Carbohydrate Research, 115 (1983) 151–163
Elsevier Science Publishers B.V., Amsterdam – Printed in The Netherlands

STRUCTURE OF INTESTINAL-MUCUS GLYCOPROTEIN FROM HUMAN POST-MORTEM OR SURGICAL TISSUE: INFERENCES FROM CORRELATION ANALYSES OF SUGAR AND SULFATE COMPOSITION OF INDIVIDUAL MUCINS

ALISON W. WESLEY*, JANET F. FORSTNER, AND GORDON G. FORSTNER

The Kinsmen Cystic Fibrosis Research Centre, Departments of Biochemistry and Paediatrics, Research Institute, Hospital for Sick Children and the University of Toronto, Toronto, Ontario M5G 1X8 (Canada)

(Received February 18th, 1982; accepted for publication in revised form, June 28th, 1982)

ABSTRACT

The carbohydrate composition of 14 human, small-intestine mucins, obtained at surgery or post-mortem, varied greatly from specimen to specimen with respect to individual sugars and average chain-length (ratio of total carbohydrate to *N*-acetylgalactosamine). Three monosaccharides, galactose, *N*-acetylglucosamine, and fucose gave good correlations with each other, and to total carbohydrate content, when expressed as a ratio to the chain-terminal *N*-acetylgalactosamine residue. In contrast, sialic acid gave a good correlation only with *N*-acetylgalactosamine. In eight specimens the molar sulfate to *N*-acetylgalactosamine ratios gave good correlation with the ratios of galactose to *N*-acetylgalactosamine, *N*-acetylglucosamine to *N*-acetylgalactosamine, and total carbohydrate to *N*-acetylgalactosamine. These results indicate that the intraspecies variability of intestinal-mucin carbohydrates arises from the interdependent addition of galactose, *N*-acetylglucosamine, fucose, and sulfate residues. Partial correlation-analysis indicated that proportions of *N*-acetylglucosamine and fucose were correlated only through a mutual dependence on galactose, suggesting that the key elongating-factors involve the addition of galactose residues. The number of sialic acid residues per oligosaccharide chain remained relatively unchanged from mucin to mucin, and this, coupled with the close correlation between the proportions of sialic acid and *N*-acetylgalactosamine, suggests that almost all sialic acid residues are bound to the core *N*-acetylgalactosamine residues in intestinal mucin. High fucose-to-sialic acid and high sulfate-to-sialic acid ratios reported in some disease states are explained as the consequence of chain elongation.

INTRODUCTION

Mucus glycoproteins (mucins) are complex structures consisting of a peptide

*Supported by Fellowships from the Medical Research Council of New Zealand and the Canadian Cystic Fibrosis Foundation.

0008-6215/83/$ 03.00

core rich in serine, threonine, and proline residues, oligosaccharide branches *O*-glycosyl-linked through an *N*-acetylgalactosamine* residue to either serine or threonine, and carbohydrate-free proteins that may be linked to the glycosylated peptides by covalent disulfide bonds[1-3].

The mucin from the small-intestinal goblet cell can be isolated from post-mortem, human intestine[4]. A radioimmunoassay using an antibody raised against the human, small-intestine mucin was developed, and it showed no difference between the antigenic reactivities of mucus glycoproteins from small intestine and colon, whether from a single or various individuals[5]. The mucin from human, small-intestine-goblet cell, even from one individual, is nevertheless heterogeneous with respect to size, density, and charge[6]. Similar heterogeneity has been noted for mucus glycoproteins isolated from other sources[7,8]. Nonhomogeneity can be explained in part by the well-known polydispersity of these macromolecules, in which there is a continuous spectrum of molecular structure centered around a mean size or compositional state. The possibility that two or more different mucin macromolecules might be secreted by one cell type has also been considered[9]. When mucus glycoproteins from many individuals are compared at post-mortem, there is also the possibility that genetic, metabolic, and pathological factors may contribute to diversity. This is especially true for the carbohydrate composition, which is currently thought to depend upon the sequential addition of monosaccharide residues to oligosaccharide chains by a series of glycosyltransferases, each one of which is potentially subject to a unique set of controls[9,10].

In this paper, we analyze the variations in carbohydrate and sulfate composition of intestinal mucins isolated at post-mortem or surgery from individuals suffering from various illnesses. Six of the subjects had cystic fibrosis. In a separate communication, we have shown that intestinal mucins from these individuals were on average denser, larger, and less sialylated than intestinal mucins from nonaffected individuals[12]. Our purpose in presenting data on the combined group of intestinal mucins is to demonstrate that certain variations in monosaccharide and sulfate composition occur in a predictable and interrelated manner, even though there are great differences between specimens. These interrelationships may be useful in understanding the variations reported in disease and in predicting favored pathways of oligosaccharide elongation.

