Biochemical characterization of the component parts of intestinal mucin from patients with cystic fibrosis

Michèle MANTLE, Gordon G. FORSTNER and Janet F. FORSTNER*

Research Institute, The Hospital for Sick Children and Department of Biochemistry, The University of Toronto, Toronto, Ontario M5G 1X8, Canada

(Received 22 February 1984/Accepted 1 August 1984)

Previous studies have shown that human small-intestinal mucin consists of high- M_r glycoproteins and a smaller S-S-bonded protein of 118 kDa. The major antigenic determinants of the mucin were associated with the large glycoproteins, but depended for stability on intact disulphide bonds, and were destroyed by digestion with Pronase. In the present study we isolated and analysed the component parts of mucin from patients with cystic fibrosis with special attention being paid to the peptide constituents. After reduction with 0.2 M β -mercaptoethanol [5 min, 100°C in 1% SDS (sodium dodecyl sulphate)], the large glycoproteins and smaller peptide with an apparent molecular size of 118 kDa were separated by equilibrium density-gradient centrifugation in CsCl, Sepharose 4B chromatography or preparative SDS/polyacrylamide-gel electrophoresis. The large glycoproteins contained about 70% of the protein of the native mucin. Digestion with Pronase resulted in a further loss of 'naked' protein (10% of the native mucin protein) from the C-terminal end of the glycoprotein peptide core, and left behind highly glycosylated proteins comprised mainly (70 mol%) of threonine, serine and proline. The 118 kDa component, which contained about 30% of the native mucin protein, consisted mainly of aspartic acid, serine, glutamic acid and glycine (40 mol%), plus threonine, proline, alanine, valine and leucine (35 mol%). Together with the 'naked' protein segment, the 118 kDa component contained most of the cysteine residues of the native mucin. Surprisingly, the peptide also contained carbohydrate ($\leq 5\%$ of the native mucin carbohydrate but 50% by weight of the 118 kDa component), which included 9 mol% mannose, suggesting the presence of N-linked oligosaccharides. The peptide exhibited strong non-covalent interactions with the high- M_r glycoproteins and a tendency to self-aggregate in the absence of dissociating agents. Our findings therefore suggest that native mucin consists of large glycoproteins capable of forming disulphide bridges from their C-terminal 'naked' (antigenic) regions to a smaller glycopeptide having an M_r of 118 000.

Recently we have demonstrated that an antibody developed in rabbits against human intestinal mucin recognizes primarily a non-glycosylated peptide region in the mucin (Mantle *et al.*, 1984). Antigenicity depended on intact disulphide bonds to stabilize the peptide, and was lost after

prolonged Pronase digestion. The fact that antigenicity depended on the peptide moiety was somewhat surprising, since the mucin contains only 15–20% protein by weight, and one might expect that components of the oligosaccharide branches would serve as major antigens (Wood *et al.*, 1980).

In the course of this investigation we discovered that human mucin contained a putative 'link' peptide of M_r 118 000 that was liberated from polymeric mucin molecules by exhaustive thiol reduction. Boiling in SDS was not effective. A similar 'link' peptide was described previously for rat

Abbreviations used: SDS, sodium dodecyl sulphate; PMSF, phenylmethanesulphonyl fluoride; NEM, *N*-ethylmaleimide; IgA, immunoglobulin A; PAS, periodic acid/Schiff; dansyl, 5-dimethylaminonaphthalene-1-sulphonyl; CF, cystic fibrosis.

* To whom correspondence and reprint requests should be directed.

intestinal mucin (Fahim *et al.*, 1983), and similar but smaller peptides have been detected in pig gastric (Pearson *et al.*, 1981) and pig small-intestinal (Mantle *et al.*, 1981) mucins. The 118000- M_r peptide in human mucin did not appear to represent the entire complement of non-glycosylated peptide, since 8–12 h of digestion with Pronase degraded the 118000- M_r peptide, but did not abolish antigenicity. The latter was lost only after prolonged proteolysis (48–72 h). After digestion, residual glycoprotein covered a wide spectrum of M_r , as judged from column chromatography on Sepharose 2B (Mantle *et al.*, 1984).

In the present study we have analysed the composition of various parts of the mucin macromolecule(s), including the glycoprotein moiety before and after thiol reduction and Pronase treatment, and the 118000- M_r component. The latter has been shown to possess a unique composition, electrophoretic and chromatographic behaviour, and the capacity for strong non-covalent interactions. These characteristics may be important for polymeric structure and gel-forming properties of mucins *in vivo*.

Methods

Human small intestine was obtained within 8 h *post mortem* from three patients with cystic fibrosis. The mucosal scrapings were collected into 15 vol. of ice-cold 0.2M-NaCl containing 0.02% NaN_3 , 5mM- Na_2EDTA , 1mM-PMSF plus or minus 10mM-NEM. The mucin was purified from supernatants of mucosal scrapings (homogenates) by equilibrium density-gradient centrifugation in CsCl (twice) followed by gel filtration on Sepharose 2B, by the method of Mantle & Allen (1981). Protein comprised 16.6% of the mucin dry weight (average for five independent mucin samples). The purity of the mucins was established by the absence of non-mucin protein bands on SDS/polyacrylamide-gel electrophoresis, and the failure to detect DNA (Hinegardner, 1971), mannose or uronic acid by g.l.c. analysis. Immunofluorescent localization in previous work (Forstner *et al.*, 1973) established the specificity of the mucin antibody for human intestinal goblet cells. Radioimmunoassay showed that there was no cross-reactivity with human serum, secretory IgA, brush-border membranes and non-mucin-producing human tissues, such as liver, muscle and heart (Qureshi *et al.*, 1979).

