Biochem. J. (1978) **173**, 569–578
Printed in Great Britain

The Isolation and Characterization of the High-Molecular-Weight Glycoprotein from Pig Colonic Mucus

By THOMAS MARSHALL and ADRIAN ALLEN*
Department of Physiology, The Medical School, University of Newcastle upon Tyne, Newcastle upon Tyne NE1 7RU, U.K.

(*Received* 8 *December* 1977)

1. A high-molecular-weight glycoprotein constitutes over 80% by weight of the total glycoprotein from water-soluble pig colonic mucus. 2. It was isolated free from nucleic acid and non-covalently bound protein by nuclease digestion followed by equilibrium centrifugation in a CsCl gradient. 3. The glycoprotein has the following composition by weight: fucose 10.4%; glucosamine 23.9%; galactosamine 8.3%; sialic acid 9.9%; galactose 20.8%; sulphate 3.0%; protein 13.3%; moisture about 10%. 4. The native glycoprotein has the high mol.wt. of 15×10^6. 5. Reduction of the native glycoprotein with 2-mercaptoethanol results in a glycoprotein of mol.wt. 6×10^6. 6. Pronase digestion removes 29% of the protein (3% of the glycoprotein) but none of the carbohydrate. 7. The molecular weight of the Pronase-digested glycoprotein is 1.5×10^6, which is halved to 0.76×10^6 on reduction with 2-mercaptoethanol. 8. The contribution of non-covalent interactions, disulphide bridges and the non-glycosylated peptide core to the quaternary structure of the glycoprotein are discussed and compared with the known structure of pig gastric glycoprotein.

Mucous secretions are characterized by their glycoprotein constituents (Gottschalk, 1972; Clamp, 1978). A glycoprotein of mol.wt. 2×10^6 is the principal determinant of the viscous and gel-forming properties of pig gastric mucus (Allen & Snary, 1972; Allen *et al.*, 1976). Although this pig gastric-mucous glycoprotein has been extensively characterized, relatively little is known of the macromolecular structure of mucous glycoproteins originating from other regions of the pig gastrointestinal tract.

The present paper describes the isolation, without proteolysis, and characterization of the principal glycoprotein of an aqueous extract of pig colonic mucus. The molecular weight (15×10^6) of the colonic glycoprotein is far higher than that of the corresponding pig gastric glycoprotein. However, the two glycoproteins are similar in that they both yield relatively low-molecular-weight subunits after Pronase digestion or treatment with mercaptoethanol.

Materials and Methods

Enzymes

Pronase (type VI), bovine pancreatic ribonuclease (type 1A) and bovine pancreatic deoxyribonuclease were from Sigma (London) Chemical Co., London S.W.6, U.K.

* To whom reprint requests should be addressed.

Preparation of mucus

The terminal 60cm of the colons of freshly slaughtered pigs were collected in ice, cut open and the mucosa was washed thoroughly under cold running tap water. Mucosal scrapings were taken with a blunt scalpel and collected in ice-cold 0.2% (w/v) NaN_3. For every 30 colons scraped the final volume of scrapings was adjusted to 1 litre and the NaN_3 concentration to 0.02% (w/v). After exhaustive dialysis against water the non-diffusible material was filtered through muslin to remove tissue debris and the water-soluble mucus separated from the water-insoluble mucus by centrifugation (10000*g* for 10min).

Analytical methods

Total hexose was measured by the orcinol method of Weimer & Moshin (1953). Values were corrected for interference by fucose, which gave 63% of the A_{540} value obtained with an equal weight of standard galactose. Nucleic acid was monitored by the A_{260} and quantified from the phosphate content measured by the method of Chen *et al.* (1956). Deoxyribose was measured by the diphenylamine method of Burton (1956) and corrected for sialic acid interference by the method of Croft & Lubran (1965). Sialic acid, after hydrolysis of the sample in 0.05M-H_2SO_4 at 80°C for 50min, was measured by the thiobarbituric acid method of Aminoff (1961), with *N*-acetyl-

neuraminic acid as a standard. Fucose was measured by the method of Gibbons (1955) and hexosamines by the method of Swann & Balazs (1966) with a Locarte amino acid analyser. Identification of monosaccharides by g.l.c. was kindly performed by Dr. M. J. Crumpton (Mill Hill) by the method of Chambers & Clamp (1971). Protein content was estimated by summation of amino acids after amino acid analysis by the method of Mahowald *et al.* (1962). The protein was hydrolysed for 24h in 6M-HCl, norleucine was included as an internal standard, and the analysis was corrected for the destruction of serine and threonine and incomplete release of leucine. Cysteine and cystine were measured by the method of Press *et al.* (1966) as described by Starkey *et al.* (1974). Sulphate was determined by the method of Clarke & Denborough (1971).

Equilibrium density-gradient centrifugation

Non-covalently bound protein was separated from glycoprotein by equilibrium density-gradient centrifugation in aqueous CsCl (Creeth & Denborough, 1970; Starkey *et al.*, 1974). The mucus (concentration 5–8mg/ml) was adjusted to a density of 1.43g/ml by the addition of 0.6g of CsCl/ml of solution and centrifuged at 1.5×10^5g and 5°C for 48h. Gradients were fractionated into eight equal fractions, and the fractions from the same position in each tube pooled and analysed.

Enzymic digestion

The water-soluble mucus (4–7mg/ml) was digested with proteinase-free bovine pancreatic ribonuclease and bovine pancreatic deoxyribonuclease by incubation for 17h (overnight) at 37°C in the presence of 0.01M-$MgCl_2$ and 0.02% (w/v) NaN_3. Denatured protein was removed from the digest by centrifugation at 10000*g* for 10min. The glycoprotein (2mg/ml) was digested with Pronase type VI by incubation for 48h at 37°C in 0.2M-ammonium acetate buffer, pH6.5, containing 0.02% (w/v) NaN_3. The initial enzyme:substrate ratio was 1:20 (w/w) and additional enzyme (0.1mg/ml) was added after 12, 24 and 36h of incubation.

