Evidence of Regio-specific Glycosylation in Human Intestinal Mucins

PRESENCE OF AN ACIDIC GRADIENT ALONG THE INTESTINAL TRACT*

Received for publication, March 12, 2003, and in revised form, September 2, 2003
Published, JBC Papers in Press, September 2, 2003, DOI 10.1074/jbc.M302529200

Catherine Robbe‡, Calliope Capon‡§, Emmanuel Maes‡, Monique Rousse†, Alain Zweibaum‡, Jean-Pierre Zanetta‡, and Jean-Claude Michalski¶

From the ‡Unité de Glycobiologie Structurale et Fonctionnelle, Unité Mixte de Recherche (UMR) CNRS/USTL 8576, IFR 118, Université des Sciences et Technologies de Lille, 59655 Villeneuve d’Ascq Cedex, France, §Laboratoire de Biochimie, Faculté des Sciences Pharmaceutiques et Biologiques de Lille 2, 59006 Lille Cedex, France, and ¶UMR INSERM U505/UPMC, Institut Biomédical des Cordeliers, 15 rue de l’Ecole de Médecine, 75006 Paris, France

Mucin glycans were isolated from different regions of the normal human intestine (ileum, cecum, transverse and sigmoid colon, and rectum) of two individuals with ALeβ blood group. A systematic study of the monosaccharides and oligosaccharide alditols released by reductive β-elimination from mucins was performed using gas chromatography, matrix-assisted laser desorption ionization time-of-flight mass spectrometry, and nuclear magnetic resonance spectroscopy techniques. Important variations were observed in the mucin-associated oligosaccharide content with an increasing gradient of sialic acid from the ileum to the colon associated with a reverse gradient of fucose. Moreover, a comparative study of the Sda/Cad and ABH blood group determinants along the gastrointestinal tract showed the same reverse distribution in the two kinds of antigens. In addition, besides their heterogeneity, sialic acids presented considerable variations in the degree of O-acetylation in relation to glycan sialylation level. These data are discussed in view of recent concepts suggesting that the oligosaccharide composition of the gut constitutes a varied ecosystem for microorganisms that are susceptible to adapt there and possess the specific adhesion system and specific enzymes able to provide a carbohydrate nutrient.

Intestinal mucins are mainly produced by goblet cells but also by enterocytes and occur both as soluble-secreted and membrane-bound forms. High glycosylation and high mass structures give the mucins gel-forming abilities and other general physical properties that have been regarded as having a protective and rheological function at mucosal surfaces (2). The identification of several different encoded mucins and an enormous repertoire of possible mucin oligosaccharides indicate that the tasks for these glycoproteins may be more subtle than their macroscopic properties suggest. Indeed, mucins may also provide precise structural information carried in determinants among their glycans and in the peptide core, which could mediate specific binding of antibodies, pathogenic microbes, and leukocytes, and may be important in host-pathogen interactions, inflammation, and cancer metastasis (3–8).

Consistent data indicate that mucin genes are expressed in a regulated cell- and tissue-specific manner in the intestine. In situ hybridization has demonstrated that MUC2 is expressed only in goblet cells in the intestine, whereas MUC3 is expressed in both goblet cells and absorptive cells (9). Moreover, a maturation gradient exists for MUC3. This mucin is expressed only on the mucosal surface and in the upper parts of the crypts in the colon. Thus, both the structural differences between MUC2 and MUC3 and the differences in their expression suggest that these two mucins may have different functions. Moreover, MUC4 is widely expressed in the gastrointestinal tract including the colon and the small intestine (10, 11), whereas MUC5B and MUC6 are weakly expressed in the goblet cells of the colon but absent in other parts of the digestive tract (12).

Glycosylation is thought to be essential for the biological functions of mucins, and it has been claimed that cell type-specific mucin expression may be related to the type of glycosylation. To date, few structural studies have dealt with the glycosylation of human intestinal mucins (13–16). Capon et al. (16) demonstrate an increasing expression of blood group Sda determinants from the mucin of a normal human-descending colon (16). Karlsson et al. (17) show glycosylation variation in rat intestinal MUC2 mucin between rat strains and between the small and large intestine, the latter being enriched in sulfated oligosaccharides and the former containing higher amounts of sialylated species (17). These different studies suggest a regio-specific glycosylation of the mucins of the human intestinal tract. Using NMR and mass spectrometry (MS) techniques, we demonstrated the presence of decreasing gradients from ileum to rectum of fucose and ABH blood group expression and of an increasing acidic gradient along the gut. These differences in glycosylation of intestinal mucins could have im-

---

*This investigation was supported in part by the CNRS (Unité Mixte de Recherche CNRS/USTL 8576; Director: Dr. Jean-Claude Michalski) and by the Ministère de la Recherche et de l’Enseignement Supérieur. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

To whom correspondence should be addressed. Tel.: 33-3-20-43-41-46; Fax: 33-3-20-4365-55; E-mail: Jean.Claude.Michalski@univ-lille1.fr.

The abbreviations used are: GalNAc, N-acetylgalactosamine; GalNAc-ol, N-acetylgalactosaminol; NeuAc, N-acetylenuraminic acid; Fuc, fucose; MALDI, matrix-assisted laser desorption ionization; TOF, time-of-flight; MS, mass spectrometry; GC, gas chromatography; H-3ax, H-3eq, H-3 axial; H-3eq, H-3 equatorial.
important implications for the specialized biological functions these molecules play in the intestine. Moreover, we demonstrated a similar glycosylation pattern in each part of the intestine between the two individuals.

EXPERIMENTAL PROCEDURES

Human Samples and Mucin Preparation

All of the immediate autopsy specimens were obtained by France Transplant Association from kidney donors according to protocols approved by the National Ethical Committee. Samples of mucosa were snap frozen in liquid nitrogen and stored in liquid nitrogen until used. The human samples from which the different parts of the gut were analyzed came from two male donors with Alεb blood group and colon type W+Z+, two antigenic specificities called W and Z by Zwiebaum et al. (18) and described in normal colonic secretions. After thawing, the tissue was kept at 4 °C and the mucosa was scraped, homogenized in distilled water (Ultra-turrax, Jankee and Kunkel, Staufffer, Germany), and centrifuged (1 h at 48,000 × g). The supernatant was heated for 1 h in a boiling water bath and further centrifuged (1 h at 48,000 × g). The resultant supernatant was dialyzed against distilled water for 2 days at 4 °C and lyophilized (19).

