Carbohydrate Research, 34 (1974) 115–137

FURTHER INVESTIGATIONS INTO THE STRUCTURE OF HUMAN GASTRIC MUCIN: THE STRUCTURAL CONFIGURATION OF THE OLIGOSACCHARIDE CHAINS

MALCOLM D. G. OATES, ALAN C. ROSBOTTOM, AND JACOB SCHRAGER
The Group Laboratory, The Royal Infirmary, Wigan, Lancashire (Great Britain)
(Received June 5th, 1973; accepted for publication in revised form January 7th, 1974)

ABSTRACT

Human gastric mucin from aspirates has been degraded by proteolysis to afford a soluble glycopolypeptide which, when purified by gel filtration on Bio-Gel P 150 and chromatography on Ecteola cellulose, was essentially free from mannose. Structural information on the oligosaccharide chains attached to the polypeptide core was obtained by mild hydrolysis with acid, Smith degradation, and degradation with sodium phosphate–borohydride, both before and after the removal of fucose by controlled, acid hydrolysis. A number of oligosaccharides were isolated and purified by gel and paper chromatography. Minimal degradation due to base-catalysed peeling reactions was experienced with the sodium phosphate–borohydride technique. The oligosaccharides have been characterised and structures proposed for the various side-chains present in human gastric mucin.

INTRODUCTION

Investigations into the carbohydrate side-chains of the glycoproteins of various human epithelial mucins and the so-called blood-group mucins from ovarian-cyst materials have been very largely confined to the latter materials due to their much greater abundance in single individuals, their relative purity in the native state which is devoid of a hostile environment, and their relative ease of final purification to furnish a reasonably pure glycoprotein for degradative work. Thus, a considerable body of data[1–4] has been accumulated particularly with regard to the serological properties, and a nearly complete side-chain structure for A, B, H, Le^a, and Le^b secretor groups[5] has been proposed. However, the structure of the side chains in other human epithelial mucins is largely unknown.

We[6,7] have shown that a purified, gastric glycopolypeptide may be readily isolated from gastric mucin, after proteolysis with pepsin, by chromatography on Bio-Gel P 150, followed by Ecteola cellullose, to remove sulphated glycopolypeptide. Generally, this contaminant was present only in small quantities and may have originated from saliva[8], since gastric mucin does not usually contain any ester sulphate[9] although saliva is richly endowed. Analysis of the glycopolypeptide showed

it to have a common, basic composition superimposed on which were additional sugar residues which determined the blood-group specificity[6,7]; 2-acetamido-2-deoxy-galactose, galactose, and fucose were associated with A, B, and H activity, respectively. These findings were in agreement with the results on ovarian-cyst fluids and it appears that the terminal sugars are identical in each type of mucin. The investigations reported herein were undertaken to determine the types of chains present in gastric mucin; the configurations of the serological end-sugars are now well known and are probably applicable to all human epithelial secretions[5].

Most of the results described in this paper were obtained from H secretors, due to their relative abundance and somewhat simpler breakdown-products compared to materials from A and B secretors.

The glycopolypeptide was (*1*) hydrolysed with hydrochloric acid under the mild conditions widely used for ovarian-cyst glycoproteins[10]; (*2*) subjected to Smith degradation[11–15]; (*3*) treated with alkali–borohydride in order to release the carbohydrate chains from the polypeptide core. Previous workers in the ovarian-cyst field[5,16,17], using the last method, observed considerable degradation of the saccharide chains by a peeling reaction. Similar results were obtained with the gastric glycopolypeptide. By varying pH, borohydride concentration, and temperature, conditions were found where the peeling reactions were virtually eliminated, but the β-elimination reaction involving the peptide–carbohydrate linkage proceeded at a reasonable rate.

MATERIALS AND METHODS

Neutral mono-, di-, and tri-saccharides were obtained from commercial sources. β-D-Gal-(1→4)-D-GlcNAc was kindly supplied by Professor W. T. J. Morgan, β-D-Gal-(1→3)-D-GlcNAc by Dr. A. Gauhe, β-D-Gal-(1→6)-D-GlcNAc by Professor E. A. Kabat, and reference methyl glycosides by Dr. Elizabeth Percival. De-Acidite G (HO^-) resin (7–9% crosslinking, 100–200 mesh) was obtained as a special batch of non-isoporous material from the Permutit Co., London. To remove any residual strong-base groups, the resin was suspended in water and washed (in a column) with several column volumes of 2M sodium carbonate until chloride free, then with water until a carbonate-free eluate was obtained. A slurry of the resin was then autoclaved at 120° for 16 h. The regeneration and autoclaving were repeated (usually twice) until no further chloride could be eluted from the resin. Final cleaning of the resin was achieved by boiling in methanol for 4 h, and then washing with several column volumes of methanol and finally with water.

G.l.c. — Monosaccharide analysis was performed[18] on a Pye Series 104 chromatograph, with the following minor modifications which gave a more stable column. The column packing was prepared by coating Chromosorb G (AWDMCS 80–100 mesh, Johns–Manville Co., New York) with a solution of 2% Apiezon L and 4% cyclohexanedimethanol adipate (Phase Separations Ltd., Cheshire) in benzene. Columns (5 ft × ~2 mm i.d.) were conditioned in the usual way and operated with a

carrier-gas flow of ~20 ml/min. Detector-response factors for the various monosaccharides were as previously found[18] and, in addition, response factors for the trimethylsilyl (TMS) derivatives of 2-carboethoxy-2-deoxyglucitol and 2-carboethoxy-2-deoxygalactitol were determined relative to that of galactitol.

Hydrolysis of the glycopolypeptide. — Hydrolysis was effected with 2M hydrochloric acid at 100° for 2–2.5 h for optimal release of neutral sugars, and 2.5–3M hydrochloric acid at 100° for 16 h for optimal release of amino sugars. To the cooled hydrolysate, a suitable amount of galactitol (internal standard) was added, and deacidification was effected with De-acidite G, as previously described[6,7], except that a final water-wash of 4 ml was used to remove all the sugars from the resin. The eluate was treated with 10% aqueous sodium hydrogen carbonate (0.5 ml) and 30% ethyl chloroformate in ethyl acetate (0.25 ml) at room temperature for 3 h with occasional vigorous shaking. Deionization was carried out as previously described, except that all the eluate was collected and a water wash of 3 ml was used to remove all the sugars from the column. The final eluate was evaporated and silylated as described previously. Solid injection allowed the use of large volumes of sample[18] for g.l.c.

Chromatography. — Whatman No. 1 paper was used with the solvent systems: (*1*) ethyl acetate–pyridine–water (10:4:3); (*2*) 1-butanol–pyridine–water (6:1:1); (*3*) pyridine–1-butanol–water (6:4:3). D-Galactose and lactose were used as reference compounds. For preparative paper chromatography (p.c.), a strip was cut from the edge of the paper and sprayed. The corresponding bands were excised and eluted with water.

Silica gel F_{254} (Merck) was used for t.l.c., and detection was effected with aniline–diphenylamine.

The methods for gel filtration, using Bio-Gel P 100 and P 150, have been described[6,7]. Columns of Bio-Gel P 4 (<400 mesh) were maintained at 60° by an external water-jacket. For oligosaccharides, initial separations were carried out on columns of 150 × 1.6 cm and further purifications on columns of 150 × 0.7 cm, by elution with 0.03M sodium chloride. Peak positions were visualised by using orcinol–sulphuric acid. Column calibration was carried out with a mixture of Blue Dextran (Pharmacia A.C.), D-galactose, lactose, and raffinose; the column flow was adjusted to optimise their separation.

Analytical procedures. — *N*-Acetylneuraminic acid was determined by the method of Warren[19], and ester sulphate as described previously[6].

Amino acids were determined on a Technicon automatic amino acid analyser after hydrolysis of the samples in 4M hydrochloric acid at 100° for 16 h in evacuated tubes. Norleucine was used as internal standard.

Mass spectrometry. — The oligosaccharides (1–3 mg) were permethylated by suspending them in *N*,*N*-dimethylformamide (0.2 ml), and adding sodium hydride (50 mg) followed by methyl iodide (0.1 ml) .After 1 h, the mixture was partitioned between water and chloroform, and the chloroform layer was washed with water (3 ×) and evaporated to dryness to furnish the permethylated oligosaccharide[20].

Mass spectrometry was carried out on an A.E.I. MS-9 instrument. The oligosaccharides isolated from the mild, acid hydrolysis were examined by Dr. D. W. Thomas at the Institut de Chemie des Substances Naturelles, Gif-sur-Yvette, and those from the sodium phosphate–sodium borohydride degradation by Dr. B. Blessington, University of Bradford.