Thiol reduction

Purified mucin samples (3mg/ml final concn.) were reduced by incubation with 0.2M- β -

mercaptoethanol in 0.18M-NaCl/0.02M-sodium barbital buffer (plus or minus 1% SDS), pH 8.6, for 5 min at 100°C. The resultant glycoproteins and protein were separated either by gel filtration on Sepharose 4B (2.5 cm $\times$ 200 cm column; Pharmacia Fine Chemicals, Uppsala, Sweden) eluted with 0.2M-NaCl containing 0.1% SDS, or by equilibrium density-gradient centrifugation in 60% (w/v) CsCl . In both cases, glycoprotein and protein fractions were pooled separately, dialysed and where necessary, SDS was extracted by the method of Henderson *et al.* (1979) before analysis. Reduced mucin glycoproteins contained 11.5% protein by weight (average for four independent mucin preparations).

Pronase digestion

Purified mucin or reduced glycoproteins (isolated free of protein) were digested at 37°C for 72 h in phosphate-buffered saline [0.1M- $\text{Na}_2\text{HPO}_4/\text{NaH}_2\text{PO}_4$ (pH 7.0)/0.1M-NaCl/0.02% (w/v) NaN_3] with Pronase (type XIV from *Streptomyces griseus*; Sigma Chemical Co., St. Louis, MO, U.S.A.). Enzyme was added at the start of the incubations (enzyme/mucin-protein ratio, 1:1, w/w), after 24 h (ratio 1:0.5) and at 48 h (final ratio 1:0.3). These conditions have previously been shown to result in complete digestion of non-glycosylated peptide in pig gastric and intestinal mucins (Scawen & Allen, 1977; Mantle *et al.*, 1981). Digested mucin glycoproteins were separated from the remaining enzyme and released peptides on Sephadex G-150 (Pharmacia Fine Chemicals), and eluted with 0.2M-NaCl containing 0.02% NaN_3 . The digested glycoproteins contained 9.7% protein by weight (average for three independent mucin preparations).

Preparative gel electrophoresis

To isolate the 118000- M_r peptide, purified mucins containing 1 mg of protein ($\sim$ 6 mg dry wt.) were reduced by 0.2M-mercaptoethanol for 5 min at 100°C in 1% SDS/0.1M-Tris/borate buffer, pH 8.6, and then applied to 7.5% polyacrylamide vertical slab gels. Electrophoresis was carried out for 17 h at 7 mA/gel in 0.2M-Tris/borate buffer, pH 8.6, containing 0.1% SDS. A 2 cm strip from the edge of the gel was stained with Coomassie Blue to locate the 118000- M_r protein band, which was then excised from the whole gel and electro-eluted with an Isco electrophoretic concentrator. The eluate was re-applied to 5.7% polyacrylamide gels, electrophoresed, and the 118000- M_r protein band was again eluted as described above. After an overnight dialysis against distilled water, SDS was removed from the 118000- M_r protein by ion-pair extraction (Henderson *et al.*, 1979) before analysis.

Analytical polyacrylamide-gel electrophoresis

SDS/polyacrylamide-gel electrophoresis was carried out by the method of Laemmli (1970) on a vertical slab gel system (16 cm × 20 cm; BRL, Gaithersburg, MD, U.S.A.). Samples (4–10 µg of mucin protein) were boiled in 1% SDS for 2–5 min and applied to each sample slot for electrophoresis. M_r assignments were made by comparison with standard protein mixtures (M_r range 14000–200000) obtained from Bio-Rad Laboratories (Richmond, CA, U.S.A.). Protein bands were detected by the sensitive silver-staining procedure of Merrill *et al.* (1981), which has a lower limit of sensitivity of 15–20 ng for ovalbumin standards.

Chemical analyses

Protein was determined by the method of Lowry *et al.* (1951), and for all compositional data presented in the Results section, by amino acid analysis of samples hydrolysed for 24 h at 110°C in 5.5 M-HCl using a Durrum D-500 automatic amino acid analyser. The cysteine content of samples was determined in separate experiments by hydrolysing under the same conditions but in the presence of 0.5 M-dimethyl sulphoxide and subsequent detection of cysteic acid (Spencer & Wold, 1969). Total glycoprotein was estimated in column eluates by the periodic acid/Schiff (PAS) assay described by Mantle & Allen (1978), with purified normal human mucin as a standard. Carbohydrates were quantified by g.l.c. of trifluoroacetate derivatives of mucin hydrolysates as described previously (Wesley *et al.*, 1983). *N*-Terminal analysis was carried out by using the dansyl chloride procedure described by Gray (1972), with identification of fluorescent spots on polyamide thin-layer plates (Woods & Wang, 1967). Blood group activity was measured by standard haemagglutination techniques using H lectin from *Ulex europaeus* (gorse), anti-(blood group A), anti-(blood group B) (Hyland Division, Travenol Laboratories, Glendale, CA, U.S.A.), and anti-(blood group Le) antisera (Ortho Diagnostic Systems, Rantan, NJ, U.S.A.).

