTKDQLVQKQESALL
Pig ACACDMEGKYLGFCTAVAMYAEACSAVGVCVSWRKPNLCVYCDYNAFGECSWHYEPCTGTVTAKTKDQLVQKQESALL
Rat ACACDMEGKYLGFCTAVAMYAEACSAVGVCVSWRKPNLCVYCDYNAFGECSWHYEPCTGTVTAKTKDQLVQKQESALL
* * * * *

Human EGGYAKCPDSAPYLDPNTMKCVLSSECSCFYNDVIPPAGGVIEDNCGRTGTCYIAGOLECS 1178 amino acids
Horse EGGYAKCPDSAPYLDPNTMKCVLSSECSCFYNDVIPPAGGVIEDNCGRTGTCYIAGOLECS 1167 amino acids
Cow EGGYAKCPDSAPYLDPNTMKCVLSSECSCFYNDVIPPAGGVIEDNCGRTGTCYIAGOLECS 1157 amino acids
Pig EGGYAKCPDSAPYLDPNTMKCVLSSECSCFYNDVIPPAGGVIEDNCGRTGTCYIAGOLECS 1178 amino acids
Rat EGGYAKCPENAPYLDPNTMKCVLSSECSCFYNDVIPPAGGVIEDNCGRTGTCYIAGOLECS 1179 amino acids
* * * * *