Reduction with sodium borohydride. — The oligosaccharide (0.1–1 mg) was dissolved in water (2 ml), sodium borohydride (5–10 mg) was added, and the solution was allowed to stand at room temperature for 24 h. Excess borohydride was destroyed by acidification to pH 5 with M hydrochloric acid, and the solution was evaporated in a vacuum desiccator over sodium hydroxide. Borate was removed from the residue by repeated evaporation of dry methanol (1 ml) therefrom. Hydrolysis and g.l.c. were carried out on the resulting, reduced oligosaccharide in the usual manner.

Periodate oxidation. — The oligosaccharide (0.1–1 mg) was dissolved in water (2 ml). Periodate reagent (0.2 ml; 0.1M with respect to sodium acetate and sodium periodate, and adjusted to pH 5 with acetic acid) was added. After storage in the dark for 16 h at room temperature, excess periodate was destroyed with potassium iodide and 4M hydrochloric acid. Sodium thiosulphate was then added to remove iodine. Hydrolysis of the oxidised oligosaccharide and g.l.c. were carried out in the usual manner.

Smith degradations. — Periodate oxidation of the glycopolypeptide was monitored[21] by diluting aliquots (10 μl) with water (3 ml) and measuring the extinction coefficients at 225 nm. Oxidation was essentially complete after 1 h.

The solution of glycopolypeptide (40 ml, 13mM with respect to galactose) was oxidised with the periodate reagent (12 ml) in the dark for 1 h. At this stage, an aliquot (0.5 ml) was removed for monosaccharide analysis by g.l.c. The residual material was treated with sodium borohydride (0.5 g) at room temperature for 5 h. 4M hydrochloric acid was then added to pH 6, and the bulk of the borate was removed by dialysis overnight. The solution was made 0.07M with respect to hydrochloric acid, heated at 80° for 30 min, neutralised with dilute sodium hydroxide, and dried *in vacuo* over sodium hydroxide. The residue was dissolved in water (1 ml) and fractionated on a column (150 × 0.7 cm) of Bio-Gel P4 to afford an excluded peak and material in the polyol region. The materials in the excluded and retarded peaks were analysed for carbohydrate by g.l.c., and for amino acids. Polyols were identified on the same g.l.c. column used for sugar analysis but at 120°. The TMS derivatives of the polyols were prepared as for the monosaccharide derivatives.

For the second Smith-degradation, a small sample of the above product was used to determine the optimal time (16–20 h) for periodate oxidation. The remainder was then treated with periodate reagent (6 ml) for 24 h. An aliquot (0.5 ml) was then reduced ($NaBH_4$, 0.3 g), for 5 h followed by addition of hydrochloric acid to pH 6, dialysis overnight, mild hydrolysis with acid, and chromatography on Bio-Gel P4 to furnish an excluded peak and a retarded peak in the polyol region. The material in these peaks was analysed as described above.

The third Smith-degradation was carried out by using periodate reagent (3 ml)

for 24 h, followed by reduction, dialysis, hydrolysis, and analysis as described above.

The fourth and fifth Smith-degradations were carried out as for the third degradation.

First and second Smith-degradations on the glycopolypeptides degraded by mild, acid hydrolysis were carried out as described above, with appropriate amounts of reagents and a 24-h oxidation period.

Methylation. — The oligosaccharides were permethylated as described above. The methyl glycosides of the permethylated sugars were obtained by treatment with 0.5M methanolic hydrogen chloride at 80° for 20 h. The neutral methyl sugars were analysed as their methyl glycosides by g.l.c. on 6% polyphenyl ether (6 ring) on Chromosorb G (AW-DMCS) at 190°, and 4% cyclohexanedimethanol adipate with 2% Apiezon L on Chromosorb G (AW-DMCS) at 170°.

Retention times relative to that of methyl 2,3,4,6-tetra-*O*-methyl-α-D-glucoside were determined.

EXPERIMENTAL AND RESULTS

The pepsinised glycoprotein. — A batch (2 l) of H-secretor gastric secretion was subjected[6,7] to autodigestion and centrifugation at 20,000 r.p.m. for 1 h, followed by chromatography on Bio-Gel P150 and Ecteola cellulose. The resulting solution of purified glycopolypeptide was clear, colourless, and odourless. A typical analysis is shown in Table I.

Four monosaccharides were present: galactose, fucose, 2-acetamido-2-deoxygalactose, and 2-acetamido-2-deoxyglucose. Ester sulphate, *N*-acetylneuraminic acid, glucose, and mannose were present only in trace amounts. The sum of threonine and serine was equimolar with the 2-acetamido-2-deoxygalactose.

Mild, acid hydrolysis. — The solution (20 ml) of glycopolypeptide from a 2-litre batch of gastric secretion was hydrolysed by making it 0.4M with respect to hydrochloric acid and then keeping it at 65° for 16 h. The solution was neutralised (sodium carbonate) and concentrated to a small volume (2–3 ml) before fractionation on a column (150 × 1.6 cm) of Bio-Gel P4 by elution with 0.03M sodium chloride. Seven fractions were collected, the g.l.c. analyses of which are shown in Table II. The positions of the oligosaccharides of lower molecular weight were identified by comparison with reference compounds. Examination of Table II shows that the majority of the fucose was eliminated as the monosaccharide, together with much smaller amounts of galactose and very little 2-acetamido-2-deoxyglucose, whereas 2-acetamido-2-deoxygalactose remained almost entirely in the material in the excluded peak.

The oligosaccharide fractions. — Fraction VII contained the materials of lowest molecular weight; its position corresponded to that of a monosaccharide. It contained 90, 26, and 17% of the total recovered fucose, galactose, and 2-acetamido-2-deoxyglucose, respectively. As expected, all the fucose and 15% of the galactose were present as free monosaccharides.

TABLE I

COMPOSITION OF THE PURIFIED GLYCOPOLYPEPTIDE (*A*) AND THE MATERIAL IN THE PEAK EXCLUDED FROM BIO-GEL P4 AFTER MILD, ACID HYDROLYSIS (*B*)

Component	A		B	
	Conc. (mM)	*Molar ratio*[a]	*Conc.* (mM)	*Molar ratio*[a]
Aspartic acid	0.37	0.6	0.34	0.5
Threonine	6.39	10.0	6.39	10.0
Serine	3.20	5.0	3.15	4.9
Glutamic acid	1.17	1.8	1.20	1.8
Glycine	1.43	2.2	1.35	2.1
Alanine	2.68	4.1	2.40	3.9
Valine	0.72	1.1	0.76	1.2
Isoleucine	0.33	0.5	0.33	0.5
Leucine	0.90	1.4	0.91	1.4
Tyrosine	0.13	0.2	0.12	0.2
β-Phenylalanine	0.58	1.0	0.52	0.9
Lysine	0.41	0.6	0.40	0.6
Histidine	0.46	0.7	0.43	0.6
Arginine	0.45	0.7	0.43	0.6
Proline	3.16	4.9	3.05	4.7
Galactose	35.60	55.80	16.00	25.00
Fucose	24.50	38.30	0.80	1.25
2-Acetamido-2-deoxyglucose	29.50	46.20	10.70	16.80
2-Acetamido-2-deoxy-galactose	10.70	16.60	9.26	14.20
N-Acetylneuraminic acid	0.10	—	—	—
Mannose	0.20	—	—	—
Glucose	0.07	—	—	—
Ester sulphate	0.02	—	—	—

[a]Threonine = 10.

The presence of these sugars and also a small amount of 2-acetamido-2-deoxyglucose was confirmed by p.c. (solvents *1* and *2*). A slower moving spot, which corresponded to a disaccharide containing galactose and 2-acetamido-2-deoxyglucose, was also observed, which reflected the disparity between the analysis of the unhydrolysed and hydrolysed material.

Fraction VI corresponded to a disaccharide, although monosaccharides (particularly galactose) were also present (Table II). Purification was effected by p.c. to give a product (10 mg) having R_{GAL} 0.62 and 0.33 in solvents *1* and *2*, respectively. Analysis indicated equimolar amounts of galactose and 2-acetamido-2-deoxyglucose (Table III). Reduction (sodium borohydride) followed by hydrolysis gave 2-acetamido-2-deoxyglucitol. Periodate oxidation of the disaccharide resulted in complete oxidation of both sugars, indicating that neither C-3 nor C-4 of the amino sugar was substituted.

