Characterization of pig colonic mucins

Fiona J. J. FOGG*, David A. HUTTON*, Kornelia JUMEL†, Jeffrey P. PEARSON*‡, Stephen E. HARDING† and Adrian ALLEN*

*Department of Physiological Sciences, Medical School, Framlington Place, The University, Newcastle upon Tyne NE2 4HH, and
†Department of Analytical Biochemistry and Food Science, Sutton Bonnington Campus, University of Nottingham, LE12 5RD, U.K.

Pig colonic mucins isolated from the adherent mucus gel in the presence of proteinase inhibitors were solubilized by homogenization and the component mucins fractionated by CsCl density-gradient centrifugation. Polymeric and reduced pig colonic mucin were both largely excluded on Sepharose CL-2B, papain-digested colonic mucin was included. The M_r values of polymeric, reduced and digested mucins were 5.5×10^6, 2.1×10^6 and 0.6×10^6 respectively. This suggests that pig colonic mucin is comprised of 2–3 subunits, each subunit containing 3–4 glycosylated regions. The intrinsic viscosities of polymeric, reduced and digested mucin were 240 $ml \cdot g^{-1}$, 100 $ml \cdot g^{-1}$ and 20 $ml \cdot g^{-1}$ respectively. Polymeric pig colonic mucin comprised 16 % protein per mg of glycoprotein and was rich in serine, threonine and proline (43 % of total amino acids). There were approx. 150 disulphide bridges and 53 free thiol groups per mucin polymer. A seventh of the protein content was lost on reduction. This protein was particularly rich in proline and the hydrophobic amino acids. Papain-digested pig colonic mucin contained 11 % protein per mg of glycoprotein and was rich in serine, threonine, glutamate and aspartate. All types of amino acids with the exception of aspartate were lost on digestion. The amino acid analysis of the proteolytically digested regions of pig colonic mucin are markedly different to the tandem repeat regions of the human mucin genes shown to be expressed in the colon.

INTRODUCTION

In the colon, mucus forms a continuous, insoluble adherent gel layer which protects the underlying mucosa from the hostile environment of the lumen [1]. The protective properties of the mucus gel are directly correlated with the polymeric structure of the component mucins [2]. The polymeric structure is stabilized by interchain disulphide bridges [3–5]. The mucin protein core consists of highly glycosylated regions (resistant to proteolysis) and regions sparsely or non-glycosylated (susceptible to proteolysis) [5–7]. Several mucin genes have now been partially sequenced [8]. Characteristic of these mucin genes is a region coding for tandemly repeated amino acids which is thought to correspond to the glycosylated region of the mucin protein core [9–12]. It has become apparent from molecular biology studies that a mucus secretion is probably composed of more than one gene product. At least three mucin genes have been shown to be expressed in the human colon, MUC2, 3 and 4 [13]. The tandem repeat regions of these three mucin gene products have considerably different amino acid compositions although all contain a large proportion of threonine, between 25 and 60 %. The tandem repeat regions of MUC3 and 4 also contain large proportions of serine (29 % and 19 % respectively). MUC2 contains little serine in its tandem repeat region (~4 %). MUC2, which is fully sequenced, is predicted to contain two glycosylated regions, a tandem repeat and a base pair repeat region, of different size and amino acid composition [14]. To date two porcine mucin genes have been partially sequenced, porcine submaxillary mucin (PSM) gene and pig gastric mucin (PGM) gene [15,16]. The tandem repeat regions of PGM and PSM both contain a larger proportion of serine than threonine. PGM also appears to have a second glycosylated region, like MUC2, not comprised of tandem repeats [16].

Earlier studies on mucins employed techniques which probably did not enable the total, intact secretion to be analysed. Pig colonic mucin previously described was isolated in the absence of proteinase inhibitors and solubilized by stirring [17]. The method of solubilization has been shown to be important in isolating mucin representative of the whole secretion [4,18]. Rat small intestinal mucin solubilized by stirring alone has been shown to account for as little as 20 % of the total secretion [19]. Homogenization or stirring in the presence of mercaptoethanol is reported to be necessary for complete solubilization of the gel [7,19]. The inclusion of proteinase inhibitors has since been shown to produce larger mucins and is therefore important for the analysis of the intact glycopeptide [20,21]. The aim of this study was to isolate and analyse secreted undegraded pig colonic mucin representative of the total adherent gel secretion.

EXPERIMENTAL

Purification of mucins

Colons from freshly slaughtered pigs were collected on ice, cut longitudinally and their contents gently flushed with cold running water. The mucus gel was removed with a microscope slide and collected into an ice-cooled 0.067 M phosphate buffer, pH 6.7. The buffer contained 3 mM PMSF, 15 mM benzamidine hydrochloride, 10 mM EDTA, 10 mM *N*-ethylmaleimide, 100 mM α-aminohexanoic acid and 1 mM iodoacetamide. Scrapings were solubilized for 1 min in a Waring blender. Soluble and insoluble material was separated by centrifugation (1 h, 8000 ***g***, 4 °C) into two fractions, soluble mucus and pellet (which was retained for study).