Blood-group activity

Blood-group-A activity was measured by the method of Kabat & Bezer (1945) over a dilution range of 2.3–100μg of glycoprotein/ml by using that dilution of antiserum which just gave complete agglutination of the erythrocyte suspension. In some assays the sialic acid was removed by incubating the glycoprotein (1mg/ml) with neuraminidase (EC 3.2.1.18; 0.1mg/ml) for 1h at 37°C in the presence of 0.02% (w/v) NaN_3. The enzyme was subsequently inactivated by heating at 56°C for $\frac{1}{2}$h and the liberated sialic acid removed by dialysis, overnight, against 2 litres of 0.85% (w/v) NaCl.

Precipitation of nucleic acid

Nucleic acid in the mucus was precipitated at acid pH by (1) addition of trichloroacetic acid at 4°C to give a final concentration of 5% or 10% (w/v), or (2) dialysis of the mucus for 4 days at 4°C against one of the following buffers: 0.05M-KCl/HCl, pH1.5, 0.05M-glycine/HCl, pH3.5, or 0.2M-sodium acetate/acetic acid, pH5.5.

After precipitation at acid pH the mucus was centrifuged at 10000*g* for 10min, and the supernatants and resuspended precipitates were exhaustively dialysed and freeze-dried.

Sedimentation and diffusion coefficients

Ultracentrifuge studies were performed in a Beckman model E analytical ultracentrifuge at a glycoprotein concentration of 2–5mg/ml in one of the following buffers: (i) 0.18M-KCl/0.02M-potassium acetate/0.02% (w/v) NaN_3, pH5.5($\eta_{rel.,25}=1.017$; $\rho_{25}=1.011$g/ml); (ii) 0.18M-KCl/0.02M-sodium barbitone/0.2M-2-mercaptoethanol/0.02% (w/v) NaN_3, pH8.5 ($\eta_{rel.,25}=1.037$; $\rho_{25}=1.013$g/ml); (iii) in some experiments 6M-guanidinium chloride ($\eta_{rel.,25}=1.62$; $\rho_{25}=1.14$g/ml) was also included in the buffers (i) and (ii).

In all cases linear relations were obtained when values of $1/s_{25,w}$ were plotted against glycoprotein concentration and the $1/s^0_{25,w}$ was obtained by extrapolation to zero concentration.

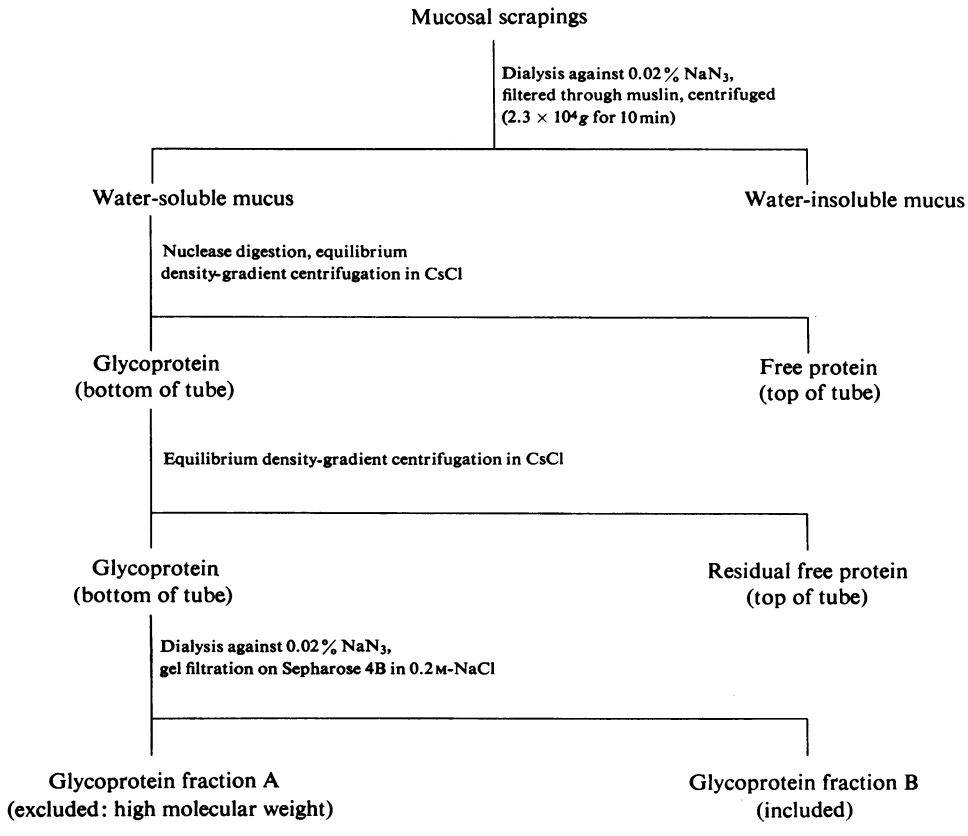
Static diffusion coefficients were calculated by the height-area method of Creeth & Pain (1967) from the spreading of boundaries formed at 25°C and 2600rev./min between solution and solvent in a double-sector synthetic-boundary cell.

Molecular weights were determined by combination of $s^0_{25,w}$ and $D^0_{25,w}$ with the Svedberg equation. The $D_{25,w}$ was not found to vary significantly with concentration and its value at a single concentration was used directly for calculating the molecular weight. The value for the partial specific volume ($\bar{v}$) used for the glycoprotein (0.64ml/g) was estimated by comparison of its protein content with that of other glycoproteins of known partial specific volume (Snary *et al.*, 1974).