Release of Oligosaccharide Alditols from Mucin by Alkaline Borohydride Treatment

The colonic mucins were submitted to β-elimination under reductive conditions (0.1 M KOH, 1 M KBH4, for 24 h at 45 °C) (20). The mixture of oligosaccharide alditois was purified by size-exclusion chromatography on a column of Bio-Gel P2 (85 × 2 cm 400 mesh, Bio-Rad), equilibrated, and eluted with water (10 ml/h) at room temperature. The oligosaccharide fractions detected by UV absorption at 206 nm were pooled for and eluted with water (10 ml/h) at room temperature. The oligosaccharide fractions were determined by gas chromatography (GC) on a 25 m × 0.32-mm CP-Sil5 CB Low bleed/MS capillary column and a Finnigan Automass II mass spectrometer. Analyses were performed in the electron impact mode of GC/MS. The GC (Chrompack France, Les Ullis, France) after methanolysis (0.5 M HCl-methanol for 24 h at 80 °C), N-acetylation, and trimethylsilylation (21, 22).

GC/MS Analysis of Sialic Acids

The glycoprotein-bound sialic acids were hydrolyzed (105 min at 80 °C in 2 N acetic acid) and lyophilized. The dry samples were then analyzed as described previously (23). For GC/MS analysis, the GC separation was performed on a Carbo Erba GC 8000 gas chromatograph equipped with a 25 m × 0.32-mm CP-Sil5 CB Low bleed/MS capillary column, a 0.25-μm film phase, and a Finnigan Automass II mass spectrometer. Analyses were performed in the electron impact mode of ionization (ionization energy, 70 eV; source temperature, 150 °C).

NMR Spectroscopy

Samples were repeatedly treated with H2O (99.97% 2H atoms, Euriso-top, CEA, Saclay, France). 1H Chemical shifts were expressed in parts/million downfield from TMS (4,4-dimethyl-4-silapentaneo- sodium salt) but were measured to reference to the internal standard acetone (δ = 2.235 ppm). Samples were analyzed in 200 × 5-mm BMS-005B Shigemi® tubes at 300 K on a Bruker AXS 400-NB spectrometer equipped with a double resonance (1H/15N) broadband inverse z-gradient probe head. NMR samples were recorded without sample spinning. The one-dimensional 1H spectra were measured using 90°-tipping angle for the pulse and 1.5 s as recycle delay between each of the 16 acquisitions of 2.4 s. The spectral width of 4006 Hz was collected in 16,384 complex data points. Two-dimensional homonuclear spectra (TOCSY) were measured using the standard Bruker pulse program and acquired in the time phase-sensitive proportionate increment method. Two-dimensional TOCSY spectra performed on oligosaccharide alditol mixtures from each intestinal segment were strictly recorded with the same parameters (pulses and delays) using 120 ms as mixing time for the MLEV-17 sequence. Moreover, the spectral width was 4006 Hz in both dimensions. 512 spectra of 4096 data points with 16 scans/time incrementation in experiment in F1 dimension were recorded giving a spectral resolution of 0.9 Hz/pt in F2 and 7.8 Hz/point in F1. All of the spectra were interpreted in comparison with spectral data published for a large number of O-linked glycans and other relevant saccharides. Overviews were given by Kamerling and Vliegenthart (24) and in Sugabase, a data base available at www.boc.chem.ruu.nl/sugabase/sugabase.html.

MALDI-MS

All of the mass spectra were acquired on a Voyager Elite (DE-STR) reflectron time-of-flight (TOF) mass spectrometer (Perseptive Biosystems, Framingham, MA) equipped with a pulsed nitrogen laser (337 nm) and a gridless delayed extraction ion source. Oligosaccharide samples were analyzed in delayed extraction mode using an accelerating voltage of 20 kV, a pulse delay time of 200 ns, and a grid voltage of 66%. Detector bias gating was used to reduce the ion current for masses below 500 Da. Between 100 and 200 scans were averaged for each mass spectrum. Oligosaccharide alditois were co-crystallized with 2,5-dihydroxybenzoic acid as matrix (10 mg/ml 2,5-dihydroxybenzoic acid in methanol/water (50/50) containing 0.1% trifluoroacetic acid supplemented with 5 mM NaCl). For all of the measurements, the “dried droplet” preparation technique was used. Typically, 1 μl of the matrix was mixed on target with 1 μl of water-dissolved oligosaccharides and allowed to dry under an air stream. They were analyzed in both positive and negative ion modes.

RESULTS

Experimental Strategy

Mucins were isolated from the different parts of normal human intestinal tract (ileum, cecum, transverse, sigmoid, and rectum) of two Alεb blood group individuals. Oligosaccharide alditois were released by base/borohydride treatment followed by desalting. The percent of saccharides that were recovered after release from the protein backbone was estimated based on monosaccharide compositional analyses of the whole mucin and of the released fraction. The yield was ~50%, a typical yield for this procedure.

After the desalting step, oligosaccharide mixtures from each part of the intestine were analyzed by MALDI-TOF mass spectrometry and NMR spectroscopy without prior derivatization or fractionation.

In the current study, 1H one-dimensional (Fig. 1) and 1H-1H TOCSY-NMR spectroscopy (Figs. 2–4) were used to obtain information on the structure of mucin O-glycans released by reductive β-elimination. All of the analyses were performed on the same total amount of oligosaccharide alditols in each fraction to make a qualitative and a “semi-quantitative comparison” of specific sequences or blood group determinants. The one-dimensional NMR spectrum of the unfraccionated mixture of O-linked glycans showed three proton areas corresponding to (i) anomeric protons, (ii) skeleton protons, and (iii) H-3 of NeuAc and methyl protons of both N-acetyl groups and Fuc residues. These three areas made it possible to compare specific sequences identified in each fraction i.e. α3 or α6-linked NeuAc, Lewis groups, Sda/Cad determinants, sulfate groups, or cores of O-glycans.

The one-dimensional NMR spectra provided high sensitivity, but some crowded areas precluded any positive identification of important signals. The TOCSY with mixing times of 120 ms resolved these signals. Characteristic cross-peak areas corresponding to well known structures allowed us to define typical oligosaccharidic sequences. Therefore, despite the complexity of the different classes of NMR signals, it was possible to distinguish many typical features of glyclosidic sequences along the intestinal tract.