Preparation of polymeric, reduced and digested mucin

Solubilized glycoprotein was purified in a CsCl density gradient (starting density 1.42 $g \cdot ml^{-1}$) and fractionated by Sepharose CL-2B column chromatography. The void volume fractions were pooled, dialysed, freeze-dried and used as polymeric mucin. Purity of mucin was determined by SDS/PAGE and staining with silver. Purified mucin was digested with papain (0.08 mg of papain/mg of glycoprotein, at 60 °C), for 48 h in 0.067 M sodium phosphate buffer, pH 6.25, containing 5 mM cysteine hydrochloride and 5 mM EDTA. Digested mucin was fractionated in

Abbreviations used: GuHCl, guanidinium chloride; PAS, periodic acid/Schiff; PGM, pig gastric mucin; PSM, porcine submaxillary mucin.
‡ To whom correspondence should be addressed.

a further CsCl gradient, starting density 1.42 g·ml^{-1}. Polymeric mucin was reduced with 0.2 M mercaptoethanol, for 24 h, at room temperature in 0.2 M Tris/HCl buffer, pH 8.0, containing 0.01 M EDTA, and was subsequently blocked overnight with 0.22 M iodoacetamide. Reduced mucin was fractionated in a further CsCl gradient, starting density 1.42 g·ml^{-1}.

Quantification

The proportion of glycoprotein solubilized by homogenization and the proportion remaining in the pellet were determined for mucin isolated both in the presence and absence of 4 M guanidinium chloride (GuHCl). The pellet of insoluble material was digested with papain and assayed for glycoprotein. The solubilized glycoprotein was purified by equilibrium centrifugation in a CsCl gradient, starting density 1.42 g·ml^{-1} (1.39 g·ml^{-1} in the presence of GuHCl); 100000 ***g***, 24 h, 4 °C using a Sorval vertical rotor. The glycoprotein-rich fractions from this gradient were pooled, digested with papain and assayed for glycoprotein.

Glycoprotein estimation

Glycoprotein was estimated by the periodic acid/Schiff (PAS) assay [22] either in solution or after blotting on to nitrocellulose sheets [23].

Gel filtration

Sepharose CL-2B, 4B, 6B and Sephacryl S-500 (1.5 cm diam., length 150 cm) columns were eluted by upward flow with 0.2 M NaCl containing 0.02 % (w/v) NaN_3 at a flow rate of 15 ml·h^{-1}.

Intrinsic viscosity

Measurements were made in 0.2 M NaCl, 0.02 % (w/v) NaN_3 using a Contraves Low Shear 30 Viscometer, 25 °C, at shear rates of between 1.7 and 128.5 s^{-1} and at mucin concentrations between 1.8 and 4.0 mg·ml^{-1}. Reduced specific viscosity (η_{sp}/c) was plotted against concentration and the intrinsic viscosity determined by extrapolation to infinite dilution (Huggins plot). Alternatively a Kraemer plot of $\ln(\eta_{rel})$ versus concentration was used which also yields the intrinsic viscosity at the intercept on the y axis.

Low-speed sedimentation equilibrium

Measurements were made in 0.01 M phosphate buffer, pH 6.8, containing 0.001 M EDTA, 0.01 % (w/v) NaN_3 and 0.03 M NaCl, using a Beckman Model E ultracentrifuge at 20 °C. Solute distributions were recorded by Rayleigh interference optics according to the intermediate speed method [24].

Light scattering

Measurements were made in 0.01 M phosphate buffer, pH 6.8, containing 0.001 M EDTA, 0.01 % (w/v) NaN_3 and 0.03 M NaCl using a Dawn F (Wyatt Technology) system [25]. This enabled the combination of size-exclusion chromatography with a classical light-scattering photometer to produce an absolute measure of the weight-average molecular mass. The procedure is as described elsewhere [26].

Amino acid analysis

Analysis was performed by the method of Carlton and Morgan [27]. Samples were hydrolysed for 24 h, at 110 °C, derivatized with 9-fluoroenylmethylchloroformate and analysed by reverse-phase HPLC.

Thiol analysis

Thiol content of mucin was measured by the method of Mantle et al. [28]. Purified mucin was reduced in 8 M urea containing 0.02 M Na_2EDTA and 0.33 M $NaBH_4$, pH 8, at 37 °C for 90 min. Thiol groups were labelled with 4,4-dithiopyridine (0.15 M final concentration) for 30 min at 20 °C, and detected at an absorbance wavelength of 324 nm. Free thiols were determined by the above procedure, but the reduction step was omitted.

RESULTS

After homogenization and centrifugation the proportion of insoluble mucin remaining in the pellet was 30 ± 14.5 % ($\bar{x} \pm$ S.E.M, $n = 3$) in the presence of GuHCl (after papain digestion, assessed by PAS) and 30 ± 5.9 % ($\bar{x} \pm$ S.E.M, $n = 3$) in the absence of GuHCl. Soluble mucin in the supernatant, accounting for ~ 70 % of the secretion, was purified by CsCl density-gradient centrifugation and fractionated by Sepharose CL-2B chromatography (recovery of PAS-positive material from these two procedures was 95 % and 80 % respectively). The high-M_r material was pooled and used for further analysis. Mucin (10 μg loading) was assessed to be free of contaminating protein bands by SDS/PAGE using silver stain, i.e. < 0.1 % of the material was contaminating non-mucin protein.

