Glycoconj J (2009) 26:397–413
DOI 10.1007/s10719-008-9186-9

Glycosylation of human fetal mucins: a similar repertoire of *O*-glycans along the intestinal tract

Catherine Robbe-Masselot · Emmanuel Maes · Monique Rousset · Jean-Claude Michalski · Calliope Capon

Received: 19 June 2008 / Revised: 28 August 2008 / Accepted: 29 August 2008 / Published online: 20 September 2008

Abstract Intestinal mucins are very high molecular weight glycoproteins secreted by goblet cells lining the crypt and the surface of the colonic mucosa. Profound alterations of mucin *O*-glycans are observed in diseases such as cancer and inflammation, modifying the function of the cell and its antigenic and adhesive properties. Based on immunohistochemical studies, certain cancer- and inflammation- associated glycans have been defined as oncofetal antigens. However, little or no chemical analysis has allowed the structural elucidation of *O*-glycans expressed on human fetal mucins. In this paper, mucins were isolated from different regions of the normal human intestine (ileum, right, transverse and left colon) of eight fetuses with A, B or O blood group. After alkaline borohydride treatment, the released oligosaccharides were investigated by nanoESI Q-TOF MS/MS (electrospray ionization quadrupole time-of-flight tandem mass spectrometry). More than 117 different glycans were identified, mainly based on core 2 structures. Some core 1, 3 and 4 oligosaccharides were also found. Most of the structures were acidic with NeuAc residues mainly α2–6 linked to the *N*-acetylgalactosaminitol and sulphate residues 3-linked to galactose or 6-linked to GlcNAc. In contrast to adult human intestinal mucins, Sda/Cad determinants were not expressed on fetal mucin *O*-glycans and the presence of an acidic gradient along the intestinal tract was not observed. Similar patterns of glycosylation were found in each part of the intestine and the level of expression of the major oligosaccharides was in the same order of magnitude. This study could help determining new oncofetal antigens, which can be exploited for the diagnosis or the treatment of intestinal diseases.

Keywords Human fetal mucins · Intestine · Glycosylation · Mass spectrometry

C. Robbe-Masselot (✉) · E. Maes · J.-C. Michalski · C. Capon
Unité de Glycobiologie Structurale et Fonctionnelle,
UMR CNRS/USTL 8576, IFR 147,
Université des Sciences et Technologies de Lille,
59655 Villeneuve d'Ascq Cedex, France
e-mail: catherine.robbe@univ-lille1.fr

M. Rousset
UMR INSERM U505/UPMC, Institut Biomédical des Cordeliers,
15 rue de l'Ecole de Médecine,
75006 Paris, France

Introduction

Mucins are very high molecular weight glycoproteins secreted by mucosae or some exocrine glands into the lumen of the respiratory, gastrointestinal and reproductive tracts [1–2]. They consist of a protein backbone containing a central domain of amino acids repeated in tandem with extensive *O*-glycosylation (up to 80% of the total weight).

Intestinal mucins are mainly produced by goblet cells, but are also found in 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, which have been regarded as having a protective and rheological function at mucosal surfaces [3]. Mucins provide a physical barrier between the underlying epithelium and luminal contents such as toxins, pathogenic bacteria, viruses and parasites [4]. It also protects the epithelium from attack by pancreatic digestive enzymes and bacterial proteases.

Consistent data indicate that mucin genes are expressed in a regulated cell- and tissue- specific manner in the intestine. MUC2 is expressed only in goblet cells in the intestine, whereas MUC3 is expressed in both goblet cells and absorptive cells [5]. Moreover, MUC4 is widely

expressed in the gastrointestinal tract including the colon and the small intestine [6–7], whereas MUC5B is weakly expressed in the goblet cells of the colon, but absent in other parts of the digestive tract [8]. MUC6 is expressed in a wide variety of digestive epithelial tissues including colon and duodenum [2]. Several structural studies have allowed the complete elucidation of the repertoire of *O*-glycans in human intestinal mucins [9–13]. Oligosaccharides are mainly based on sialylated core 3 structures (GlcNAcβ1-3 [NeuAcα2-6]GalNAcβ1-Ser/Thr) and carried specific determinants such as blood group antigens recovered exclusively in the ileum and the proximal colon and Sda/Cad blood group determinants found in the distal part of the colon. These studies have demonstrated the establishment of an acidic gradient along the adult intestinal tract and a high level of expression of specific glycans for each intestinal region, which may be directly related to the specialized biological functions played by these molecules in the human intestine. It has been proposed that mucin *O*-glycans create an enormous repertoire of potential binding sites for microorganisms that could explain the region-specific colonization of bacteria in the intestine.

In fetal intestine, mucus secretion begins as early as 10 weeks after gestation [14]. This affords protection against bile and swallowed amniotic fluid, and contributes to meconium formation. Two studies have demonstrated both temporal and spatial differences in the expression of MUC2, MUC4 and MUC5AC in the intestine of embryos and fetuses, compared with normal adults, suggesting a role for these genes in the normal development of fetal intestine. MUC2 expression is cell lineage dependent, since it is expressed only in mature goblet cells. MUC4 expression seems to be linked to the epithelial cell position along the jejunum to colon axis and MUC5AC expression is stage dependent as it is closely related to intestinal development. No structural information is currently available on the repertoire of *O*-glycans in fetal human intestinal mucins, with the exception of a few studies on human meconium glycoproteins [15–18]. However, using immunohistochemical techniques, certain cancer- and inflammation- associated glycans, such as Tn (GalNAcβ1-Ser/Thr), TF (Galβ1-3GalNAcβ1-Ser/Thr) and their sialylated counterparts, have been defined as oncofetal antigens since a neo-expression of these glycans, normally restricted to embryonic tissues, was observed in cancerous tissues [19].

In this paper, we have used nanoESI Q-TOF MS/MS to investigate the pattern of glycosylation of human fetal intestinal mucins purified from different regions of the gut: right, transverse and left colon and ileum, from eight fetuses between 24 and 40 weeks of gestation and with blood group A, B or O. The results demonstrate a remarkable degree of heterogeneity in human fetal mucins. Above all there is a similar repertoire of *O*-glycans between individuals and at the different parts of the intestine. This suggests that region-specific glycosylation of intestinal mucins is acquired after birth and depends on environmental factors such as nutrition or microflora colonization. By comparing glycosylation of intestinal mucins isolated from fetuses, normal adult and patients suffering with different pathologies such as inflammatory bowel diseases or cancer, it may be now possible to identify new oncofetal antigens, which can be useful for both the diagnosis or the treatment of such diseases.

Materials and methods

Human fetal samples and mucin preparation Tissues from eight human fetuses were obtained after spontaneous or therapeutic abortion according to protocols approved by the National Ethical Committee. Tissues were collected by Dr. A. Zweibaum. Samples of mucosa were snapfrozen in liquid nitrogen and stored in liquid nitrogen until used. The human samples used to analyse the different parts of the gut were obtained from three female and five male donors aged 24 to 40 weeks of gestation, with A, B or O Le^b blood group and colon type W+Z+, two antigenic specificities, called W and Z by Zweibaum *et al*. [20] and described in normal colonic secretions. All fetuses were ABH secretors; blood group activity was detected by haemagglutination inhibition tests and confirmed by immunofluorescence on small intestine sections. Briefly, crude colon extracts at 10 mg/ml in saline, were used for haemagglutination inhibition tests in the following systems: A anti-A, B anti-B, Le^a anti-Le^a and Le^b anti-Le^b with human serum. For H inhibiting activity we used the lectin *Ulex Europus*. The inhibition of the different reagents was done at specific agglutinin titers of 1/16.

After thawing, the tissue was kept at 4°C, homogenized in distilled water (Ultra-turrax, Jankee and Kunkel, Stauffer, 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 [21].

Human adult samples and mucin preparation All 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 snapfrozen 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 2 male donors with ALe^b blood group and colon type W+Z+. Mucins were purified as described above.

Slot-blotting of mucin proteins Polyclonal antibodies (LUM2-3, LUM5B-3, LUM5-1 and LUM6-1) used in this study were kind gifts from Dr. Ingemar Carlstedt, University of Lund, Sweden. Aliquots of purified mucins were slot-blotted onto nitrocellulose membranes. The membranes were then blocked with 0.5% (*w*/*v*) skimmed milk in TBS containing 0.1% (*v*/*v*) Tween 20 (blocking solution) for 1 h and incubated with LUM2-3 antiserum (1:1,000), LUM5B-3 (1:1,000), LUM5-1 (1:1,000) and LUM6-1 (1:1,000) diluted in blocking solution. Bound antibody was detected by incubation with horseradish peroxidase-conjugated anti-rabbit (1:2,000) antibody in blocking solution for 1 h followed by the ECL Western detection kit. All incubations were carried out at room temperature.

Hydrolysis of sialic acids and derivatization with DMB (4,5-methylene dioxy1,2-phenylene diamine dihydro-chloride) The glycoprotein-bound sialic acids were hydrolyzed [2 h at 80°C in H_2O/TFA 0.1% (*v*/*v*)] and lyophilized. Lyophilized samples were then heated at 50°C for 2 h in the dark in 7 mM DMB, 1 M 2-mercaptoethanol and 18 mM sodium hydrosulfite in 1 N TFA [22].

High performance liquid chromatography of derivatized sialic acids DMB derivatized sialic acids were separated by RP-HPLC using a C18 column (C18 Hypergold, Alltech, 250×4.6 mm, particle size 5 μm). Elution was achieved in isocratic mode, using a mixture of 9/7/84 of acetonitrile/methanol/water over 40 min, at 1 ml/min flow. Derivatized sialic acids were detected by UV spectroscopy at 373 nm using an UVD 170U detector (Dionex, Sunnyvale, CA, USA).

Release of oligosaccharide alditols from mucin by alkaline borohydride treatment The intestinal mucins were submitted to β-elimination under reductive conditions (0.1 M KOH, 1 M KBH_4 for 24 h at 45°C [23]). The mixture of oligosaccharide alditols was purified by size exclusion chromatography on a column of Bio-Gel P2 (85×2 cm ID, 400 mesh, Bio-Rad, Richmond, CA, USA) equilibrated and eluted with water (10 ml/h) at room temperature. The oligosaccharide fractions, detected by UV absorption at 206 nm, were pooled for structural analysis.

Monosaccharide analysis The monosaccharide composition of the mucins and oligosaccharide alditol fractions were determined by gas chromatography (GC) on a Shimadzu gas chromatograph equipped with a 25 m× 0.32 mm CP-Sil5 CB Low bleed/MS capillary column, 0.25 μm film phase (Chrompack France, Les Ullis, France) after methanolysis (0.5 M HCl–methanol for 24 h at 80°C), *N*-reacetylation and trimethylsilylation [24–25].