EXPERIMENTAL

Materials. — Thirteen specimens were obtained from human, small intestine within 12 h post-mortem. One specimen was obtained at operation for a blind-loop syndrome, when a segment of jejunum was removed. The mucosa of the bowel was normal macroscopically and histologically. The age of the 14 subjects ranged from one

*All residues have the D-configuration, except for L-fucose residues.

day to 90 years. Six of the subjects had cystic fibrosis, the remainder had died from a variety of diseases, none of which involved the small intestine.

Goblet-cell mucin was isolated as previously described by applying the post-microsomal, supernatant solutions to columns of Sepharose 4B[6]. The mucin was eluted in the void volume. The pooled mucin was further purified by digestion with a nuclease enzyme and chromatography on a column (99 × 3.5 cm) of Sepharose 2B, as follows. The nucleic acid content was determined from the ratio of optical densities[13] at 280 and 260 nm. Ribonuclease (EC 3.1.27.5, 5 units, bovine pancreas, P-L Biochemicals Inc., Milwaukee, WI 53205) and deoxyribonuclease (EC 3.1.21.1, 5 units, bovine pancreas, P-L Biochemicals) were added for each μg of nucleic acid, and the digestion was performed for 24 h at room temperature in 0.1M dipotassium hydrogenphosphate–potassium dihydrogenphosphate buffer, pH 7.4, containing 10mM magnesium chloride and 0.02% sodium azide. Following digestion, the samples were centrifuged at 10 000 *g* for 10 min, and the supernatant solution (30 mL) containing 15 mg of protein was applied to a column of Sepharose 2B. All of the carbohydrate-containing fractions that appeared at or near the void volume were pooled, dialyzed, concentrated by lyophilization, and resuspended in ice-cold, distilled water. In four samples, the void-volume peak obtained from the Sepharose 4B column was centrifuged at 30 000*g* for 30 min, and the supernatant solution dialyzed, concentrated by lyophilization, and resuspended in ice-cold, distilled water without nuclease digestion. As subsequent measurements revealed no evidence of contamination in these four samples, they were treated in an identical fashion for the analysis of data.

Methods. — The purification of the mucin was monitored by carbohydrate analysis. In addition, the mucins were examined by sodium dodecyl sulfate (SDS)–polyacrylamide gel electrophoresis, and proteins detected by the Coomassie Blue or silver stain[14]. The mucins were also subjected to analytical, density-gradient ultracentrifugation as described previously[6].

Amino acids were determined with a Durrum Amino Acid Analyser, after hydrolysis with 6M hydrochloric acid, for 16 or 22 h, at 110° *in vacuo*. In 10 samples, the monosaccharides galactose, fucose, and *N*-acetylneuraminic acid were determined by g.l.c. of the trifluoroacetate derivatives of the methyl glycosides, after methanolysis[15] with 0.5M methanolic hydrogen chloride for 20 h at 80°, using inositol as the internal standard. In four samples, the g.l.c. method of Clamp *et al.*[16] was employed. *N*-Acetylgalactosamine and *N*-acetylglucosamine were measured with a Durrum Amino Acid Analyser after hydrolysis with 4M hydrochloric acid, for 7 h at 100°. Sulfate groups were determined by the method of Mende and Whitney[17]. The washing procedure was modified slightly by introducing descending (rather than ascending) chromatography, for 36 h, to ensure adequate removal of uncomplexed barium-133 ions. Protein content was determined by the method of Lowry *et al.*[18], and hexose content by the anthrone procedure[19]. Linear regression analyses, simple correlation, and partial correlation coefficients were determined by standard methods[20].

Blood-group activity in each sample was tested by standard haemagglutination

techniques, using Anti-A, Anti-B, Anti-Le^b (Ortho Diagnostics, Toronto) antiserum, and Anti-H lectin *Ulex Europeus* (Hyland Laboratories, Toronto). Five mucin specimens were ABH negative and Le^{a+} (nonsecretors) blood-group active. Three specimens were blood-group A, one blood-group B, three blood-group H, and eight, Le^{b+} blood-group active. One specimen, the initial mucin isolated, was not tested for blood-group activity, but was derived from a blood-group O individual; whether it is H or Le blood-group type is unknown

RESULTS

A typical Sepharose 2B profile is shown in Fig. 1. In some samples, the protein and carbohydrate peaks were almost symmetrical, whereas other specimens showed a considerable shoulder or trailing peak. There was no difference between CF and other specimens in the general profile produced from either the Sepharose 4B or 2B columns. All the materials that contained carbohydrate were combined for analysis.

The amino acid analysis (Table I) of the 14 mucin samples showed a profile typical of mucins, with threonine, serine, and proline forming 34–58% (average 45%) of the total amino acids. The coefficients of variation for amino acids with more than 5 per 100 residues varied from 13.28 to 30.96%. When the mean protein-value was calculated by summation of the amino acids determined by amino acid analysis from 18 separate assays, fair agreement was found with the protein content determined by the Lowry method[18], suggesting that the variation was not due to analytical loss. For 10 of the samples, assuming complete recoveries from carbohydrate and amino acid analyses, the mean amino acid content was $22.4 \pm 4.5\%$, and the mean carbohydrate content was $77.6 \pm 4.5\%$. Most of the preparations were free of detectable extraneous protein by sodium dodecylsulfate–polyacrylamide-gel electrophoresis and analytical cesium chloride-gradient ultracentrifugation. Trace amounts of low-molecular weight, nonmucin protein and low-density glycoproteins were present.