Monosaccharide Composition of the Mucins

Purified mucins isolated from each part of the intestine were analyzed for their carbohydrate composition. It should be noted that the mannose determined in some of the oligosaccharide fractions suggested the presence of N-linked glycans in the mucins. Because N-linked oligosaccharides are not usually released by the alkaline conditions used here, the detected mannose probably represented unreleased N-linked glycopeptides remaining.
FIG. 1. One-dimensional $^1$H NMR spectra of the total mixture of oligosaccharide alditols from ileum (I), cecum (C), transverse (T), sigmoid (S), and rectum (R). Spectra were recorded at 300 K and pH 6.8. Expansion of two regions of one-dimensional NMR spectra ($\delta = 5.30-4.20$ ppm and $\delta = 3.00-1.00$ ppm) with assignments of the major resolved signals is indicated. Signals marked with asterisks did not originate from carbohydrate material.
after the cation-exchange step. Moreover, the lack of contamination by non-mucin glycoproteins has been assessed by SDS-PAGE electrophoresis. With the exception of high molecular components (over 300 kDa), no other protein was detected (data not shown). It should also be stressed that free GalNAc-ol, resulting from a single GalNAc residue attached to Ser or Thr, could not be separated from the high amount of salt used during the reductive β-elimination procedure and was completely lost by gel filtration.

As indicated in Table I, the monosaccharide composition of the different mucins and oligosaccharide alditols fractions presented interesting features. Indeed, GC analyses showed an increasing gradient in the molar ratio of sialic acid from the ileum to the rectum, whereas the reverse situation was observed for Fuc.

Unreduced GalNAc was detected by gas chromatography (Table I), indicating that some of the GalNAc residues were localized in a non-reducing position in the glycans. Because

![Extended regions of two-dimensional TOCSY NMR spectra of oligosaccharide alditol mixture isolated from each part of the intestinal part showing H-5/H-6 cross-peaks of fucosyl residues. Extended parts: F2, δ = 5.00–4.30 ppm; F1, δ = 3.00–2.50 ppm. Assignments of the major resolved signals are indicated.](https://www.jbc.org/Downloaded from)
GalNAc-ol is always situated at the reducing position of oligosaccharide alditols, the molar ratio relative to this compound produced important information. Indeed, these results suggested the presence of A blood group and/or Sda/Cad antigens. Calculated average oligosaccharide chain length of donor 1 was substantially longer than that of donor 2, especially for the

![Extended regions of two-dimensional TOCSY NMR spectra of oligosaccharide alditol mixture isolated from each part of the intestinal part showing the NeuAc H-3ax/H-3eq correlation. Extended parts: F2, δ = 2.10–1.50 ppm; F1, δ = 3.00–2.50 ppm. Assignments of the major resolved signals are indicated.](http://www.jbc.org/Downloaded from)
FIG. 4. Extended regions of two-dimensional TOCSY NMR spectra of oligosaccharide alditol mixture isolated from each part of the intestinal part showing spin systems of 3-O-sialylated Gal, 3-O-sulfated Gal, and 6-O-sulfated GlcNAc. Extended parts: F2, \( \delta = 4.80 - 3.40 \) ppm; F1, \( \delta = 4.80 - 4.10 \) ppm. Assignments of the major resolved signals are indicated. Plain circle, H-3 and H-4 of sulfated Gal; plain square, H-3 of sialylated Gal; and dotted circle, H-5, H-6, and H-6' correlation of 6-O-substituted GalNAc-ol.
two last parts of the intestine: 14 monosaccharides versus 8 in the sigmoid and 18 versus 8 in the rectum. These oligosaccharide lengths are longer than those estimated by MS (see below). This discrepancy could be due in part to the MALDI-MS method, which favors detection of low masses, but it could also be attributed to a possible underestimation of the GalNAc-ol during monosaccharide analyses. Nevertheless, the relative differences in the lengths of oligosaccharides between the two donors suggested the presence of repeated polylactosamine structures in donor 1.

Analysis of the Sialic Acid Diversity

The analysis of sialic acids (Table II) showed a large heterogeneity in the different portions of the gut. Throughout, Neu5Ac was the major compound. Among the O-acetylated sialic acids, 9-O-acetylation was the most important followed by 7-O-acetylation, whereas 8-O-acetylation was detected only in the sigmoid colon. Three different triacylated sialic acids were present: Neu5,8,9Ac₃, Neu5,7,9Ac₃, and Neu5,8Ac₂,9-lactyl. A special concentration of O-acetylated sialic acids was found in the rectum (>40% of total sialic acids), the level being approximately twice much as in the other gut areas. The same type of profiles was observed for the two human specimens, although the detailed sialic acid pattern was different. This variation was not related to technical problems because repetitive analyses of the same samples gave reproducible results but was actually the reflection of individual variations as previously observed for the sialic acids of the human erythrocyte membranes (25). It should be strengthened that Neu5Ac9-lactyl, Neu5Gc, and 3-deoxy-D-glycero-D-galacto-nonulosonic acid were never detected in these human samples, but a trace of Neu4,5Ac₂ was recovered in one of the rectum samples. The presence of this sialic acid in human was already demonstrated by guest on July 24, 2018http://www.jbc.org/Downloaded from

TABLE I

Monosaccharide composition of the native mucins and oligosaccharides isolated from the different mucins after a β-elimination procedure

| Donor 1 | T² | C | T² | S | R² | Donor 2 | T² | C | T² | S | R² |
|---------|----|---|----|---|----|---------|----|---|----|---|----|
| Fuc     | 1.1| 0.8| 0.65| 0.7| 0.6| 2.0     | 1.8| 1.7| 1.5| 1.4|
| Gal     | 1.3| 1.4| 1.5| 2.5| 3.9| 2.4     | 3.0| 3.5| 4.6| 6.4|
| Man     | 0.2| 0.3| 0.2| 0.2| 0.1| 0.3     | 0.4| 0.5| 0.35| 0.2|
| GlcNAc  | 1.2| 1.5| 1.6| 2.5| 3.8| 1.7     | 2.6| 2.7| 3.9| 5.3|
| GalNAc  | 1.0| 1.0| 1.0| 1.0| 1.0| 0.65    | 0.65| 0.8| 0.7| 0.5|
| NeuAc   | 0.3| 0.7| 1.1| 1.6| 2.6| 0.15    | 0.8| 1.0| 1.8| 2.7|
| GalNAc1 | 1.0| 1.0| 1.0| 1.0| 1.0| 1.0     | 1.0| 1.0| 1.0| 1.0|

² The molar ratio of the different monosaccharides was calculated on the basis of one GalNAc residue for the native mucins and on the basis of one GalNAc1 residue for the oligosaccharides.