Physical analysis of pig colonic mucus

Pig polymeric colonic mucin was 57 % excluded on Sepharose CL-2B. While reduced mucin was 44 % excluded (K_{av} 0.08) there was a shift of material into the included volume. Digested mucin produced a broad included peak (K_{av} 0.5) with a small proportion of material excluded (3 %) (Figure 1). On Sepharose CL-4B and 6B digested mucin was largely excluded with an included shoulder. None of the gel-filtration profiles of digested mucin showed clearly separated species.

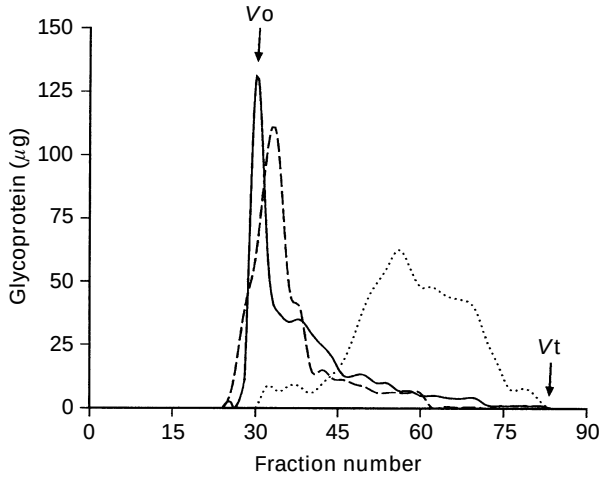

Figure 1 Sepharose CL-2B gel-filtration profiles of polymeric, reduced and digested pig colonic mucin

Polymeric (full line), reduced (broken line) and proteolytically digested (dotted line) mucin [1 mg in 2 ml of 0.2 M NaCl/0.02% (w/v) NaN_3] was eluted by upward flow with the above solvent from a Sepharose CL-2B column (150 cm × 1.5 cm). Fractions (2 ml) of the eluate were analysed for glycoprotein [22].

Table 1 M_r determination of pig colonic mucins, by multi-angle light scattering (M.A.L.S.) and sedimentation equilibrium (S.E.) and intrinsic viscosity measurements [η]

Mucin	Sample	M.A.L.S.	S.E.	$\bar{x} \pm$ S.E.M. (M.A.L.S.)	$[\eta]$ ($ml \cdot g^{-1}$)
Polymer	1	6.0×10^6	6.0×10^6		
	2	5.7×10^6		$5.5 \pm 0.3 \times 10^6$	240
	3	4.9×10^6			
Reduced	1	1.3×10^6	0.8×10^6		
	2	1.7×10^6		$2.1 \pm 0.5 \times 10^6$	100
	3	3.3×10^6			
Digested	1	0.5×10^6	0.5×10^6		
	2	0.6×10^6		$0.6 \pm 0.04 \times 10^6$	20
	3	0.7×10^6			

The weight-average molecular mass (M_w) of three samples of polymeric, reduced and digested mucin was determined by multi-angle light scattering using the Dawn F system (Table 1). On this system samples are passed through a size-exclusion column prior to light scattering. Polymeric pig colonic mucin eluted as a single broad excluded peak, M_w $(5.5 \pm 0.3) \times 10^6$ ($\bar{x} \pm$ S.E.M.). Reduced pig colonic mucin eluted as a broad peak extending from the excluded volume into the included volume, M_w $(2.1 \pm 0.5) \times 10^6$. Papain-digested pig colonic mucin eluted as a broad included peak, M_w $(0.6 \pm 0.04) \times 10^6$.

The weight-average molecular mass (M_w^o) of one of the samples of polymeric, reduced and digested pig colonic mucin was also determined by low-speed sedimentation equilibrium in addition to light scattering (Table 1). M_w^o of this sample of polymeric mucin was determined to be 6×10^6 (6×10^6 by light scattering). The M_w^o of reduced colonic mucin was determined to be 0.8×10^6 (1.3×10^6 by light scattering). The M_w^o of digested colonic mucin was determined to be 0.5×10^6 (0.6×10^6 by light scattering). The point-average molecular mass (M_w) and Z average molecular mass (M_z) of papain-digested mucin were determined to be 0.5×10^6.

The intrinsic viscosities of polymer, reduced and digested pig colonic mucin were calculated to be 240, 100 and 20 $ml \cdot g^{-1}$ respectively by extrapolation of a Kraemer and Huggins plot to infinite dilution (Table 1). The intrinsic viscosity values were in agreement with the molecular mass distribution.