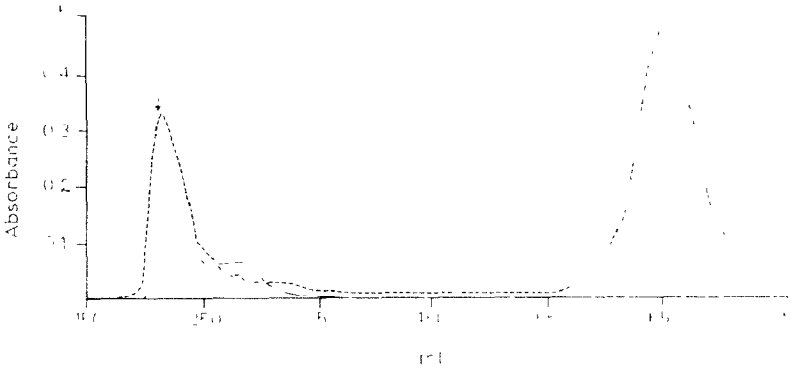

Fig. 1. Chromatography on Sepharose 2B of the peak excluded from a Sepharose 4B column (nuclease-treated): absorbance at 280 nm (– – – – –), and at 260 nm (anthrone reaction reagent for total hexose ——). V_0, void volume.

TABLE I

AMINO ACID COMPOSITION OF SMALL-INTESTINE MUCIN[a]

Amino acids (residues per 100 residues)	*Mean*	*SD*	*CV[b] (%)*
Asp	6.12	± 1.46	23.82
Thr	24.19	3.95	16.32
Ser	9.80	1.30	13.28
Glu	7.25	1.62	22.31
Pro	11.34	2.29	20.23
Gly	7.28	1.22	16.69
Ala	5.39	1.67	30.96
½ Cyst	Trace		
Val	4.96	0.92	18.49
Met	0.77	0.45	57.97
Ile	3.42	0.68	19.77
Leu	5.25	1.05	20.03
Tyr	2.16	0.76	35.85
Phe	2.53	1.12	44.15
His	2.42	0.88	36.31
Lys	3.13	0.87	27.67
Arg	2.97	1.13	37.99

[a]One mg of protein according to Lowry *et al.*[18] = 1.05 ± 0.29 (± SD) mg of amino acid protein.
[b]Coefficient of variation.

TABLE II

CARBOHYDRATE COMPOSITION OF SMALL-INTESTINE MUCIN

Carbohydrate component	*Mol. per 100 mol*		*Molar ratio relative to GalNAc*	
	Mean S.D.	*Range*	*Mean S.D.*	*Range*
Fuc	21.09 ± 6.68	8.96–30.77	1.73 ± 1.07	0.29–4.00
Gal	34.34 ± 12.69	23.43–42.80	2.73 ± 1.87	0.75–5.68
GlcNAc	20.40 ± 6.02	10.55–31.70	1.67 ± 1.07	0.40–3.22
Sialic acid	10.32 ± 7.45	1.40–26.26	0.65 ± 0.36	0.21–1.34
GalNAc	15.98 ± 7.45	6.60–30.11		
Man	0.37	0.0 – 2.6		

however, in 3 preparations, particularly from younger subjects, but these accounted for less than 5% of the total protein.

The mean content of individual sugars (mol per 100 mol and molar ratio relative to GalNAc) is shown in Table II. Mannose was present in trace amounts only (under 0.6%), with the exception of one preparation from a newborn child in which the mannose concentration was 2.6 per 100 mol. The standard deviation for all analyses was quite large. For mol per 100 mol results, the highest values for

TABLE III

RELATIVE PROPORTIONS OF TOTAL CARBOHYDRATE, *N*-ACETYLGALACTOSAMINE, AND SERINE AND THREONINE RESIDUES IN SMALL-INTESTINE MUCIN

Ratios (residues/residues)	*Mucin*	*Mean*	*S.D.*	*Range*
GalNAc to Thr + Ser	All subjects	0.98	0.40	0.43 - 1.97
	Less blood-group A	0.91	0.33	0.43 - 1.50
Total carbohydrate to GalNAc	All subjects	7.78	3.82	3.32-15.0
	Less blood-group A	8.47	4.04	3.32-15.0

TABLE IV

CORRELATION MATRIX FOR MONOSACCHARIDES

Mol/100 mol

↓ \ →	*GalNAc*	*Gal*	*GlcNAc*	*Fucose*
Gal	-0.84[a]			
GlcNAc	0.35	0.45		
Fuc	-0.58	0.18	-0.10	
Sialic acid	0.62[c]	-0.63[b]	0.78[c]	0.43