Results

Previous studies showed that purified intestinal mucins from patients with cystic fibrosis are the same as control mucins with respect to amino acid profile, electrophoretic, antigenic and chromatographic behaviour. Although inter-individual variations were present, the significant statistical difference in CF samples was their enrichment in galactose, *N*-acetylglucosamine and fucose (Wesley *et al.*, 1983). Since the major focus of the present study was on mucin protein components, we felt that CF samples would provide data relevant to

Fig. 1. Sepharose 2B chromatography of purified human intestinal mucin

Mucin samples (3 mg) were applied to Sepharose 2B columns (1.5 cm × 180 cm) before (●) and after (■) reduction in 0.2 M-β-mercaptoethanol at 100°C for 5 min, and were eluted with 0.2 M-NaCl containing 0.02% NaN₃. Fractions were analysed for glycoprotein by the method of Mantle & Allen (1978). V_0 is the void volume. V_1 was at 360 ml.

both normal and CF mucins. For most of the experiments, we chose three patient samples that were in sufficient quantity to permit compositional analyses before and after various treatments. This avoided misinterpretations due to minor individual variations in mucin composition.

As was true in previous experiments using pig small-intestinal mucin (Mantle *et al.*, 1981), exhaustive thiol reduction resulted in the partial breakdown of polymeric human mucins into smaller heterogeneous glycoproteins eluted as a broad peak within the included volume of Sepharose 2B (Fig. 1). The same mucin is shown before (track 2) and after (track 3) reduction on SDS/7.5% (w/v)-polyacrylamide gels (Fig. 2). Most of the mucin barely penetrated the gel before reduction and gave only a faint smear of silver stain near the origin. After reduction, a dark band was found at a position corresponding to M_r 118000, along with a faint smear below. The 118000- M_r band was similar to a peptide of the same mobility earlier detected in rat goblet-cell mucin (Fahim *et al.*, 1983).

Important aspects of the physical and compositional features of mucin components were detected by using three different separation approaches.

(a) Equilibrium density-gradient ultracentrifugation in CsCl

Reduced mucin was mixed with CsCl (60%, w/v) and centrifuged at 105000g for 48 h at 5°C. The resulting gradient was divided into eight equal parts,

Fig. 2. SDS/polyacrylamide-gel electrophoresis of human intestinal mucin

Electrophoresis was carried out by the method of Laemmli (1970) on 7.5% (w/v)-acrylamide running gels with 5.7% stacking gels. Bands were detected by silver staining (Merril *et al.*, 1980). 1, Standard marker proteins; 2, native human mucin; 3, reduced native mucin; 4, low-density fraction (1,2,3) from CsCl gradient; 5, glycoprotein fraction (4-8) from CsCl gradient; 6, Sepharose 4B-excluded peak of reduced mucin; 7, 118000- M_r peptide isolated by preparative gel electrophoresis; 8, same as 7, but after removal of SDS; 9, Sephadex G-150-excluded peak of Pronase-digested mucin. All samples except that in gel 2 were reduced in 0.2M-mercaptoethanol (100°C, 2min) before application. Samples contained 5-7 μ g of protein. ▲ indicates artefacts caused by the use of mercaptoethanol.

each of which was dialysed before analysis. Over 95% of the reduced glycoproteins (assessed by PAS assay) formed a broad peak on the gradient (fractions 4-8; density 1.35-1.52g/ml). This peak was devoid of the 118000- M_r component when examined by SDS/polyacrylamide-gel electrophoresis (Fig. 2, track 5).

Compositional analyses are shown in Table 1 for one representative mucin (patient T.B., mucin blood group A, Le^{b+}) and its CsCl fractions after reduction. As expected, removal of the 118000- M_r component left the resulting glycoprotein fraction (Table 1, column b) enriched in total carbohydrate relative to protein. There were also small, but variable, changes in the carbohydrate profile. However, certain changes in the amino acid profile occurred consistently, including slight increases in

threonine and serine, and decreases in aspartic acid, cysteine, leucine, phenylalanine and arginine. The low-density zone of the CsCl gradient (pooled fractions 1, 2 and 3) was quite unique (Table 1, column c), containing proportionately much less threonine and proline, and more of most other amino acids, especially aspartic acid, glutamic acid, glycine, alanine, leucine and arginine. The enrichment in cysteine supports the idea that the 118000- M_r peptide forms disulphide bonds with other parts of the mucin. Interestingly, the low-density zone also contained carbohydrate (50% by weight) although the carbohydrate/protein ratio was only one-tenth of that measured in the glycoprotein fractions (Table 1, column b). Most surprisingly, the low-density fraction contained about 7 mol% of mannose, a sugar that is normally thought not to be associated with purified mucin structures, and has not been detected (except in trace amounts) in our laboratory in native mucins purified from over 20 (normal and CF) human intestinal samples. Thus isolation of the 118000- M_r component seems simultaneously to have concentrated a mannose-containing oligosaccharide or glycopeptide.