Chemical analysis of pig colonic mucins

Pig polymeric colonic mucin contained 15% serine, 15% threonine and 13% proline. Of the remaining residues 20% were acidic (Asp and Glu), 19% were hydrophobic (Ala, Val, Ile, Leu and Phe), 11% were basic (His, Arg and Lys) and glycine comprised 8%. The protein content was 16% by freeze-dried weight (Table 2). Following reduction a seventh of the protein content was lost. Approximately half of the proline and the hydrophobic amino acids were lost. The proportion of serine (15%) remained similar to that of the polymer. The proportion of threonine (19%) and the acidic amino acids (26%) increased. Protein content was 13% by freeze-dried weight. On digestion approx. one-third of the protein content was lost. A large proportion of proline and the hydrophobic amino acids were lost. The proportion of serine (15%) remained similar to that of the polymer. The proportion of threonine (18%) increased as did the proportion of acidic amino acids (24%). The protein content was 11% by freeze-dried weight.

Pig colonic mucin not solubilized following homogenization and present in the pellet after centrifugation was papain digested. The amino acid analysis of the papain-digested pellet was the same as the solubilized mucin.

Amino acid analysis was also performed on papain-digested colonic mucin following gel filtration on Sepharose CL-4B. The yield of mucin from the column was approx. 60% ($n = 3$). The analysis of mucin in pooled fractions 28–45 was found to be the same as analysis of pooled fractions 46–70 (V_0 fraction 30, V_t fraction 80). In digested mucin pooled after column chromatography there was a much greater proportion of serine (27%) than threonine (7%). There was also a large proportion of acidic amino acids (35%). The percentage of protein was 9% by freeze-dried weight (Table 2). The analysis of digested mucin pooled after column chromatography was different to that of digested mucin prior to chromatography. In particular the proportion of threonine decreased and that of the acidic amino acids increased

Table 2 Amino acid composition of pig colonic mucins

Amino acid	Polymer ($\mu g \cdot mg^{-1}$; $\bar{x} \pm$ S.E.M., $n = 6$)	Reduced ($\mu g \cdot mg^{-1}$; $\bar{x} \pm$ S.E.M., $n = 5$)	[a]Digested ($\mu g \cdot mg^{-1}$; $\bar{x} \pm$ S.E.M., $n = 5$)	[b]Post column digested ($\mu g \cdot mg^{-1}$; $\bar{x} \pm$ S.E.M., $n = 3$)
His	5.4 ± 0.6	6.1 ± 0.8	4.4 ± 0.8	2.2 ± 0.3
Arg	6.2 ± 0.4	5.3 ± 1.7	4.3 ± 0.8	2.1 ± 0.3
Ser	23.8 ± 1.9	19.9 ± 2.2	17.0 ± 0.9	23.5 ± 1.2
Asp	11.1 ± 2.4	13.1 ± 0.9	12.2 ± 2.1	14.2 ± 2.8
Glu	20.4 ± 3.5	21.1 ± 1.3	14.5 ± 2.3	20.0 ± 3.8
Thr	23.1 ± 2.5	25.2 ± 0.9	19.6 ± 1.5	6.5 ± 0.5
Gly	11.9 ± 1.5	13.5 ± 2.1	7.7 ± 1.9	7.0 ± 0.5
Ala	10.4 ± 1.2	6.4 ± 1.1	5.0 ± 0.9	5.9 ± 0.4
Pro	20.5 ± 3.7	8.8 ± 1.3	9.9 ± 1.8	5.0 ± 0.2
Val	6.1 ± 0.7	3.0 ± 0.6	2.6 ± 0.8	5.1 ± 0.3
Ile/Leu/Phe	13.0 ± 2.6	7.0 ± 0.6	7.8 ± 1.7	1.9 ± 0.3
Lys	5.4 ± 0.3	3.1 ± 0.2	4.4 ± 0.9	1.6 ± 1.3
Total protein	155.7 ± 8.5	133.7 ± 5.8	111.8 ± 5.2	95.6 ± 4.4

[a] Analysis of papain-digested mucin fractionated by CsCl gradient.
[b] Analysis of papain-digested pig colonic mucin pooled after Sepharose CL-4B column chromatography.

Table 3 Thiol composition of pig colonic mucins

Abbreviation: n.d., not determined.

Mucin	Total thiols (nmol · mg^{-1}; $\bar{x} \pm$ S.E.M.)	Free thiols (nmol · mg^{-1}; $\bar{x} \pm$ S.E.M.)	No. of Cys residues	Percentage of Cys
Polymer ($n = 5$)	64.3 ± 4.4	9.6 ± 0.6	353 (150 S-S bridges)	4
Reduced ($n = 3$)	56.3 ± 2.2	n.d.	118	4
Digested ($n = 2$)	16.8 ± 0.4	n.d.	11	2

after chromatography. This suggests the loss of material on the column rich in threonine and poor in acidic amino acids, which is to some extent confirmed by the recovery of only about 60 % of the material.