Fuc = L-fucose; Gal = D-galactose; Man = D-mannose; GlcNAc = N-acetyl-D-glucosamine; NeuAc = N-acetylneuraminic acid; GalNAc = N-acetylgalactosamine.

MATERIALS AND METHODS

MALDI-TOF Analysis of Oligosaccharide Alditols: a Regular Gradient of Glycans from Ileum to Rectum

Desalted oligosaccharide alditols were directly analyzed by MALDI-TOF MS after reductive β-elimination. Based on the mass of the compounds obtained by MS analysis, 71 different oligosaccharide alditols were detected in gut mucins (Table III). This number is certainly underestimated because MALDI-TOF MS only shows molecular ions and does not distinguish among isomer oligosaccharides. This observation is especially important because the complexity and number of potential isomers grow quickly as a function of the number of sugar residues. Taking into account the number of isomers, it was likely that mucins carried several hundred different oligosaccharide structures.

The ileum presented the specificity of having larger mass compounds characterized by a very high degree of fucosylation and without sialic acids. In fact, 43 of the 52 oligosaccharide alditols found in this part of the gut were fucosylated, 7 of them having a single sialic acid residue, and only 6 compounds were neither fucosylated nor sialylated. None of them was sulfated. Moreover, 29 oligosaccharide alditols were specifically found in ileum but not in other intestinal areas.

In agreement with the monosaccharide composition, oligosaccharides recovered in the other areas were less fucosylated and contained more sialic acid and sulfate residues. Indeed, in the cecum, 10 of the 29 oligosaccharides were sialylated and/or sulfated, whereas in the distal colon, more than half of the structures was acidic.

Only small variations in the oligosaccharide repertoire were observed between the two individuals, mainly for the highly fucosylated glycans recovered in ileum. Because the sensitivity of the MALDI-TOF mass spectrometer decreases with the number of sugar residues, it is difficult to know whether some large oligosaccharides could be missed in these analyses or whether they are not present in the fraction. Regardless, these results suggested that glycosylation of intestinal mucins between individuals with the same blood group seemed to be quite similar.

As illustrated in Table III, O-linked glycans formed an almost regular gradient from the ileum to the rectum. Indeed, with the exception of a few compounds common to all parts of
the gut, all of the other compounds were found in one, two, three, or four adjacent areas without interruption. Independent of the mass of the compounds, this scheme was always respected. For example, the situation in which compounds present both in the ileum and in the rectum were absent from the three intermediate regions was never found. Although the absolute demonstration of this phenomenon would require the study of shorter portions of the gut, the absence of any exception to the rule strongly suggests that this was the actual situation.

**NMR Spectroscopy**

Unfractionated mixtures of O-glycans were analyzed by one-dimensional and two-dimensional TOCSY proton NMR spectroscopy. Despite the complexity, different classes of NMR signals (Table IV) could be readily identified based on previously published spectra of purified reference compounds (24).

**GalNAc Residues**—The results obtained from anomic protons (Fig. 1) and their correlations enabled us to identify three types of α-linked GalNAc residues already described in glycans from several other sources. One was a GalNAc residue α-linked to GalNAc-ol, characteristic of core type 5 (GalNAc/H9251–3GalNAc-ol). This core described previously by Feeney et al. (27) in human meconium (27) was found in all of the fractions. The two other GalNAcs identified corresponded to GalNAc that was linked to a GalNAc-ol because they presented the structural reporter sequence. This identification of their H-4 atom resonance enabled us to go back to their downfield shifted H-1 proton along the fucosyl spin system. Moreover, some typical sequences including Fuc residues gave typical H-6/H-5 correlation and especially Fuc units, which were included in Lewis group determinants (24). In this case, Fuc H-5 was significantly downfield-shifted and resonated at approximately 4.8 ppm, whereas in other cases, these protons were upfield-shifted at approximately δ = 4.3 ppm.

From these experiments, several fucosylated determinants were identified and characterized (Figs. 1 and 2 and Table IV). The Leα determinants (Galβ1–4(Fucα1–3)(GlcNAcβ) and Leα determinants (Galβ1–3(Fucα1–4)(GlcNAcβ) were increasingly expressed along the intestinal tract, whereas the Leβ determinant ([Fucα1–2Galβ1–3(Fucα1–4)(GlcNAcβ)] was found exclusively in both ileum and cecum. The same situation was observed for the H-1 determinant ([Fucα1–2Galβ1–3(GlcNAcβ) and the A blood group determinant [GalNAcα1–3(Fucα1–2)Galβ], which were found only in ileum and cecum. Moreover, the Leβ determinant was never detected.

**NeuAc Residues**—NeuAc residues were much more abundant in the last parts of the intestine, in agreement with monosaccharide analysis and MALDI-TOF analyses. They were easily identified in the one-dimensional NMR and two-dimensional TOCSY spectra (Figs. 1 and 3) by their typical chemical shifts and their H-3ax and H-3eq correlations. The cross-peaks H-3ax/H-3eq at approximately δ = 1.70/2.73 ppm were unambiguously attributable to an α2–6 sialic acid linked to a GalNAc-ol because they presented the structural reporter groups described in a large variety of compounds (16, 32). The details of the sequence, in which these α2–6-linked sialic acids were involved, were resolved by the observation of the typical set of H-5, H-6, and H-6′ protons of GalNAc-ol at δ = 4.18, 3.84, and 3.48 ppm, respectively, demonstrating the presence of sialylated core 3, NeuAcα2–6 (GlcNAcβ1–3)GalNAc-ol (33, 34).