Attempts to demonstrate, in the low-density fraction, a 118000- M_r band on SDS/polyacrylamide-gel electrophoresis were unsuccessful, even when the sample was applied in high concentrations (5 μ g) and electrophoresed under reducing conditions (Fig. 2, track 4). Instead we observed silver-stained material both at the top of the gel and in a broad smear below. We suspected that the 118000- M_r glycopeptide had self-aggregated after dialysis to remove CsCl, or that it had formed strong non-covalent association with minor amounts of contaminating glycoproteins that did not enter the gel. Accordingly, a different approach was tried for the separation of mucin components.

(b) Sepharose 4B column chromatography in 0.1% SDS

Previous work with rat intestinal mucin (Fahim *et al.*, 1983) suggested that the 118000- M_r 'link' peptide could be separated from larger mucin glycoproteins after reduction by using SDS in the column eluant. Therefore reduced human mucin (patient C.A.M., mucin blood group A, Le^{b+}) was applied in 0.1% SDS to Sepharose 4B and eluted with 0.2M-NaCl in 0.1% SDS. The eluate was grouped into four fractions representing the void volume, two partially included fractions, and the totally included volume of the column. Over 70% of the PAS-positive material was eluted in the void volume (fraction 1), and this peak was free of the 118000- M_r band on SDS/polyacrylamide-gel electrophoresis (Fig. 2, track 6). The other frac-

Table 1. *Chemical analyses of native mucin and CsCl fractions after reduction*

Samples of purified native mucin (patient T.B., blood group A, Le^{b+}) were reduced in 0.2M-mercaptoethanol for 5 min at 100°C and separated in a 60% (w/v) CsCl gradient. Values in (a) are means $\pm$ s.d. Values in parentheses are the numbers of independent analyses. Abbreviations used: ND, not detected; CHO/prot., carbohydrate/protein ratio. Amino acids were quantified on a Durrum amino acid analyser after hydrolysis in 5.5M-HCl for 24 h *in vacuo*. Cysteine was estimated as cysteic acid from separate samples hydrolysed in the presence of dimethyl sulfoxide. Carbohydrate was analysed by g.l.c.

Component	Composition (mol/100 mol)		
	(a)	(b)	(c)
	Native mucin (5)	Reduced glycoprotein (fractions 4-8) (2)	Low-density fraction (fractions 1-3) (2)
Asp	5.2 $\pm$ 0.6	3.5	9.8
Thr	26.7 $\pm$ 0.7	31.3	8.4
Ser	9.2 $\pm$ 0.7	11.8	9.1
Glu	5.5 $\pm$ 0.4	7.2	10.9
Pro	14.8 $\pm$ 1.6	15.1	7.0
Gly	5.7 $\pm$ 0.2	4.7	12.4
Ala	4.3 $\pm$ 0.5	4.6	7.2
↓Cys	2.6	1.5	3.7
Val	5.4 $\pm$ 0.4	5.2	5.6
Met	0.5	0.5	0.4
Ile	4.2 $\pm$ 0.5	4.6	2.8
Leu	4.7 $\pm$ 0.5	3.9	7.0
Tyr	1.4 $\pm$ 0.2	1.0	2.0
Phe	4.5 $\pm$ 0.7	1.6	3.9
His	2.0 $\pm$ 0.6	1.7	1.5
Lys	1.6 $\pm$ 0.1	1.2	3.8
Arg	1.7 $\pm$ 0.3	0.6	4.5
Fuc	22.8 $\pm$ 1.4	23.1	20.0
Gal	22.9 $\pm$ 1.9	27.4	25.2
GlcNAc	24.6 $\pm$ 3.6	21.8	22.3
GalNAc	21.7 $\pm$ 2.1	17.6	15.8
Sialic acid	8.0 $\pm$ 2.5	10.1	9.9
Man	ND	ND	6.8
CHO/prot. (w/w) ...	6.13	10.8	1.2

tions, 2, 3, and 4, contained 18, 7 and 1% of the PAS-positive material respectively, and on SDS/polyacrylamide-gel electrophoresis all contained a 118000- M_r band, as well as material at the top of the gel. The most intense 118000- M_r band was noted in fraction 3 (not shown). Chemical analyses of fraction 1, and the corresponding unfractionated (native) mucin, are presented in Table 2. Although not identical with the reduced glycoprotein fraction from CsCl gradients (Table 1, patient T.B., column b), fraction 1 of Sepharose 4B (Table 2, column b) was very similar and showed the same difference in profile relative to its own untreated (native) mucin. That is, reduction with mercaptoethanol caused an increase in the carbohydrate/protein ratio, and similar changes in amino acid profile, consistent with the loss of less-glycosylated protein. The other 4B fractions gave compositions indicative of mixtures of glycoproteins and the 118000- M_r component, and are

therefore not presented. We judged that Sepharose 4B columns were inadequate for full separation of the 118000- M_r glycopeptide from larger glycoproteins, despite the use of 0.1% SDS in the column eluant.