Sialic acid	0.05	0.06	0.21	0.06
Fuc	0.90[c]	0.85[c]	0.75	
GlcNAc	0.92[c]	0.92[c]		
Gal	0.99[c]			

↑ \ →	*Total carbohydrate*	*Gal*	*GlcNAc*	*Fucose*

Mol/mol GalNAc

[a]p 0.01 [b]p 0.05. [c]p 0.001.

galactose, *N*-acetylglucosamine, and fucose were 2–3 times greater than the lowest values, but for sialic acid the ratio was ~18. Components were also expressed as a molar ratio to *N*-acetylgalactosamine in order to relate the number of residues to the number of oligosaccharide chains. Content of individual monosaccharides still varied greatly, but the order of variability was changed. Sialic acid now varied the least (6.3 times), whereas fucose varied to the greatest extent (13.7 times).

The ratio of *N*-acetylgalactosamine to the sum of threonine and serine residues (see Table III), which reflects the degree of glycosylation of the mucin peptides with oligosaccharide side-chains, was 0.98 ±0.4 (±SD). This ratio was not appreciably altered by excluding the three specimens having blood-group A activity (Table III). The average length of oligosaccharide chains was estimated by the ratio of total carbohydrate to *N*-acetylgalactosamine residues. The mean was 7.8, varying from

3 to 15. Neither the mean nor the range changed significantly when blood-group A specimens were excluded (Table III).

The range of values for individual monosaccharides and for chain length indicates that there was a high degree of intersample heterogeneity. To determine whether the variation was entirely random, or whether organizational patterns could be detected, correlation coefficients were determined for pairs of analytical results (Table IV). They revealed a strong negative correlation between galactose and *N*-acetylgalactosamine. In contrast, sialic acid correlated positively with *N*-acetylgalactosamine, and negatively with galactose and *N*-acetylglucosamine. When monosaccharide residues were expressed as ratios to *N*-acetylgalactosamine, strong positive correlations were observed for all three neutral monosaccharides, and each correlated positively with the chain length (ratio of total carbohydrate to *N*-acetylgalactosamine residues). Sialic acid failed to correlate with the chain length or any of the neutral monosaccharides.

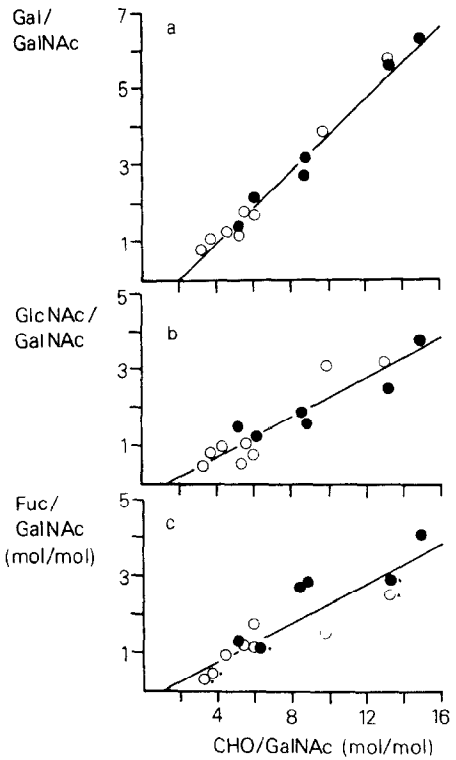

Fig. 2. Ratio of total carbohydrate (CHO) to *N*-acetylgalactosamine content related to: (a) Galactose to *N*-acetylgalactosamine, (b) *N*-acetylglucosamine to *N*-acetylgalactosamine, and (c) fucose to *N*-acetylgalactosamine; (●) CF mucins, and (○) non-CF mucins. The regression equations were (a) $y = 0.48\,x - 1.04$, $r = 0.99$, $p < 0.001$; (b) $y = 0.26\,x - 0.31$, $r = 0.92$, $p < 0.001$; and (c) $y = 0.25\,x - 0.24$, $r = 0.90$, $p < 0.001$. The samples from nonsecretors are indicated by a dot.

Figure 2 shows the linear-regression lines for the molar ratios of individual neutral monosaccharide to *N*-acetylgalactosamine against the molar ratios of total carbohydrate to *N*-acetylgalactosamine. The ratios of galactose to *N*-acetylgalactosamine, *N*-acetylglucosamine to *N*-acetylgalactosamine, and fucose to *N*-acetylgalactosamine increased in a linear fashion with increasing chain-length. The values of the CF samples, indicated by the closed circles, tended to fall within the upper range, but were not exclusive to it. They clearly extended trends established by the values of the non-CF samples, and therefore do not appear to represent a separate group of results subject to independent controls.