Fractionation of the oligosaccharide alditols by HPLC The oligosaccharide alditols released from each part of the fetal intestine were subjected to fractionation by HPLC (Dionex Chromeleon System) on a primary amino-bonded silica column (Supelcosyl, LC-NH_2, 4.6×250 mm, Supelco, Bellefonte, CA, USA) using a mixture of acetonitrile/H_2PO_4K·30 mM/H_2O (75:0:25, by vol.) to (50:50:0, by vol.) in 60 min with a flow rate of 1 ml/min. Oligosaccharides were detected by UV spectroscopy at 200 nm using an UVD 170U detector (Dionex).

Electrospray mass spectrometry (nanoESI-MS/MS) All analyses were performed on a Q-STAR Pulsar quadrupole time-of-flight (Q-q-TOF) mass spectrometer (Applied Biosystems/MDS Sciex, Toronto, Canada) fitted with a nanoelectrospray ion source (Protana, Odense, Denmark). Oligosaccharides dissolved in water (60 pmol/μl) were acidified by addition of an equal volume of methanol/0.1% formic acid and sprayed from gold-coated "medium length" borosilicate capillaries (Protana). A potential of −800 V was applied to the capillary tip and the focusing potential was set at −100 V, the declustering potential varying between −60 and −110 V. For the recording of conventional mass spectra, time-of-flight data were acquired by accumulation of ten MCA (multiple channel acquisition) scans over mass ranges of *m*/*z* 400–2,000. In the collision-induced dissociation (CID) tandem MS analyses, multiple charged ions were fragmented using nitrogen as collision gas (5.3×10^{-5} Torr), the collision energy varying between −40 and −90 eV to obtain optimal fragmentation. The CID spectra were recorded on the orthogonal TOF analyzer over a range of *m*/*z* 80–2,000. Data acquisition was optimized to supply the highest possible resolution and the best signal-to-noise ratio even in the case of low abundance signals. Typically, the full width at half maximum (FWHM) was 7,000 in the measured mass ranges. External calibration was performed prior to each measure using a 4 pmol/μl solution of taurocholic acid in acetonitrile/water (50:50, *v*/*v*) containing 2 mM of ammonium acetate.

Nuclear magnetic resonance (NMR) spectroscopy Samples were repeatedly treated with ^2H_2O (99.97% ^2H atoms, Euriso-top, CEA, Saclay, France). ^1H chemical shifts were expressed in parts/million downfield from 4, 4-dimethyl-4-silapentanoate sodium salt (TMSP), but were measured by reference to the internal standard acetone (δ=2.225 ppm). Samples were analysed in 200×5-mm BMS-005B Shigemi® tubes at 300 K on a Bruker® 9.4 T spectrometer equipped with a double resonance (^1H/X) broadband inverse z-gradient probe head. NMR samples were recorded without sample spinning.

Results

Mucins were isolated from the different parts of normal human intestinal tract (ileum, right colon, transverse and left colon) of eight fetuses from 24 to 40 weeks of gestation (Table 1).

Slot blotting of mucin proteins We have studied the expression of four mucin proteins: MUC2, MUC5AC, MUC5B and MUC6 in eight fetuses at different gestational ages in order to establish a possible relationship between their pattern of expression and the repertoire of glycosylation of fetal mucins. MUC2 proteins were detected from all fetuses, whereas no reactivity was observed with the anti-MUC5AC antibody in any of the fetuses (data not shown). These results were in agreement with a previous study published by Buisine *et al.* [26] showing that MUC2 gene was expressed as early as nine weeks of gestation and that the developmental switch from a fetal to an adult pattern of MUC2 gene expression occurs at around 25 weeks. On the contrary, after 12 weeks, MUC5AC was not detected in any region of the intestine.

Reactivity for MUC5B was recovered in all fetuses, whereas MUC6 was expressed only in fetuses at around 29 weeks of gestation (*i.e.* individuals 1, 2, 5, 6 and 8). In the previous study [26], where fetuses ranged in age from 6.5 to 27 weeks of gestation, MUC5B and MUC6 were not observed in fetal intestine by *in situ* hybridization. Our results may suggest a late expression of these two proteins in the fetal intestine.

Analysis of the sialic acid diversity Hydrolyzed sialic acids were derivatized with DMB to produce 3-substituted 6,7 methylene dioxy-2 (1H)-quinoxalinone fluorescent derivatives [27]. These compounds were then separated and analysed by RP-HPLC. Samples from human adult intestinal mucins were treated under the same conditions in order to compare the level of *O*-acetylation found in adult and fetal mucins. As shown in Table 2, many different types of sialic acids were recovered in both adult and fetal mucins and the most abundant component in each sample was the Neu5Ac, as expected. Some Neu5Gc was also recovered in these human mucins. This result is in agreement with recent studies demonstrating that free *N*-glycolyl neuraminic acid, originating from dietary sources, can be incorporated into human cells [28]. Among the di-acetylated sialic acids, Neu5,9Ac_2 was the major constituent in both mucins whereas Neu5,7Ac_2 was only slightly expressed. In fetal mucins, the level of expression of Neu5,8Ac_2 decreased from ileum to distal colon, whereas the reverse gradient was observed for adult mucins. This Neu5,8Ac_2 was found at higher levels in fetuses than in adult intestine. Neu5,9Ac_2 and Neu5,7(8)Ac_3 were more common in proximal colon than in small intestine and distal colon. In contrast to adult mucins, where around 30% of total sialic acids were *O*-acetylated, only approximatively 20% of total sialic acids are *O*-acetylated in fetuses.

Monosaccharide composition of the mucins Purified mucins isolated from each part of the intestine were analysed for their carbohydrate composition together with oligosaccharides released by base/borohydride treatment followed by desalting. The percent ages of saccharides that were recovered after release from the protein backbone were estimated based on monosaccharide compositional analyses of the whole mucin and of the released fraction. The yield was approximately 50%, a typical yield for this procedure.

As shown in Table 3, the level of sialic acid slightly increased from ileum to distal colon. Similar variations were also observed for Gal and GlcNAc residues. In contrast to human adult intestinal mucins [12], a decreasing gradient in the molar ratio of fucose was not found in fetal mucins. Fetuses analysed in this study were of blood group A, B or O. Interestingly, a higher level of unreduced GalNAc was detected in fetuses with blood group A (individuals 1, 2, 4, 5 and 6): around 1 GalNAc per GalNAcol versus 0.3 GalNAc per GalNAcol for B and O blood groups individuals. This result suggests that fetal intestinal mucins do express blood group determinants and is in agreement with the fact that all individuals are secretor positive.

Table 1 Blood group status, age and part of the human intestine of the eight fetuses included in the study

Donor	ABO blood group	Sex	Age (weeks)	Ileum	Right colon	Transverse	Left colon
1	A	F	30	+	+	+	+
2	A	M	40	+	+	+	+
3	O	M	24	+	Whole colon		
4	A	F	29	+	Whole colon		
5	A	M	40	+	+	+	+
6	A	M	29	+	−	−	−
7	B	F	24	+	+	−	+
8	O	M	40	+	+	−	+

Table 2 Sialic acid composition of the mucins from different parts of the gut

Sialic acid	Ileum	Cecum	Transverse	Rectum
Donor 1-A[a]				
Neu5Gc	0	0.3	0.3	1.8
Neu5Ac	85	77.1	83.1	82.4
$Neu5,7Ac_2$	0.9	1.3	1.3	1
Neu5Gc9Ac	0	0	0	0
$Neu5,8Ac_2$	5.9	4.9	4.8	4
$Neu5,9Ac_2$	6.7	14.7	9.3	10.3
$Neu5Gc7(8),9Ac_2$	0.3	0.3	0.4	0.2
$Neu5,7(8),9Ac_3$	1.2	1.4	0.8	0.3
Donor 7-B[b]				
Neu5Gc	0.5	0.5		0.4
Neu5Ac	84.7	80.7		85.3
$Neu5,7Ac_2$	0.5	1.7		0.7
Neu5Gc9Ac	0	0		0
$Neu5,8Ac_2$	5	4.1		3.6
$Neu5,9Ac_2$	6.4	8.6		7.3
$Neu5Gc7(8),9Ac_2$	0.5	0.3		0.4
$Neu5,7(8),9Ac_3$	2.4	3.1		2.3
Donor 8-O[c]				
Neu5Gc	0.2	0.3		0.5
Neu5Ac	81.8	83.6		83.8
$Neu5,7Ac_2$	0.6	0.3		0.4
Neu5Gc9Ac	0	0		0
$Neu5,8Ac_2$	5.2	1.8		1.9
$Neu5,9Ac_2$	7.1	8.9		8.5
$Neu5Gc7(8),9Ac_2$	0.6	0.2		0.5
$Neu5,7(8),9Ac_3$	4.5	4.9		4.4
Donor adult				
Neu5Gc	0.35	3.1	1.5	0.6
Neu5Ac	84.1	65.5	71.5	71
$Neu5,7Ac_2$	–	2.1	1.2	1.2
Neu5Gc9Ac	0.45	0.4	0.3	0.6
$Neu5,8Ac_2$	3	3.9	4.2	4.65
$Neu5,9Ac_2$	10	14.8	14.8	16.9
$Neu5Gc7(8),9Ac_2$	0.6	4.2	2.2	0.75
$Neu5,7(8),9Ac_3$	1.5	6	4.3	4.3

Results are expressed as percentage of total sialic acids in each sample
[a] Donor 1 with blood group A
[b] Donor 7 with blood group B
[c] Donor 8 with blood group O

HPLC fractionation Mass spectrometry is not a quantitative method, since signal intensity is mainly dependent on the charge state and the size of the molecule. Thus, oligosaccharides released from each part of the intestine were fractionated by HPLC on a primary amino-bounded column, which was useful for the determination of the relative abundance of major oligosaccharides. Fractions collected from donor 2 with blood group A were analyzed by NMR spectroscopy and nano ESI Q-TOF MS in order to identify the different oligosaccharides. The amount of *O*-glycans in each fraction was semi-quantified by integration of the individual peak areas in the chromatograms. Detection of glycans by UV at 206 nm is not as sensitive as MS and weakly expressed oligosaccharides cannot be identified. As shown in Table 4 and Fig. 1, major oligosaccharides in each fraction were neutral and corresponded mainly to structures based on core 2 (ions at *m*/*z* 773, 919, 960, 1,122 and 1,268 on MS spectra). Expression of these glycans represented around 25% of total major oligosaccharides.

Some differences in the level of expression of oligosaccharides were observed along the intestinal tract. For example, sialyl Tn (ion at *m*/*z* 537 in the positive ion mode) was more expressed in ileum and distal colon than in the proximal colon. On the contrary, sialyl TF (ion at *m*/*z* 699 in the positive ion mode) was more expressed in the proximal colon than in ileum and distal colon. The sialylated core 5 (NeuAcα2-6[GalNAcβ1-3]GalNAcol) was the major oligosaccharide in ileum: around 20%, whereas it represented only 4–5% of total expression of glycans in other parts of the intestine. Interestingly, the ion at *m*/*z* 737 in the positive ion mode, carrying blood group A determinant was mainly expressed in transverse and proximal colon.