Results

Purification of the high-molecular-weight glycoprotein from pig colonic mucus

The mucosal scrapings were exhaustively dialysed against water, filtered through muslin and the water-soluble mucus and water-insoluble mucus separated



Scheme 1. *Preparation of the high-molecular-weight glycoprotein A from pig colonic mucus*

by centrifugation (Scheme 1). The water-soluble mucus constituted 72% of the freeze-dried weight of the total mucosal scrapings and consisted predominantly of protein (72%) with nucleic acid (12%) and glycoprotein (about 5%, Table 1).

Equilibrium density-gradient centrifugation of the water-soluble mucus in an aqueous CsCl gradient (Starkey *et al.*, 1974) separated non-covalently bound protein ($\rho < 1.45$ g/ml, fractions 1–4, top of the tube) from the glycoprotein ($\rho > 1.50$ g/ml, fractions 7 and 8, bottom of tube, Fig. 1). Analysis of the glycoprotein fractions 7 and 8 (Table 1) showed that although they were considerably enriched in glycoprotein, they contained over 60% nucleic acid as measured by phosphate analysis; 30% of this nucleic acid was DNA. Unsuccessful attempts to separate the glycoprotein from the nucleic acid either before or after fractionation in the CsCl density gradient included: (i) gel filtration on Sepharose 4B, where it was found that glycoprotein and nucleic acid were both excluded; (ii) precipitation with cold 5% trichloroacetic acid, when substantial amounts of glycoprotein were precipitated with the nucleic acid; and (iii) further fractionation of the glycoprotein/nucleic acid fraction 8 by a second equilibrium density-gradient centrifugation in CsCl at an initial starting density of 1.59–1.70 g/ml (in this case no satisfactory separation was achieved between the lighter glycoprotein and the more dense nucleic acid); (iv) dialysis of the water-soluble mucus against 0.05 M-KCl/HCl, pH 1.5, or 0.05 M-glycine/HCl, pH 3.5, which resulted in almost complete precipitation of both glycoprotein and nucleic acid. After dialysis against 0.2 M-acetate buffer, pH 5.5, 90% of the glycoprotein in the water-soluble mucus remained in solution, and all but 11% of nucleic acid was precipitated. However, the non-precipitated glycoprotein when further purified in a CsCl density gradient still contained 12.6% (w/w) of nucleic acid.

The best method to remove nucleic acid from the glycoprotein was by enzymic digestion with nucleases free from proteolytic activity. Digestion of the crude water-soluble mucus with deoxyribonuclease and ribonuclease for 17 h at 37°C (see the Materials and

Table 1. *Chemical composition of the glycoprotein fractions from pig colonic mucus*
All phosphate was assumed to be nucleic acid. Protein content was calculated from the amino acid content. Where values are marked by an asterisk the basic amino acids arginine, lysine and histidine are not included.

	Composition (% by weight of freeze-dried glycoprotein)				
	Water-soluble mucus	CsCl-fractionated mucous glycoprotein/ nucleic acid	Glycoprotein A		
				G.l.c.	Pronase digested
Fucose	1.0	2.8	6.5	10.4	6.8
Glucosamine	1.6	7.0	25.3	23.9	27.0
Galactosamine	0.4	3.0	9.7	8.3	10.9
Sialic acid	0.6	3.4	3.2	9.9	3.4
Galactose	1.0	10.0	22.7	20.8	25.5
Mannose	—	—	—	0.6	—
Glucose	—	—	—	0.7	—
Sulphate	—	1.5	3.0	—	—
Phosphate	3.0	15.5	0.6	—	—
Nucleic acid	12.4	63.6	—	—	—
Protein	72.0*	5.7*	13.3 (11.9*)	—	8.4*

Methods section) produced a flocculent precipitate of denatured protein, which was removed by centrifugation. Exhaustive dialysis of the supernatant yielded a glycoprotein-containing solution with no detectable nucleic acid. In contrast, nuclease digestion of the glycoprotein/nucleic acid fraction 8 recovered after centrifugation in the CsCl gradient resulted in only 20–30% of the nucleic acid becoming diffusible. Why the nucleic acid was more resistant to nuclease digestion after CsCl fractionation is not known. For large-scale preparation of the colonic glycoproteins, nuclease digestion of the water-soluble mucosal scrapings was followed by removal of the non-covalently bound protein by equilibrium centrifugation in a CsCl gradient (see the Materials and Methods section and Scheme 1). When the glycoprotein fraction was re-centrifuged in a second aqueous CsCl gradient its protein content was decreased further, from 14 to 11.9% (excluding the basic amino acids) of the freeze-dried weight. Thus to ensure effective removal of the non-covalently bound protein, two successive fractionations in a CsCl density gradient were included in the standard preparation method (Scheme 1). The glycoprotein recovered from the second CsCl density gradient was exhaustively dialysed against water and fractionated by gel filtration on a Sepharose 4B column. Some 84% of the freeze-dried glycoprotein recovered after gel filtration on a Sepharose 4B column was eluted in the excluded volume of the column and designated glycoprotein fraction A; the remainder (16%), of relatively low molecular weight, was included.

Thus the principal glycoprotein constituent of the water-soluble pig colonic mucus is the high-molecular-weight glycoprotein fraction A, which was the material used for all further studies described.

Chemical analysis of the high-molecular-weight glycoprotein fraction A

Glycoprotein fraction A is a sulphated glycoprotein consisting of about 77% carbohydrate and 13.3% protein by freeze-dried weight of glycoprotein (moisture by difference was about 10%, Tables 1 and 2). Its carbohydrate constituents were identified by g.l.c. as glucosamine, galactosamine, galactose, fucose and sialic acid, together with traces of glucose and mannose. Uronic acid was not detected. Values for fucose and sialic acid content were higher when determined by g.l.c. than those obtained by standard colorimetric measurement, whereas the corresponding values for galactose and the amino sugars were in close agreement. Threonine, serine and proline were the predominant amino acids, comprising 53% of the total amino acid residues: of the remainder 10% were acidic, 23% aliphatic, 11% basic and 3% aromatic.