(c) *Isolation of the 118000- M_r glycopeptide by preparative gel electrophoresis*

In an effort to avoid contamination by large glycoproteins, the 118000- M_r glycoprotein was purified by sequential preparative gel electrophoresis. Native mucins (from patients T.B. and C.A.M.) were reduced with 0.2M-mercaptoethanol in 1% SDS and electrophoresed on 7.5%-polyacrylamide gels. The 118000- M_r band was electroeluted and re-electrophoresed on 5.7% gels. After elution of the band and dialysis, a single 118000- M_r band was detected on SDS/7.5%-polyacrylamide-gel electrophoresis (Fig. 2, track 7). The amino acid profiles of the purified band from both

Table 2. *Chemical analyses of human intestinal mucin and its major glycoprotein component before and after Pronase digestion*
 Native mucin (patient C.A.M. blood group A, Le^{b+}) was reduced as described in Table 1 and glycoprotein recovered in fraction 1 of Sepharose 4B eluted with 0.2M-NaCl in 0.1% SDS. The same glycoprotein (column *b*) was digested for 72 h at 37°C with Pronase and harvested in the excluded volume of Sephadex G-150 (column *c*). Values are means $\pm$ s.d.. Values in parentheses are the numbers of independent analyses. Abbreviations are as for Table 1.

Component	Composition (mol/100mol)		
	(a) Native mucin (9)	(b) Reduced glycoprotein (fraction 1 of Sepharose 4B) (3)	(c) Fraction (b) after Pronase digestion (2)
Asp	5.5 $\pm$ 0.5	3.3 $\pm$ 0.2	1.6
Thr	27.2 $\pm$ 1.8	33.3 $\pm$ 2.0	44.0
Ser	8.0 $\pm$ 0.8	8.5 $\pm$ 0.3	8.5
Glu	6.2 $\pm$ 0.8	6.2 $\pm$ 0.6	4.6
Pro	13.7 $\pm$ 0.8	16.4 $\pm$ 0.5	16.8
Gly	6.0 $\pm$ 0.6	6.0 $\pm$ 0.8	4.3
Ala	4.4 $\pm$ 0.8	3.4 $\pm$ 0.7	2.8
½Cys	2.3	1.4	0.5
Val	5.3 $\pm$ 0.4	4.7 $\pm$ 0.6	4.0
Met	0.5	0.5	0.4
Ile	4.4 $\pm$ 0.7	4.4 $\pm$ 0.2	4.3
Leu	5.2 $\pm$ 0.8	4.1 $\pm$ 0.6	2.6
Tyr	1.9 $\pm$ 0.2	1.2 $\pm$ 0.4	0.8
Phe	3.3 $\pm$ 0.9	2.2 $\pm$ 0.7	2.4
His	1.6 $\pm$ 0.2	1.3 $\pm$ 0.8	0.8
Lys	2.3 $\pm$ 0.2	1.5 $\pm$ 0.1	0.6
Arg	2.2 $\pm$ 0.4	1.5 $\pm$ 0.1	1.0
Fuc	23.1 $\pm$ 2.4	20.0 $\pm$ 1.8	24.0
Gal	20.9 $\pm$ 1.1	23.7 $\pm$ 1.9	19.3
GlcNAc	18.7 $\pm$ 2.5	19.5 $\pm$ 0.8	17.5
GalNAc	25.4 $\pm$ 2.4	27.3 $\pm$ 0.5	26.5
Sialic acid	11.9 $\pm$ 2.6	9.5 $\pm$ 1.1	12.7
Man	ND	ND	ND
CHO/prot. (w/w) ...	4.43	6.17	7.2

patient samples are shown in Table 3. The two samples were identical within experimental error; both were very different from their respective native mucins or reduced glycoproteins in Tables 1 and 2 and, importantly, both were similar in profile to the low-density fraction from CsCl gradients present in Table 1, column (c). As in the latter case, the 118000- M_r glycopeptide contrasted with the large glycoproteins by having much less threonine and proline, an increase in most other amino acids, and a smaller amount of carbohydrate. No *N*-terminal amino acid was detected. Mannose was again present in significant amounts, suggesting that the 118000- M_r glycopeptide may contain at least some *N*-linked oligosaccharide chains. The carbohydrate was calculated to represent less than 5% of the total (native) mucin carbohydrate, a feature that probably explains why mannose is not detected in carbohydrate analyses of native mucins.

Our difficulties in isolating the 118000- M_r glycopeptide in 'pure' form indicated that it had a ten-

dency to self-aggregate if dissociating agents were removed. As a partial test of this idea, the purified 118 kDa component was dialysed as usual and then ion-pair-extracted to remove SDS. On re-subjection to SDS/polyacrylamide-gel electrophoresis (Fig. 2, track 8) the 118 kDa band became a broad smear, and silver-stained material also appeared at the top of the gel. This finding suggests self-aggregation, and probably explains why the glycopeptide isolated in CsCl gradients did not run 'true' on gels after removal of CsCl by dialysis.