A similar pattern of glycosylation in the different regions of fetal intestinal mucins After the desalting step, oligosaccharide mixtures isolated from the different parts of fetal intestine were analysed directly by nanoESI Q-TOF tandem mass spectrometry, without prior derivatization or fractionation step [29]. Fragment annotations applied in this study were based upon the suggested nomenclature by Domon and Costello [30] and by Karlsson *et al.* [31]. The fragment ions obtained were mainly A_i, B_i, Y_j, Z_j ions. Moreover, an α suffix was used to designate cleavages in the 6-linked branch and a β suffix for cleavage in the 3-linked branch from the GalNAcol.

Sequencing of oligosaccharides was acquired in both positive and negative ion modes to obtain complete elucidation of the structure for each mucin oligosaccharide. Positive ion mode is useful for the determination of the core-type oligosaccharide and the position of fucose, whereas negative ion mode produces characteristic cross-ring cleavages necessary for differentiation between isomers and identification of NeuAc (*N*-acetylneuraminate) and sulfate substitution. For example, a specific $^{0,2}A_i$ cleavage with concomitant loss of water from a GlcNAc substituted on C-4 allowed to differentiate between type 1 (Galβ1-3GlcNAc) and type 2 (Galβ1-4GlcNAc) chains. Some characteristic ions also allowed assigning unambiguously the presence of certain terminal structures such as the Lewis determinants and the Sda/Cad antigens. However, in complex oligosaccharides, diagnostic ions are sometimes of lower intensity, rendering the complete elucidation of the structure difficult or impossible.

Table 3 Monosaccharide composition of the native mucins and oligosaccharides isolated from the different mucins of three donors with blood group A, B or O after a β-elimination procedure

	Fuc	Gal	GalNAc	GlcNAc	NeuAc	GalNAcol
Native mucin populations[a]						
1-A[c]						
Ileum	0.9	1.3	1.0	1.1	0.5	–
Right colon	0.9	1.6	1.0	1.6	0.7	–
Transverse	0.94	2.2	1.0	1.9	0.6	–
Left colon	1.0	2.2	1.0	1.9	0.6	–
7-B[d]						
Ileum	1.1	2.2	1.0	1.6	0.7	–
Right colon	0.9	2.1	1.0	1.9	1.0	–
Left colon	0.5	2.1	1.0	1.9	1.0	–
8-O[e]						
Ileum	1.9	2.7	1.0	2.3	0.5	–
Right colon	1.7	2.8	1.0	2.5	0.6	–
Left colon	1.5	2.7	1.0	2.4	0.7	–
Released oligosaccharides[b]						
2-A						
Ileum	1.7	2.6	1.1	2.2	0.9	1.0
Right colon	1.9	4.1	1.2	3.3	1.4	1.0
Transverse	1.75	3.2	0.7	2.7	0.9	1.0
Left colon	2.4	4.8	1.0	4.2	1.2	1.0
7-B						
Ileum	2.2	4.0	0.8	2.9	0.9	1.0
Right colon	1.6	3.9	0.7	2.9	1.0	1.0
Left colon	0.8	3.2	0.3	2.6	1.3	1.0
8-O						
Ileum	2.1	3.0	0.3	2.5	0.6	1.0
Right colon	1.7	3.6	0.4	3.1	0.7	1.0
Left colon	1.8	3.4	0.4	2.9	0.7	1.0

[a] The molar ratio of the different monosaccharides was calculated on the basis of one GalNAc residue for native mucins
[b] The molar ratio of the different monosaccharides was calculated on the basis of one GalNAcol residue for the oligosaccharides released from the native mucin populations
[c] Donor 1 with blood group A
[d] Donor 7 with blood group B
[e] Donor 8 with blood group O

Figure 2 represents the MS spectra acquired in negative ion mode for the total oligosaccharide fractions from the four mucin populations (ileum, right colon, transverse and left colon) from individual 2 with blood group A. Most of the ions could be related to $[M-H]^-$ ions of the theoretical oligosaccharide structures. The spectra illustrate a remarkable structural diversity: the true heterogeneity is likely to be even larger, since a substantial number of the molecular ions may represent oligosaccharide species with isomeric structures. To avoid confusion, the mass values used in the text, Tables and illustrated fragmentations are mainly nominal masses. As shown in Tables 4, 5 and 6 and Fig. 2, the pattern of glycosylation for the four samples was quite similar and showed significant heterogeneity with more than 117 different structures identified by MS analysis. Most of them were neutral oligosaccharides, mainly based on core 2 structures (Galβ1-3[GlcNAcβ1-6]GalNAcol). Numerous fucosylated oligosaccharides carried the blood group H determinant (Fucα1-2Galβ1). Cores 1, 3 and 4 were also detected in these neutral mucins. Acidic glycans were mainly based on sialylated core 3 structures: GlcNAcβ1-3[NeuAcα2-6]GalNAcol, as found in human adult intestinal mucins [9–13]. NeuAc residues were found α2–6 linked to the initial GalNAc and α2–3 linked to peripheral galactose residues. They were easily differentiated on MS/MS spectra by characteristic diagnostic ions at *m/z* 306 and 513 for NeuAcα2-6 and 408 or 611 for NeuAcα2-3 linked to a Gal residue. Sulfate groups were linked to the C-6 of GlcNAc (identified in the MS/MS spectra by diagnostic ions at *m/z* 199 and 282) or to the C-3 of peripheral Gal (identified by ion at *m/z* 241). Fucose was present in a large variety of terminal linkages that included blood group H as well as Le^a (Galβ(1-3[Fucα1-4]GlcNAc), Le^b (Fucα1-2Galβ1-3[Fucα1-4]GlcNAc), Le^x (Galβ(1-4

Table 4 Estimated abundance of the major oligosaccharide structures identified in fetal mucins of donor 2 with blood group A

Fraction	Retention time	$[M-H]^-$	$[M+Na]^+$	% I	% Right C	% TC	% Left C
1	20.4	425	449	5.4	1.1	2.3	3.3
2	21.6	384	408	1.9	0.7	2.6	1.6
		425	449				
3	23.3	530	554	0.9	1.4	2.0	–
4	28.4	587	611	0.2	9.3	2.9	1.0
6	35.5	587	611	–	3.6	1.7	1.3
7	36.8	587	611	0.8	1.4	2.3	0.8
8	38.6	513	537	7.4	4.5	2.3	5.9
9	40.7	675	699	1.8	6.4	4.2	0.3
10	44.9	733	757	1.0	0.2	4.1	7.0
11	46.4	733[a]	757[a]	2.0	0.2	1.0	4.1
12	50.7	716[b]	740[b]	18.8	5.2	4.8	4.3
13	53.1	716[c]	740[c]	6.6	14.4	6.3	4.6
		733	757				
14	55.6	675	699	3.3	3.6	5.7	0.8
		733[d]	757[d]				
15	62.3	749	773	11.7	12.3	16.2	17.4
		895	919				
		936	960				
17	69.3	1,041	1,065	0.5	n.d.	1.8	3.1
18	71.6	1,040	1,064	0.8	2.8	3.4	8.6
20	75.3	1,227	1,251	7.7	5.8	8.5	6.9
21	77.0	966	990	0.4	1.9	3.2	–
		1,186	1,210				
		1,227	1,251				
		1,243	1,267				
		1,323	1,347				
22	78.3	1,098	1,122	13.2	12.6	8.0	5.7
		1,244	1,268				
23	80.3	1,331	1,355	0.2	1.0	0.8	0.2
		1,389	1,413				
		1,447	1,471				
24	80.7	1,331	1,355	4.4	1.0	1.4	3.4
26	84.2	1,331	1,355	4.3	3.3	7.3	3.7
		1,372	1,396				
		1,406	1,430				
29	89.4	1,315	1,339	1.1	1.3	4.8	4.1
		1,477	1,501				
		1,795	1,819				
30	92.2	1,729	1,753	3.0	n.d.	n.d.	3.5
		2,135	2,159				
31	93	1,639	1,663	1.3	n.d.	n.d.	3.4
		1,729	1,753				
		1,843	1,867				

The relative percentage of each oligosaccharide was calculated based on peak intensity in HPLC spectra. In fractions containing more than one compound, the percentage noted corresponds to the sum of the compounds present in the fraction

nd not determined

[a] Structure carrying a blood group A determinant

[b] Ion at *m/z* 716 corresponding to a sialylated core 5

[c] Ion at *m/z* 716 corresponding to a sialylated core 3

[d] Structure carrying a Le^x determinant

[Fucα1-3]GlcNAc) and Le^y (Fucα1-2Galβ1-4[Fucα1-3] GlcNAc) determinants.

As expected, some described oncofetal antigens were present in fetal intestinal mucins, such as the TF antigen (Galβ1-3GalNAcol) at *m/z* 408 in the positive ion mode, the sialyl TF (Galβ1-3[NeuAcα1-6]GalNAcol) at *m/z* 675 and the sialyl-Tn (NeuAcα2-6GalNAcol) at *m/z* 513 in the negative ion mode.