The glycoprotein had blood-group-A activity of 1.25 μg/0.1 ml, which was increased by treatment with neuraminidase to a value of 0.63 μg/0.1 ml.

Physical analysis of the high-molecular-weight glycoprotein fraction A

The physical parameters of glycoprotein fraction A are summarized in Table 3. On sedimentation-velocity analysis the glycoprotein fraction A gave a single peak, hypersharp at high concentrations and polydisperse at low (Fig. 2). Sedimentation was strongly concentration dependent and the $s^0_{25,w}$ was determined by extrapolation to zero concentration (Fig. 3). The $s^0_{25,w}$ values for three completely separate preparations of the colonic glycoprotein were 76.4,

92.8 and 91.6S respectively. A plot of $1/D$ against $c/2$ showed the static diffusion coefficient to be independent of concentration (c). From a combination of $s^0_{25,w}$ and $D^0_{25,w}$ in the Svedberg equation the molecular weight of the glycoprotein was estimated at 15×10^6 ($s^0_{25,w}=87$S). Sedimentation-velocity analysis of the glycoprotein in 6M-guanidinium chloride yielded a single peak of $s^0_{25,w}=52$S. However, a corresponding value of molecular weight was not determined, owing to the problem of measuring the diffusion coefficient in 6M-guanidinium chloride and the estimated length of time, 10 days, to perform an equilibrium run.

After dialysis against 0.2M-2-mercaptoethanol the glycoprotein sedimented as a single component of $s^0_{25,w}=34$S, with a mol.wt. of 6×10^6. An almost identical sedimentation coefficient, $s^0_{25,w}=33$S, was obtained for glycoprotein that had been dialysed against mercaptoethanol containing 6M-guanidinium chloride (Table 3). After reduction with mercaptoethanol and alkylation with iodoacetic acid, the *S*-carboxymethylcysteine content of the glycoprotein was estimated as 0.011 μmol/mg; this is equivalent to 166 residues of cysteine per molecule of glycoprotein.

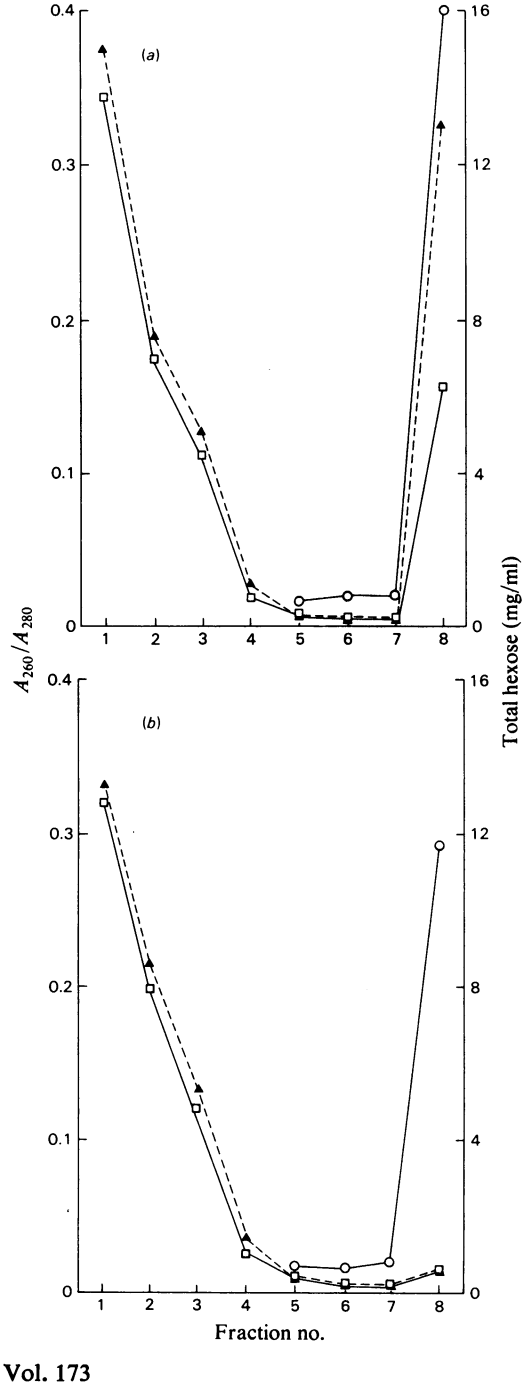


Fig. 1. *Equilibrium centrifugation of pig colonic mucus in a CsCl gradient (a) before nuclease digestion and (b) after nuclease digestion*

The starting density of the CsCl was 1.42 g/ml. Density after centrifugation was 1.38 g/ml for fraction 1 and 1.56 g/ml for fraction 8. ▲, Nucleic acid (A_{260}); □, Protein (A_{280}); ○, total hexose (orcinol method). The values for fractions 1–4 are not included because of interference from turbidity owing to high protein concentrations.

Table 2. *Amino acid composition of the high-molecular-weight glycoprotein A from pig colonic mucus*

Compositions are expressed as μmol/mg of freeze-dried material.

	Undigested glycoprotein (μmol/mg)	Digested glycoprotein (μmol/mg)	Digested glycoprotein (% loss)
Asp	0.055	0.015	74
Thr	0.278	0.272	6
Ser	0.140	0.129	11
Glu	0.061	0.017	73
Pro	0.171	0.146	18
Gly	0.063	0.037	43
Ala	0.060	0.039	37
Val	0.042	0.013	69
Ile	0.030	0.012	60
Leu	0.052	0.014	73
Tyr	0.016	Trace	—
Phe	0.021	0.006	71
His	0.048	—	—
Lys	0.030	—	—
Arg	0.040	—	—

Table 3. *Physical parameters of the high-molecular-weight glycoprotein A from pig colonic mucus*
For experimental details see the Materials and Methods section; 0.02M-potassium acetate/0.02% (w/v) NaN_3, pH5.5, was included in buffers marked * and 0.02M-sodium barbitone/0.02% (w/v) NaN_3, pH8.5, was included in buffers containing mercaptoethanol (†). Results represent the average values of two completely separate pooled preparations of glycoprotein, except for the values for the native glycoprotein in 0.18M-KCl, which are averages of the values from three preparations. The static diffusion coefficient, $D_{25,w}$, for the native glycoprotein in 0.18M-KCl was shown to be independent of concentration.