The ratio of fucose to *N*-acetylgalactosamine residues showed more scatter than either ratios for the other two hexoses. This is most likely explained by the variability introduced by the nonsecretor status of five of the samples (dots). The correlation coefficient for the secretor samples alone was 0.97. When the slopes of the regression lines were compared, the slope for the ratios of *N*-acetylglucosamine to *N*-acetylgalactosamine *vs.* the ratios of total carbohydrate to *N*-acetylgalactosamine residues was 0.26, and for the ratios of galactose to *N*-acetylgalactosamine residues 0.48. The slope for the ratio of fucose to *N*-acetylgalactosamine *vs.* the ratio of total carbohydrate to *N*-acetylgalactosamine residues was 0.25 if all samples were included, but 0.33 for secretors alone. These results imply that the galactose content of the mucins increases at almost twice the rate of the fucose and *N*-acetylglucosamine content during chain elongation.

The molar ratio of galactose to *N*-acetylgalactosamine positively correlated with the molar ratios of *N*-acetylglucosamine to *N*-acetylgalactosamine and fucose to *N*-acetylgalactosamine. These last two ratios also positively correlated with one another (Table IV), but this correlation was entirely dependent upon a common association with the ratio of galactose to *N*-acetylgalactosamine. The partial correlation of the ratios of *N*-acetylglucosamine to *N*-acetylgalactosamine and fucose to *N*-acetylgalactosamine holding the ratio of galactose to *N*-acetylgalactosmine constant was −0.144.

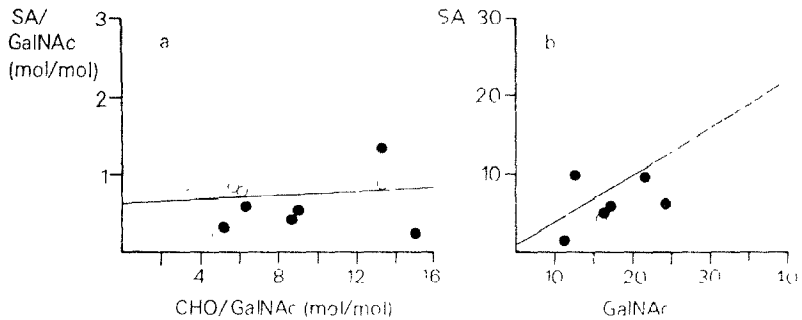

Fig. 3. (a) Total carbohydrate (CHO) to *N*-acetylgalactosamine ratio *vs.* sialic acid to *N*-acetylgalactosamine ratio ($r = 0.05$, not significant). (b) Molar proportion of sialic acid *vs.* molar proportion of *N*-acetylgalactosamine ($r = 0.62$ *x* 0.47, $r = 0.62$, $p < 0.05$). (●) CF mucins, and (○) non-CF mucins.

TABLE V

SULFATE CONTENT

Sulfate	*Mean*	*S.D.*	*Range*
μMol/mg of protein	1.042	± 0.365	0.378–1.488
Mol/100 mol of total carbohydrate	5.8	± 1.3	4.1 –7.6
Mol/mol of GalNAc	0.45	± 0.23	0.17 –0.93
Mol/mol of GlcNAc	0.33	± 0.09	0.22 –0.49
Mol/mol of Gal	0.19	± 0.05	0.16 –0.27

When the molar ratio of sialic acid to *N*-acetylgalactosamine was plotted against chain length (Fig. 3a), no obvious correlation was observed, and the results were not influenced by the CF samples. The content of sialic acid per oligosaccharide chain therefore remained unchanged, despite the wide variation in average chain-length. As shown in Table IV, the ratio of sialic acid to *N*-acetylgalactosamine does not correlate with any of the neutral ratios for monosaccharides, but the content of sialic acid correlates with that of *N*-acetylgalactosamine (mol per 100 mol). This relationship is illustrated in Fig. 3b.

The sulfate content was determined in 8 specimens (4 CF). As shown in Table V, the mean values were ~1 μmol of sulfate/mg of protein and 6 mol/100 mol of carbohydrate. Strong correlations were obtained between the molar ratios of sulfate to *N*-acetylgalactosamine, and total carbohydrate to *N*-acetylgalactosamine ($r = 0.86$, $p < 0.001$), galactose to *N*-acetylgalactosamine ($r = 0.88$, $p < 0.001$), and *N*-acetylglucosamine to *N*-acetylgalactosamine ($r = 0.94$, $p < 0.001$); no correlation was observed with the ratio of sialic acid to *N*-acetylgalactosamine ($r = 0.04$). Fig. 4 illustrates the correlation between the molar ratios of sulfate to *N*-acetylgalactos-