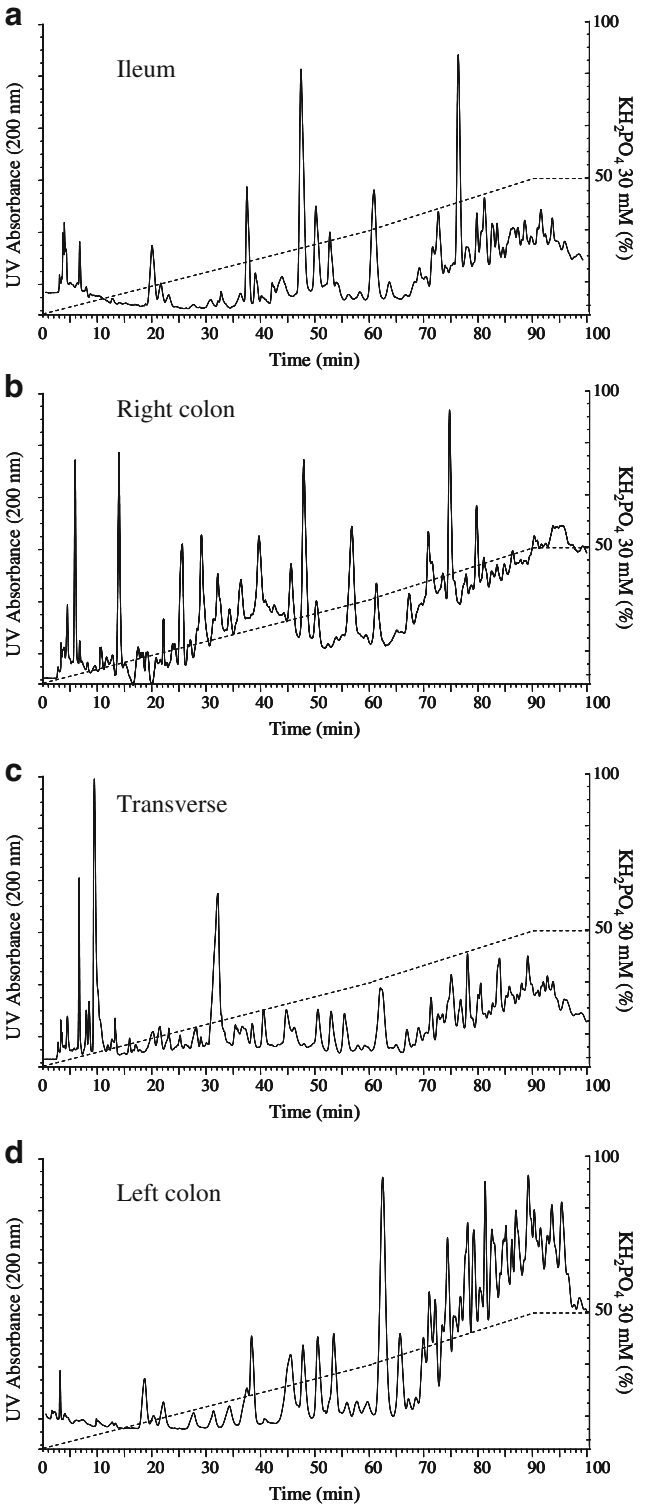
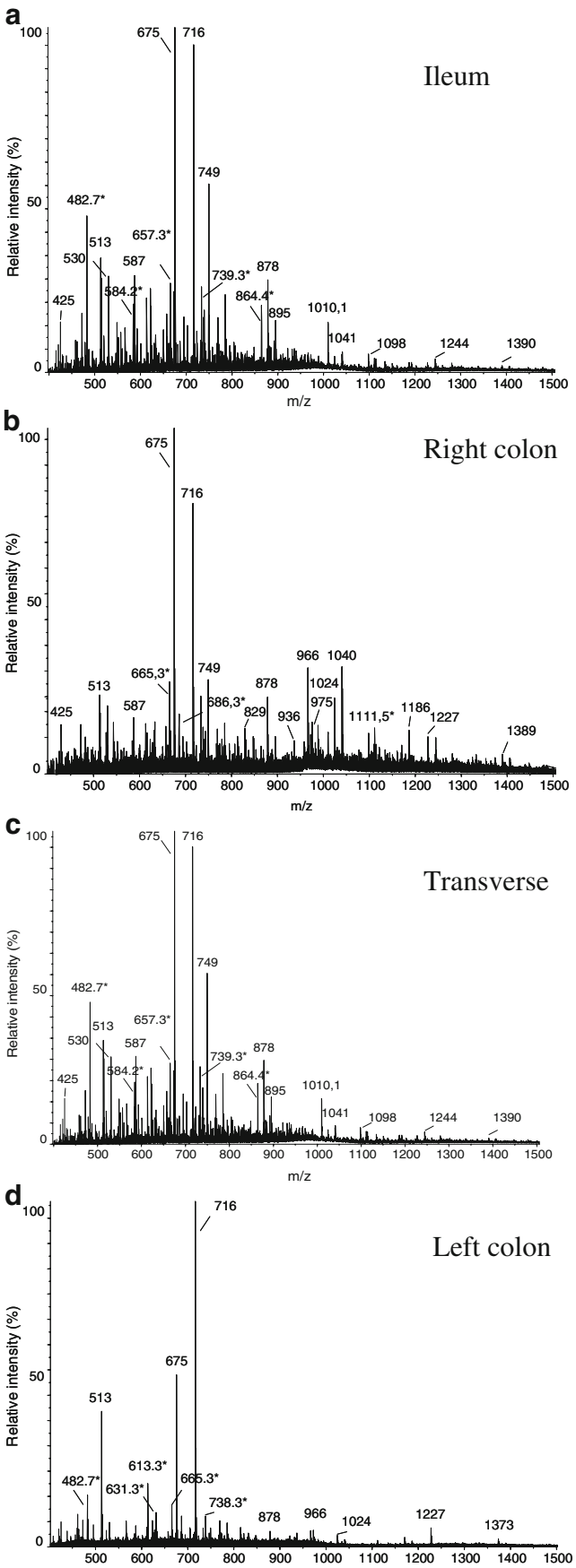


Fig. 1 HPLC chromatograms of fetal mucins from ileum (**a**), right colon (**b**), transverse (**c**) and left colon (**d**) from donor 2 with blood group A. Oligosaccharide alditols were separated on a primary amino-bonded column eluted with gradient described under "Materials and methods"

Expression of blood group antigens in fetal intestinal mucins In contrast to human adult intestinal mucins, where blood group antigens are expressed only in the ileum and the proximal colon [12], these specific determinants were recovered from each part of the fetal intestine. Because mass spectrometry does not distinguish between a terminal GalNAc (blood group A determinant) or GlcNAc residue and between a Galα1-linked or a Galβ1-linked (blood group B determinant), oligosaccharides released from each part of the intestine, from individuals 2 (blood group A) and 7 (blood group B), were fractionated by HPLC and the fractions collected were analysed by NMR spectroscopy. As shown in Fig. 3 and Table 7, this physical method allowed us to confirm the expression of blood group antigens in colonic fetal mucins. For example, we report here the NMR data of the oligosaccharide at *m/z* 733 in the negative ion mode, with the blood group A structure. Starting from ^{1}H chemical shifts (δ), three anomeric protons have been identified at 5.393, 5.189 and 4.716 ppm, showing the presence of three monosaccharides and one GalNAcol residue. This last unit is easily identified by its classical spin system (H2/H1,1′, H5/H6,6′, H3/H4) read on the 2D spectrum and the presence of a *N*-acetyl group at 2.048 ppm. On the basis of its proton chemical shifts reported in Table 7 and the literature we can confirm that GalNAcol is *O*3-substitued by a β-Gal*p* residue forming the classical core 1. The β-Gal*p* is identified by its typical anomeric chemical shifts at 4.716 ppm and its spin system read on the COSY spectrum (Fig. 2).

The 2D-COSY spectrum allowed us to determine the four first vicinal coupling constants (^{3}J) defining the configuration of each residue like following example. Starting from the anomeric proton at δ 5.189, the correlated H2 was identified at 4.243 ppm as a quadruplet showing a large vicinal coupling constant ($^{3}J_{H2,H3}$). This H2 is correlated with a H3 proton at 3.922 as a quadruplet defining a small $^{3}J_{H3,H4}$, finally this proton is itself correlated with H4 at 4.019 defining as a pseudosinglet with a small $^{3}J_{H4,H5}$. On the basis on these four coupling constants defined as small, large, small and small, this monosaccharide has an α-*Galacto* configuration [32]. Moreover, because the H2 is highly deshielded, this indicates that H2 is carried by a carbon substituted by a *N*-acetamido group (δ 2.046), confirming the presence of an α-GalNAc residue. The chemical shifts of H3 and H4 confirmed that GalNAc is in an unreduced terminal position. Finally, the proton at 5.393 ppm is identified as a terminal α-Fuc residue on the basis of its δH1, δH2 and the presence of methyl group at δ 1.236.

Taking into account the mass results, sugar composition analysis and structural reported groups defined by Kamerling and Vliegenthart [33], the oligosaccharide has the following sequence: GalNAc(α1-3)[Fuc(α1-2)]Gal(β1-3)



◀ **Fig. 2** Glycosylation pattern of fetal mucins from ileum (**a**), right colon (**b**), transverse (**c**) and left colon (**d**) from donor 1. MS spectra of the total oligosaccharides acquired in the negative ion mode $[M-H]^-$. Signals marked with *asterisks* referred to dicharged ions $[M-2H]^{2-}$

GalNAcol. This sequence is already reported in the sweet-DB database under the Linucs number 639 [34] and is widely quoted in the literature [35].

Discussion

Several studies have analysed the glycosylation of human meconium glycoproteins, showing expression of multiple blood group activities and oligosaccharides based on core 1, 2, 3, 4 and 5 structures [15–18]. However, to our knowledge, no study has dealt with the glycosylation of fetal mucins along the gut. In the present paper, we provide the first detailed analysis of normal fetal mucin glycans originating from ileum, proximal colon, transverse and distal colon. More than 117 different oligosaccharides were identified in such mucins, mainly based on core 2 structures. The data agree with previous work on meconium glycoproteins in that all the structures described before have been recovered in this study. Many more structures have been identified in the present work. This can be readily explained by the use of nanoESI mass spectrometry methodology, which is known to be more sensitive than NMR or FAB-MS.

This study shows a very large repertoire of *O*-glycans in fetal mucins. In contrast to adult intestinal mucins, the same structures are identified in each part of the fetal intestinal tract and only slight variation in the level of expression of certain glycans is observed. Small individual variability between the eight fetuses studied has been observed, concerning mainly structures expressing blood group antigens and this can be easily explained by the fact that all individuals studied were secretor positive. The limited individual variability has previously been described in adult intestinal mucins even in distal colon, where no expression of blood group antigens occurs [13]. This is in contrast to human salivary or cervix mucins, where for example blood group and secretor status are partly responsible of a high diversity [36, 37]. No explanation for the preserved glycosylation between individuals in the intestine, whereas an extensive variability is observed for mucins at other locations, has been proposed until now. However, it may be speculated that glycans originating from mucins of different tissue origins will play different roles. Whereas glycans from intestinal mucins are believed to help in the selection for a specific commensal flora, which is similar between individuals [38], oligosaccharides from respiratory or salivary mucins are necessary to trap a broad diversity of pathogenic microorganisms, thereby impeding infection of the underly-

Table 5 Proposed neutral oligosaccharide structures identified in fetal intestinal mucins by nanoESI MS/MS in positive ion mode