Pronase digestion of reduced and non-reduced mucins

Examination of peptide areas in human mucin was extended by focussing on areas that do not disappear simply with reduction alone. That is, Pronase digestion removes non-glycosylated antigenic protein, which remains with mucin glycoproteins even after removal of the 118 kDa glycopeptide (Mantle *et al.*, 1984). The reduced glycoprotein fraction harvested from the excluded volume of Sepharose 4B (Table 2, column *b*) was

Table 3. Chemical analyses of the 118000- M_r 'link' (glyco) peptide of human mucin

The 118000- M_r peptide was isolated from mucin samples of two patients by two sequential preparative-gel-electrophoresis and electro-elution steps. Values represent the means for two independent analyses.

Component	Patient ...	Composition (mol/100 mol)	
		T.B.	C.A.M.
Asp		10.2	10.2
Thr		7.9	7.2
Ser		8.0	10.2
Glu		11.5	11.5
Pro		6.9	7.0
Gly		11.7	13.2
Ala		8.0	8.5
Cys		4.2	3.2
Val		5.5	4.8
Met		0.9	1.2
Ile		2.8	2.8
Leu		7.2	5.8
Tyr		2.2	2.7
Phe		3.3	2.6
His		3.1	2.1
Lys		3.8	3.7
Arg		2.8	3.3
Fuc		29.6	24.0
Gal		21.9	24.1
GlcNAc		18.9	20.8
GalNAc		14.3	14.2
Sialic acid		5.4	8.0
Man		9.9	8.9
CHO/prot. (w/w) ...		0.99	1.15

digested for 72 h with Pronase, and then separated from the enzyme and degradation products by gel filtration on Sephadex G-150. The PAS-positive material emerged in the excluded volume, although approx. 10% of the protein was lost. No bands, including the 118000- M_r glycopeptide, were detected on SDS/7.5% polyacrylamide-gel electrophoresis (Fig. 2, track 9). *N*-Terminal analysis of the reduced glycoproteins, pre- and post-Pronase treatment, produced arginine in both cases. Failure to generate new *N*-terminals suggests that Pronase hydrolysed amino acids at the C-terminus of the protein core. The composition of the remaining (digested) glycoproteins (Table 2, column c) indicates no significant change in carbohydrate profile (from column b). Major changes were seen in the amino acid profile, however, including an increase in threonine and a decrease in most other amino acids. The digested glycoprotein now consisted almost entirely of serine, threonine and proline (70.3 mol%); therefore the (antigenic) peptide region destroyed by Pronase is probably enriched in aspartic acid,

Table 4. Chemical composition of human mucin before and after Pronase digestion

Purified native mucin from patient KL (blood group H,Le^{b+}) was digested at 37°C for 72 h with Pronase (final enzyme/protein ratio 1:0.3). Glycoprotein (column b) was harvested in the excluded volume of Sephadex G-150. Abbreviation used: ND, not detected.

Component	Composition (mol/100 mol)	
	(a)	(b)
	Native mucin (6)	Pronase-digested (3)
Asp	5.4±0.4	1.7±0.3
Thr	29.4±1.6	40.5±2.0
Ser	7.7±1.3	10.4±0.4
Glu	5.9±0.3	4.3±0.7
Pro	13.8±1.2	18.2±2.5
Gly	6.4±0.3	7.3±2.7
Ala	4.1±0.4	2.1±0.6
Val	4.4±0.4	3.7±0.4
Met	0.7	0.7±0.2
Ile	4.1±0.7	4.5±0.7
Leu	5.6±1.1	2.2±0.2
Tyr	2.0±0.3	0.8±0.2
Phe	4.4±0.7	1.0±0.5
His	2.1±0.6	1.1±0.3
Lys	2.4±0.2	0.9±0.2
Arg	1.6±0.3	0.6±0.1
Fuc	27.7±1.1	30.6±2.1
Gal	29.8±0.7	29.0±1.4
GlcNAc	24.8±2.0	20.8±2.2
GalNAc	11.3±1.0	12.1±1.3
Sialic acid	6.4±1.0	7.5±1.0
Man	ND	ND

glutamic acid, glycine, alanine, leucine, tyrosine, histidine, lysine and cysteine.

Another native mucin (patient K.L., blood group H,Le^{b+}) was subjected to Pronase digestion without prior reduction, and the resultant glycoprotein (PAS-positive material) was again harvested in the excluded volume of Sephadex G-150. Chemical analyses (Table 4) show that the amino acid profile of Pronase-digested glycoprotein (column b) closely resembles that of the Pronase-digested reduced glycoprotein of patient C.A.M. (Table 2, column c). Both glycoprotein samples contained arginine as the only *N*-terminal amino acid before and after digestion. There was thus no indication that native mucin (containing the 118000- M_r glycopeptide, and stabilized by disulphide bonds) was more resistant to 72 h Pronase digestion than reduced mucin (i.e., reduced before enzymic treatment).

Discussion

The present study reveals that purified human intestinal mucins possess at least three protein

Fig. 3. Relationship between mucin components

components: a 'core' peptide of the highly glycosylated glycoprotein(s), a non-glycosylated peptide region probably located at the C-terminal end of the core peptide, and the 118000- M_r glycopeptide released by thiol reduction. The relationship between the various components are shown in Fig. 3.