FIGURE 3

Human	LNPAHNGRVCSTWGDVHYKTFDGDVFRFPGLCNVVFSEHCRAAYEDFNVQLRRGLVGSRPVVTRVVIKAQGLVLKASNGS
Horse	VNPAHNGRVCSTWGDVHYKTFDGDVFRFPGLCNVVFSEHCRAAYEDFNVQLRRGLVGSRPVVTRVVIKAQGLVLKASNGS
Cow	LNAHNGRVCSTWGDVHYKTFDGDVFRFPGLCNVVFSEHCRAAYEDFNVQLRRGLVGSRPVVTRVVIKAQGLVLKASNGS
Rat	LHLAHNGRVCSTWGDVHYKTFDGDVFRFPGLCNVVFSEHCRAAYEDFNVQLRRGLVGSRPVVTRVVIKAQGLVLKASNGS
	* * * * *
Human	VLINGQRLEELPYSRTGLLVEQSGDYIKVSIKRLVLTFLWNGEDSALLELDPKYANQTCGLCGDFNGLPAFNEFYAHNARLT
Horse	ILINGQRLEELPYSRTGLLVEQSGDYIKVSIKRLVLTFLWNGEDSALLELDPKYANQTCGLCGDFNGLPAFNEFYAHNARLT
Cow	VLINGWRLEELPYSRTGLLVEQSGDYIKVSIKRLVLTFLWNGEDSALLELDPKYANQTCGLCGDFNGLPAFNEFYAHNARLT
Rat	VLVNGHREELPYSRTGLLVEQSGDYIKVSIKRLVLTFLWNGEDSALLELDPKYANQTCGLCGDFNGLPAFNEFYAHNARLT
	* * * * *
Human	PLOFCNLOKLDGPTEQCQDPLPLP-AGNCTDEEGICHRLLGPAFAEACHALVDSTAYLAACAODLCRCPTCPCATFVEYS
Horse	PLOFCNLOKLDGPTEQCQDPLPLP-ADNCTDEEGICHRLLGPAFAEACHALVDSTAYLAACAODLCRCPTCPCATFVEYS
Cow	PLOFCNLOKLDGPTEQCQDPLPLP-ADNCTDEEGICHRLLGPAFAEACHALVDSTAYLAACAODLCRCPTCPCATFVEYS
Rat	PLOFCNLOKLDGPTEQCQDPLPLP-ADNCTDEEGICHRLLGPAFAEACHALVDSTAYLAACAODLCRCPTCPCATFVEYS
	* * * * *
Human	ROCAHAGGQPRNWRCPLELCPRTCPLNQHOEGGSPCTDTCSPNQRAQLCEDHCVDDGCFPPGSTVLDDITHSGCLPLGQC
Horse	ROCAHAGGQPRNWRCPLELCPRTCPLNQHOEGGSPCTDTCSPNQRAQLCEDHCVDDGCFPPGSTVLDDITHSGCLPLGQC
Cow	ROCAHAGGQPRNWRCPLELCPRTCPLNQHOEGGSPCTDTCSPNQRAQLCEDHCVDDGCFPPGSTVLDDITHSGCLPLGQC
Rat	ROCAHAGGQPRNWRCPLELCPRTCPLNQHOEGGSPCTDTCSPNQRAQLCEDHCVDDGCFPPGSTVLDDITHSGCLPLGQC
	* * * * *
Human	PCTHGGRTYSPGTSFNNTCSSTCSGGLWQCQDPLCPGTCSPVQGGASHISTYDEKLYDLHGDCSYVLSKKCADSSFTVWAE
Horse	PCTHGGRTYSPGTSFNNTCSSTCSGGLWQCQDPLCPGTCSPVQGGASHISTYDEKLYDLHGDCSYVLSKKCADSSFTVWAE
Cow	PCTHGGRTYSPGTSFNNTCSSTCSGGLWQCQDPLCPGTCSPVQGGASHISTYDEKLYDLHGDCSYVLSKKCADSSFTVWAE
Rat	PCTHGGRTYSPGTSFNNTCSSTCSGGLWQCQDPLCPGTCSPVQGGASHISTYDEKLYDLHGDCSYVLSKKCADSSFTVWAE
	* * * * *
Human	LRKCCLTNDNENCLKAVTSL-DGGDTAIRVQADGGVFLNSIYTLPLSAANITLFTPSSEFIVVOTGLGLQLVLQVLPLM
Horse	LRKCCLTNDNENCLKAVTSL-DGGDTAIRVQADGGVFLNSIYTLPLSAANITLFTPSSEFIVVOTGLGLQLVLQVLPLM
Cow	LRKCCLTNDNENCLKAVTSL-DGGDTAIRVQADGGVFLNSIYTLPLSAANITLFTPSSEFIVVOTGLGLQLVLQVLPLM
Rat	LRKCCLTNDNENCLKAVTSL-DGGDTAIRVQADGGVFLNSIYTLPLSAANITLFTPSSEFIVVOTGLGLQLVLQVLPLM
	* * * * *
Human	OVFVRLDPAHOGOMCGLCGNFNONOADDFTALSGVVEATGAFAANTWRAOACANARNSFETPCSLSVENENYARHWCSR
Horse	OVFVRLDPAHOGOMCGLCGNFNONOADDFTALSGVVEATGAFAANTWRAOACANARNSFETPCSLSVENENYARHWCSR
Cow	OVFVRLDPAHOGOMCGLCGNFNONOADDFTALSGVVEATGAFAANTWRAOACANARNSFETPCSLSVENENYARHWCSR
Rat	OVFVRLDPAHOGOMCGLCGNFNONOADDFTALSGVVEATGAFAANTWRAOACANARNSFETPCSLSVENENYARHWCSR
	* * * * *
Human	LTDPSAFAFRCHSIINPKPFHNSCMFDTNCERSEDCALCALSSYVHACAKGVOLSDWRDGVCTHYMONCPKRSQRYAYV
Horse	LTDPSAFAFRCHSIINPKPFHNSCMFDTNCERSEDCALCALSSYVHACAKGVOLSDWRDGVCTHYMONCPKRSQRYAYV
Cow	LTDPSAFAFRCHSIINPKPFHNSCMFDTNCERSEDCALCALSSYVHACAKGVOLSDWRDGVCTHYMONCPKRSQRYAYV
Rat	LTDPSAFAFRCHSIINPKPFHNSCMFDTNCERSEDCALCALSSYVHACAKGVOLSDWRDGVCTHYMONCPKRSQRYAYV