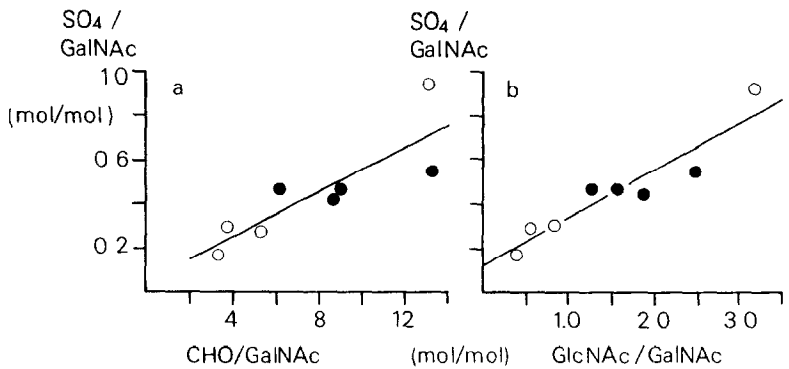

Fig. 4. (a) Ratio of sulfate to *N*-acetylgalactosamine *vs.* ratio of total carbohydrate (CHO) to *N*-acetylgalactosamine ($y = 0.05\ x + 0.05$, $r = 0.85$, $p < 0.001$). (b) Ratio of sulfate to *N*-acetylgalactosamine *vs.* ratio of *N*-acetylglucosamine to *N*-acetylgalactosamine ($y = 0.216\ x + 0.13$, $r = 0.94$, $p < 0.001$).

amine, and total carbohydrate to *N*-acetylgalactosamine (Fig. 4a) and *N*-acetylglucosamine to *N*-acetylgalactosamine (Fig. 4b). Positive correlations were obtained in both cases. The sulfate concentration per chain therefore increased linearly with the total size. Once again CF and non-CF samples behaved as part of a continuous spectrum.

In summary, these results indicate that all galactose, *N*-acetylglucosamine, fucose, and sulfate contents increase linearly per oligosaccharide chain as the size of the chain increases. The *N*-acetylglucosamine and fucose contents are strongly correlated, but through a major dependency on the galactose content. In contrast, the sialic acid content depends primarily upon the number of *N*-acetylgalactosamine residues, and hence upon the number of oligosaccharide chains.

DISCUSSION

Oligosaccharides isolated from mucus glycoproteins from various organs and species have shown extreme heterogeneity, varying in length from single, unsubstituted *N*-acetylgalactosamine residues to elongated, branched oligosaccharides having more than 20 sugars[21–27], for which many of the structural details are still uncertain. The sequence of these chains is thought to depend completely on a competition between specific glycosyltransferases having highly specific substrate-specificities[10,11], but this is still conjectural, as few of the key transferases involved in the branching and elongation of the oligosaccharide chains have been isolated or characterized.

Our data indicate that intestinal mucins from various individuals of one outbred species (human) have an extremely variable carbohydrate composition, with at least a four-fold range in the total number of monosaccharides per oligosaccharide chain. Some of this variation may reflect the inclusion of six specimens from patients having CF, but the range of values in samples from non-CF patients was almost as large. We have attempted to determine whether recognizable organizational-patterns could be detected through a relatively simple, statistical analysis of the data. This appears likely as it is evident that much of the apparent heterogeneity of the human intestinal mucin stems from a variable elongation of the average, oligosaccharide chain-length without a significant change in sialic acid content. This behavior has certain implications for the structure of mucus glycoproteins. As the sialic acid content is quite independent of chain length, one may conclude that the sialic acid residues are not added to the oligosaccharide chain during chain elongation. Sialic acid may be incorporated into the mucin oligosaccharide by at least three processes[24,26,28–30], which link sialic acid to the core *N*-acetylgalactosamine residue[24,28,29] and to more distal sugars in the oligosaccharide chain[11,26,30]. The close correlation between sialic acid and *N*-acetylgalactosamine content in human intestinal mucin suggests that, of these three alternatives, the majority of the sialic acid residues are attached to the core *N*-acetylgalactosamine residues. If so, a fixed proportion of these core residues must be sialylated in all intestinal mucins, regardless of the chain length or the apparent heterogeneity. In ovine submaxillary mucin,

transfer of a sialic acid to an *N*-acetylgalactosamine residue effectively prevents chain elongation[29]. If this condition held in intestinal mucin, one would have to conclude that elongation occurred only in nonsialylated oligosaccharides; however, sialic acid is not denied access to *N*-acetylgalactosamine residues if the latter are substituted β-(1→3) with galactose residues. This is even true for the α-Fuc-(1→3)-β-Gal-(1→3)-GalNAc sequence, although the affinity of sialic acid for *N*-acetylgalactosamine is greatly reduced by the presence of fucose[28]. Thus, transferase specificities do permit elongation of chains provided that sialic acid is not the first sugar added to an *N*-acetylgalactosamine residue.