Sequence/composition of oligosaccharide alditols	[M+Na]+	Ileum			Right colon			Left colon		
		1-A	7-B	8-O	1-A	7-B	8-O	1-A	7-B	8-O
Gal→3GalNAcol	408	+	+	+	+	+	+	+	+	+
GlcNAc→3GalNAcol	449	+	+	+	+	+	+	+	+	+
GalNAc→3GalNAcol	449	+	+	+	+	+	+	+	+	+
Fuc→2Gal→3GalNAcol	554	+	+	+	+	+	+	+	+	+
Gal→3/4GlcNAc→3GalNAcol	611	+	+	+	+	+	+	+	+	+
Gal→3(GlcNAc→6)GalNAcol	611	+	+	+	+	+	+	+	+	+
GlcNAc→3Gal→3GalNAcol	611	+	+	+	+	+	+	+	+	+
(Fuc→2)Gal→GlcNAc→3GalNAcol	757	+	+	+	+	+	+	+	+	+
Gal→3(Fuc→4)GlcNAc→3GalNAcol	757	+	+	+	+	+	+	+	+	+
GlcNAc→3(Fuc→2)Gal→3GalNAcol	757	+	+	+	+	+	+	+	+	+
GalNAc→3(Fuc→2)Gal→3GalNAcol	757	+	−	−	+	−	−	+	−	−
(Fuc→2)Gal→(GlcNAc→6)GalNAcol	757	+	+	+	+	+	+	+	+	+
Gal→3(Gal→4GlcNAc→6)GalNAcol	773	+	+	+	+	+	+	+	+	+
Gal→4GlcNAc→3Gal→3GalNAcol	773	+	+	+	+	+	+	+	+	+
(Fuc→2)Gal→3(Fuc→4)GlcNAc→3GalNAcol	903	+	+	+	+	+	+	+	−	+
(Fuc→)GlcNAc→3(Fuc→2)Gal→3GalNAcol	903	+	−	−	+	+	+	+	−	−
(Fuc→2)Gal→3[(Fuc3/4)GlcNAc→6]GalNAcol	903	+	−	−	+	−	−	+	−	−
2 Gal, GlcNAc, Fuc, GalNAcol	919	+	+	+	+	+	+	+	+	+
(Fuc→2)Gal→4GlcNAc→3Gal→3GalNAcol	919	+	+	+	+	+	+	+	−	+
(Fuc→2)Gal→3(Gal→4GlcNAc→6)GalNAcol	919	+	+	+	+	−	+	+	+	+
Gal→3[(Fuc→2)Gal→4GlcNAc→6]GalNAcol	919	+	+	+	+	−	+	+	−	+
Gal→3[Gal→(Fuc3/4)GlcNAc→6]GalNAcol	919	+	+	+	+	−	+	+	+	+
αGal→3Gal→3/4(Fuc3/4)GlcNAc→3GalNAcol	919	−	+	−	−	+	−	−	+	−
Gal, 2 HexNAc, Fuc, GalNAcol	960	+	+	+	+	+	+	+	+	+
Gal, 2 GlcNAc, Fuc, GalNAcol (core 2)	960	+	−	+	+	−	+	+	+	+
Gal, 2 GlcNAc, Fuc, GalNAcol (core 4)	960	+	+	+	+	+	+	+	−	+
HexNAc(Fuc→2)Gal→GlcNAc→3GalNAcol	960	+	+	+	+	+	nd	+	+	nd
2 Gal, 2 HexNAc, GalNAcol	976	+	−	+	+	−	+	+	−	+
GlcNAc→Gal→3(Gal→4GlcNAc→6)GalNAcol	976	+	−	+	+	−	nd	+	−	nd
2 Gal, HexNAc, 2 Fuc, GalNAcol	1,065	+	+	+	+	+	+	+	+	+
Gal→(Fuc3/4)GlcNAc→(Fuc2)Gal→3GalNAcol	1,065	+	−	+	+	−	nd	+	−	nd
(Fuc→2)Gal→GlcNAc→(Fuc2)Gal→3GalNAcol	1,065	+	−	+	+	−	nd	+	−	nd
Gal→3[(Fuc→2)Gal→3(Fuc→4)GlcNAc→6] GalNAcol	1,065	+	−	−	+	−	+	+	−	+
(Fuc→2)Gal→3[(Fuc→2)Gal→4GlcNAc→6] GalNAcol	1,065	+	−	+	+	−	+	+	−	+
GalNAc→3(Fuc→2)Gal→3(Fuc→4) GlcNAc→3GalNAcol	1,106	+	−	−	+	−	−	+	−	−
2 Gal, 2 HexNAc, Fuc, GalNAcol	1,122	+	−	+	+	+	+	+	−	+
Gal→GlcNAc→(Fuc2)Gal→GlcNAc→3GalNAcol	1,122	+	−	nd	+	nd	nd	+	−	nd
GlcNAc→(Fuc→2)Gal→3(Gal→GlcNAc→6) GalNAcol	1122	+	−	nd	+	nd	nd	+	−	nd
Gal→3[HexNAc→(Fuc→2)Gal→GlcNAc→6] GalNAcol	1,122	+	−	nd	+	nd	nd	+	−	nd
(Fuc→2)Gal→GlcNAc→3(Gal→GlcNAc→6) GalNAcol	1,122	+	−	nd	+	nd	nd	+	−	nd
Gal→3[Gal→GlcNAc→3Gal→GlcNAc→6)] GalNAcol	1,138	+	−	nd	+	nd	nd	+	nd	nd
Gal→GlcNAc→3Gal→3(Gal→GlcNAc→6) GalNAcol	1,138	+	−	nd	+	nd	nd	+	nd	nd
Gal→3GlcNAc→Gal→3GlcNAc→Gal→ 3GalNAcol	1,138	+	−	nd	+	nd	nd	+	nd	nd
αGal→Gal→3GlcNAc→Gal→3GlcNAc→ 3GalNAcol	1138	−	−	−	−	+	−	−	+	−

Table 5 (continued)

Sequence/composition of oligosaccharide alditols	[M+Na]+	Ileum			Right colon			Left colon		
		1-A	7-B	8-O	1-A	7-B	8-O	1-A	7-B	8-O
2 Gal, HexNAc, 3 Fuc, GalNAcol	1,211	+	−	+	+	−	+	+	−	+
(Fuc→2)Gal→3[(Fuc2)Gal→(Fuc3/4)GlcNAc→6] GalNAcol	1,211	+	−	+	+	−	+	+	−	+
(Fuc→2)Gal→(Fuc3/4)GlcNAc→(Fuc→2)Gal→ 3GalNAcol	1,211	+	−	+	+	−	+	+	−	+
2 Gal, 2 HexNAc, 2 Fuc, GalNAcol	1,268	+	+	+	+	+	+	+	+	+
HexNAc→(Fuc→2)Gal→GlcNAc→(Fuc2)Gal→ 3GalNAcol	1,268	+	nd	+	+	nd	+	+	+	+
(Fuc→2)Gal→3[HexNAc→ (Fuc→2)Gal→GlcNAc→6]GalNAcol	1,268	+	nd	+	+	nd	+	+	+	+
(Fuc→2)Gal→3GlcNAc→ 3[(Fuc→2)Gal→4GlcNAc→6]GalNAcol	1,268	−	nd	+	−	nd	+	−	+	+
3 Gal, 2 HexNAc, Fuc, GalNAcol	1,284	+	+	+	−	+	+	+	+	+
2 Gal, 3 HexNAc, 1 Fuc, GalNAcol	1,325	+	−	+	+	+	+	+	+	−
3 Gal, 3 HexNAc, GalNAcol	1,341	+	+	+	+	+	+	+	+	+
2 Gal, 2 HexNAc, 3 Fuc, GalNAcol	1,414	+	−	+	+	−	+	+	+	+
3 Gal, 2 HexNAc, 2 Fuc, GalNAcol	1,430	+	+	+	−	+	+	+	+	+
2 Gal, 3 HexNAc, 2 Fuc, GalNAcol	1,471	+	+	−	−	−	−	+	+	−
3 Gal, 3 HexNAc, Fuc, GalNAcol	1,487	+	−	+	+	+	+	+	+	+
4Gal, 3HexNAc, GalNAcol	1,503	+	+	+	+	+	+	+	+	+
2 Gal, 4 HexNAc, 1 Fuc, GalNAcol	1,528	+	+	+	+	+	+	+	+	+
3 Gal, 2 HexNAc, 3 Fuc, GalNAcol	1,576	+	−	+	−	−	+	−	+	+
2 Gal, 3 HexNAc, 3 Fuc, GalNAcol	1,617	−	−	+	+	+	+	+	+	+
3 Gal, 3 HexNAc, 2 Fuc, GalNAcol	1,633	+	−	−	−	−	−	−	+	−
4 Gal, 3 HexNAc, Fuc, GalNAcol	1,649	+	+	−	−	−	−	−	+	−
2 Gal, 4 HexNAc, 2 Fuc, GalNAcol	1,674	−	−	+	−	+	+	+	+	−
2 Gal, 4 HexNAc, 3 Fuc, GalNAcol	1,820	−	+	+	−	−	−	−	+	−
3 Gal, 4 HexNAc, 2 Fuc, GalNAcol	1,836	−	+	+	−	+	+	−	+	+

ing epithelia. The individual variation of glycosylation of these other mucins has been linked with predisposition to certain bacterial infections such as by *Helicobacter pylori*.

Sialylated core 5 was identified as the major oligosaccharide in the ileum. This result is in agreement with human adult mucins, where this structure represented around 24% of total glycosylation in adult ileum and less than 10% in other part of the intestinal tract [13]. No known biological activity is associated with this structure. Core 5 was previously described by Kurosaka *et al.* [39] in a human rectal adenocarcinoma glycoprotein and by Feeney *et al.* [16] in human meconium. It has long been considered as a tumour marker. However, our results suggest that core 5 glycans are structures recovered in normal human fetal and adult intestinal mucins and are not diagnostic of a pathological situation. It may be hypothesized that core 5 synthase is more active in ileum than in colon. Sialylated core 5 may be a specific ligand for bacteria expressed mainly in the small intestine.

In contrast to human adult intestinal mucins, no increasing gradient of sialic acid or decreasing gradient of fucose was observed from ileum to distal colon [12–13]. The major glycans identified were neutral. Several hypotheses may explain the lower sialylation observed in fetal mucins. First, charged residues such as sulfate and sialic acid are partly responsible for the characteristic viscoelastic properties of mucins [40–42] and play a role in water and electrolyte transport in the distal colon of adults. In fetuses, intestine is functional, but does not need to play its digestive role since nutrients are directly provided by the placenta. Swallowed amniotic fluid moves through the gastrointestinal tract from oesophagus to stomach and small intestine. In the large bowel the majority of fluid (water) is absorbed along with electrolytes, glucose, urea and hormones. This process contributes to fetal nutrition and prepares the gastrointestinal tract for its postnatal function. A second hypothesis explaining the lower sialylation in fetal intestine may be the fact that sialic acids, as well as sulfate residues, are implicated in protection against bacterial attack and confer resistance to protease degradation [43]. However, fetal intestine is sterile since colonization of the gastrointestinal tract of newborn infants starts immediately after birth and occurs within a few days.