On comparison (t.l.c.) with β-D-Gal-(1→3)-D-GlcNAc, β-D-Gal-(1→4)-D-GlcNAc, and β-D-Gal-(1→6)-D-GlcNAc, the disaccharide was shown to be (1→6)-linked. The mass spectrum of the permethylated disaccharide contained peaks at

TABLE II

G.L.C. ANALYSIS OF THE FRACTIONS OBTAINED FROM BIO-GEL P4 AFTER MILD, ACID HYDROLYSIS OF THE GLYCOPOLYPEPTIDE.[a]

	GalNAc (mM)	*GlcNAc* (mM)	*Galactose* (mM)	*Fucose* (mM)
Fraction I (Excluded peak)	9.26	10.70	16.0	0.80
Fraction II	0.34	1.20	1.24	0.11
Fraction III	0	1.00	1.62	0.13
Fraction IV (Tetrasaccharide zone)	0	0.92	0.93	0.18
Fraction V (Trisaccharide zone)	0	0.85	1.03	0.21
Fraction VI (Disaccharide zone)	0	2.90	3.05	0.13
Free monosaccharides	0	0.16	0.75	0.08
Fraction VII (Monosaccharide zone)	0	3.80	8.23	18.40
Free monosaccharides	0	0.33	5.33	18.74
Total	9.60	22.46	32.18	20.78

[a]Hydrolytic conditions are described in the text.

TABLE III

ANALYSIS OF THE LOWER OLIGOSACCHARIDES ISOLATED AFTER MILD, ACID HYDROLYSIS OF THE GLYCOPOLYPEPTIDE

Oligosaccharide[a]	R_{GAL}	*Solvent*	*Galactose*		*GlcNAc*	
			(mM)	*Molar ratio*	(mM)	*Molar ratio*
Fraction IV (Tetrasaccharide)	0.22	*3*	0.37	1	0.33	0.89
After $NaIO_4$			0.18	0.48	0.35	0.95
After $NaBH_4$			0.37	1	0.16	0.44
Fraction V (Trisaccharide)	0.34	*3*	0.60	1	0.30	0.50
After $NaIO_4$			0	—	0.31	0.51
After $NaBH_4$			0.31	0.51	0.28	0.47
Fraction VI (Disaccharide)	0.62	*1*	0.40	1	0.38	0.95
	0.33	*2*				
After $NaIO_4$			0	—	0	—
After $NaBH_4$			0.42	1	0	—

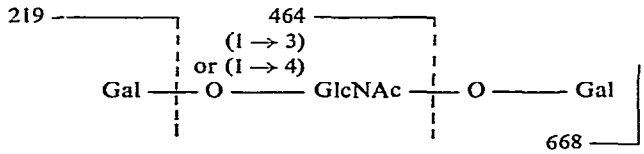
[a]Reaction conditions are described in the text.

m/e 129 and 142, indicating[22] that C-3 and C-4 were substituted by methoxyl groups. Peaks at *m/e* 88 and 101 verified the presence of methoxyl groups at C-2, C-3, and C-4 of the galactosyl residue. An intense peak at *m/e* 219 was considered to be due to the presence of an oxonium ion originating from the loss of the C-1 substituent. Peaks at *m/e* 187 (219−32) and 155 (187−32) were formed by loss of one or two molecules

of methanol; similarly, peaks at 260 (291 − MeO) and 246 (291 − CH_2OMe) arose from the 2-acetamido-2-deoxyglucose residue. Thus, substitution at C-1 of the galactosyl residue and C-6 of 2-acetamido-2-deoxyglucose was confirmed. A comparison of this disaccharide with that obtained from Fraction VII showed them to be identical.

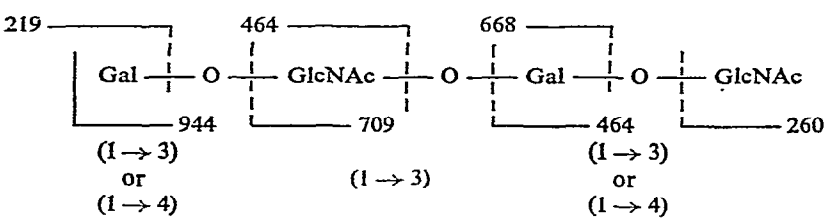
Fraction V corresponded to a trisaccharide. Purification by p.c. gave a major product (2 mg) having R_{GAL} 0.34 (solvent *1*). A small amount of β-D-Gal-(1→6)-D-GlcNAc was also isolated. Analysis of the major component revealed a galactose–2-acetamido-2-deoxyglucose ratio of 2:1. Reduction with sodium borohydride resulted in the loss of 1 mol. of galactose, and the formation of 1 mol. of galactitol. Periodate oxidation eliminated the galactose completely but not the 2-acetamido-2-deoxyglucose.

Mass spectrometry of the permethylated derivative produced intense peaks at *m/e* 219, 464, and 668. The evidence was consistent with a trisaccharide of the following sequence:



The low yield of this trisaccharide prevented further investigations.

Fraction IV corresponded to a tetrasaccharide. Purification by p.c. (solvent *3*) gave a major product (4 mg) having R_{GAL} 0.22. Analysis indicated equimolar amounts of galactose and 2-acetamido-2-deoxyglucose (Table III). Reduction with sodium borohydride resulted in the loss of half the 2-acetamido-2-deoxyglucose, and the formation of an equimolar amount of 2-acetamido-2-deoxyglucitol. Periodate oxidation resulted in the loss of half of the galactose. These results are indicative of an unbranched tetrasaccharide having one 2-acetamido-2-deoxyglucose residue at the reducing end, and a further one linked through C-3 or C-4. One galactose residue must be linked at C-3 in order to survive periodate oxidation. Further information on the sequence was obtained by mass spectrometry on the permethylated tetrasaccharide Peaks at *m/e* 219, 260, 464, 668, 709, 913, and 944 were consistent with the following structure:



Fractions II and III were not amenable to further purification by p.c. Purification was achieved on small, high-efficiency columns of Bio-Gel P4 and resulted in three distinct fractions (Table IV). The fraction of lowest molecular weight contained 2-acetamido-2-deoxyglucose and galactose in the ratio of 2:3. Reduction with sodium borohydride resulted only in the loss of one third of the galactose. Periodate oxidised two-thirds of the galactose only, indicating one galactose residue to be substituted at C-3, and two 2-acetamido-2-deoxyglucose residues at C-3 or C-4.

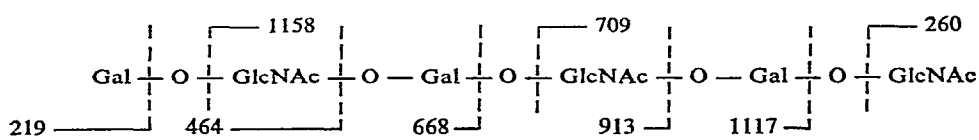
TABLE IV

ANALYSIS OF THE FRACTIONS OBTAINED BY RECHROMATOGRAPHY OF FRACTIONS II AND III ON BIO-GEL P4

Oligosaccharide	*Galactose*		*GlcNAc*	
	mM	*Molar ratio*	*mM*	*Molar ratio*
Fraction 1 (Pentasaccharide)	1.20	3	0.80	2
After $NaIO_4$	0.44	1	0.76	2
After $NaBH_4$	0.87	2	0.84	2
Fraction 2 (Hexasaccharide)	1.87	3	1.67	3
After $NaIO_4$	1.25	2	1.74	3
After $NaBH_4$	1.80	3	1.15	2
Fraction 3 (Heptasaccharide)	1.04	4	0.71	3
After $NaIO_4$	0.80	3	0.75	3
After $NaBH_4$	0.82	3	0.74	3

The mass spectrum of the permethylated pentasaccharide contained peaks at *m/e* 219, 464, 668, and 913 consistent with the structure Gal-[1→3(4)]-GlcNAc-(1→3)-Gal-[1→3(4)]-GlcNAc-(1→?)-Gal.

The fraction of intermediate molecular weight was obtained in the largest quantity (2 mg), and analysis indicated that it contained equimolar amounts of 2-acetamido-2-deoxyglucose and galactose. Borohydride reduced one third of the 2-acetamido-2-deoxyglucose, and periodate oxidised one third of the galactose, indicating two 2-acetamido-2-deoxyglucose residues to be substituted at C-3 or C-4, and two galactose residues at C-3. The mass spectrum of the permethylated hexasaccharide contained peaks at *m/e* 219, 260, 464, 668, 709, 913, 1117, and 1158, which are consistent with the following structure:



The fraction of highest molecular weight had a galactose–2-acetamido-2-deoxyglucose ratio of 4:3. Periodate oxidised, and borohydride reduced, one-fifth of the galactose residues, indicating a heptasaccharide structure. The mass spectrum of the

permethylated derivative was uninformative, but peaks at *m/e* 709 and 260 suggested an oligosaccharide structure containing an even number of units and terminating in 2-acetamido-2-deoxyglucose and galactose. Further investigation was precluded because of lack of material.

The excluded peak. — Reference to Table II shows that over 90% of the original amino acids were recovered in the material excluded during chromatography on Bio-Gel P4, indicating that little degradation of the polypeptide core had occurred. There was concomitant recovery of 90% of the 2-acetamido-2-deoxygalactose. Approximately half of the original galactose and 2-acetamido-2-deoxyglucose were also recovered in the material in the excluded peak.