The chemical shifts of the H-3ax/H-3eq correlation observed at δ = 1.60/2.76 ppm in the three last gut segments perfectly matched those observed for a terminal sialic acid linked to Galβ2 unit (35), forming the terminal NeuAcα2–3Galβ sequence. This

---

**Table II**

| Sialic acid composition of the mucins from different parts of the gut |
|------------------------|------------------------|------------------------|------------------------|------------------------|
| Donor 1                | Ileum                  | Cecum                  | Transverse             | Sigmoid                | Rectum                 |
|                        | %                      | %                      | %                      | %                      | %                      |
| Neu5Ac                 | 56.24                  | 56.95                  | 59.37                  | 58.27                  | 32.18                  |
| Neu5Ac8Me              | 1.18                   | 4.32                   | 2.42                   | 3.50                   | 4.32                   |
| Neu5Ac1,7L             | 21.22                  | 20.13                  | 14.03                  | 9.28                   | 25.56                  |
| Neu5,9Ac2              | 14.44                  | 13.23                  | 17.03                  | 22.35                  | 22.62                  |
| Neu                   | 0.00                   | 1.20                   | 2.09                   | 1.13                   | 0.00                   |
| Neu5,7Ac2              | 1.35                   | 1.83                   | 3.04                   | 0.00                   | 2.45                   |
| Neu5,8Ac2              | 0.00                   | 0.00                   | 0.00                   | 0.76                   | 0.00                   |
| Neu5,8,9Ac3            | 0.32                   | 1.17                   | 1.02                   | 0.00                   | 1.51                   |
| Neu5,8Ac9Lt            | 0.32                   | 1.17                   | 1.02                   | 2.36                   | 1.51                   |
| Neu5,7,9Ac2            | 4.53                   | 0.00                   | 0.00                   | 2.36                   | 12.61                  |
| Acylated*              | 20.64                  | 17.40                  | 22.11                  | 27.83                  | 40.70                  |
| Donor 2                |                        |                        |                        |                        |                        |
| Neu5Ac                 | 55.52                  | 70.16                  | 72.41                  | 68.37                  | 36.24                  |
| Neu5Ac8Me              | 0.80                   | 2.57                   | 1.84                   | 1.18                   | 0.98                   |
| Neu5Ac1,7L             | 20.37                  | 11.83                  | 7.91                   | 5.73                   | 19.45                  |
| Neu5,9Ac2              | 15.76                  | 11.58                  | 12.64                  | 19.19                  | 22.29                  |
| Neu                   | 0.00                   | 0.00                   | 0.00                   | 1.13                   | 0.40                   |
| Neu5,7Ac2              | 1.68                   | 1.71                   | 2.60                   | 2.77                   | 2.34                   |
| Neu5,8Ac2              | 0.00                   | 0.00                   | 0.00                   | 0.00                   | 0.00                   |
| Neu5,8,9Ac3            | 0.46                   | 1.09                   | 1.35                   | 1.40                   | 1.90                   |
| Neu5,8Ac9Lt            | 0.50                   | 1.05                   | 1.25                   | 1.36                   | 1.86                   |
| Neu5,7,9Ac2            | 0.91                   | 0.00                   | 0.00                   | 14.42                  | 42.93                  |
| Acylated*              | 23.31                  | 15.43                  | 17.84                  | 24.73                  | 42.93                  |

* Acylated corresponds to the percentage of the sum of all sialic acids having at least one O-acetyl group (O-acetyl and/or O-lactyl).
Deduced monosaccharide compositions of the major oligosaccharides identified by MALDI-TOF analysis in the two donors (in lightgray), in donor 1 (in medium gray) and in donor 2 (in dark gray)