Since the total mucin protein content fell from an average of 16.6 to 11.5% (by weight) after reduction and only one broad band at 118000kDa was generated, we assume that the 118000- M_r glycopeptide accounts for approx. 30% of the total mucin protein. The glycopeptide showed some very unusual characteristics, including an amino acid profile that was very different from that of the 'core' peptide of the highly glycosylated glycoproteins. There was proportionately much less serine, threonine plus proline (25 as against 60%), more cysteine, and much more polar and hydrophobic residues in the 118kDa component. These latter residues may, in part, account for the strong non-covalent interactions exhibited between the 118kDa glycopeptide and the larger mucin glycoproteins during separation. The same interactions may also be responsible for self-aggregation in the absence of denaturing agents. The amino acid profile of the glycopeptide closely resembles that of a 118000- M_r peptide in rat intestinal mucin (Fahim *et al.*, 1983), a 70000- M_r peptide in pig gastric mucin (Pearson *et al.*, 1981), and a 70000- M_r proposed 'linker' peptide of human sputum mucin (Tabachnik *et al.*, 1981), all of which are released from their native mucins by reduction.

Since the 118000- M_r glycopeptide contains a relatively high content of cysteine (4mol%), it would have the capacity to form a large number of disulphide bonds. For example, if half of the M_r of 118000 is accounted for by protein (as indicated by

the carbohydrate/protein ratios in Table 3), there would be approx. 600 amino acids per molecule, of which 24 are cysteine. Thus a maximum of 24 disulphide bonds could be formed with other mucin components or 12 intramolecular bonds. With this capacity we suggest that the 118000kDa glycopeptide plays a major linkage role in the polymerization of mucins and formation of gel structures [as suggested for gastric mucin by Allen *et al.* (1982)].

Unlike reports from other laboratories, our preparations of the 118kDa component appear to contain considerable carbohydrate with a difference in profile from sugars in the rest of the mucin, since there is a higher molar ratio of *N*-acetylglucosamine/*N*-acetylgalactosamine and significant amounts of mannose. Mannose has been detected, but only in trace amounts, in several previously described mucins, including human (Oemrawsingh & Roukema, 1974), rat (Malinowski & Herp, 1981) and monkey (Levine *et al.*, 1978) salivary mucins, and human bronchial mucin (Creeth *et al.*, 1977; Tabachnik *et al.*, 1981). There are at least two possible interpretations of the carbohydrate composition of our 118kDa component. (a) This component is actually a glycopeptide bearing both *O*- and *N*-linked oligosaccharide branches, the latter accounting for the presence of mannose and enrichment in *N*-acetylglucosamine. If this is true, one implication is that mucins (or components of them) are 'mixed' glycoproteins and their biosynthesis is regulated, in part, by the same processing reactions known to govern *N*-linked oligosaccharide formation (Schachter & Williams, 1982). It is also worth noting that if the 118kDa component is a glycopeptide containing 50% (w/w) carbohydrate (Table 3), the electrophoretic mobility is

likely to be anomalous and the present M_r assignment may be erroneously high. (b) The 118 000- M_r component may be a peptide contaminated by separate non-mucin (serum, membrane) glycopeptides. In view of the extensive purification of the mucin, this explanation seems inadequate. Since mucin glycoproteins are very heterogeneous in composition and size, however, some of the 'contamination' may be due to mucin glycoproteins themselves, with M_r values similar to that of the 118 kDa peptide. Unfortunately no *N*-terminal residues could be detected to judge the number of peptides present in the 118 kDa fraction, and thus other separation methods are now required to determine whether peptides and carbohydrates are covalently attached or belong to different molecules.

The second protein component of interest in the human mucin is a fraction that was not released by reduction but was destroyed on subsequent Pronase digestion. Reduced glycoproteins contained 11.5% protein, whereas after Pronase treatment they contained an average of 9.7% by weight. Thus the Pronase-sensitive 'naked' peptide accounts only for about 10% of the mucin protein. The lack of new *N*-terminals after Pronase treatment suggests that the peptide is located at the *C*-terminus of the core of mucin glycoproteins. In earlier studies, we showed that the antigenic determinants of mucin reside in this fraction, and that full antigenicity requires intact disulphide bonds (Mantle *et al.*, 1984). Judging from the loss of cysteine residues after Pronase treatment (Table 2, column c), there are cysteine residues in the antigenic peptide region but hardly any in the remainder of the glycoprotein core peptides. Thus disulphide bonds probably bridge the antigenic *C*-terminal peptide regions and the 118 kDa glycopeptide in intact mucin polymers. This means that the antigenic peptide as well as the 'link' glycopeptide play an important role in the formation of mucin polymers, and by inference, in mucus gel formation.

The size of the antigenic peptide could not be measured directly. However, since it represents only 10% of mucin protein, it is likely to be relatively small. In a native mucin with a probable M_r of 2×10^6 (Forstner *et al.*, 1973; Mantle *et al.*, 1981) and only 16.6% protein by weight, the antigenic peptide would have an M_r of about 33 000. The latter could be a single peptide or be distributed as a series of smaller peptides located at the ends of many glycoprotein monomeric units in native intestinal mucin (Mantle *et al.*, 1984).