Further mild hydrolysis with acid. — In an attempt to obtain further information on the more-resistant carbohydrate chains, a second acid hydrolysis, using conditions identical to those of the first, was carried out on the material in the excluded peak. The hydrolysate was worked up and fractionated as before to afford seven fractions. (Table V). Considerable degradation occurred, resulting in the appearance of large amounts of the four monosaccharides, whereas the amounts of oligosaccharide material were too small to permit characterisation. A considerable quantity of glycopolypeptide (*X*) still remained in the excluded peak, although some loss of the polypeptide core and 50% of the 2-acetamido-2-deoxygalactose had occurred. However, it was noteworthy that only 10% of the original galactose and 2-acetamido-2-deoxyglucose remained. The glycopolypeptide recovered had ratios of 2-amino-2-deoxygalactose–2-amino-2-deoxyglucose–galactose of approximately 3:2:4.

TABLE V

G.L.C. ANALYSIS OF THE FRACTIONS OBTAINED FROM BIO-GEL P4 AFTER THE SECOND MILD, ACID HYDROLYSIS OF THE GLYCOPOLYPEPTIDE[a]

	GalNAc (mM)	*GlcNAc* (mM)	*Galactose* (mM)	*Fucose* (mM)
Fraction I (Excluded peak)	2.43	1.80	3.30	0
Fraction II	0.30	0.18	0.37	0
Fraction III	0.11	0.13	0.21	0
Fraction IV	0.60	0.12	0.18	0
Fraction V	0.50	0.17	0.20	0
Fraction VI	0.14	0.93	0.61	0
Free monosaccharides	0	0.33	0	0
Fraction VII	0.80	3.80	5.81	0.48
Free monosaccharides	0.32	1.41	3.80	0.48

[a]Hydrolytic conditions are described in the text.

Degradation of the glycopolypeptide by β-elimination. — The glycopolypeptide was treated with 0.1M sodium hydroxide at 65° for 48 h in the presence of excess sodium borohydride (3%). The reaction mixture was then acidified to pH 6 with hydrochloric acid and evaporated, and the residue was freed from borate by repeated distillation of dry methanol therefrom, and analysed for sugars and amino acids. The

results for several glycopolypeptides (Table VI) show that most of the 2-amino-2-deoxygalactose disappears with the concomitant appearance of 2-amino-2-deoxygalactitol (up to 85% but not reproducible); only small losses of the other sugars occurred. All the amino acids were recovered (with the exception of threonine and serine), α-aminobutyric acid (derived from threonine) appeared, and an increase in alanine was noted. The amounts recovered were variable and always less than the amounts of threonine and serine destroyed, suggesting that the reduction with sodium borohydride was incomplete.

TABLE VI

CHEMICAL ANALYSIS[a] OF THREE H-SECRETOR GLYCOPOLYPEPTIDES BEFORE (*A*) AND AFTER (*B*) TREATMENT WITH ALKALINE BOROHYDRIDE

	1		*2*		*3*	
	A	B	A	B	A	B
Aspartic acid	0.06	0.06	0.04	0.04	0.04	0.04
Threonine	0.70	0.06	0.50	0.06	0.54	0.06
Serine	0.37	0.07	0.27	0.06	0.30	0.05
Glutamic acid	0.11	0.12	0.07	0.07	0.08	0.07
Glycine	0.15	0.16	0.13	0.14	0.13	0.13
Alanine	0.20	0.43	0.19	0.43	0.21	0.35
Valine	0.08	0.08	0.07	0.07	0.07	0.06
Isoleucine	0.02	0.02	0.03	0.03	0.04	0.04
Leucine	0.07	0.07	0.06	0.06	0.07	0.06
Tyrosine	Trace	Trace	0.03	0.03	0.02	0.03
β-Phenylalanine	Trace	Trace	0.04	0.03	0.04	0.03
Lysine	Trace	Trace	0.04	0.03	0.05	0.04
Histidine	Trace	Trace	0.04	0.04	0.05	0.04
Arginine	Trace	Trace	0.05	0.05	0.06	0.06
Proline	0.36	0.35	0.29	0.30	0.32	0.30
α-Aminobutyric acid	0	0.38	0	0.26	0	0.26
Galactose	3.90	3.78	2.62	2.28	3.05	2.87
Fucose	2.84	2.53	2.11	2.06	2.37	2.30
2-Acetamido-2-deoxyglucose	2.93	2.90	2.11	2.04	2.25	2.16
2-Acetamido-2-deoxygalactose	1.05	0.13	0.67	0.03	0.78	0.032
2-Acetamido-2-deoxygalactitol	0	0.80	0	0.63	0	0.65

[a]All values are expressed as mmoles/l.

Smith degradation of the glycopolypeptide. — In the first degradation, the initial rate of periodate oxidation was very rapid, and the oxidation was essentially complete within an hour. During this period, all the fucose was destroyed and 50% of the galactose. The amino sugars and the more-abundant amino acids of the polypeptide core were not oxidised. On completion of the degradation, glycerol was the only polyol formed in appreciable amounts; the sugar values remained the same as after the initial oxidation.

In the second degradation, the rate of periodate oxidation was somewhat lower than in the first and was complete after about 16 h; the end of the reaction was

difficult to determine and differentiate from the over-oxidation. Consequently, oxidation was allowed to proceed for one day. After this time, the galactose remained unoxidised, but ~30% of the 2-amino-2-deoxygalactose and 50% of the 2-amino-2-deoxyglucose had been oxidised. On completion of the degradation, these values were unchanged and glycerol was the only polyol liberated.

A similar rate of oxidation occurred in the third as in the first degradation, and ~33% of the galactose only was destroyed. On completion of the degradation, glycerol was the only polyol liberated.

In the fourth degradation, very little oxidation occurred, ~20% of the 2-amino-2-deoxygalactose was destroyed, and a small amount of glycerol was released. Attempts to further oxidise the polymeric material recovered at this stage were unsuccessful; the final ratios of sugars remaining unoxidised were 2-amino-2-deoxygalactose–2-amino-2-deoxyglucose–galactose 1:3:3.

A typical Smith-degradation result is shown in Table VII.

TABLE VII

SMITH-DEGRADATION DATA (MOL.) OF THE H-SECRETOR GASTRIC GLYCOPOLYPEPTIDE (*A*) AND OF THE NON-RETARDED GLYCOPOLYPEPTIDE (*B*) RECOVERED FROM BIO-GEL P4 AFTER TWO HYDROLYTIC DEGRADATIONS[a]

	Threonine	*Serine*	*GalN*	*GlcN*	*Galactose*	*Fucose*
Glycopolypeptide (*A*)	2	1	3.3	9.4	12.8	8.8
First cycle	2	1	3.2	9.5	6.3	—
Second cycle	2	1	2.2	5.4	6.7	—
Third cycle	2	1	2.1	4.9	4.6	—
Fourth cycle	2	1	1.7	4.9	4.7	—
Fifth cycle	2	1	1.6	5.0	4.9	—
Glycopolypeptide (*B*)	2	1	2.4	1.8	3.3	—
First cycle	2	1	2.0	1.8	0.3	—
Second cycle	2	1	Trace	Trace	Trace	—

[a]Hydrolytic conditions are described in the text.

Smith degradation of the glycopolypeptide after partial degradation with acid. — Smith degradation of the product *X* described above resulted in the loss of 90% of galactose and 10% of the 2-amino-2-deoxygalactose. No loss of amino acids or 2-amino-2-deoxyglucose was observed. The net result of the oxidation was a glycopolypeptide (*Y*) containing equal amounts of 2-amino-2-deoxygalactose and 2-amino-2-deoxyglucose, together with a small amount of residual galactose.

When *Y* was treated with sodium hydroxide and sodium borohydride as described above, 65% of the 2-amino-2-deoxygalactose was destroyed, yielding an almost equivalent amount of 2-amino-2-deoxygalactitol. There was no appreciable loss of 2-amino-2-deoxyglucose, and the amino acids remained unchanged except for losses of threonine and serine, and concomitant appearance of rather less than equivalent amounts of α-aminobutyric acid and alanine. Chromatography of the reaction mixture on Bio-Gel P4 indicated the presence of a disaccharide but the

material was not characterised. When *Y* was subjected to a second Smith-degradation, a polypeptide virtually devoid of sugars was obtained (Table VII).

Fucose-free glycopolypeptide. — (*a*) *Preparation.* This material was obtained as an excluded peak after chromatography on Bio-Gel P4 of the product obtained after mild, acid hydrolysis (0.1M HCl, 6 h, 80°) of a 2-litre batch of H-secretor gastric glycopolypeptide. The analytical data are shown in Table VIII. Small losses of sugars other than fucose also occurred, and it was possible to characterise β-D-Gal-(1→6)-D-GlcNAc in addition to galactose.