	* * * * *
Human	VDACOETCRGLSEADVTCSVSFVVDGCTCPAGTFLNDAGACVPAQECPCYAHCTVLAPGEVVDHDEGAVCSCTGGKLSCL
Horse	VDACOETCRGLSEADVTCSVSFVVDGCTCPAGTFLNDAGACVPAQECPCYAHCTVLAPGEVVDHDEGAVCSCTGGKLSCL
Cow	VDACOETCRGLSEADVTCSVSFVVDGCTCPAGTFLNDAGACVPAQECPCYAHCTVLAPGEVVDHDEGAVCSCTGGKLSCL
Rat	VDACOETCRGLSEADVTCSVSFVVDGCTCPAGTFLNDAGACVPAQECPCYAHCTVLAPGEVVDHDEGAVCSCTGGKLSCL
	* * * * *
Human	GASLOKSTGCAAPMVYLDGNSSSAGTPGAECCLRSCHTLVDGCFSTHCVSGCVCPGLVSDSGSGGIAEEDCPVHNATY
Horse	GASLOKSTGCAAPMVYLDGNSSSAGTPGAECCLRSCHTLVDGCFSTHCVSGCVCPGLVSDSGSGGIAEEDCPVHNATY
Cow	GASLOKSTGCAAPMVYLDGNSSSAGTPGAECCLRSCHTLVDGCFSTHCVSGCVCPGLVSDSGSGGIAEEDCPVHNATY
Rat	GASLOKSTGCAAPMVYLDGNSSSAGTPGAECCLRSCHTLVDGCFSTHCVSGCVCPGLVSDSGSGGIAEEDCPVHNATY
	* * * * *
Human	KPGETIRVDCNCTCRNRRWECSSHRLLGLGTCVAYGDGHFTFDGDRYSFEGSCYILAQDYC-GDNTTHGTFRTVTENIP
Horse	KPGETIRVDCNCTCRNRRWECSSHRLLGLGTCVAYGDGHFTFDGDRYSFEGSCYILAQDYC-GDNTTHGTFRTVTENIP
Cow	KPGETIRVDCNCTCRNRRWECSSHRLLGLGTCVAYGDGHFTFDGDRYSFEGSCYILAQDYC-GDNTTHGTFRTVTENIP
Rat	KPGETIRVDCNCTCRNRRWECSSHRLLGLGTCVAYGDGHFTFDGDRYSFEGSCYILAQDYC-GDNTTHGTFRTVTENIP
	* * * * *
Human	CCTGCTTCSKAIFLVESEYELILOEGTFKAVARCGEGDLPYKIRYMGIFLVIETH-GMAVSWDRKTSVFIRLHQDYKGRV
Horse	CCTGCTTCSKAIFLVESEYELILOEGTFKAVARCGEGDLPYKIRYMGIFLVIETH-GMAVSWDRKTSVFIRLHQDYKGRV
Cow	CCTGCTTCSKAIFLVESEYELILOEGTFKAVARCGEGDLPYKIRYMGIFLVIETH-GMAVSWDRKTSVFIRLHQDYKGRV
Rat	CCTGCTTCSKAIFLVESEYELILOEGTFKAVARCGEGDLPYKIRYMGIFLVIETH-GMAVSWDRKTSVFIRLHQDYKGRV
	* * * * *
Human	CGLCGNFDDNAINDFTRSRSVVGDALEFGNSWKLSPSCPDALAPKDPCTANPFRKSWAQKQCSILHGPTFAACRSQVDS
Horse	CGLCGNFDDNAINDFTRSRSVVGDALEFGNSWKLSPSCPDALAPKDPCTANPFRKSWAQKQCSILHGPTFAACRSQVDS
Cow	CGLCGNFDDNAINDFTRSRSVVGDALEFGNSWKLSPSCPDALAPKDPCTANPFRKSWAQKQCSILHGPTFAACRSQVDS
Rat	CGLCGNFDDNAINDFTRSRSVVGDALEFGNSWKLSPSCPDALAPKDPCTANPFRKSWAQKQCSILHGPTFAACRSQVDS
	* * * * *
Human	TKYEAACVNDACACDSGGDCECFCTAVAAAYACACHDAGLCVSWRTFDTCPFLFCDFYNPHGGCEWHYQPCGAPCLKTCRNP
Horse	TKYEAACVNDACACDSGGDCECFCTAVAAAYACACHDAGLCVSWRTFDTCPFLFCDFYNPHGGCEWHYQPCGAPCLKTCRNP
Cow	TKYEAACVNDACACDSGGDCECFCTAVAAAYACACHDAGLCVSWRTFDTCPFLFCDFYNPHGGCEWHYQPCGAPCLKTCRNP
Rat	TKYEAACVNDACACDSGGDCECFCTAVAAAYACACHDAGLCVSWRTFDTCPFLFCDFYNPHGGCEWHYQPCGAPCLKTCRNP
	* * * * *
Human	SGHCLVDLPGLGECYPCPCPSOPFFNEDQMCKVAQC-GCYDKDGNYYDVGARVPTAENCQSCNCTPFGGICAHSLAETC
Horse	SGHCLVDLPGLGECYPCPCPSOPFFNEDQMCKVAQC-GCYDKDGNYYDVGARVPTAENCQSCNCTPFGGICAHSLAETC
Cow	SGHCLVDLPGLGECYPCPCPSOPFFNEDQMCKVAQC-GCYDKDGNYYDVGARVPTAENCQSCNCTPFGGICAHSLAETC
Rat	SGHCLVDLPGLGECYPCPCPSOPFFNEDQMCKVAQC-GCYDKDGNYYDVGARVPTAENCQSCNCTPFGGICAHSLAETC
	* * * * *
Human	TYEDRTYSYQDVIYNTDGLGACLAIAICGNGTIIRKAVACPGTPATTPTFTTAWVPHSTT1257 amino acids
Horse	TYEGKTYGEEVIYNTDGLGACLAIAICGNGTIIRKAVACPGTPATTPTFTTAWVPHSTT1259 amino acids
Cow	TYEGRTYAGDVIYNTDGLGACLAIAICRDNGTIVRAVECPGTLFKTPTFTTSTAAPPSTT940 amino acids
Rat	IYEGRTYNYNDVIYNTDGLGACLAIAICKDNGTIKRTTECEPEVPSTTPTFTTTLTPLTTN1258 amino acids
	* * * * *