| Assigned composition | m/z<sup>a</sup> | I<sup>b</sup> | C<sup>c</sup> | T<sup>d</sup> | S<sup>e</sup> | R<sup>f</sup> |
|----------------------|----------------|------------|-------------|-------------|------------|------------|
| GalNAc, HexNAc       | 449            |            |             |             |            |            |
| GalNAc, NeuAc        | 559            |            |             |             |            |            |
| GalNAc, HexNAc, Hex  | 611            |            |             |             |            |            |
| GalNAc, HexNAc, Hex, Fuc | 757         |            |             |             |            |            |
| GalNAc, HexNAc, NeuAc | 762           |            |             |             |            |            |
| GalNAc, HexNAc, 2 Hex | 773            |            |             |             |            |            |
| GalNAc, 2 HexNAc, Hex | 814            |            |             |             |            |            |
| GalNAc, HexNAc, Hex, NeuAc | 924         |            |             |             |            |            |
| GalNAc, HexNAc, Hex, Fuc, NeuAc | 1070     |            |             |             |            |            |
| GalNAc, 2 HexNAc, 2 Hex, Fuc | 1122       |            |             |             |            |            |
| GalNAc, 2 HexNAc, 2 Hex | 976            |            |             |             |            |            |
| GalNAc, HexNAc, Hex, Fuc, S | 859         |            |             |             |            |            |
| GalNAc, HexNAc, Hex, Fuc, NeuAc, S | 1172       |            |             |             |            |            |
| GalNAc, 3 HexNAc, 4 Hex | 1503          |            |             |             |            |            |
| GalNAc, HexNAc, 2 Hex, S | 873           |            |             |             |            |            |
| GalNAc, HexNAc, Hex, NeuAc, S | 1036        |            |             |             |            |            |
| GalNAc, 2 HexNAc, Hex, NeuAc | 1127         |            |             |             |            |            |
| GalNAc, 3 HexNAc, 3 Hex | 1341          |            |             |             |            |            |
| GalNAc, 2 HexNAc, 2 NeuAc | 1440         |            |             |             |            |            |
| GalNAc, 2 HexNAc, 2 Hex | 1138          |            |             |             |            |            |
| GalNAc, 3 HexNAc, 2 Hex | 1179          |            |             |             |            |            |
| GalNAc, HexNAc, 2 Hex, NeuAc, S | 1188       |            |             |             |            |            |
| GalNAc, HexNAc, Hex, Fuc, 2 NeuAc | 1383        |            |             |             |            |            |
| GalNAc, 2 HexNAc, 3 Hex, 2 Fuc, 2 S | 1634      |            |             |             |            |            |
| GalNAc, 3 HexNAc, 4 Hex, S | 1905          |            |             |             |            |            |
| GalNAc, 2 HexNAc, 3 Hex, S | 1240          |            |             |             |            |            |
| GalNAc, 3 HexNAc, 2 Hex, S | 1281          |            |             |             |            |            |
| GalNAc, 3 HexNAc, 2 Hex, S | 1440          |            |             |             |            |            |
| GalNAc, HexNAc, Hex, 2 Fuc | 903           |            |             |             |            |            |
| GalNAc, 2 HexNAc, 1 Hex, 2 Fuc | 1196        |            |             |             |            |            |
| GalNAc, HexNAc, Hex, 2 Fuc, NeuAc | 1216        |            |             |             |            |            |
| GalNAc, HexNAc, 2 Hex, Fuc | 919           |            |             |             |            |            |
| GalNAc, 2 HexNAc, Hex, Fuc | 960           |            |             |             |            |            |
| GalNAc, HexNAc, 2 Hex, 2 Fuc | 1065          |            |             |             |            |            |
| GalNAc, 3 HexNAc, Hex, Fuc | 1163          |            |             |             |            |            |
| GalNAc, 2 HexNAc, 2 Hex, 2 Fuc | 1268         |            |             |             |            |            |
| GalNAc, 3 HexNAc, Hex, Fuc, NeuAc | 1273        |            |             |             |            |            |
| GalNAc, 3 HexNAc, 2 Hex, Fuc | 1325          |            |             |             |            |            |
| GalNAc, 2 HexNAc, 2 Hex, 3 Fuc | 1414         |            |             |             |            |            |
| GalNAc, 2 HexNAc, Hex, 2 Fuc, NeuAc | 1419        |            |             |             |            |            |
| GalNAc, 3 HexNAc, 4 Hex, 2 Fuc | 1471          |            |             |             |            |            |
| GalNAc, 3 HexNAc, Hex, 2 Fuc, 3 Fuc | 1617         |            |             |             |            |            |
| GalNAc, 3 HexNAc, Hex, 3 Fuc | 1017          |            |             |             |            |            |
| GalNAc, HexNAc, 2 Hex, 3 Fuc | 1211          |            |             |             |            |            |
| GalNAc, 2 HexNAc, 3 Hex, 1 Fuc | 1294          |            |             |             |            |            |
| GalNAc, 2 HexNAc, 3 Hex, 2 Fuc | 1430          |            |             |             |            |            |
| GalNAc, 3 HexNAc, 3 Hex, 1 Fuc | 1487          |            |             |             |            |            |
| GalNAc, 4 HexNAc, 2 Hex, Fuc | 1528          |            |             |             |            |            |
| GalNAc, 3 HexNAc, 3 Hex, 2 Fuc | 1633          |            |             |             |            |            |
| GalNAc, 4 HexNAc, 2 Hex, 2 Fuc | 1674          |            |             |             |            |            |
| GalNAc, 4 HexNAc, 3 Hex, Fuc | 1690          |            |             |             |            |            |
| GalNAc, 3 HexNAc, 3 Hex, 3 Fuc | 1779          |            |             |             |            |            |
| GalNAc, 3 HexNAc, 4 Hex, 2 Fuc | 1795          |            |             |             |            |            |
| GalNAc, 4 HexNAc, 2 Hex, 3 Fuc | 1820          |            |             |             |            |            |
| GalNAc, 4 HexNAc, 3 Hex, 2 Fuc | 1836          |            |             |             |            |            |
| GalNAc, 4 HexNAc, 4 Hex, Fuc | 1852          |            |             |             |            |            |
| GalNAc, 4 HexNAc, 2 Hex, 3 Fuc | 1982          |            |             |             |            |            |
| GalNAc, 4 HexNAc, 4 Hex, 2 Fuc | 1998          |            |             |             |            |            |
| GalNAc, 5 HexNAc, 3 Hex, 2 Fuc | 2039          |            |             |             |            |            |
| GalNAc, 4 HexNAc, 3 Hex, 4 Fuc | 2128          |            |             |             |            |            |
| GalNAc, 5 HexNAc, 3 Hex, 3 Fuc | 2185          |            |             |             |            |            |
| GalNAc, 4 HexNAc, 3 Hex, 4 Fuc | 2290          |            |             |             |            |            |
| GalNAc, 5 HexNAc, 3 Hex, 4 Fuc | 2331          |            |             |             |            |            |
| GalNAc, 5 HexNAc, 4 Hex, 3 Fuc | 2347          |            |             |             |            |            |
| GalNAc, 5 HexNAc, 3 Hex, Fuc | 2493          |            |             |             |            |            |
| GalNAc, 5 HexNAc, 4 Hex, 3 Fuc | 2550          |            |             |             |            |            |
| GalNAc, 5 HexNAc, 5 Hex, 3 Fuc | 2539          |            |             |             |            |            |
| GalNAc, 5 HexNAc, 4 Hex, 4 Fuc | 2696          |            |             |             |            |            |
| GalNAc, 5 HexNAc, 5 Hex, 4 Fuc | 2801          |            |             |             |            |            |
| GalNAc, 5 HexNAc, 4 Hex, 5 Fuc | 2858          |            |             |             |            |            |
| GalNAc, 5 HexNAc, 5 Hex, 5 Fuc | 2923          |            |             |             |            |            |

<sup>a</sup> Monoisotopic m/z values refer to \([M+(n-1)H+nNa]^+\), where the number of sodium adducts proportionally increased with the sialic acid or sulfate content.

<sup>b</sup> Ileum.

<sup>c</sup> Cecum.

<sup>d</sup> Transverse.

<sup>e</sup> Sigmoid.

<sup>f</sup> Rectum.
observation was confirmed by typical shifts corresponding to the H-3 and H-4 proton downfield shifts at δ = 4.10 and 3.93 ppm, respectively, of a Gal residue substituted in position 3 by a NeuAc unit. As shown in Table IV, this sequence was increasingly expressed along the descending intestinal tract.

In donor 1 only, evidence for NeuAcα2–3, with the characteristics of Sda/Cad determinants was detected by cross-peak signal at 1.93/2.68 ppm in the TOCSY spectra with a mixing time of 120 ms. This sequence was absent in ileum but was recovered in all of the other fractions of donor 1 with an increasing gradient of expression in the descending part of the human gut. A very low amount of Sda/Cad determinant was detected and only in the rectum part of donor 2 by NMR. This result was surprising because Morton et al. (36) showed that Sda antigen was immunologically present in ~98% of human colon. Nevertheless, MALDI-TOF MS analysis revealed the presence of two compounds at m/z 1127 and 1440, respectively, which could match with the presence of the Sda/Cad epitope. In fact, taken together, these results suggested that Sda/Cad was present in the two donors with an underexpression in donor 2.