TABLE VIII

ANALYSIS OF THE H-SECRETOR GLYCOPOLYPEPTIDE BEFORE (*A*) AND AFTER (*B*) MILD, ACID HYDROLYSIS TO REMOVE FUCOSE, AND SUBSEQUENT FRACTIONATION ON BIO-GEL P4

	GalNAc (mM)	*GlcNAc* (mM)	*Galactose* (mM)	*Fucose* (mM)
Glycopolypeptide (*A*)	4.70	13.80	16.70	11.31
Glycopolypeptide (*B*) (Excluded peak)	4.70	9.40	9.50	0.78
Disaccharide zone	0	2.04	2.82	0.46
Free monosaccharides	0	0	0	0
Monosaccharide zone	0	0.74	3.10	10.20
Free monosaccharides	0	0	2.80	13.40

(*b*) *Degradation of the glycopolypeptide with sodium phosphate–sodium borohydride.* Treatment of glycoproteins with sodium hydroxide and sodium borohydride results[5,16,17] in base-catalysed peeling reactions of the released oligosaccharides, thereby affording saccharide chains in various stages of degradation. Two competing reactions occur after the base-catalysed β-elimination reaction: (*1*) The reduction of 2-acetamido-2-deoxygalactose and of the two unsaturated amino acids (to α-aminobutyric acid and alanine); (*2*) alkali-catalysed degradation of reducing 2-acetamido-2-deoxygalactose residues and of the polypeptide core, particularly in the vicinity of the unsaturated linkages. Preliminary experiments showed that reaction *2* was greatly pH-dependent, as was the β-elimination reaction which commenced at pH ~10. At pH 11.7, the rate of reaction *2* was greatly decreased, whereas β-elimination occurred at a reasonable rate. The rate of reaction *1* was increased by using a large excess of sodium borohydride at an elevated temperature.

The glycopolypeptide was made 0.1M with respect to trisodium phosphate (pH 11.7), and three aliquots of sodium borohydride (0.5M each) were added during 7 days at 37°. The mixture was then brought to pH 6–7 with dilute hydrochloric acid, evaporated to 2 ml, and fractionated on a column (150 × 1.6 cm) of Bio-Gel P4. Seven fractions were obtained, the lower ones of which were identified by comparison with known oligosaccharides (Fig. 1). Analysis of the seven fractions and the excluded peak showed that full recovery of 2-amino-2-deoxyglucose, galactose, and fucose was obtained, and that the loss of 2-amino-2-deoxygalactose was compensated for by the appearance of an equivalent amount of 2-amino-2-deoxygalactitol, indicating that little or no "peeling" reactions had occurred (Table IX).

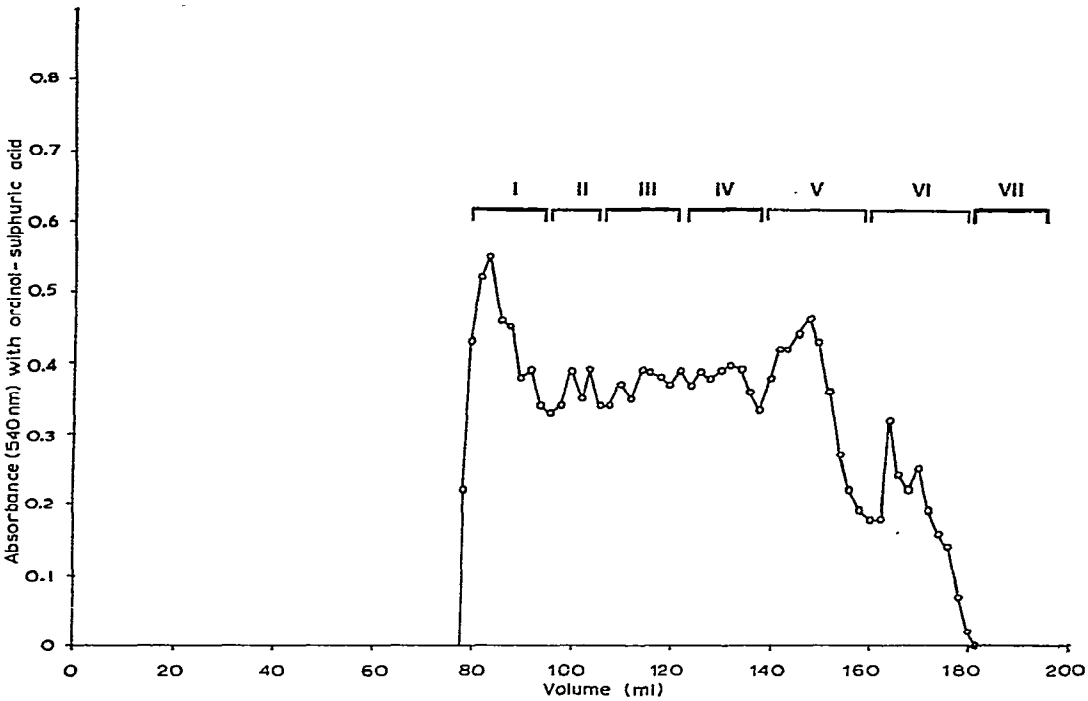


Fig. 1. Chromatography on Bio-Gel P4 of "fucose-free" H-secretor gastric glycopolypeptide after $Na_3PO_4/NaBH_4$ degradation (see text). Blue dextran mainly elutes at 85 ml, raffinose at 156 ml, lactose at 170 ml, and galactose at 190 ml.

TABLE IX

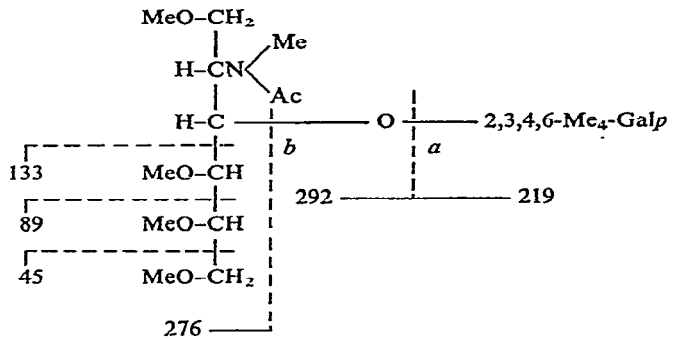
G.L.C. ANALYSIS OF THE FRACTIONS FROM A LARGE COLUMN OF BIO-GEL P4 AFTER THE β-ELIMINATION REACTION[a]

	GalNAc (mM)	*GalNAcol*[b] (mM)	*GlcNAc* (mM)	*Galactose* (mM)	*Fucose* (mM)
Fraction I (Excluded peak)	0.40	0	1.55	1.70	0.19
Fraction II	0.07	0.05	0.65	0.64	0.10
Fraction III	0.06	0.17	0.80	0.85	0.11
Fraction IV	0.06	0.26	0.92	1.04	0.09
Fraction V	0.14	1.70	1.90	3.00	0.23
Fraction VI	0.11	0.70	2.20	1.60	0.16
Fraction VII	0	0.25	0.05	0.04	0
Free monosaccharides	0	0.20	0.02	0	0
Total recovery	0.84	3.33	8.09	8.87	0.88

[a]Using $Na_3PO_4/NaBH_4$ at pH 11.7; reaction conditions are described in the text. [b]2-Acetamido-2-deoxy-D-galactitol.

Fraction VII contained only a small amount of 2-acetamido-2-deoxygalactitol.

Fraction VI corresponded approximately to a disaccharide unit, but contained three different monosaccharides. It was therefore subjected to further purification on Bio-Gel P4. The major peak contained a disaccharide composed of equimolar amounts of galactose and 2-amino-2-deoxygalactitol (Table X). Each sugar residue in the disaccharide was oxidised by periodate. The mass spectrum of the permethylated derivative indicated the following structure:



Peaks at *m/e* 480, 466, 422, and 378 occur by loss from the molecular ion of OMe, CH_2OMe, $H(CHOMe)_2$, and $H(CHOMe)_3$ groups, respectively, *i.e.*, losses occur at C-4, C-5, and C-6 of the 2-amino-2-deoxygalactitol residue, thus indicating substitution at C-3. A peak at *m/e* 276 arises from the 2-acetamido-2-deoxygalactitol moiety formed by cleavage at site *b*. A peak at *m/e* 260 is accounted for by loss of MeOH from the 2-acetamido-2-deoxygalactitol. Both these ions are breakdown products characteristic of this type of compound. As expected, galactosyl peaks occur at *m/e* 219 and 187. Smaller peaks at *m/e* 88 and 101 verify the presence of OMe at C-2, C-3, and C-4 of the galactosyl residue.

Further confirmation was obtained by p.c. (solvent *2*), when material (14 mg) was obtained having an $R_{LACTOSE}$ value (1.36) which was identical to that cited[23] for β-D-Gal-(1→3)-GalNAcol.