Table 6 Proposed acidic oligosaccharides structures identified in fetal intestinal mucins by nanoESI MS/MS in negative ion mode

Sequence/composition of oligosaccharide alditols	$[M-H]^-$	Ileum			Right colon			Left colon		
		1-A	7-B	8-O	1-A	7-B	8-O	1-A	7-B	8-O
Oligosaccharides with one NeuAc residue										
NeuAc→6GalNAcol	513	+	+	+	+	+	+	+	+	+
Gal→3(NeuAc→6)GalNAcol	675	+	+	+	+	+	+	+	+	+
GlcNAc→3(NeuAc→6)GalNAcol	716	+	+	+	+	+	+	+	+	+
GalNAc→3(NeuAc→6)GalNAcol	716	+	+	+	+	+	+	+	+	+
Gal→3GlcNAc→3(NeuAc→6)GalNAcol	878	+	nd	+	+	+	+	+	+	+
GlcNAc→3Gal→3(NeuAc→6)GalNAcol	878	+	nd	+	+	+	+	+	+	+
(NeuAc→3)Gal→(GlcNAc→6)GalNAcol	878	+	nd	+	+	+	+	+	+	+
(NeuAc→3)Gal→4GlcNAc→3GalNAcol	878	+	nd	+	+	+	+	+	+	+
(NeuAc→3)Gal→4(Fuc→3)GlcNAc→3GalNAcol	1,024	+	−	+	+	+	+	+	+	+
HexNAc→(Fuc2)Gal→3(NeuAc→6)GalNAcol	1,024	+	−	+	+	+	+	+	+	+
Gal→3(Fuc→4)GlcNAc→3(NeuAc→6)GalNAcol	1,024	+	−	+	+	+	+	+	+	+
Gal→4(Fuc→3)GlcNAc→3(NeuAc→6)GalNAcol	1,024	+	−	+	+	+	+	+	+	+
(Fuc→2)Gal→4GlcNAc→3(NeuAc→6)GalNAcol	1,024	+	+	+	+	+	+	+	+	+
(Fuc→2)Gal→3(Fuc→4)GlcNAc→3(NeuAc→6) GalNAcol	1,170	+	nd	nd	+	+	+	+	nd	+
HexNAc→3Gal→3(Fuc→4)GlcNAc→ 3(NeuAc→6)GalNAcol	1,227	+	nd	−	+	+	+	+	−	+
HexNAc→3(Fuc→2)Gal→3GlcNAc→ 3(NeuAc→6)GalNAcol	1,227	+	nd	−	+	+	+	+	−	+
(NeuAc→3)Gal→4GlcNAc→3Gal→ GlcNAc→3GalNAcol	1,243	−	nd	−	+	+	+	+	+	+
HexNAc→3(Fuc→2)Gal→4(Fuc→4)GlcNAc→ 3(NeuAc→6)GalNAcol	1,373	nd	nd	−	+	+	+	+	+	−
3 Gal, 2 GlcNAc, 2 Fuc, NeuAc, GalNAcol	1,697	+	+	+	+	+	+	+	+	+
NeuAc→3Gal→3[(Fuc→2)Gal→4GlcNAc→ 3Gal→3(Fuc→3)GlcNAc→6]GalNAcol	1,697	+	nd	nd	+	nd	+	+	nd	nd
Oligosaccharides with one sulfate residue										
(SO_3^-)3Gal→4GlcNAc→3GalNAcol	667	+	+	+	+	+	+	+	+	+
Gal→4(SO_3^-)GlcNAc→3GalNAcol	667	+	+	+	+	+	+	+	+	+
(Fuc→2)Gal→4(SO_3^-)GlcNAc→3GalNAcol	813	+	nd	nd	+	nd	nd	+	nd	nd
(SO_3^-)3Gal→4(Fuc→3)GlcNAc→3GalNAcol	813	+	nd	nd	+	nd	nd	+	nd	nd
(Fuc→2)Gal→3[(SO_3^-)GlcNAc→6]GalNAcol	813	+	nd	nd	+	nd	nd	+	nd	nd
(SO_3^-)3Gal→4GlcNAc→Gal→3GalNAcol	829	+	nd	nd	+	nd	nd	+	nd	nd
Gal→4(SO_3^-)3GlcNAc→Gal→3GalNAcol	829	+	nd	nd	+	nd	nd	+	nd	nd
Gal→3[(SO_3^-)3Gal→4GlcNAc→6]GalNAcol	829	+	nd	nd	+	nd	nd	+	nd	nd
Gal→3[Gal→4(SO_3^-)GlcNAc→6]GalNAcol	829	+	nd	nd	+	nd	nd	+	nd	nd
Gal→3[(SO_3^-)3Gal→4(Fuc→3)GlcNAc→6] GalNAcol	975	+	nd	−	+	−	−	+	−	nd
(SO_3^-)3Gal→4(Fuc→3)GlcNAc→3Gal→ 3GalNAcol	975	+	nd	−	+	−	−	+	−	nd
Gal→3[(Fuc→2S)Gal→4(SO_3^-)GlcNAc→6] GalNAcol	975	+	nd	−	+	−	−	+	−	nd
Gal→4(SO_3^-)3GlcNAc→3(Fuc→2)Gal→ 3[Gal→4(Fuc→3)GlcNAc→6]GalNAcol	1,486	+	nd	nd	+	−	nd	+	nd	nd
(SO_3^-)3Gal→4(Fuc→3)GlcNAc→3Gal→ 3[Gal→4(Fuc→3)GlcNAc→6]GalNAcol	1,486	+	nd	nd	+	−	nd	+	nd	nd
Oligosaccharides with two acidic residues										
(NeuAc→3)Gal→3(NeuAc→6)GalNAcol	966	+	+	+	+	+	+	+	+	+
(SO_3^-)3Gal→4(Fuc→3)GlcNAc→3(NeuAc→6) GalNAcol	1,104	−	−	−	+	+	+	+	+	+
(NeuAc→3)Gal→3[(SO_3^-)3Gal→4GlcNAc→6] GalNAcol	1,120	−	−	−	+	+	+	+	+	+

Table 6 (continued)

Sequence/composition of oligosaccharide alditols	$[M-H]^-$	Ileum			Right colon			Left colon		
		1-A	7-B	8-O	1-A	7-B	8-O	1-A	7-B	8-O
(NeuAc→3)Gal→3[Gal→4(SO_3^-)GlcNAc→6] GalNAcol	1,120	−	−	−	+	+	+	+	+	+
(NeuAc→3)Gal→4GlcNAc→3(NeuAc→6) GalNAcol	1,169	+	+	+	+	+	+	+	+	+
(NeuAc→3)Gal→4(Fuc→3)GlcNAc→ 3(NeuAc→6)GalNAcol	1,315	−	−	−	+	+	+	+	+	+
(SO_3^-)3Gal→4GlcNAc→3Gal→4GlcNAc→ 3(NeuAc→6)GalNAcol	1,323	−	−	−	+	+	+	+	+	+
(NeuAc→3)Gal→3(NeuAc→3Gal→GlcNAc→6) GalNAcol	1,331	+	+	+	+	+	+	+	+	+
GalNAc→4(NeuAc→3)Gal→4GlcNAc→ 3(NeuAc→6)GalNAcol	1,372	−	−	−	−	+	+	+	+	+
GalNAc→4(NeuAc→3)Gal→3GlcNAc→ 3(NeuAc→6)GalNAcol	1,372	−	−	−	−	+	+	+	+	+
(SO_3^-)3Gal→4(Fuc→3)GlcNAc→3Gal→ 4GlcNAc→3(NeuAc→6)GalNAcol	1,469	+	−	−	+	+	+	+	nd	+
(NeuAc→3)Gal→3(NeuAc→3Gal→ (Fuc→3/4)GlcNAc→6)GalNAcol	1,477	+	−	+	+	+	+	+	+	+
3 Gal, 2 HexNAc, NeuAc, SO_3^-, GalNAcol	1,485	+	−	+	+	−	+	+	−	+
(SO_3^-)3Gal→4(Fuc→3)GlcNAc→3Gal→ 3[(SO_3^-)3Gal→4(Fuc→3)GlcNAc→6] GalNAcol	1,566	+	−	−	+	−	nd	+	−	nd
(SO_3^-)3Gal→4(Fuc→3)GlcNAc→3Gal→ 4(Fuc→3)GlcNAc→3(NeuAc→6)GalNAcol	1,615	+	+	+	+	+	+	+	+	+
2 Gal, 2 HexNAc, Fuc, 2 NeuAc, GalNAcol	1,680	−	−	−	+	+	+	+	+	+
2 Gal, 2 HexNAc, 2 Fuc, 2 NeuAc, GalNAcol	1,826	−	−	−	+	+	+	+	+	+

Using four polyclonal antibodies for MUC2, MUC5AC, MUC5B and MUC6, we have shown that MUC2 and MUC5B glycoproteins are expressed throughout the length of the intestinal tract in the eight fetuses included in the study. MUC6 mucin was recovered only in intestine of fetuses after 29 weeks of gestation and MUC5AC was not found in any fetal intestine. In fetal gut, mucus secretion begins as early as 10 weeks after gestation [14] and provides resistance to ingested bile and amniotic fluid, leading to the characteristic composition of meconium. Two previous studies have described developmental expression of intestinal mucin genes: Chambers *et al.* [44] analysed MUC1 and MUC2 gene expression in four fetuses aged between 12 and 19 weeks of gestation and Buisine *et al.* [26] utilized *in situ* hybridization to determine the pattern of expression of MUC2, MUC3, MUC4, MUC5B, MUC5AC and MUC6 in 32 human embryos and fetuses between 6.5 and 27 weeks of gestation. MUC2 is the major fetal and adult colonic mucin. MUC2 gene is expressed as early as 9 weeks of gestation and the developmental switch from a fetal to an adult pattern of MUC2 gene expression occurs at around 23 weeks in the small intestine and 27 weeks for the colon [26].

MUC5AC is mainly expressed in the gastric and respiratory mucosae of normal human adults and is not found in the intestinal tract except in pathological situation such as colorectal cancer. Because MUC5AC is selectively expressed in early fetal life (between 8 and 12 weeks of gestation) and in cancer, it has been suggested that this mucin plays a role in intestinal epithelial cell differentiation, both in fetal development and in tumorigenesis. MUC6, in addition to MUC5AC, is not normally expressed in normal human small and large intestine. Several studies have demonstrated the presence of this mucin in colorectal neoplasia [45]. Because our results show an expression of MUC6 after 29 weeks of gestation in fetuses, it may be hypothesized that MUC6 play a similar role to MUC5AC in intestinal cell differentiation in ontogenesis and oncogenesis. MUC5B was not observed in either embryonic or fetal intestine [26]. This is in contrast to our findings, but may be explained by the fact that fetuses in our work were older. MUC5B has been reported at the base of the colonic crypts in the human adult colon [46], and it has been speculated that this mucin play a role in cytoprotection, protecting the colonic epithelial cells from bile acids [3].