Sulfate Residues—The sulfation degree of oligosaccharide aldolts could be observed through the TOCSY spectrum correlations starting from the Galβ H-1 signal (Fig. 4). The presence of a 3-O-sulfated Gal unit could be inferred from the observation of particular sets of H-3 (δ = 4.32 ppm) and H-4 (δ = 4.27 ppm) spin system starting from H-1 (δ = 4.60 ppm). As shown in Table IV, this structural motif was scarcely expressed in the ileum but was increased in other parts in proportion to the expression of sialylated oligosaccharide aldolts. These downfield chemical shifts were always identical in such cases (37, 38). Moreover, in sigmoid and rectum parts of donor 2, two correlated protons were observed at δ = 4.35 and 4.23 ppm, corresponding to H-6 and H-6’ of a 6-O-sulfated GlcNAc. These data confirmed unambiguously a 6-O-sulfation on a core 2 GlcNAcβ residue.

N-Acetyl Group—1H one-dimensional spectra, through N-acetyl signals between 1.95 and 2.11 ppm (Fig. 1), confirmed the presence of the different hexosamine residues such as GlyNAc-ol, αGlcNAc or βGlcNAc, GlcNAc, and αNeuAc. Because of the high level of NAc signals, it was not possible to attribute each one but a particular downfield shift at δ = 2.105 ppm could be unambiguously attributable to N-acetyl group from GlcNAcβ included in the Fucα1–2Galβ1–3GlcNAcβ1 sequence (32, 39–41). This sequence, in fact, directly β-linked to the GalNAc-ol unit, forming an elongation of core 3.

**DISCUSSION**

Few structural studies have dealt with the glycosylation of mucins along the human gut, probably because of the difficulty in obtaining necropsy material and in performing complex analyses on relatively small amounts of human material. Nevertheless, progress in the field of glycan analysis and especially in the field of mass spectrometry has made it possible to confront this essential problem with relatively relevant technologies.

A combination of the results obtained from the different analysis (GC, MALDI-TOF, and NMR analyses) showed clearly the presence of different gradients along the intestinal tract. Indeed, monosaccharide analysis allowed the demonstration of an increasing gradient of NeuAc with a decreasing gradient of Fuc from the ileum to the rectum. These results were confirmed by NMR spectroscopy and MALDI-MS analyses, showing a decreasing expression of Fuc along the descending intestinal tract and showing an increasing expression of sialylated structures.

Moreover, NMR analysis of unfractionated mixtures of O-glycans allowed a semi-quantitative comparison of specific sequences such as ABH blood group determinants or Sda/Cad epitopes. Thus, NMR analysis allowed the demonstration of an increasing gradient of Sda/Cad blood group antigen associated with a decreasing gradient of ALeβ blood group epitope. Because mucin samples analyzed in this study were obtained from individuals with blood type ALeβ, the present data indicated that only the first two parts of the intestine carried the A, B, and O blood group structures. These results were confirmed by NMR analysis of α-linked GalNAc residues, showing that

| Assignments of signals (ppm) | Estimated abundance |
|-----------------------------|---------------------|
|                            | ileum D1 | D2 | Transverse D1 | D2 | Sigmoid D1 | D2 | Rectum D1 | D2 |
| α2-3 linked NeuAc          | 2.75-2.78 | 1.79-1.81 | ± | + | ++ | ++ | ++ | ++ | ++ |
| α2-3 linked NeuAc in Sda/Cad | 2.63-2.67 | 1.91-1.93 | - | - | - | - | - | - | - |
| α2-6 linked NeuAc          | 2.72-2.75 | 1.68-1.71 | ++ | ++ | ++ | ++ | ++ | ++ | ++ |

| O-3 Sialylated Gal         | 4.53-4.59 | 3.5-3.6 | 4.08-4.13 | 3.92-3.95 | ± | - | + | ++ | ++ | ++++ |
| O-3 Sulfated Gal           | 4.56-4.61 | 3.7-3.8 | 4.31-4.33 | 4.26-4.28 | - | - | - | ++ | ++ | ++++ |
| O-6 Sulfated GlcNAc        | 4.4-4.4 | 4.3-4.2 | - | - | - | - | - | - | - |

| α1-2 linked Fuc in H group | 5.15-5.35 | 3.76-3.84 | 4.26-4.32 | 1.18-1.29 | + | + | + | + | - | - |
| α1-2 linked Fuc in Le group| 5.27-5.29 | 3.72-3.82 | 4.34-4.39 | 1.28-1.29 | ++ | ++ | ++ | + | - | - |
| α1-3 linked Fuc in Le group| 5.11-5.15 | 3.69-3.71 | 4.81-4.82 | 1.13-1.19 | ++ | ++ | ++ | + | - | - |
| α1-4 linked Fuc in Le group| 5.01-5.06 | 3.80-3.81 | 4.82-4.86 | 1.15-1.20 | ++ | ++ | ++ | + | - | - |
| α1-4 linked Fuc in Le group| 5.02-5.04 | 3.81 | 4.86-4.87 | 1.27 | ++ | ++ | ++ | + | - | - |
| α3-3 linked GalNAc core 5  | 5.06-5.07 | 4.22 | 3.92 | 4.04 | + | + | + | + | + | + |
| GalNAc in A group          | 5.18-5.23 | 4.24 | 3.90-3.93 | 4.0 | + | + | + | + | - | - |
| GalNAc in A Le group       | 5.23-5.27 | 4.17-4.18 | 3.97 | 3.97 | + | + | - | - | - | - |

* Assignments were given taking into account the minimal and maximal chemical shifts observed in all of the experiments.
* Chemical shifts corresponding to α1-2-linked fucosyl residues carried by type 1 or type 2 lactosamine chains.
the GaINAc unit implicated in blood group A and ALe^b sequences were recovered only in ileum and cecum, in agreement with previously published data (5). Indeed, blood group A, B, and H antigens have already been shown to be present in proximal but absent in the distal colon.

Finally, two-dimensional TOCSY spectra showed that the relative level of both 3-O-sulfated Gal and NeuAc increased proportionally along the intestinal tract. Consequently, glycans in intestinal mucins become more and more acidic, hence establishing an acidic gradient along the gut.