Fraction VI also afforded a minor peak on rechromatography, which corresponded to a trisaccharide. Analysis showed it to contain equimolar amounts of galactose, 2-amino-2-deoxyglucose, and 2-amino-2-deoxygalactitol, all of which were susceptible to periodate. Due to the small amount (2 mg) of material isolated, the mass spectrum of the permethylated derivative was not ideal; however, peaks were obtained indicative of a molecular weight higher than that of a disaccharide.

Fraction V was subjected to rechromatography on Bio-Gel P4, and afforded one major (13 mg) and two very minor peaks. The material in the major peak was non-reducing and contained 2-amino-2-deoxygalactitol, 2-amino-2-deoxyglucose, and galactose in the ratios 1:1:2. Only the 2-amino-2-deoxyglucose moiety survived periodate oxidation. After hydrolysis of the trisaccharide for 16 h at 100° with 2M acetic acid, 2-acetamido-2-deoxyglucose and β-D-Gal-(1→4)-D-GlcNAc were identi-

TABLE X

G.L.C. ANALYSES[a] OF THE OLIGOSACCHARIDES OBTAINED BY FRACTIONATION OF FRACTIONS II–VI (Na_3PO_4/$NaBH_4$ DEGRADATION) ON COLUMNS OF BIO-GEL P4 AND P6

	Oligosaccharide	*GalNAcol*	*GlcNAc*	*Galactose*
Fraction VI				
	a. (Di)	7.10 (1)	0 (–)	7.25 (1)
After $NaIO_4$	a.	0 (–)	0 (–)	0 (–)
	b. (Tri)	2.23 (1)	2.00 (1)	2.10 (1)
After $NaIO_4$	b.	0 (–)	0.35 (–)	0 (–)
Fraction V				
	a. (Tri)	1.60 (1)	1.80 (1)	2.00 (1)
	b. (Tetra)	7.50 (1)	7.80 (1)	15.00 (2)
After $NaIO_4$	b.	0 (–)	7.00 (1)	0.84 (–)
	c. (Mixture)	1.50	2.25	2.80
Fraction IV				
	a. (Penta)	1.08 (1)	2.00 (2)	2.00 (2)
After $NaIO_4$	a.	0 (–)	0.94 (1)	0.87 (1)
	b. (Hexa)	2.40 (1)	5.20 (2)	8.00 (3)
After $NaIO_4$	b.	0 (–)	4.60 (2)	3.00 (1)
	c. (Mixture)	0.33	0.76	1.29
After $NaIO_4$	c.	0	0.72	0.83
Fraction III				
	a. (Octa)	1.60 (1)	5.60 (3)	6.50 (4)
After $NaIO_4$	a.	0 (–)	5.50 (3)	5.00 (3)
Fraction II				
	a. (Deca)	0.075 (1)	0.44 (4)	0.60 (5)
After $NaIO_4$	a.	0 (–)	0.50 (4)	0.40 (4)
	b. (Dodeca)	0.19 (1)	1.00 (5)	1.42 (6)
After $NaIO_4$	b.	0 (–)	0.96 (5)	0.86 (4)

[a]Mmoles/litre; molar ratios to the nearest whole number in brackets.

fied (t.l.c., solvent *3*). Methylation analysis afforded 2,3,4,6-tetra-*O*-methylgalactose as the sole neutral-sugar derivative (Table XI).

The mass spectrum of the permethylated derivative was interpreted on the basis of the following structure:

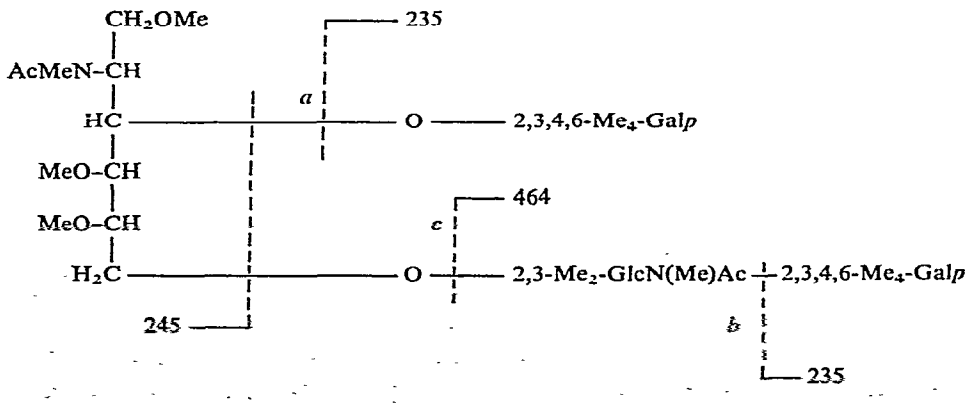


TABLE XI

G.L.C. OF NEUTRAL METHYL SUGARS FROM PERMETHYLATED OLIGOSACCHARIDES ISOLATED AFTER $NaBH_4/Na_3PO_4$ DEGRADATION OF THE GLYCOPOLYPEPTIDES

Oligosaccharides	*Elution times relative to methyl 2,3,4,6-tetra-O-methyl-α-D-glucoside*		
	Column A[a]	*Column B*[a]	*Identity of methyl sugars*
Fraction VI a (Di)	1.09 (sh), 1.17	1.10 (sh), 1.21	2,3,4,6-Me_4-Gal
Fraction V b (Tetra)	1.10 (sh), 1.16	1.11 (sh), 1.23	2,3,4,6-Me_4-Gal
Fraction IV b (Hexa)	1.07 (sh), 1.15, 1.53 (m), 1.74	1.09 (sh), 1.22, 1.99 (m), 2.24	2,3,4,6-Me_4-Gal; 2,4,6-Me_3-Gal
Fraction III a (Octa)	1.08 (sh), 1.15, 1.50 (m), 1.71	1.10 (sh), 1.22, 2.00 (m), 2.27	2,3,4,6-Me_4-Gal; 2,4,6-Me_3-Gal
Fucose-containing pentasaccharide	0.55, 1.10 (sh), 1.17, 1.62, 1.98 (m)	0.55, 1.11 (sh), 1.21, 2.13, 3.03 (m)	2,3,4,6-Me_4-Gal; 3,4,6-Me_3-Gal; 2,3,4-Me_3-Fuc
Reference compounds:			
2,3,4,6-Me_4-Galactose	1.02 (sh), 1.15	1.11 (sh), 1.22	
2,4,6-Me_3-Galactose	1.52 (m), 1.68	2.00 (m), 2.26	
3,4,6-Me_3-Galactose	1.58, 2.00 (m)	2.12, 3.00 (m)	
2,3,4-Me_3-Fucose	0.55	0.56	

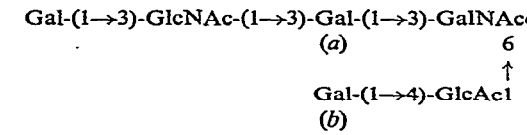
[a] *A*, polyphenyl ether (6%) on Chromosorb G (AW.DMCS) at 190°. *B*, cyclohexanedimethanol adipate (4%)+Apiezon L (2%) on Chromosorb G (AW.DMCS) at 170°. Minor peaks are indicated as: sh = shoulder, m = medium.

Peaks at *m/e* 960, 929 (960−MeO), 915 (960−CH_2OMe), and 871 [960−H(CHOMe)$_2$] were present. Splitting at sites (*a*) or (*b*), *i.e.*, loss of either galactosyl residue, gives a peak at *m/e* 725. Loss of both galactosyl residues leaves an ion having *m/e* 522. The peak at *m/e* 464 can be attributed to the GlcNAc–Gal fragment which would result from a scission at (*c*); linkage (*a*) is assigned the (1→3) position by analogy with the isolated disaccharide; linkage (*b*) is assigned the (1→4) position due to isolation of the (1→4)-linked disaccharide after acetic acid hydrolysis, and linkage (*c*) is assigned the (1→6) position due to the results obtained from sequential periodate oxidation of the product obtained from gastric glycoprotein by mild, acid hydrolysis.

Confirmation of this structure was obtained by p.c. (solvent *3*), since the $R_{LACTOSE}$ value (0.40) was in good agreement with that (0.41) reported[24].

Fraction IV was rechromatographed on Bio-Gel P4 to afford two minor and one major fractions. The major fraction corresponded to a hexasaccharide (19 mg), which contained galactose, 2-amino-2-deoxyglucose, and 2-amino-2-deoxygalactitol in the ratios of 3:2:1. Periodate oxidation destroyed two-thirds of the galactose and all of the reduced amino sugar. Hydrolysis with 2M acetic acid at 100° for 16 h gave four main degradation-products, which were identified (t.l.c., solvent *3*) as galactose, 2-acetamido-2-deoxyglucose, β-D-Gal-(1→4)-D-GlcNAc (R_{GAL} 1.13), and β-D-Gal-(1→3)-D-GlcNAc (R_{GAL} 1.37).