Polymeric mucin contained 64.3 ± 4.4 nmol of total thiols · mg^{-1} including 9.6 ± 0.6 nmol of free thiols · mg^{-1} ($\bar{x} \pm$ S.E.M., $n = 5$) equivalent to (assuming an M_r of 5.5×10^6) 353 cysteine residues, ~4 % of the total protein content. Of these cysteine residues 300 were involved in disulphide bridges, i.e. 150 disulphide bridges per polymer and 53 free cysteines. Reduced pig colonic mucin contained 56.3 ± 2.2 nmol of total thiols · mg^{-1} ($n = 3$) equivalent to (assuming an M_r of 2.1×10^6) approx. 118 cysteine residues, ~4 % of the total protein content. Papain-digested pig colonic mucin contained 16.8 ± 0.4 nmol of total thiols · mg^{-1} ($n = 2$) (assuming an M_r of 0.6×10^6). Each digested subunit therefore has approx. 11 cysteine residues, ~2 % of the total protein content (Table 3).

DISCUSSION

Many mucins have been characterized previously [5]. However, it has become apparent that methods used during the isolation and solubilization of these mucins probably did not enable the whole secretion to be analysed [19,29]. In this study we have used brief homogenization as the method of solubilization of the colonic mucus gel. Intestinal mucus cannot be effectively solubilized by low-shear stirring even in GuHCl, and recent studies on rat intestinal mucin [19] and human colonic mucin [30] have shown that up to 80 % of the mucin remains insoluble. The mucus gel can be completely solubilized by stirring in reducing agents such as dithiothreitol [19]; however, reduction of the disulphide bridges destroys the polymeric structure of the mucin.

Therefore, in this study we used homogenization, rather than reduction or stirring in GuHCl, in order to solubilize as much of the secreted pig colonic mucin as possible while still essentially preserving the polymeric structure of the mucin [29].

Brief homogenization, used in this study, which breaks non-covalent gel-forming interactions, does involve strong shear forces which could break some covalent bonds in these large molecules. One study provides evidence that shear will rupture pig gastric mucin (a mucin which unlike colonic mucin can be extensively solubilized by stirring in GuHCl) resulting in a 5-fold drop from the very large M_r of 44×10^6 to 8×10^6 [31]. In contrast, another study, also isolating in the presence of proteolytic inhibitors, showed no difference in mass range [$(6–58) \times 10^6$] when stirring and homogenization were compared [29].

It is also relevant that proteoglycans and proteoglycan aggregates, M_r values of 2.5×10^6 and 100×10^6 respectively, are not damaged by isolation using high shear as long as proteolysis is inhibited [32]. It cannot be categorically ruled out that shear has damaged the mucin polymer and that the size of the native molecule in the gel secretion is larger than that seen in the isolated mucin. However, the large mass of the mucin isolated here (5.5×10^6) and the large mass of the reduced subunit (2.1×10^6), a similar mass to reduced subunits isolated from other gastrointestinal mucins without the use of shear [5], suggest no substantial disruption of covalent bonds has occurred.

Homogenization solubilized about 70 % of the secretion and the remaining insoluble 30 % when solubilized by proteolysis had the same amino acid analysis as the heavily glycosylated region from the soluble mucin, implying the same identity at the highly glycosylated core level. Therefore in this study we are analysing pig colonic mucin representative of the whole adherent mucus gel secretion.

Fragmentation of pig colonic mucin followed the same general pattern as that of other mucins [33]. Polymeric and reduced colonic mucin were excluded on Sepharose CL-2B, papain-digested mucin was included. Gel-filtration profiles of papain-digested pig colonic mucin on Sepharose CL-2B, 4B and 6B suggested size heterogeneity; however, no individual species were discernible by this method. This is in contrast to rat small intestinal mucin which has been separated into two different-sized glycosylated regions of M_r 355000 and 650000 on Sephacryl S-500 [19].

The weight-average molecular masses determined by light scattering were: polymeric colonic mucin, 5.5×10^6; reduced mucin, 2.1×10^6; and digested mucin, 0.6×10^6. The M_r values of mucins determined by light scattering were in general agreement with those produced by sedimentation equilibrium. Polymeric and reduced mucin were heterogeneous. This could be due to a number of factors: (1) it has become apparent that mucin secretions are comprised of a mixture of gene products; (2) different mucin polymers are comprised of different numbers of subunits or different-sized mucin subunits; (3) heterogeneity of the mucins could also be explained by selective proteolytic nicking of some of the protein cores *in vivo*, although care was taken to avoid proteolysis *in vitro*. The values for intrinsic viscosities for polymeric, reduced and digested mucin followed the same trend as the relative molecular masses.

Pig colonic polymeric mucin described in the present study is smaller than other polymeric mucins such as human cervical mucin (M_r 11×10^6) and pig gastric mucin (M_r 44×10^6) which were isolated in the presence of inhibitors and also GuHCl, and solubilized by stirring [21,34]. Human cervical mucin, however, has a similar M_r for reduced and digested subunit, i.e. $(2–2.5) \times 10^6$ and 0.4×10^6 respectively [35].

The polymeric structure of pig colonic mucin follows the same general pattern as has been described for other mucins. Pig colonic polymer (M_r 6×10^6) is composed of two or three subunits (M_r 2×10^6). Since there was no detectable loss of carbohydrate on proteolysis, each subunit would appear to comprise 3–4

glycosylated regions (M_r 6×10^5). This arrangement of repeating glycosylated regions in a mucin subunit is a model that has previously been reported for several mucins, including human cervical mucin and human middle ear mucin [18,36]. Recently a human mucin gene MUC5AC has been described with a predicted gene product of repeating glycosylated and non- or sparsely glycosylated regions [37].