	$s^0_{25,w}$ (S)	$10^{-8} \times D_{25,w}$ ($cm^2 \cdot s^{-1}$)	$10^{-6} \times$ Mol.wt. (s^0, D^0)	Frictional ratio (f/f_0)
Native glycoprotein				
0.18M-KCl*	86.9±8.2	3.79±0.11	15.1	4.2
0.2M-Mercaptoethanol/0.18M-KCl†	33.7±1.7	3.73±0.02	6.0	5.8
6M-Guanidinium chloride/0.18M-KCl*	52.3±0.5	—	—	—
6M-Guanidinium chloride/0.2M-mercaptoethanol†	33.0±0.5	—	—	—
Pronase-digested glycoprotein				
0.18M-KCl*	20.2±0.3	8.80±0.4	1.51	4.2
0.2M-Mercaptoethanol†	13.5±0.0	11.80±0.2	0.76	3.5

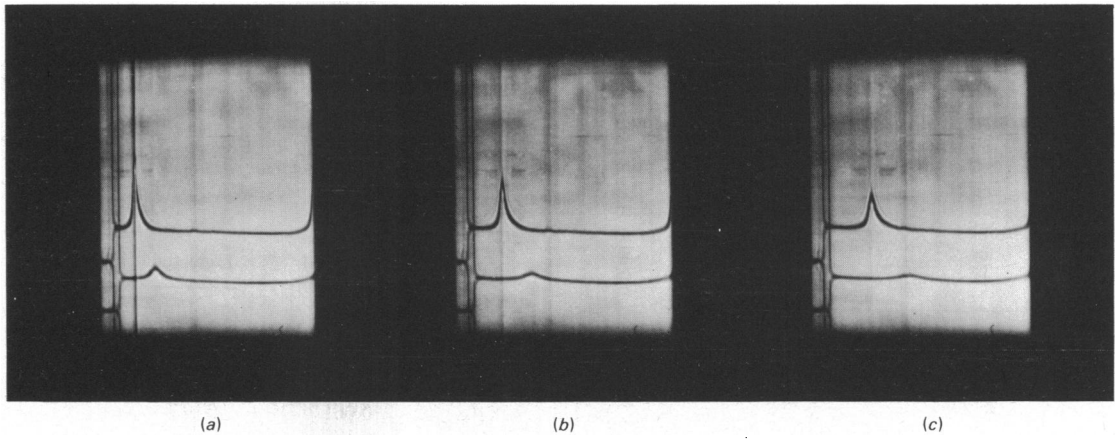


Fig. 2. *Sedimentation velocity of purified colonic glycoprotein A* ($s^0_{25,w} = 87\,S$; *mol.wt.* 15×10^6)
Schlieren optical patterns were obtained 4min (*a*), 8min (*b*) and 12min (*c*) after reaching maximum speed. Concentration of glycoprotein 4mg/ml (upper pattern) and 2mg/ml (lower pattern).

Pronase digestion of the high-molecular-weight glycoprotein fraction A

Glycoprotein A was digested with Pronase as described in the Materials and Methods section and separated from the digested peptides and added Pronase by equilibrium density-gradient centrifugation in aqueous CsCl. A comparison of the chemical composition of the glycoprotein before and after Pronase digestion is given in Tables 1 and 2. Digestion was characterized by an increase in carbohydrate content (from 67.4 to 73.6% by freeze-dried weight), indicating that carbohydrate was conserved. In contrast there was a fall in protein content from 11.9 to 8.4% (basic amino acids not measured in either case) of the freeze-dried weight. Threonine, serine and proline comprised 78% of the total amino acid residues after digestion, compared with only 60% before (Table 2), and there was a correspondingly greater loss of the other amino acids. On sedimentation-velocity analysis the digested glycoprotein (Table 3 and Fig. 3) gave a single peak of $s^0_{25,w} = 20S$, which in combination with D^0 produced a mol.wt. of 1.5×10^6. After dialysis against 0.2M-2-mercaptoethanol, sedimentation analysis of the Pronase-digested glycoprotein again gave a single peak of $s^0_{25,w} = 13.5S$ and a mol.wt. of 0.76×10^6.

Discussion

The water-soluble mucus of pig colon was isolated by centrifugation of the non-diffusible fraction of the mucosal scrapings, a method similar to that previously described for pig gastric mucus (Snary & Allen, 1971).

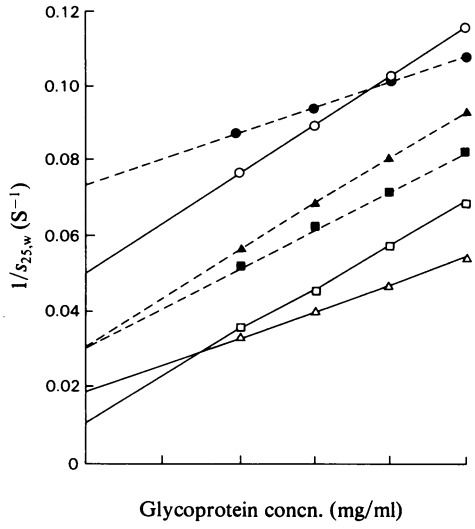


Fig. 3. *Sedimentation-velocity analysis of the purified colonic glycoprotein A after digestion with Pronase or reduction with 2-mercaptoethanol*

□, Native glycoprotein; △, glycoprotein in guanidinium chloride; ○, Pronase-digested glycoprotein; ■, glycoprotein in mercaptoethanol; ▲, glycoprotein in mercaptoethanol and guanidinium chloride; ●, Pronase-digested glycoprotein in mercaptoethanol.