O-acetylation is one of the most frequent modifications of sialic acids. Different functions have been attributed to *O*-acetyl sialic acids in intestine [47]. For example, elevated *O*-acetylation level has been recovered in colon of four

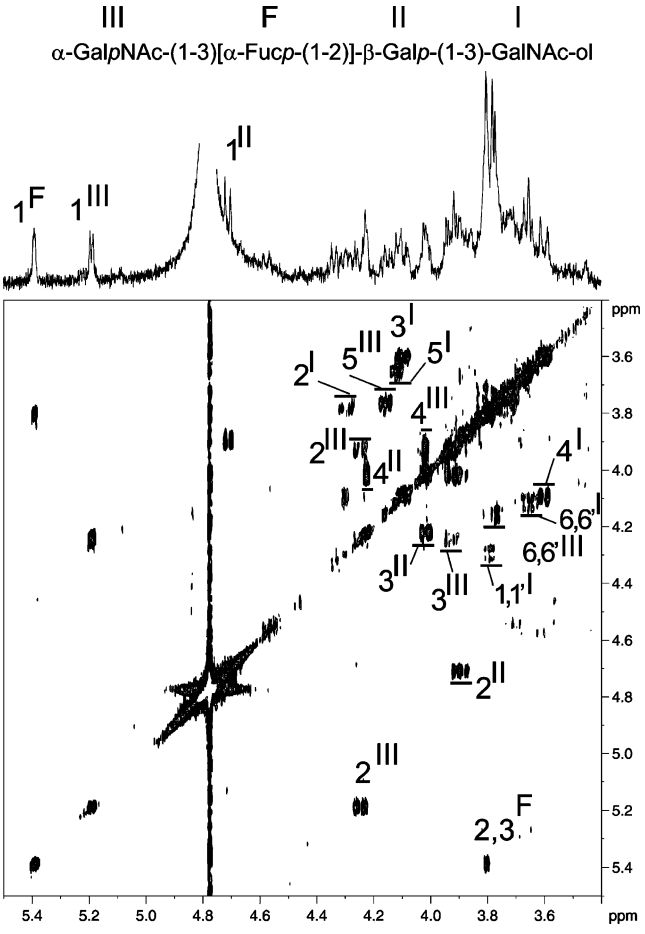


Fig. 3 Part of 2D-COSY90 spectrum of oligosaccharide 11 obtained after HPLC fractionation showing the spin system of each monosaccharide unit. This oligosaccharide corresponds to the ion at *m/z* 757 and carries a blood group A determinant

fetuses, three of them with a bacterial infection. Thus it has been suggested that *O*-acetylation appears in response to bacterial colonization. Our results on the diversity of sialic acids in the different regions of the adult intestine are in agreement with data already published [12, 43]. The level of *O*-acetylated sialic acids was more elevated in adults (around 30–35%) than in fetuses (around 20%). *O*-acetylation substitution is influenced by environmental stimuli and exposure to bacterial products [47]. This data may account for the variation observed between adults and fetuses.

Less *O*-acetylation was found in small intestine than in colon and supports the proposal that modification by *O*-acetylation is correlated with significant bacterial colonization. Di- and tri-*O*-acetylated sialic acids were more expressed in proximal colon than in ileum and distal colon. This result suggests that *O*-acetylated sialic acids in the cecum play a protective role in the barrier against microorganisms derived from the small intestine. Indeed, specific strains, which are able to de-*O*-acetylate sialic acids can be more competitive than others in developing colonies at this intestinal location.

A low level of Neu5Gc and its *O*-acetylated derivatives were recovered in human intestine. Neu5Ac and Neu5Gc are the two major sialic acids in most mammalian cells [48–49]. However, Neu5Gc was long thought to be absent from healthy human tissues [50]. Indeed, humans are genetically unable to synthesize Neu5Gc due to an exon deletion/frameshift mutation in the human *CMA*H gene [51–52]. Our results are in agreement with recent studies [28, 53], detecting small amounts of Neu5Gc in epithelial and endothelial cells of normal humans and various carcinomas. They have demonstrated that free Neu5Gc, originating from dietary sources, can be incorporated in human tissues. It is efficiently delivered to the cytosol utilizing fluid pinocytosis and specific lysosomal transporters [28].

Using physical methods (*i.e.* NMR and mass spectrometry), we have demonstrated that blood group A, B and H determinants are expressed in each part of the fetal intestine. This result is in agreement with earlier immunohistochemistry and immunofluorescence studies [54–55]. This is in contrast with adult human intestine, where these particular antigens are recovered only in the proximal colon. Because cancers of the distal colon frequently re-express blood group structures, they are considered as oncofetal antigens at this organ site.

Very little expression of Sda/Cad blood group antigens (GalNAcβ1-4[NeuAcα2–3]Galβ1-3/4GlcNAc-) was recovered in the distal colon of fetuses 1 and 8, both at 40 weeks of gestation. No expression of this determinant was found in small intestine or proximal colon of any fetuses. In a previous study, we have shown that Sda antigen-like structures are prominent features of glycans from the

Table 7 Proton chemical shifts (ppm) of monosaccharides from oligosaccharide with the blood group A structure corresponding to the ion at *m/z* 757

Residue	Proton						
	H1	H2	H3	H4	H5	H6	–CH3 from NAc group
Ga1NAc-ol	~3.79	4.299	4.092	3.597	4.116	~3.65	2.048
β-Galp	4.716	3.895	4.016	4.223	n.d.	n.d.	–
α-Gal*p*NAc	5.189	4.243	3.922	4.019	n.d.	n.d.	2.046
α-Fuc*p*	5.393	3.80	3.80	n.d.	4.336	1.236	–

mucin of human adult descending colon [11]. Another study has demonstrated that activities of β1-4-*N*-acetylgalactosaminyltransferase (β1-4GalNAc transferase) and α2–3 sialyltransferase, both enzymes involved in the biosynthesis of the Sda blood group antigen, are exclusively found in epithelial cells of the rat large intestine and occur preferentially at a specific stage of cell differentiation and during postnatal maturation. Indeed, β1-4GalNAc transferase is practically absent at birth, increases slowly in the first days of life and then rapidly after weaning [56]. β1-4GalNAc transferase activity dramatically decreases in human colon carcinomas and colon carcinoma cell lines and regulates sialyl-Lewis^x, a colorectal cancer marker [57]. In colon cancer cell lines the enzyme increases with differentiation and when over expressed leads to reduced metastasis through reduction of sialyl-Lewis^x expression [58]. Little information is available on the role of the Sda antigen, however it has been suggested that microbial infections operate as selective agents for the high frequency of Sda antigen. The addition of the GalNAcβ1-4 residue to modify the 3′-sialyllactosamine may prevent the binding of certain *E. coli* strains to the cells, with a consequent decrease in susceptibility to infection [59].

Taking these data together and allowing for the fact that only mucin glycans from fetuses at 40 weeks of gestation expressed Sda/Cad antigens, our results support the correlation of this antigen with ontogenesis and colonization of intestine by microorganisms.

In conclusion, our results show a remarkable diversity of oligosaccharides in human fetal intestinal mucins. In contrast to adults, no gradient of oligosaccharide structures were found in the different parts of the intestine and a similar pattern of glycosylation was observed. These results suggest that region-specific glycosylation of human intestine is acquired after birth, probably due to bacterial colonization and gut postnatal absorptive and digestive functions.

Acknowledgments We greatly appreciate the gift of fetal intestinal mucins collected by Dr. A. Zweibaum and the gift of polyclonal antibodies (LUM2-3, LUM5B-3, LUM5-1 and LUM6-1) by Dr Ingemar Carlstedt. 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 Mass Spectroscopy facility used in this study was funded by the European Community (FEDER), the Région Nord-Pas de Calais (France), the CNRS and the Université des Sciences et Technologies de Lille.

References

1. Thornton, D.J., Rousseau, K., McGuckin, M.A.: Structure and function of the polymeric mucins in airways mucus. Annu. Rev. Physiol. **70**, 459–486 (2008). doi:10.1146/annurev.physiol.70.113006.100702
2. Linden, S.K., Sutton, P., Karlsson, N.G., Korolik, V., McGuckin, M.A.: Mucins in the mucosal barrier to infection. Nat. Mucosal Immunol. **1**, 183–197 (2008). doi:10.1038/mi.2008.5
3. Atuma, C., Strugala, V., Allen, A., Holm, L.: The adherent gastrointestinal mucus gel layer: thickness and physical state in vivo. Am. J. Physiol. Gastrointest. Liver Physiol. **280**, G922–G929 (2001)
4. Hollingsworth, M.A., Swanson, B.J.: Mucins in cancer: protection and control of the cell surface. Nat. Rev. Cancer **4**, 45–60 (2004). doi:10.1038/nrc1251
5. Chang, S.K., Dohrman, A.F., Basbaum, C.B., Ho, S.B., Tsuda, T., Toribara, N.W., *et al.*: Localization of mucin (MUC2 and MUC3) messenger RNA and peptide expression in human normal intestine and colon cancer. Gastroenterology **107**, 28–36 (1994)
6. Porchet, N., Nguyen, V.C., Dufosse, J., Audie, J.P., Guyonnet-Duperat, V., Gross, M.S., *et al.*: Molecular cloning and chromosomal localization of a novel human tracheo-bronchial mucin cDNA containing tandemly repeated sequences of 48 base pairs. Biochem. Biophys. Res. Commun. **175**, 414–422 (1991). doi:10.1016/0006-291X(91)91580-6
7. Audie, J.P., Janin, A., Porchet, N., Copin, C., Gosselin, B., Aubert, J.P.: Expression of human mucin genes in respiratory, digestive, and reproductive tracts ascertained by in situ hybridization. J. Histochem. Cytochem. **41**, 1479–1485 (1993)
8. De Bolos, C., Garrido, M., Real, F.X.: MUC6 apomucin shows a distinct normal tissue distribution that correlates with Lewis antigen expression in the human stomach. Gastroenterology **109**, 723–734 (1995). doi:10.1016/0016-5085(95)90379-8
9. Podolsky, D.K.: Oligosaccharide structures of human colonic mucin. J. Biol. Chem. **260**, 8262–8271 (1985)
10. Podolsky, D.K.: Oligosaccharide structures of isolated human colonic mucin species. J. Biol. Chem. **260**, 15510–15515 (1985)
11. Capon, C., Maes, E., Michalski, J.C., Leffler, H., Kim, Y.S.: Sd^a-antigen-like structures carried on core 3 are prominent features of glycans from the mucin of normal human descending colon. Biochem. J. **358**, 657–664 (2001)
12. Robbe, C., Capon, C., Maes, E., Rousset, M., Zweibaum, A., Zanetta, J.P., *et al.*: Evidence of regio-specific glycosylation in human intestinal mucins. J. Biol. Chem. **278**, 46337–46348 (2003). doi:10.1074/jbc.M302529200
13. Robbe, C., Capon, C., Coddeville, B., Michalski, J.C.: Structural diversity and specific distribution of *O*-glycans in normal human mucins. Biochem. J. **384**, 307–316 (2004)
14. Stauffer, A., Lallemand, A., Gaillard, D.: Mucin histochemistry of the digestive tract in the human fetus. Gastroenterol. Clin. Biol. **14**, 561–566 (1990)
15. Hounsell, E.F., Lawson, A.M., Feeney, J., Gooi, H.C., Pickering, N.J., Stoll, M.S., *et al.*: Structural analysis of the O-glycosidically linked core-region oligosaccharides of human meconium glycoproteins which express oncofoetal antigens. Eur. J. Biochem. **148**, 367–377 (1985). doi:10.1111/j.1432-1033.1985.tb08848.x
16. Feeney, J., Frenkiel, T.A., Hounsell, E.F.: Complete 1H-NMR assignments for two core-region oligosaccharides of human meconium glycoproteins, using 1D and 2D methods at 500 MHZ. Carbohydr. Res. **152**, 63–72 (1986). doi:10.1016/S0008-6215(00)90287-8
17. Capon, C., Cache, P., Leroy, Y., Strecker, G., Montreuil, J., Fournet, B.: Isolation of the major O-glycosidically linked oligosaccharides obtained by alkaline borohydride degradation of human meconium glycoproteins. J. Chromatogr. A **425**, 35–45 (1988). doi:10.1016/0378-4347(88)80004-5
18. Capon, C., Leroy, Y., Wieruszeski, J.M., Ricart, G., Strecker, G., Montreuil, J., *et al.*: Structures of O-glycosidically linked oligosaccharides isolated from human meconium glycoproteins. Eur. J. Biochem. **182**, 139–152 (1989). doi:10.1111/j.1432-1033.1989.tb14810.x