Evidence from sedimentation equilibrium and gel filtration indicated that the papain-digested colonic mucin (ratio of M_z:M_w ~ 1) while polydisperse was not heterogeneous [38]. The polydispersity presumably primarily reflects differences in carbohydrate chain composition and length. These results suggest pig colonic mucin essentially contains one size of glycosylated species (M_r 0.6×10^6). This is unlike the gene product of MUC2 expressed in human colon which contains two different-sized glycosylated regions; a tandem repeat region and a base pair repeat region [14]. However, as yet there is no evidence that both of these glycosylated regions are resistant to proteolysis. The glycosylated region of pig colonic mucin (M_r 0.6×10^6) is smaller than predicted for the MUC2 tandem repeat region, which is comprised of ~ 100 tandem repeats of 23 amino acids (predicted M_r 1×10^6, assuming 80% glycosylation), but larger than the MUC2 base pair repeat region (predicted M_r 0.2×10^6). Interestingly PSM has a tandem repeat region containing 25 repeats of 81 amino acids [15]. The M_r of this tandem repeat region would be 0.4×10^6 (predicted from the gene sequence and assuming 50% glycosylation). However, the M_r of secreted digested pig submaxillary mucin has been determined to be much smaller (M_r 0.14×10^5) [39] and interestingly similar to the size of an individual glycosylated 81-amino-acid tandem repeat (predicted M_r 0.16×10^5). This points to sites susceptible to proteolysis within the tandem repeat region of some mucin gene products.

It is widely accepted that interchain disulphide bridges are essential for the polymerization of mucins [33,40]. Pig polymeric colonic mucin contained 353 thiol groups (~ 4% of the protein core). There were approx. 150 disulphide bridges (either inter or intra) per polymer and 53 free thiols. Reduced mucin contained ~ 9% less thiol groups than polymer, approx. 118 per reduced subunit. The cysteine content of pig colonic mucin is similar to that reported for the MUC2 subunit, which is also ~ 4%. A small proportion of thiols remained after papain digestion, approx. 11 per digested subunit (~ 2% of the total protein content). The loss of 74% of thiols on digestion indicated that the majority of cysteine residues were located in the regions of the protein core which are susceptible to proteolysis.

The results show that the subunits of colonic mucin are linked by disulphide bridges between the non- or sparsely glycosylated regions of the peptide core. Both the N-terminus and C-terminus of several mucin genes, including MUC2 (expressed in human colon), MUC5AC and FIMB1 (frog integumentary mucin), contain cysteine residues which would allow polymerization at either end [14,37,41]. MUC2 and FIMB1 also show sequence similarity, especially with respect to positioning of cysteine residues, with von Willebrand factor which polymerizes N-terminus to N-terminus and C-terminus to C-terminus [14,42]. In contrast other mucin genes such as MUC3 (expressed in human colon) appear to have epidermal growth factor-like domains at their C-terminal ends which are not thought to be able to form interdisulphide bridges [43,44].

Pig polymeric colonic mucin contained in total 43% by freeze-dried weight serine, threonine and proline. Serine, threonine and proline are also found in high proportions in other well-characterized polymeric mucins. Human gastric mucin comprised 46% serine, threonine and proline and human cervical mucin 43% [21,45]. Pig colonic polymeric mucin contained a higher proportion of aspartate and glutamate (20%) than human cervical mucin and human gastric mucin (both 13%) [21,45].

On reduction of pig colonic mucin the percentage protein decreased to 13% (a loss of approx. one-seventh of the total protein content). Approx. 57% of proline and 45% of the hydrophobic amino acids were lost on reduction. Protein lost on reduction of mucins is now generally thought to be originally part of the mucin protein core and the result of selective proteolysis either *in vivo* or *in vitro* [46,47]. In contrast to pig colonic mucin, protein lost on reduction of other mucins is characterized by a high content of acidic amino acids as well as hydrophobic amino acids [48]. The C- and N-termini of mucin genes so far sequenced contain a high proportion of hydrophobic amino acids [14,47].

On digestion of pig colonic mucin approx. one-third of the protein content was lost. The proportion of serine remained similar to that of the polymer (15%) and the proportion of threonine increased (18%). The proportion of proline decreased by one-third to 9%. The loss of proline on digestion is not characteristic of other well-characterized mucins. In other well-characterized secretions the proportion of serine, threonine and proline increases on digestion. In human cervical mucin their proportion increases from 43% to 60% [35], and in human gastric mucin from 43% to 52% [45]. The proportion of acidic amino acids in digested pig colonic mucin (24%) was similar to that in the polymer and this is in contrast to other mucins such as human cervical mucin and human gastric mucin where a large proportion of these amino acids are lost on digestion.