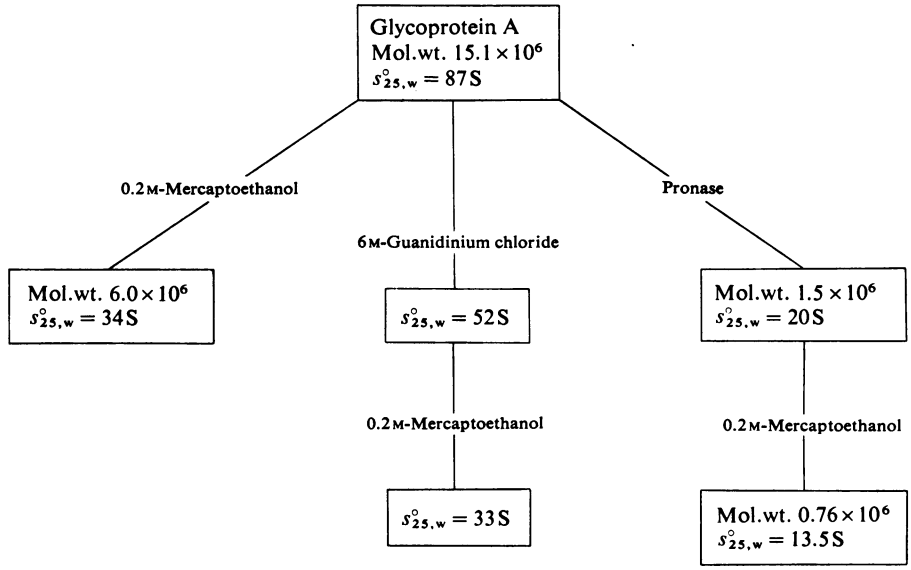
It consisted predominantly of protein (72% dry wt.) together with nucleic acid (12.4% dry wt.) and glycoprotein (about 6% dry wt.). There is no evidence to suggest that nucleic acid is a primary constituent of colonic mucus, and its most likely source is thought to be disrupted epithelial and/or bacterial cells. Pig gastric mucosa, which is visibly more resistant to scraping than colonic mucosa and possesses a relatively sparse bacterial flora (Gorbach, 1971), yields a mucus which is much richer in glycoprotein (60–70% dry wt.), contains less protein (33–48% dry wt.) and is devoid of nucleic acid (Starkey *et al.*, 1974). Nucleic acid has also been reported in other mucosal preparations, for example proteinase digests of human gastric mucosa (Hough & Jones, 1972) and sheep colonic mucosa (Kent & Marsden, 1963), as well as aqueous extracts of rat small intestine (Bella & Kim, 1972).

Non-covalently bound protein was separated from the glycoprotein by density gradient centrifugation in a CsCl gradient (Fig. 1). Unfortunately, equilibrium density-gradient centrifugation of pig colonic mucus in aqueous CsCl, although successful in removing non-covalently bound protein ($\rho < 1.4$ g/ml), yields a glycoprotein preparation ($\rho > 1.5$ g/ml) heavily contaminated with 60% (by freeze-dried weight) of nucleic acid. All subsequent attempts to separate these two components by gel filtration, nuclease digestion, trichloroacetic acid precipitation or re-fractionation in aqueous CsCl at a higher starting density proved unsuccessful. However, after nuclease digestion of the crude mucus with proteinase-free ribonuclease and deoxyribonuclease the nucleic acid was completely removable by dialysis and the glycoprotein preparation obtained could then be separated from non-covalently bound protein by equilibrium density-gradient centrifugation in aqueous CsCl (Scheme 1).

Satisfactory purification of the glycoprotein from the non-covalently bound protein was only achieved after two successive fractionations in an aqueous CsCl density gradient. This was due to the very high free protein content of the crude mucus. The high-molecular-weight glycoprotein fraction A was then separated from a low-molecular-weight glycoprotein fraction by gel filtration on Sepharose 4B.

Glycoprotein fraction A contained 85% carbohydrate (assuming 10% moisture; Table 1) with a sugar analysis typical of these glycoproteins from mucous secretions (Table 1; Gottschalk, 1972; Allen, 1978). The glycoprotein contained substantial amounts of both ester sulphate and sialic acid, two anionic groups that have been shown to occur together in a number of gastrointestinal glycoproteins, including sheep colonic-mucus glycoprotein (Kent & Marsden, 1963) and the glycoprotein purified from Pronase-digested pig colonic mucosa (Inoue & Yosizawa, 1966). Traces of glucose and mannose were detected in the purified glycoprotein fraction A by g.l.c. analysis. Although not usually associated with glycoproteins from mucus secretions, small amounts (less than 1%) of these sugars are also found in purified glycoproteins from pig gastric mucus (M. Scawen & A. Allen, unpublished work) and other mucous secretions (J. Clamp, personal communication). Such sugars may be indicative of small amounts of contaminant polysaccharides, or they might reflect heterogeneity in the sugar chains of the colonic-mucus glycoprotein molecules. The colonic glycoprotein fraction A also exhibited blood-group-A activity similar to pig gastric mucus glycoprotein (Snary & Allen, 1971), pig submaxillary-gland glycoprotein (Carlson, 1968) and rat small-intestinal glycoprotein (Bella & Kim, 1972). After treatment with neuraminidase the blood-group-A activity of the colonic-mucus glycoprotein increased, presumably owing to the exposure of additional *N*-acetylgalactosamine residues to which sialic acid was covalently attached and/or because of conformational changes in the molecule.