The only known structure, in mucus glycoprotein, that absolutely prevents addition of sialic acid to the core *N*-acetylgalactosamine residue is the presence of a β-(1→6)-linked *N*-acetylglucosamine residue. This structure has been described for several gastric mucins[21,27,31], although not for small intestine mucin. The upper limit of oligosaccharide substitutions with sialic acid could be set by the number of core β-D-GlcNAc-(1→6)-D-GalNAc substitutions. The relationship between core GlcNAc→GalNAc and sialic acid→GalNAc structures is presumably determined by the competition between *N*-acetylglucosamine- and sialyl-transferases, but indirect control may be of equal importance. Williams *et al.*[32] have shown, for example, that the canine β-D-GlcNAc-(1→6)-D-GalNAc-transferase is very active with a β-D-Gal*p*-(1→3)-D-GalNAc→protein substrate, but completely inactive with unsubstituted D-GalNAc→protein. Thus, the addition of β-D-GlcNAc-(1→6) to the core *N*-acetylgalactosamine residues, and, hence, the exclusion of sialic acid may be highly influenced by the activity of a galactosyltransferase.

The important consideration in this context, however, is that whatever the mechanism, there is much less variation in sialic acid concentration per oligosaccharide chain among individuals than one might expect from a totally random process, as if this particular feature was subjected to relatively rigorous control.

In human intestinal mucins, *N*-acetylglucosamine, galactose, fucose, and sulfate contents increase in a linear fashion as the ratio of total carbohydrate to core *N*-acetylgalactosamine increases. The concentration of fucose correlated with *N*-acetylglucosamine only through a common association with galactose. The important suggestion, from this data, is that elongation depends, at least for human, small-intestinal-mucus glycoproteins, on the initial addition of galactose residues. The observation that the proportion of galactose residues increased, as the chains elongated, at a rate that was 1.5–2 times that of either the fucose or *N*-acetylglucosamine residues is consistent with this interpretation. Since ~2 mol of galactose were added per mol of fucose and *N*-acetylglucosamine, one might question whether the addition of a single galactose residue could serve as the signal for the development of a branched sequence consisting of *N*-acetylglucosamine, galactose, and fucose. The addition of β-D-Gal-(1→3) to D-GalNAc to form a more suitable substrate for the β-D-GlcNAc-(1→6)-D-GalNAc-transferase constitutes one example of such a mechanism that could be duplicated at a more distal site in the chain.

The number of fucose residues also points to increased branching as chains

elongate. In the intestinal mucus glycoproteins with the longest-average chain-length, the fucose to *N*-acetylgalactosamine ratio rises to 4 (Fig. 2). Fucose is generally found at the nonreducing terminal-end of oligosaccharide chains as the blood-group H or Lewis determinant, and the number of fucose residues per nonreducing terminal does not exceed two[10]. From Fig. 2, however, it is clear that the molar ratios of fucose to *N*-acetylgalactosamine exceeded this amount in 6 of the 14 mucin preparations, rising to a maximum of 4 in a patient having secretor status and long-average oligosaccharide chains. The simplest explanation for this observation is that fucose residues were added to chain branches rather than to successive sections of a single oligosaccharide-chain.

Finally, these observations may be helpful in interpreting the reports of sulfomucin to sialomucin and high fucose to sialic acid ratios in various disease states affecting the gastrointestinal tract. By analogy with the CF mucins, an increase in the fucose- or sulfate-to-sialic ratio would be highly suggestive of the production of mucins having elongated, branched chains. Conversely, an increase in sialic acid relative to sulfate groups would be likely to indicate a mucin having relatively short oligosaccharide-chains. Thus, in adenocarcinoma of the colon where sialomucins predominate[33], one might expect to find mucins that are relatively low in total carbohydrate. This expectation is supported by the study of Gold and Miller[34], who found that mucins extracted from adenocarcinoma contained a higher proportion of sialic acid and less carbohydrate than mucins from normal tissue. It is also possible that these generalizations can be extended beyond the gastrointestinal tract. Lamblin *et al.*[35] have shown, for example, that the average chain-length increases in bronchial-mucus fractions as the sulfate to sialic acid ratio increases. Bhaskar and Creeth[7] separated ovarian-cyst mucins as to buoyant density, and also showed that the fractions of highest density were most heavily glycosylated and also enriched with fucose relative to sialic acid.

Although the relative ratios of fucomucins, sulfomucins, and sialomucins must depend to a large extent on competing transferase activities, it is important to acknowledge that other possibilities might exist. Gold *et al.*[9] have provided good evidence for the presence, in colonic mucus, of two immunologically distinct glycopeptide fractions that differ with respect to serine, threonine, proline, and alanine content. One is a potential sialomucin and the other a fucomucin. If, in fact, intestinal mucin consists of two molecules with unique peptide chains, the balance between the two may depend on direct genetic control of peptide synthesis. It is interesting that the total variation in molar proportions of serine plus threonine plus proline, which can be calculated from the data of Gold and Miller[34] with respect to nine colonic, post-mortem and surgical mucins isolated from different individuals, was as large as that found in our samples, ranging from 35 to 61%. Some of this heterogeneity might be explained by proteolytic loss of amino acids from the nonglycosylated portions of the molecule prior to extraction, but it is equally possible that the variation stems from a mixture of different mucin glycopeptides.