Methylation analysis afforded 2,3,4,6-tetra-*O*-methylgalactose and a smaller amount of 2,4,6-tri-*O*-methylgalactose as the only two methylated neutral-sugar derivatives. The above evidence suggests a hexasaccharide of the following structure:



The terminal β-D-Gal-(1→3)-β-D-GlcNAc unit has been attached to galactose (*a*) rather than to galactose (*b*) due to the results obtained on sequential periodate-oxidation of the glycoprotein.

Fraction III was rechromatographed on Bio-Gel P6 to afford one major fraction (13 mg), which contained galactose, 2-amino-2-deoxyglucose, and 2-amino-2-deoxygalactitol in the ratios 4:3:1. Periodate oxidation destroyed 27% of the galactose and all of the 2-amino-2-deoxygalactitol. After hydrolysis with 2M acetic acid at 100° for 16 h, four products could be identified (t.l.c., solvent *3*) as galactose, 2-acetamido-2-deoxyglucose, β-D-Gal-(1→4)-D-GlcNAc (R_{GAL} 1.15), and β-D-Gal-(1→3)-D-GlcNAc (R_{GAL} 1.36). A further spot (R_{GAL} 0.77) suggested the presence of β-D-Gal-(1→6)-D-GlcNAc. Methylation analysis afforded 2,3,4,6-tetra-*O*-methylgalactose and 2,4,6-tri-*O*-methylgalactose as the only neutral-sugar derivatives.

The above evidence suggests an octasaccharide similar to that previously described but with the addition of an extra Gal→GlcNAc unit.

Fraction II was rechromatographed on Bio-Gel P6 to afford two peaks, in addition to some material which was excluded. Analyses of the material in the two

peaks tentatively suggested a decasaccharide and a dodecasaccharide having galactose, 2-amino-2-deoxyglucose, and 2-amino-2-deoxygalactitol in the ratios of 5:4:1 and 6:5:1, respectively. Periodate oxidation resulted in no loss of the 2-amino-2-deoxyglucose but complete destruction of the 2-amino-2-deoxygalactitol in both oligosaccharides. Loss of galactose was ~25% for the decasaccharide but increased to 33% for the dodecasaccharide. The low yield of material prevented further investigations.

Degradation of the glycopolypeptide with sodium phosphate–sodium borohydride. — This degradation was carried out as described above, using 2-litre batches of H-secretor glycopolypeptide. The reaction products were fractionated on Bio-Gel P4 and rechromatographed either on Bio-Gel P4 or P6. As with the "fucose-free" material, no loss of sugars occurred and all the 2-amino-2-deoxygalactose that disappeared was recovered as 2-amino-2-deoxygalactitol. However, as expected, the complexity of the mixture of oligosaccharides obtained was increased due to the presence of fucose in all fractions isolated, thereby rendering purification difficult.

Nevertheless, a reasonably pure pentasaccharide was isolated by rechromatography on Bio-Gel P4 and shown to consist of 2-amino-2-deoxygalactitol, 2-amino-2-deoxyglucose, galactose, and fucose in the ratios 1:1:2:1. Periodate oxidation destroyed all the sugars except 2-amino-2-deoxyglucose, indicating that the fucose failed to protect any of the galactose residues. After hydrolysis with 2M acetic acid at 100° for 16 h, four products could be identified (t.l.c.) as galactose, fucose, 2-acetamido-2-deoxyglucose, and β-D-Gal-(1→4)-D-GlcNAc. Methylation analysis afforded 2,3,4,6-tetra-*O*-methylgalactose, 3,4,6-tri-*O*-methylgalactose, and 2,3,4-tri-*O*-methylfucose (Table XI). Thus, the pentasaccharide consisted of the tetrasaccharide characterised after the removal of fucose by mild treatment with acid, with a fucose residue attached to one of the galactose residues at C-2.

DISCUSSION

Preliminary analysis of the principal gastric glycopolypeptide shows that it is a substance of high molecular weight, consisting of a polypeptide core to which are attached by *O*-glycosidic linkages, through both serine and threonine, a large number of carbohydrate prosthetic groups, the linking sugar being 2-acetamido-2-deoxygalactose.

The side chains consist of four sugars, namely, 2-acetamido-2-deoxygalactose, 2-acetamido-2-deoxyglucose, galactose, and fucose, as found in the blood-group substances from ovarian-cyst fluids[5]. However, a much greater degree of uniformity was noted with the gastric glycoprotein. It became clear that all the glycoproteins isolated had a common, basic composition of 2-acetamido-2-deoxyglucose, 2-acetamido-2-deoxygalactose, galactose; and fucose in the ratios of 3:1:4:2, this being the simplest structure found, and one which endowed the glycoprotein with Lewis^a activity. Superimposed on this basic structure were additional sugar residues which endowed the material with one or more additional, serological activities, for example, a further fucose residue was found in material with H activity, and further galactose and 2-acetamido-2-deoxygalactose residues with B and A activity, respectively. These

findings were in agreement with the composition of the ovarian-cyst glycoprotein, except for the much more precise analytical compositions noted in each blood group.

The structural investigations were largely confined to glycoproteins carrying H-secretor activity, as these were the simplest of the more readily available secretor groups.

Mild, acid hydrolysis of the H-secretor glycopolypeptides resulted in the formation of a series of linear oligosaccharides consisting of alternating sequences of 2-acetamido-2-deoxyglucose and galactose. The almost complete loss of fucose and its appearance as a monosaccharide suggested that its location in the glycoprotein was both at the end of the chains and also as single branch-points along the main chains. The 2-acetamido-2-deoxygalactose residues were much more resistant to hydrolysis than the other sugars, and were largely recovered in the excluded peak associated with the polypeptide moiety of the glycopolypeptide. This obversation indicated its role as the linking sugar between the polypeptide moiety and the prosthetic groups. In fact, examination of a large number of H-secretor glycopolypeptides showed that the summation of the threonine and serine values equalled that of 2-acetamido-2-deoxygalactose.

The yields of the various oligosaccharides obtained was limited due to the large amounts of monosaccharides formed during the hydrolysis; nevertheless, di-, tetra-, penta-, and hexa-saccharides were isolated. An oligosaccharide of higher molecular weight, which was isolated in very small amounts, was probably a hepta- or octa-saccharide. Periodate oxidation of these saccharides showed that all the substituted galactose residues were (1→3)-linked, and the substituted 2-acetamido-2-deoxyglucose residues were either (1→3) or (1→4)-linked.

It was noteworthy that a relatively large amount of Gal-(1→6)-GlcNAc was isolated, but no evidence was obtained for this (1→6) link in any of the higher oligomers. Similarly, no disaccharides containing (1→3) or (1→4) links could be detected.

Further mild hydrolysis with acid of the residual (excluded) material of high molecular weight isolated from the initial hydrolysis described above gave results which indicated the possibility of a link between the two amino sugars in the vicinity of the linking sugar. Considerable breakdown of the polypeptide core was also noted.

The H-secretor glycopolypeptide was then subjected to a series of Smith degradations[11–15]. Initial periodate oxidation resulted in the loss of the fucose and approximately half of the galactose residues. The two amino sugars were not attacked. After reduction, glycerol was the only polyol isolated, indicating that the neutral sugars occupied terminal positions. The complete loss of fucose was consistent with the results of the mild, acid hydrolysis, in confirming its position as a branched unit. The relatively large loss of galactose indicates that some side chains must be branched.

In the second Smith-degradation, 30% of the 2-acetamido-2-deoxygalactose and 45% of the 2-acetamido-2-deoxyglucose residues were attacked. It thus seems likely that the galactose originally present protected the two amino sugars during the first degradation, as the number of amino sugar residues attacked in the second

degradation was the same as the number of galactose residues attacked initially. This can be explained[15] if the galactose residues are linked either at C-3 or C-4 of the 2-acetamido-2-deoxy sugar units. This observation agrees with that concerning the linkages in the oligosaccharides formed by mild, acid hydrolysis. It is interesting to note that one third of the 2-acetamido-2-deoxygalactose residues were attacked. This result indicates that there is a considerable number of short chains present in the glycoprotein, since 2-acetamido-2-deoxygalactose residues are known to be attached to the core.

The third Smith-degradation destroyed a further 50% of the galactose residues, which must have been preserved from previous attack by a (1→3)-linkage involving 2-acetamido-2-deoxyglucose, again confirming the results of the mild, acid-degradation experiment.