The amino acid analysis of eluted pig digested mucin following Sepharose CL-4B gel filtration was markedly different from that of the digested mucin applied to the column. This eluted digested mucin contained lower proportions of threonine (7%) and proline (5%) and higher proportions of serine (26%) and aspartate and glutamate (35%) than the original digested mucin. This points to a loss on the column of material rich in threonine and proline, representing ~ 40% of the material applied ($n = 3$). This loss, apparently during chromatography, is possibly due to precipitation and points to the presence of more than one species in the original mucin fraction. Attempts to elute this mucin fraction, containing a high proportion of threonine and proline, from the column with 1% (w/v) SDS and 6 M GuHCl were unsuccessful.

The amino acid composition of pig digested colonic mucin before or after gel chromatography was very different to the amino acid composition of the two predicted glycosylated regions of MUC2 expressed in human colon and also different to the tandem repeat regions of the other human MUC genes. The tandem and base pair repeat regions of MUC2 both contain substantial proportions of threonine (60% and 49% respectively) and proline (36% and 21% respectively) and little serine (4% and 10% respectively) [14]. It is of course possible that the pig equivalent of human MUC2 has a completely different tandem repeat. The tandem repeat regions of all genes so far thought to be secreted in the human colon contain a higher total proportion of serine, threonine and proline (57–80%) than pig digested colonic mucins (42%). The proportion of threonine in the human tandem repeat regions is always equal to or greater than the proportion of serine [8]. However, porcine submaxillary mucin has a similar proportion of serine, threonine and proline in its tandem repeat region to that found in digested pig colonic mucin (42%). Both porcine submaxillary mucin and porcine gastric mucin contain a greater proportion of serine than threonine, in their tandem repeat regions. The high proportion of aspartate and glutamate found in digested pig colonic mucins is also not typical of the predicted tandem repeat region from

human mucin genes. However, a high proportion of aspartate has been reported in rat intestinal mucin gene tandem repeat region (16%) and high proportions of glutamate have been reported in the tandem repeat regions of frog integumentary mucin genes FIMA1 and FIMB1 (11% and 18% respectively) [41,49,50].

The support of Reckitt and Colman (U.K.) and the M.R.C. is gratefully acknowledged.