Figure 4

Figure 5

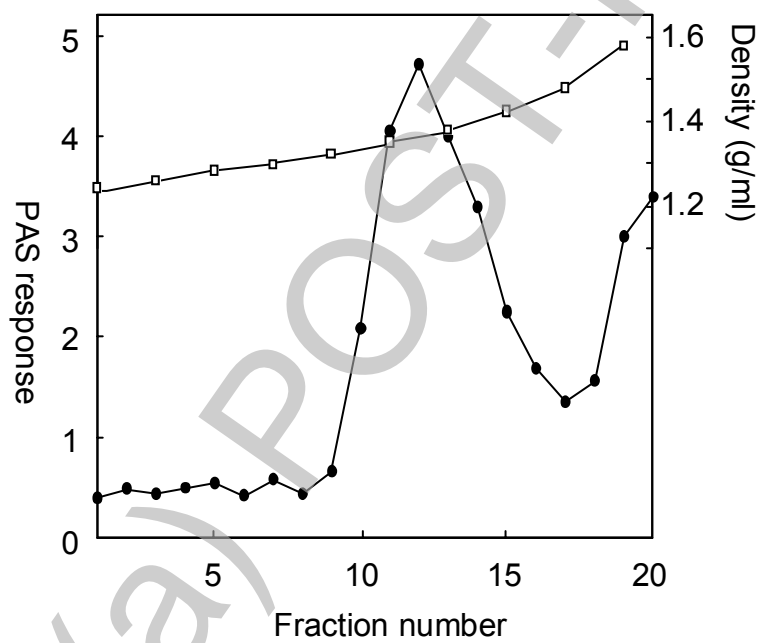


Figure 6