The core peptide or backbone of the highly glycosylated glycoprotein regions represents the main protein component of human mucin, accounting for 60% of the protein by weight and bearing over 90% of the total carbohydrate. The

peptide backbone consists almost exclusively of serine, threonine and proline (>70%), enough therefore to hold an oligosaccharide chain on every second amino acid and in close proximity to a proline residue. The average molar ratio carbohydrate/*N*-acetylgalactosamine for three independent mucins of blood group H₁Le^{b+} was found recently in our laboratory to be 8.3, suggesting that the average chain length of the oligosaccharides is eight sugar residues. This density of carbohydrate on the core peptide of the mucin probably explains the resistance of these molecules to digestion by Pronase.

The presence of intact disulphide bonds in the mucin did not prevent destruction of the 118 000- M_r glycopeptide or the 'naked' antigenic peptide during 72 h of Pronase digestion. *In vivo* this may be important in permitting the early steps of degradation of native mucins by pancreatic proteolytic enzymes, or by bacterial and/or epithelial-cell proteinases lower in the intestinal tract. By the same token the absence of proteinases (as in cystic fibrosis with pancreatic insufficiency) would be expected to retard mucin degradation and may account for excessive mucus gel accumulation in the intestinal lumen. Increased bacterial or inflammatory-cell proteinases (as in infections or inflammatory bowel disease) would be expected to degrade mucin prematurely, thereby weakening the protective mucus blanket over the intestinal epithelium.

References

- Allen, A., Bell, A., Mantle, M. & Pearson, J. P. (1982) in *Mucin in Health and Disease* (Chantler, E. N., Elder, J. B. & Elstein, M., eds.), vol. 2, pp. 115–133, Plenum Publishing Corp., New York
- Creeth, J. M., Bhaskar, K. R., Horton, J. P., Das, I., Lopez-Vidriero, M. T. & Reid, L. (1977) *Biochem. J.* **167**, 557–569
- Fahim, R. E. F., Forstner, G. G. & Forstner, J. F. (1983) *Biochem. J.* **209**, 117–124
- Forstner, J. F., Jabbal, I. & Forstner, G. G. (1973) *Can. J. Biochem.* **51**, 1154–1166
- Gray, W. R. (1972) *Methods Enzymol.* **25**, 121–138
- Henderson, L. E., Oroszlan, S. & Konigsberg, W. (1979) *Anal. Biochem.* **93**, 153–157
- Hinegardner, R. T. (1971) *Anal. Biochem.* **39**, 197–201
- Laemmli, U. K. (1970) *Nature (London)* **227**, 680–685
- Levine, M. J., Herzberg, M. C., Ellison, S. A., Shomers, J. P. & Sadowski, G. A. (1978) *Comp. Biochem. Physiol.* **60B**, 423–431
- Lowry, O. H., Rosebrough, N. J., Farr, A. L. & Randall, R. J. (1951) *J. Biol. Chem.* **193**, 265–275
- Malinowski, C. E. & Herp, A. (1981) *Comp. Biochem. Physiol.* **69B**, 605–609
- Mantle, M. & Allen, A. (1978) *Biochem. Soc. Trans.* **6**, 607–609
- Mantle, M. & Allen, A. (1981) *Biochem. J.* **195**, 267–275

- Mantle, M., Mantle, D. & Allen, A. (1981) *Biochem. J.* **195**, 277-285
- Mantle, M., Forstner, G. G. & Forstner, J. F. (1984) *Biochem. J.* **217**, 159-167
- Merril, C. R., Goldman, D., Sedman, S. A. & Ebert, M. M. (1981) *Science* **211**, 1437-1438
- Oemrawsingh, I. & Roukema, P. A. (1974) *Arch. Oral Biol.* **19**, 753-759
- Pearson, J. P., Allen, A. & Parry, S. (1981) *Biochem. J.* **197**, 135-162
- Qureshi, R., Forstner, G. G. & Forstner, J. F. (1979) *J. Clin. Invest.* **64**, 1149-1156
- Scawen, M. & Allen, A. (1977) *Biochem. J.* **163**, 363-368
- Schachter, H. & Williams, D. (1982) in *Mucus in Health and Disease* (Chantler, E. N., Elder, J. B. & Elstein, M., eds.), vol. 2, pp. 3-28, Plenum Publishing Corp., New York
- Spencer, R. L. & Wold, S. (1969) *Anal. Biochem.* **32**, 185-190
- Tabachnik, N. F., Blackburn, P. & Cerami, A. (1981) *J. Biol. Chem.* **256**, 7161-7165
- Wesley, A. W., Forstner, J. F., Qureshi, R., Mantle, M. & Forstner, G. G. (1983) *Pediatr. Res.* **17**, 65-69
- Wood, E., Hounsell, E. S., Langhorne, J. & Feizi, T. (1980) *Biochem. J.* **187**, 711-718
- Woods, K. R. & Wang, K. T. (1967) *Biochim. Biophys. Acta* **133**, 369-370