REFERENCES

1 J. F. Forstner, *Digestion*, 17 (1978) 234–263.
2 J. P. Pearson, A. Allen, and S. Parry, *Biochem. J.*, 197 (1981) 155–162.
3 M. Mantle, D. Mantle, and A. Allen, *Biochem. J.*, 195 (1981) 277–285.
4 I. Jabbal, D. I. C. Kells, G. Forstner, and J. Forstner, *Can. J. Biochem.*, 54 (1976) 707–716.
5 R. Qureshi, G. G. Forstner, and J. F. Forstner, *J. Clin. Invest.*, 64 (1979) 1149–1156.
6 J. F. Forstner, I. Jabbal, R. Qureshi, D. I. C. Kells, and G. G. Forstner, *Biochem. J.*, 181 (1979) 725–732.
7 K. R. Bhaskar and J. M. Creeth, *Biochem. J.*, 143 (1974) 669–679.
8 J. M. Creeth, K. R. Bhaskar, J. R. Horton, I. Das, M. Lopez-Vidriero, and L. Reid, *Biochem. J.*, 167 (1977) 557–569.
9 D. V. Gold, D. Shochat, and F. Miller, *J. Biol. Chem.*, 256 (1981) 6354–6358.
10 W. M. Watkins, *Adv. Hum. Genet.*, 10 (1980) 1–136.
11 T. A. Beyer, J. I. Rearick, J. C. Paulson, J. P. Prieels, J. E. Sadler, and R. L. Hill, *J. Biol. Chem.*, 254 (1979) 12531–12541.
12 A. Wesley, J. Forstner, R. Qureshi, M. Mantle, and G. Forstner, *Pediatr. Res.*, in press.
13 E. Layne, *Methods Enzymol.*, 3 (1957) 447–454.
14 B. R. Oakley, D. R. Kirsch, and N. R. Morris, *Anal. Biochem.*, 105 (1980) 361–363.
15 J. P. Zanetta, W. C. Brenkenridge, and G. Vincendon, *J. Chromatogr.*, 69 (1972) 291–304.
16 J. R. Clamp, *Biochem. Soc. Trans.*, 5 (1977) 1693–1695.
17 T. J. Mende and P. L. Whitney, *Anal. Biochem.*, 84 (1978) 570–573.
18 O. H. Lowry, N. J. Rosebrough, A. L. Farr, and R. J. Randall, *J. Biol. Chem.*, 193 (1951) 265–275.
19 R. G. Spiro, *Methods Enzymol.*, 8 (1966) 3–26.
20 G. W. Snedecor and W. G. Cochran, *Statistical Methods*, 6th edn., Iowa State University Press, Ames, Iowa, 1973.
21 V. A. Derevitskaya, N. P. Arbatsky, and N. K. Kochetkov, *Eur. J. Biochem.*, 86 (1978) 423–437.
22 M. D. G. Oates, A. L. Rosbottom, and J. Schrager, *Carbohydr. Res.*, 34 (1974) 115–137.
23 G. Lamblin, M. Lhermitte, A. Boersma, P. Roussel, and V. Reinhold, *J. Biol. Chem.*, 255 (1980) 4595–4598.
24 B. L. Slomiany, V. L. N. Murty, and A. Slomiany, *J. Biol. Chem.*, 255 (1980) 9719–9723.
25 H. E. Carlsson, G. Sundblad, A. Hammarström, and J. Lönngren, *Carbohydr. Res.*, 64 (1978) 181–188.
26 A. Slomiany and B. L. Slomiany, *J. Biol. Chem.*, 253 (1978) 7301–7306.
27 W. Newman and E. A. Kabat, *Arch. Biochem. Biophys.*, 172 (1976) 535–550.
28 J. E. Sadler, J. I. Rearick, and R. L. Hill, *J. Biol. Chem.*, 254 (1979) 9534–9541.
29 D. M. Carlson, E. J. McGuire, G. W. Jourdian, and S. Roseman, *J. Biol. Chem.*, 248 (1973) 15763–15773.
30 D. Aminoff, M. M. Baig, and W. D. Gathmann, *J. Biol. Chem.*, 254 (1979) 1788–1793.
31 E. F. Hounsell, M. Fukuda, M. E. Powell, T. Feizi, and S. Hakomori, *Biochem. Biophys. Res. Commun.*, 92 (1980) 1143–1150.
32 D. Williams, G. Longmore, K. L. Matta, and H. Schachter, *J. Biol. Chem.*, 255 (1980) 11253–11261.
33 M. I. Filipe, *Gut*, 10 (1969) 577–586.
34 D. V. Gold and F. Miller, *Cancer Res.*, 38 (1978) 3204–3211.
35 G. Lamblin, M. Lhermitte, J. J. Lafitte, M. Filliat, P. Degand, and P. Roussel, *Bull. Eur. Physiopathol. Respir.*, 13 (1977) 175–190.