The amino acid composition of the colonic-mucus glycoprotein was characteristically high in threonine, serine and proline content (53% of the total amino acids present) and particularly low in the aromatic amino acids, tyrosine and phenylalanine (3% of the total amino acids present). The Pronase-digested



Scheme 2. *Summary of the physical properties of the high-molecular-weight glycoprotein A from pig colonic mucus*

colonic-mucous glycoprotein contained all the carbohydrate present in the undigested molecule (Table 1), but there was a loss of 29.4% of the protein (3.5% by freeze-dried weight of the glycoprotein). Protein loss on digestion of the colonic-mucous glycoprotein was characterized by a considerable loss of the content of all amino acids except threonine, serine and proline, which constituted 78% of the remaining amino acids present in the digested glycoprotein. These results demonstrate the presence of two types of regions in the protein core within the glycoprotein molecule. One type of peptide region rich in the amino acids serine, threonine, and proline is heavily glycosylated and resistant to proteolysis. The other regions of the peptide core, which have a less distinctive amino acid content, are free of carbohydrate and susceptible to hydrolysis by Pronase. Such glycosylated and non-glycosylated regions of the protein core have been found in pig gastric-mucous glycoprotein (Kristiansen & Porath, 1968; Starkey *et al.*, 1974; Scawen & Allen, 1977) and ovarian-cyst glycoprotein (Donald, 1973). Because of the very different method of preparation it is difficult to make a comparison between our Pronase-digested colonic glycoprotein and that of Inoue & Yosizawa (1966); qualitatively the two glycoproteins are the same, but there are quantitative differences in both the protein and carbohydrate contents.

The sedimentation coefficient ($s^{0}_{25,w} = 87 \pm 10$S) and molecular weight ($15.1 \times 10^6 \pm 1.5 \times 10^6$) of glycoprotein fraction A are far higher than any of the corresponding values previously reported for glycoproteins of gastrointestinal mucus; for example, the molecular weights of the glycoproteins from pig gastric mucus (Snary & Allen, 1971), rat intestinal mucus (Bella & Kim, 1972), rat goblet-cell mucus (Forstner *et al.*, 1973) and human gastric juice (Schrager, 1970) are all in the region of 2.0×10^6. However, it is uncertain whether the huge size of the colonic glycoprotein may be attributed to the native glycoprotein of the mucus or alternatively an aggregation product arising from non-covalent interactions between smaller glycoprotein molecules. If the latter were the case then aggregation would appear to be specific, since three completely separate large-scale preparations yielded glycoprotein that on sedimentation-velocity analysis produced a single peak of similar size and molecular weight in each case. It is also noteworthy that 6M-guanidinium chloride, an effective disaggregating agent, did not drastically decrease the sedimentation coefficient of the glycoprotein; the $s^{0}_{25,w}$ value in guanidinium chloride of 52S is still considerably higher than that for other gastrointestinal-mucous glycoproteins in this solvent, for example pig gastric glycoprotein (mol.wt. 2×10^6, $s^{0}_{25,w} = 26$S; Snary *et al.*, 1974) and rat intestinal glycoprotein (mol.wt. 2×10^6, $s_{25,w} = 12.9$S; Forstner *et al.*, 1973).

Treatment of the colonic glycoprotein fraction A with mercaptoethanol or Pronase produced subunits of lower molecular weight (Table 3; Scheme 2). In the presence of mercaptoethanol glycoprotein fraction A sedimented as a single component of

mol.wt. 6.0×10^6 (Table 3). After Pronase digestion the molecular weight of the native glycoprotein was decreased to 1.5×10^6, and in the presence of mercaptoethanol the molecular weight of the Pronase-digested glycoprotein was further decreased to 0.76×10^6. On sedimentation-velocity analysis all these subunits gave single peaks under a unimodal, although broad, envelope. Thus a series of lower-molecular-weight subunits of distinct molecular size can be obtained by reduction or proteolytic digestion of the native colonic glycoprotein A.

The action of Pronase and mercaptoethanol in producing these subunits could be due to the splitting of either inter-subunit or intra-subunit peptide and disulphide bridges. Studies with pig gastric glycoprotein have shown that it consists of four subunits of the same size linked by disulphide bridges between the non-glycosylated regions of the protein core (Snary *et al.*, 1970). Mercaptoethanol and Pronase split the native glycoprotein into subunits of mol.wt. 5×10^5, and these subunits are resistant to further attack by both mercaptoethanol and Pronase (Allen & Snary, 1972; Scawen & Allen, 1977). By analogy, the action of mercaptoethanol and Pronase on pig colonic glycoprotein would be to split covalent bonds linking glycoprotein subunits in the native molecule of 15×10^6 mol.wt. The presence of non-glycosylated regions of the protein core in the colonic glycoprotein would be consistent with such an explanation.

It is possible, however, that the action of mercaptoethanol and Pronase on the colonic glycoprotein leads to intramolecular changes that result in separation of the glycoprotein subunits previously held together in the native glycoprotein by non-covalent interactions. However, the effect of 6M-guanidinium chloride, a strong non-covalent-bond breaker, would suggest that the interactions between the subunits of the colonic glycoprotein are unusually strong non-covalent interactions or else covalent bonds. As discussed above, the $s^0_{25,w}$ of the native glycoprotein in 6M-guanidinium chloride, 52S, is very high and in keeping with a molecule of several million molecular weight. Moreover, there was no change in the $s^0_{25,w}$ values for the mercaptoethanol-reduced (6×10^6 mol.wt.) and Pronase-digested glycoprotein (1.5×10^6 mol.wt.) when transferred to 6M-guanidinium chloride. This points to the absence of changes in molecular weight and therefore absence of disruptable non-covalent linkages between the glycoprotein units making up these glycoproteins. The smallest subunit (mol.wt. 670000) of the colonic glycoprotein was obtained after mercaptoethanol treatment of the Pronase-digested glycoprotein and was relatively close in size to the subunit of the pig gastric glycoprotein (500000mol.wt.). Hill *et al.* (1977) have shown that native sheep submaxillary glycoprotein (about 600000 mol.wt.) is dissociated by 2M-NaCl into subunits, but, owing to the close proximity of the carbohydrate side chains, these non-covalent linkages in the protein core are resistant to solutions of 7M-guanidinium chloride. Pig colonic or gastric glycoproteins are different from sheep submaxillary glycoprotein, however, in that they have disulphide bridges and a non-glycosylated protein core that is accessible to proteolytic digestion, and the gastric glycoprotein does not dissociate in 2.5M-NaCl (Snary *et al.*, 1974; Allen, 1978).