REFERENCES

1 Allen, A. and Hoskins L. C. (1988) in Diseases of the Colon and Rectum, pp. 65–94, Williams and Wilkins, Baltimore, MD
2 Sellers, L. A., Allen, A., Morris, E. R. and Ross-Murphy, S. B. (1988) Carbohydr. Res. **178**, 93–110
3 Snary, D., Allen, A. and Pain, R. H. (1970) Biochem. Biophys. Res. Commun. **21**, 844–851
4 Allen, A. (1981) in Physiology of the Gastrointestinal Tract, pp. 617–639, Raven Press, New York
5 Carlstedt, I., Sheehan, J. K., Corfield, A. P. and Gallagher, J. T. (1985) Essays Biochem. **20**, 40–76
6 Donald, A. S. R. (1973) Biochim. Biophys. Acta **317**, 420–436
7 Scawen, M. and Allen, A. (1977) Biochem. J. **163**, 363–368
8 Gendler, S. J. and Spicer, A. P. (1995) Annu. Rev. Physiol. **57**, 607–634
9 Gum, J. R., Byrd, J. C., Hicks, J. W., Toribara, N. W., Lamport, D. T. A. and Kim, Y. S. (1989) J. Biol. Chem. **266**, 22733–22738
10 Gum, J. R., Hicks, J. W., Swallow, D. M., Lagace, R. L., Byrd, J. C., Lamport, D. T. A., Siddiki, B. and Kim, Y. S. (1990) Biochem. Biophys. Res. Commun. **171**, 407–415
11 Gross, M. S., Guyonnet-Duperat, V., Porchet, N., Berheim, A., Aubert, J. P. and Nguyen, V. C. (1992) Ann. Genet. **35**, 21–24
12 Bobek, L., Tsai, H., Biesbrock, A. R. and Levine, M. J. (1993) J. Biol. Chem. **268**, 20563–20569
13 Audie, J. P., Janin, A., Porchet, M. C., Copin, B., Gosselin, B. and Aubert, J. P. (1993) J. Histochem. Cytochem. **41**, 1479–1485
14 Gum, J. R., Hicks, J. W., Toribara, N. W., Siddiki, B. and Kim, Y. S. (1994) J. Biol. Chem. **269**, 2440–2446
15 Eckhardt, A. E., Timpte, C. S., Abernethy, J. L., Zhao, Y. and Hill, R. L. (1991) J. Biol. Chem. **266**, 9678–9686
16 Turner, B. S., Bhaskar, K. R., Hadzopoulou-Cladras, M., Specian, R. D. and LaMont, J. T. (1995) Biochem. J. **308**, 89–96
17 Marshall, T. and Allen, A. (1978) Biochem. J. **173**, 569–578
18 Carlstedt, I. and Sheehan, J. K. (1984) Ciba Found. Symp. **109**, 109–157
19 Carlstedt, I., Herrmann, A., Karlsson, H., Sheehan, J., Fransson, L. and Hansson, G. C. (1993) J. Biol. Chem. **268**, 18771–18781
20 Hutton, D. A., Allen, A. and Pain, R. H. (1983) Biochem. Soc. Trans. **11**, 764–765
21 Carlstedt, I., Lindgren, H., Sheehan, J. K., Ulmsten, U. and Wingerup, L. (1983) Biochem. J. **211**, 12–23
22 Mantle, M. and Allen, A. (1978) Biochem. Soc. Trans. **6**, 607
23 Ayre, D., Hutton, D. A. and Pearson, J. P. (1994) Anal. Biochem. **219**, 373–375
24 Creeth, J. M. and Harding, S. E. J. (1982) Biochem. Biophys. Methods. **7**, 25–34
25 Wyatt, P. J. (1993) Anal. Chemica Acta **272**, 1–40
26 Jumel, K., Fiebrig, I. and Harding, S. E. (1996) Int. J. Biol. Macromol., in the press
27 Carlton, J. E. and Morgan, W. T. (1989) in Techniques in Protein Chemistry (Hugli, T. E., ed.), pp. 266–270, Academic Press, San Diego, CA
28 Mantle, M., Stewart, G., Zayas, G. and King, M. (1990) Biochem. J. **266**, 597–604
29 Allen, A., Pearson, J. P., Hutton, D. A., Mall, A. H., Coan, R. M. and Sellers, L. A. (1989) in Mucus and Related Topics. Symposia of the Society for Experimental Biology XLIII (Chantler, E. N. and Ratcliffe, N. A., eds.), pp. 241–248, The Company of Biologists Ltd., University of Cambridge, Cambridge
30 Herrmann, A., Lindell, G., Nordman, H. and Carlstedt, I. (1995) Biochem. Soc. Trans. **23**, 535S
31 Carlstedt, I. and Sheehan, J. K. (1984) Biochem. Soc. Trans. **12**, 615–617
32 Paulsson, M. and Heinegard, D. (1979) Biochem. J. **183**, 539–545
33 Sheehan, J. K. and Carlstedt, I. (1989) in Mucus and Related Topics. Symposia of the Society for Experimental Biology XLIII (Chantler, E. N. and Ratcliffe, N. A., eds.), pp. 289–316, The Company of Biologists Ltd., University of Cambridge, Cambridge
34 Carlstedt, I. and Sheehan, J. K. (1983) Proc. VIIth Int. Symp. Glycoconjugates, pp. 580–581, Rahms i Lund, Sweden
35 Carlstedt, I., Lindgren, H. and Sheehan, J. K. (1983) Biochem. J. **20**, 40–76
36 Fitzgerald, J. E., Green, G. G. R., Stafford, F. W., Birchall, J. P. and Pearson, J. P. (1987) Clin. Chim. Acta **169**, 281–298
37 Guyonnet Duperat, V., Audie, J. P., Debaileul, V., Laine, A., Buisine, M. P., Galiegue-Zouitina, S., Pigny, P., Degund, P., Aubert, J. P. and Porchet, N. (1995) Biochem. J. **305**, 211–219
38 Harding, S. E. (1989) Adv. Carbohydr. Biochem. **47**, 345–381
39 Sellers, L. A., Allen, A., Morris, E. R. and Ross-Murphy, S. B. (1988) Biochem. J. **239**, 147–153
40 Allen, A. (1989) in Handbook of Physiology (Forte, J. G., ed.), pp. 359–382, Am. Physiol. Soc., Oxford University Press, New York
41 Probst, J. C., Gertzen, E. M. and Hoffman, W. (1990) Biochemistry **29**, 6240–6242
42 Mayadas, T. N. and Wagner, D. D. (1989) J. Biol. Chem. **264**, 13497–13503
43 Kim, Y. S., Hicks, J. W., Ho, J. L., Swallow, D. and Gum, Jr., J. R. (1994) 3rd International Conference on Carcinoma Associated Mucins p. 15 (abstr.)
44 Hoffman, W. and Hauser, F. (1993) Trends Biochem. Sci. **18**, 239–243
45 Dekker, J., Aelmans, P. H. and Strous, G. J. (1991) Biochem. J. **277**, 423–427
46 Forstner, J. F. and Forstner, G. G. (1994) in Physiology of the Gastrointestinal Tract, 3rd edn., (Johnson, L. R., ed.), pp. 1255–1283, Raven Press, New York
47 Xu, G., Huan, L., Khatri, I. A., Wang, D., Bennick, A., Fahim, R. E. F., Forstner, G. G. and Forstner, J. F. (1992) J. Biol. Chem. **267**, 5401–5407
48 Pearson, J. P., Allen, A. and Parry, S. (1981) Biochem. J. **197**, 155–162
49 Gum, J. R., Hicks, J. W., Lagace, R. E., Byrd, J. C., Toribara, N. W., Siddiki, B., Fearner, F. J., Lamprot, D. T. A. and Kim, Y. S. (1991) J. Biol. Chem. **266**, 22733–22738
50 Hoffman, W. (1988) J. Biol. Chem. **263**, 7686–7690

Received 6 November 1995/5 February 1996; accepted 15 February 1996