Comparison of the O-glycosylation patterns of intestinal mucins for the two donors showed minor differences. These differences mainly concerned the expression of Sda/Cad antigen and the chain length of distal colonic mucin. However, because the major conclusions concerning the different gradients of fucosylation, sialylation, sulfation, and acylation of sialic acids were basically the same, it is tempting to speculate regarding preserved pattern of glycosylation between individuals.

Among the sialylated oligosaccharides, the Sda/Cad blood group antigen determinant was increasingly expressed along the descending intestinal tract and was absent in ileum. These results are in agreement with previous enzymatic, immunohistochemical, and structural studies. Indeed, the Sda antigen has been identified as mainly expressed in cells of the large intestine and distal kidney (36). In normal human colon mucin glycans, this determinant was found as a major structural sequence. These differences were recovered only in ileum and cecum, in agreement with previously published data (5). Indeed, blood group A, B, and H antigens have already been shown to be present in proximal but absent in the distal colon.

The biological significance of diversity in the oligosaccharide side chains of secretory mucins is still not really understood. Obviously, general properties of mucin-bound oligosaccharides include protease resistance, large water-holding capacity, and high charge density caused by sialic acid and sulfate residues (46). Furthermore, it is generally accepted that numerous oligosaccharide side chains maintain the extended random coil conformation of the molecule, which together with high molecular mass is responsible for the characteristic viscoelastic properties of mucins (47–49). It is often stated that particularly charged residues such as sulfate and sialic acid are important in this respect. In this paper, studies on the relative level of sulfation and sialylation showed that glycans became more and more acidic, hence establishing an increasing acidic gradient along the intestinal tract. In this way, it seems that charged residues play a role in water and electrolyte transport in the distal part of the colon.

The problem posed by the glycan composition of the human gut is essentially related to host/pathogenic and host/symbiotic interactions. In this field, the human gut is a good model not only because of human pathologies, but in terms of symbiotic interactions, the common bacterial population is a shield for pathogen invasions. The problem of host/bacterial interactions in the gut is complex because it involves the fixation of the bacteria on a specific site (different bacteria are not randomly distributed in the gut) and the formation of colonies in specific regions of the human gut. Because oligosaccharides often function as attachment sites for bacteria and viruses (50), it was proposed that diverse mucin oligosaccharides represent a mosaic of potential ligands designed to trap a broad diversity of microorganisms in the mucus layer, thereby impeding infection of the underlying epithelia (51). In this context, a prominent role has been envisaged for the terminal residues of the glycan chains including sialic acid and galactose, which are the primary targets for interaction with microorganisms (51–54). All together, our results show a high degree of diversity in the expression of glycans in the different parts of the intestine, the final result being of course the demonstration of an enormous repertoire of potential binding sites for microorganisms that could explain the regio-specific colonization of bacteria in the human intestinal tract. Indeed, recent studies have shown that host-microbial interactions shape the nutrient environment of the mammalian intestine (55–62). Thus, carbohydrate availability probably helps determine the niche that a given microbe is able to occupy.

The results of our study on the diversity of sialic acids in the different areas of the intestine are in agreement with data already published. Indeed, in a study performed on the colon, Corfield et al. (63) reported that O-acetylated sialic acids represented ~24% of total sialic acids. These data obtained on the whole human colon are very close to our results with the exception of the rectum for which the level of O-acetylated sialic acid was twice that of the other regions. The biological significance of the heterogeneity of sialic acids remains largely speculative, but several works have reported that a number of intestinal bacterial strains bind to sialic acids (63). The number of different sites could be greatly increased considering that bacterium adhesion systems recognize more complex glycans than single monosaccharide. This was perfectly demonstrated for Shiga-like toxins recognizing GB3 or GB4 glycosphingolipids with high affinity. Furthermore, several authors have demonstrated that the acetylation of sialic acids decreases the rate of cleavage of sialic acids from mucins by bacterial neuraminidases. A study performed on 23 isolates of human fecal bacteria (63) shows that the presence of two or more O-acetyl groups on sialic acids inhibits enteric bacterial sialidases but the production of sialate O-acetylesterase decreases this probability that mucin oligosaccharides bearing O-acetylated sialic acids should be protected from degradation. Nevertheless, the higher level of O-acetylated sialic acids in the rectum compared with other parts of the gut sustains the view that specific strains able to de-O-acetylate sialic acids can be more competitive than others in developing colonies in this area.

In conclusion, this study provides evidence of several different gradients of O-linked glycans in the mucins of the human gut. The rectal-sialylated oligosaccharides are significantly more O-acetylated than in any other gut area. Therefore, assuming that enterocytic mucins constitute an ecological system for the formation of bacterial strain colonies as proposed first by other researchers and second by using our data, it can be hypothesized that a large variety of bacterial strains having specific O-acetylesterasers, arylesterasers, and glycosulfatases can find specific fixation sites.

Acknowledgments—We thank Colette Richet for help with sialic acid analysis and Alexandra Sommerville-Tavernier for careful rereading of the paper.

REFERENCES
1. Gendler, S. J., and Spicer, A. P. (1995) Annu. Rev. Physiol. 57, 657–684
2. Forstner, J. F., and Forstner, G. G. (1994) in Physiology of the Gastrointestinal Tract (Johnson, L. R., ed) pp. 1245–1283, Raven Press Ltd, New York
3. Van Klinken, B. J., Dekker, J., Buller, H. A., and Eermering, A. W. (1995) Am. J. Physiol. 269, 615–627
4. Rosen, S. D. (1999) Am. J. Pathol. 155, 1013–1020
5. Kim, Y. S. (1998) Ref. J. Med. 47, 10–18

Downloaded from http://www.jbc.org/ by guest on July 24, 2018
Evidence of Regio-specific Glycosylation in Human Intestinal Mucins: PRESENCE OF AN ACIDIC GRADIENT ALONG THE INTESTINAL TRACT
Catherine Robbe, Calliope Capon, Emmanuel Maes, Monique Rousset, Alain Zweibaum, Jean-Pierre Zanetta and Jean-Claude Michalski

J. Biol. Chem. 2003, 278:46337-46348.
doi: 10.1074/jbc.M302529200 originally published online September 2, 2003

Access the most updated version of this article at doi: 10.1074/jbc.M302529200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 61 references, 17 of which can be accessed free at http://www.jbc.org/content/278/47/46337.full.html#ref-list-1