It is evident that the colonic glycoprotein fraction A, differs significantly from the glycoprotein of pig gastric mucus. The colonic glycoprotein contains more sialic acid and less fucose and galactose than the gastric glycoprotein. However, the most striking difference between the glycoproteins is that the colonic glycoprotein has a mol.wt. of 15×10^6, whereas that of the gastric glycoprotein is 2×10^6. There have been no previous reports of gastrointestinal glycoproteins of such high molecular weight, however, the dithiothreitol-reduced glycoprotein from bovine cervical mucus (Meyer *et al.*, 1977) has a similar mol.wt. (5.3×10^6) to the mercaptoethanol-reduced colonic glycoprotein described above. Although the contribution of non-covalent interactions to the very-high-molecular-weight pig colonic glycoprotein remains uncertain, it is evident from our studies, that, in common with pig gastric glycoprotein and some other mucus glycoproteins (Allen, 1977, 1978), disulphide bridges and non-glycosylated regions of the protein core are an integral part of the quaternary structure of the native glycoprotein.

We thank the Science Research Council for financial support.

References

Allen, A. (1977) in *Mucus in Health and Disease* (Elstein, M. & Parke, D. V., eds.), pp. 283–299, Plenum Press, London
Allen, A. (1978) *Br. Med. Bull.* **34**, 28–33
Allen, A., Pain, R. H. & Robson, T. R. (1976) *Nature (London)* **264**, 88–89
Allen, A. & Snary, D. (1972) *Gut* **13**, 666–672
Aminoff, D. (1961) *Biochem. J.* **81**, 384–392
Bella, A. Jr. & Kim, Y. S. (1972) *Arch. Biochem. Biophys.* **150**, 679–689
Burton, K. (1956) *Biochem. J.* **62**, 315–323
Carlson, D. M. (1968) *J. Biol. Chem.* **243**, 616–626
Chambers, R. E. & Clamp, J. R. (1971) *Biochem. J.* **125**, 1009–1018
Chen, P. S., Toribara, T. Y. & Warner, H. H. (1956) *Anal. Chem.* **28**, 1756–1758
Clamp, J. (ed.) (1978) *Br. Med. Bull.* **34**, 1–94
Clarke, A. E. & Denborough, M. A. (1971) *Biochem. J.* **121**, 811–816

Creeth, J. M. & Denborough, M. A. (1970) *Biochem. J.* **117**,879–891
Creeth, J. M. & Pain, R. H. (1967) *Prog. Biophys. Mol. Biol.* **17**, 217
Croft, D. N. & Lubran, M. (1965) *Biochem. J.* **95**, 612–620
Donald, A. S. R. (1973) *Biochem. Biophys. Acta* **317**, 420–436
Forstner, J. F., Jabbal, I. & Forstner, G. G. (1973) *Can. J. Biochem.* **51**, 1154–1166
Gibbons, M. N. (1955) *Analyst (London)* **80**, 268–271
Gorbach, S. L. (1971) *Gastroenterology* **60**, 1110–1129
Gottschalk, A. (ed.) (1972) *Glycoproteins*, Elsevier, Amsterdam and New York
Hill, H. D., Reynolds, J. A. & Hill, R. L. (1977) *J. Biol. Chem.* **252**, 3791–3798
Hough, L. & Jones, J. V. S. (1972) *Carbohydr. Res.* **23**, 1–16
Inoue, S. & Yosizawa, Z. (1966) *Arch. Biochem. Biophys.* **117**, 257–265
Kabat, E. A. & Bezer, A. E. (1945) *J. Exp. Med.* **82**, 207–212
Kent, P. W. & Marsden, J. C. (1963) *Biochem. J.* **87**, 38P–39P
Kristiansen, T. & Porath, J. (1968) *Biochem. Biophys. Acta* **158**, 351–357
Mahowald, T. A., Noltmann, E. A. & Kaby, S. A. (1962) *J. Biol. Chem.* **237**, 1138–1145
Meyer, F. A., Vered, J. & Sharon, N. (1977) in *Mucus in Health and Disease* (Elstein, M. & Parke, D. V., eds.), pp. 239–249, Plenum Press, London
Press, E. M., Piggot, P. J. & Porter, R. R. (1966) *Biochem. J.* **99**, 356–366
Scawen, M. & Allen, A. (1977) *Biochem. J.* **163**, 363–368
Schrager, J. (1970) *Gut* **11**, 450–456
Snary, D. & Allen, A. (1971) *Biochem. J.* **123**, 845–853
Snary, D., Allen, A. & Pain, R. H. (1970) *Biochem. Biophys. Res. Commun.* **40**, 844–851
Snary, D., Allen, A. & Pain, R. H. (1974) *Biochem. J.* **141**, 641–646
Starkey, B. J., Snary, D. & Allen, A. (1974) *Biochem. J.* **141**, 633–639
Swann, D. A. & Balazs, E. A. (1966) *Biochim. Biophys. Acta* **130**, 112–129
Weimer, H. E. & Moshin, J. R. (1953) *Am. J. Tuberc. Pulm. Dis.* **68**, 594–602