The fourth Smith-degradation destroyed a further 20% of the 2-acetamido-2-deoxygalactose residues, which must have been preserved from previous attack by (1→3) or (1→4)-linkages involving galactose. This result again shows that there are a miscellany of chains present in the gastric glycopolypeptide and that the constant, basic composition represents an average. Further attempts to oxidise the glycopolypeptide were unsuccessful, and its composition was 2-acetamido-2-deoxygalactose, 2-acetamido-2-deoxyglucose, and galactose in ratios 1:3:3, corresponding to a heptasaccharide.

Smith degradation was carried out on the glycopolypeptide of high molecular weight excluded from Bio-Gel P4 and obtained after two mild acid-hydrolyses. This material initially contained 2-acetamido-2-deoxygalactose, 2-acetamido-2-deoxyglucose, and galactose, in addition to the polypeptide moiety.

The first Smith-degradation removed nearly all the galactose and a small amount of 2-acetamido-2-deoxygalactose, leaving a residue of nearly equal amounts of the two hexosamines. In the second Smith-degradation, both hexosamines were attacked, thereby furnishing the polypeptide core virtually denuded of sugars. The fact that both hexosamines were oxidised during the second degradation indicates the presence of GlcNAc-(1→6)-GalNAc, since any other linkage would have protected the 2-acetamido-2-deoxygalactose residues. Similarly, during the first degradation, both hexosamines must have been protected by (1→3) or (1→4)-linkages involving galactose. Further proof of the implication of 2-acetamido-2-deoxygalactose as the sugar linked to the core was shown by the β-elimination reaction on the glycopolypeptide obtained after the first Smith-degradation, whereby 2-acetamido-2-deoxygalactose was lost and there was concomitant appearance of 2-acetamido-2-deoxygalactitol attached to 2-acetamido-2-deoxyglucose.

The glycopolypeptide was degraded by an alkaline borohydride procedure modified to reduce "peeling" reactions of the released oligosaccharides to a very low level. In order to reduce the complexity of the oligosaccharides, mild hydrolysis with acid was carried out to remove fucose prior to alkaline degradation. This technique furnished a number of oligosaccharides which were characterised, both by classical means and, in some cases, by mass spectrometry.

The tri- and penta-saccharides isolated in small amount were not considered to be present as prosthetic groups in the native glycopolypeptide but to have arisen by loss of a galactose residue during the hydrolysis of the fucose residues. This was confirmed by the fact that periodate oxidation on various gastric glycopolypeptides does not cause loss of 2-acetamido-2-deoxyglucose. Thus, it appears that there are present, as prosthetic side-chains, di-, tetra-, hexa-, octa-, and very probably deca- and dodeca-saccharides. This observation is in agreement with the findings of Kabat[25] on the glycoproteins of ovarian-cyst fluids, where prosthetic groups shorter than the main chain were found. The disaccharide was Gal-(1→3)-GalNAc and the higher oligomers appeared to be built up from this by the progressive addition of Gal-(1→3/4)-GlcNAc units. The relative amounts of these groups can be judged from the yields of the various oligosaccharides formed during the β-elimination reaction and also the loss of 2-acetamido-2-deoxygalactose in the second and fourth Smith-degradations. For example, there must be ~30% of di- and tetra-saccharide chains and 20% of the hexasaccharide chain; the residual 50% must be composed of octa, deca, and dodeca chains. From these results, it seems likely that there are approximately equal numbers of each type of chain, and their ratios must be constant to account for the fixed, basic composition of the native glycoprotein. It appears that branching at GalNAc occurs at the tetrasaccharide stage, and the units then continue to be added until an octasaccharide is formed. The octasaccharide presents some puzzling features. First, the periodate-oxidation result seems to be unexpectedly low, suggesting that there is only one susceptible galactose residue, whereas, by analogy with the hexasaccharide, there should be two. Second, the appearance of a (1→6)-linked disaccharide of galactose and 2-acetamido-2-deoxyglucose can only be explained if the latter residue is protected from periodate by substitution at position 3 or 4. It should also be noted that substantial amounts of the (1→6)-linked disaccharide were isolated by mild, acid hydrolysis of the original glycopolypeptide but, again, all the 2-acetamido-2-deoxyglucose present was resistant to periodate oxidation.

Another puzzling feature to be considered is the resistance of the glycopolypeptide to further oxidation after four Smith-degradations, resulting in what appears to be a resistant heptasaccharide side-chain. This resistance to oxidation disappears after mild, acid hydrolysis.

Blocking of the Smith degradation was also found by Lloyd and Kabat[25] working with ovarian-cyst material and was attributed to Schiff-base formation between the aldehyde groups generated by periodate oxidation and the amino groups of the polypeptide moiety; these links would then be reduced by borohydride[26] to give a structure stable to acid. This interpretation would not account, however, for some of the results concerning the octasaccharide.

The results of methylation analysis after β-elimination of fucose-containing H-secretor glycopolypeptide suggests that fucose is generally attached to galactose at C-2. Analysis of the original glycopolypeptide indicates that the fucose must be attached to ~75% of the galactose residues. Examination of the pentasaccharide obtained suggests that only one fucose residue is attached to galactose, but as chain

lengthening occurs the numbers of fucose and galactose residues increase at the same rate.

There are a number of points of similarity between the structure proposed for the gastric glycopolypeptide and that for the blood-group substance isolated from ovarian-cyst fluid. Branching appears to occur at the 2-acetamido-2-deoxygalactose residues in each polymer. Branching at galactose also occurs at the distal end of the chain carrying the ABO serologically active carbohydrates, and there is no reason to doubt that very similar or identical structures are present in both glycoproteins.

Points of difference seem to be the generally longer chains present in the gastric glycopolypeptide, and the greater degree of uniformity in the analytical results of the glycopolypeptide.

ACKNOWLEDGMENTS

We express our thanks to Dr. D. W. Thomas and Dr. B. Blessington for carrying out the mass-spectrometric analysis and for helpful discussion.

REFERENCES

1 G. SCHIFFMAN, E. A. KABAT, AND W. THOMPSON, *Biochemistry*, 3 (1964) 113.
2 K. O. LLOYD, E. A. KABAT, E. J. LAYUG, AND F. GRUEZO, *Biochemistry*, 5 (1966) 1489.
3 I. A. F. L. CHEESE AND W. T. J. MORGAN, *Nature (London)*, 191 (1961) 149.
4 W. P. ASTON, A. S. R. DONALD, AND W. T. J. MORGAN, *Biochem. J.*, 107 (1968) 861.
5 K. O. LLOYD, E. A. KABAT, AND E. LICERIO, *Biochemistry*, 7 (1968) 2976.
6 J. SCHRAGER AND M. D. G. OATES, *Digestion*, 3 (1970) 231.
7 J. SCHRAGER AND M. D. G. OATES, *Gut*, 12 (1971) 559.
8 J. SCHRAGER AND M. D. G. OATES, *Arch. Oral. Biol.*, 16 (1971) 287.
9 G. B. J. GLASS, H. MORI, AND T. PALMER, *Digestion*, 2 (1969) 124.
10 R. H. COTÉ AND W. T. J. MORGAN, *Nature (London)*, 178 (1956) 1171.
11 Y. C. LEE AND R. MONTGOMERY, *Arch. Biochem. Biophys.*, 95 (1961) 263.
12 R. W. JEANLOZ, *Medicine*, 45 (1964) 363.
13 R. G. SPIRO, *J. Biol. Chem.*, 239 (1964) 567.
14 W. E. MARSHALL AND J. PORATH, *J. Biol. Chem.*, 240 (1965) 209.
15 A. T. DE B. ANDREWS, G. M. HERRING, AND P. W. KENT, *Biochem. J.*, 111 (1969) 621.
16 T. J. PAINTER, W. M. WATKINS, AND W. T. J. MORGAN, *Nature (London)*, 206 (1965) 594.
17 K. O. LLOYD AND E. A. KABAT, *Biochem. Biophys. Res. Commun.*, 16 (1964) 385.
18 M. D. G. OATES AND J. SCHRAGER, *J. Chromatogr.*, 28 (1967) 232.
19 L. WARREN, *J. Biol. Chem.*, 234 (1959) 1971.
20 J. S. BRIMACOMBE, B. D. JONES, M. STACEY, AND J. J. WILLARD, *Carbohyd. Res.*, 2 (1966) 167.
21 G. O. ASPINALL AND R. J. FERRIER, *Chem. Ind. (London)*, (1957) 1216.
22 D. W. THOMAS, personal communication.
23 G. VICARI AND E. A. KABAT, *J. Immunol.*, 102 (1969) 821; *Biochemistry*, 9 (1970) 3414.
24 B. ANDERSON, L. ROVIS, AND E. A. KABAT, *Arch. Biochem. Biophys.*, 148 (1972) 394.
25 K. O. LLOYD AND E. A. KABAT, *Proc. Nat. Acad. Sci. U. S. A.*, 61 (1968) 1470.
26 G. E. MEANS AND R. E. FEENEY, *Biochemistry*, 7 (1968) 2192.