19. Byrd, J.C., Bresalier, R.S.: Mucins and mucin binding proteins in colorectal cancer. Cancer Metastasis Rev. **23**, 77–99 (2004). doi:10.1023/A:1025815113599
20. Zweibaum, A., Oriol, R., Dausset, J., Marcelli-Barge, A., Ropartz, C., Lanset, S.: Definition in man of a polymorphic system of the normal colonic secretions. Tissue Antigens **6**, 121–128 (1975)
21. Lesuffleur, T., Roche, F., Hill, A.S., Lacasa, M., Fox, M., Swallow, D.M., *et al*.: Characterization of a mucin cDNA clone isolated from HT-29 mucus-secreting cells. The 3′ end of MUC5AC? J. Biol. Chem. **270**, 13665–13673 (1995). doi:10.1074/jbc.270.23.13665
22. Klein, A., Diaz, S., Ferreira, I., Lamblin, G., Manzi, A.E.: New sialic acids from biological sources identified by a comprehensive and sensitive approach: liquid chromatography-electrospray ionization-mass spectrometry (LC-ESI-MS) of SIA quinoxalinones. Glycobiology **7**, 421–432 (1997). doi:10.1093/glycob/7.3.421
23. Carlson, D.M.: Structures and immunochemical properties of oligosaccharides isolated from pig submaxillary mucins. J. Biol. Chem. **243**, 616–626 (1968)
24. Kamerling, J.P., Gerwig, G.J., Vliegenthart, J.F., Clamp, J.R.: Characterization by gas–liquid chromatography-mass spectrometry and proton-magnetic-resonance spectroscopy of pertrimethylsilyl methyl glycosides obtained in the methanolysis of glycoproteins and glycopeptides. Biochem. J. **151**, 491–495 (1975)
25. Montreuil, J., Bouquelet, S., Debray, H., Fournet, B., Spik, G., Strecker, G.: In: Chaplin, M.F., Kennedy, J.F. (eds.) Carbohydrate Analysis: A Practical Approach, pp. 143–204. IRL Press, Oxford (1986)
26. Buisine, M.P., Devisme, L., Savidge, T.C., Gespach, C., Gosselin, B., Porchet, N., Aubert, J.P.: Mucin gene expression in human embryonic and fetal intestine. Gut **43**, 519–524 (1998)
27. Hara, S., Yamaguchi, M., Takemori, Y., Furuhata, K., Ogura, H., Nakamura, M.: Determination of mono-O-acetylated N-acetylneuraminic acids in human and rat sera by fluorometric high-performance liquid chromatography. Anal. Biochem. **179**, 162–166 (1989). doi:10.1016/0003-2697(89)90218-2
28. Bardor, M., Nguyen, D.H., Diaz, S., Varki, A.: Mechanism of uptake and incorporation of the non-human sialic acid N-glycolylneuraminic acid into human cells. J. Biol. Chem. **280**, 4228–4237 (2005). doi:10.1074/jbc.M412040200
29. Robbe, C., Capon, C., Coddeville, B., Michalski, J.C.: Diagnostic ions for the rapid analysis by nano-electrospray ionization quadrupole time-of-flight mass spectrometry of *O*-glycans from human mucins. Rapid Commun. Mass Spectrom. **18**, 412–420 (2004). doi:10.1002/rcm.1352
30. Domon, B., Costello, C.E.: A systematic nomenclature for carbohydrate fragmentations in FAB-MS/MS spectra of glycoconjugates. Glycoconj. J. **5**, 397–409 (1988). doi:10.1007/BF01049915
31. Karlsson, N.G., Karlsson, H., Hansson, G.C.: Sulphated mucin oligosaccharides from porcine small intestine analysed by four-sector tandem mass spectrometry. J. Mass Spectrom. **31**, 560–572 (1996). doi:10.1002/(SICI)1096-9888(199605)31:5<560::AID-JMS331>3.0.CO;2-0
32. Koerner, T.A.W., Prestegard, J.H., Yu, R.K.: Oligosaccharide structure by two-dimensional proton nuclear magnetic resonance spectroscopy. Methods Enzymol. **138**, 38–59 (1987). doi:10.1016/0076-6879(87)38006-1
33. Kamerling, J.P., Vliegenthart, J.F.G.: In: Berliner, L., Reben, J. (eds.) Biological Magnetic Resonance, vol. 10, pp. 1–287. Plenum, New York (1992)
34. Sweet-DB: http://www.dkfz.de/spec/glycosciences.de/sweetdb/
35. Dua, V.K., Rao, B.N.N., Wu, S.S., Dube, V.E., Bush, C.A.: Characterization of the oligosaccharide alditols from ovarian cyst mucin glycoproteins of blood group A using high pressure liquid chromatography (HPLC) and high field 1H NMR spectroscopy. J. Biol. Chem. **261**, 1599–1608 (1986)
36. Thomsson, K.A., Prakobphol, A., Fisher, S.J., Leffler, H., Reddy, M.S., Levine, M.J., *et al*.: The O-glycans of the salivary mucin MG1 (MUC5B) are large and diverse and differ from the MG2 (MUC7) mucin oligosaccharides. Glycobiology **12**, 1–14 (2002). doi:10.1093/glycob/12.1.1
37. Andersch-Björkman, Y., Thomsson, K.A., Holmén Larsson, J.M., Ekerhovd, E., Hansson, G.C.: Large-scale identification of proteins, mucins and their *O*-glycosylation in the endocervical mucus during the menstrual cycle. Mol. Cell. Proteomics **6**, 708–716 (2007). doi:10.1074/mcp.M600439-MCP200
38. Gill, S.R., Pop, M., DeBoy, R.T., Eckburg, P.B., Turnbaugh, P.J., Samuel, B.S., *et al*.: Metagenomic analysis of the human distal gut microbiome. Science **312**, 1355–1359 (2006). doi:10.1126/science.1124234
39. Kurosaka, A., Nakajima, H., Funakoshi, I., Matsuyama, M., Nagayo, T., Yamashina, I.: Structures of the major oligosaccharides from a human rectal adenocarcinoma glycoprotein. J. Biol. Chem. **258**, 11594–11598 (1983)
40. Jentoft, N.: Why are proteins O-glycosylated? Trends Biochem. Sci. **15**, 291–294 (1990). doi:10.1016/0968-0004(90)90014-3
41. Van der Reijden, W.A., Veerman, E.C.I., Nieuw Amerongen, A.V.: Shear rate dependent viscoelastic behavior of human glandular salivas. Biorheology **30**, 141–152 (1993)
42. Corfield, A.P.: The glycobiology of mucins in the human gastrointestinal tract. In: Sansom, C.A.D.O. (ed), Glycobiology, pp. 248–260. Scion, Bloxham, UK (2007)
43. Corfield, A.P., Wagner, S.A., Clamp, J.R., Kriaris, M.S., Hoskins, L.C.: Mucin degradation in the human colon: production of sialidase, sialate O-acetylesterase, N-acetylneuraminate lyase, arylesterase, and glycosulfatase activities by strains of fecal bacteria. Infect. Immun. **60**, 3971–3978 (1992)
44. Chambers, J.A., Hollingsworth, M.A., Trezise, A.E., Harris, A.: Developmental expression of mucin genes MUC1 and MUC2. J. Cell Sci. **107**, 413–424 (1994)
45. Bartman, A.E., Sanderson, S.J., Ewing, S.L., Niehans, G.A., Wiehr, C.L., Evans, M.K., *et al*.: Aberrant expression of MUC5AC and MUC6 gastric mucin genes in colorectal polyps. Int. J. Cancer **80**, 210–218 (1999). doi:10.1002/(SICI)1097-0215(19990118)80:2<210::AID-IJC9>3.0.CO;2-U
46. Van Klinken, B.J., Dekker, J., Van Gool, S.A., Van Marle, J., Büller, H.A., Einerhand, A.W.: MUC5B is the prominent mucin in human gallbladder and is also expressed in a subset of colonic goblet cells. Am. J. Physiol. **274**, G871–G878 (1998)
47. Angata, T., Varki, A.: Chemical diversity in the sialic acids and related keto acids. An evolutionary perspective. Chem. Rev. **102**, 439–469 (2002). doi:10.1021/cr000407m
48. Howard, R.J., Reuter, G., Barnwell, J.W., Schauer, R.: Sialoglycoproteins and sialic acids of Plasmodium knowlesi schizont-infected erythrocytes and normal rhesus monkey erythrocytes. Parasitology **92**, 527–543 (1986)
49. Muchmore, E.A., Diaz, S., Varki, A.: A structural difference between the cell surfaces of humans and the great apes. Am. J. Phys. Anthropol. **107**, 187–198 (1998). doi:10.1002/(SICI)1096-8644(199810)107:2<187::AID-AJPA5>3.0.CO;2-S
50. Traving, C., Schauer, R.: Structure, function and metabolism of sialic acids. Cell. Mol. Life Sci. **54**, 1330–1349 (1998). doi:10.1007/s000180050258
51. Chou, H.H., Takematsu, H., Diaz, S., Iber, J., Nickerson, E., Wright, K.L., *et al*.: A mutation in human CMP-sialic acid hydroxylase occurred after the Homo-Pan divergence. Proc. Natl. Acad. Sci. USA **95**, 11751–11756 (1998). doi:10.1073/pnas.95.20.11751
52. Irie, A., Koyama, S., Kozutsumi, Y., Kawasaki, T., Suzuki, A.: The molecular basis for the absence of N-glycolylneuraminic acid

in humans. J. Biol. Chem. 273, 15866–15871 (1998). doi:10.1074/jbc.273.25.15866
53. Tangvoranuntakul, P., Gagneux, P., Diaz, S., Bardor, M., Varki, N., Varki, A., *et al.*: Human uptake and incorporation of an immunogenic nonhuman dietary sialic acid. Proc. Natl. Acad. Sci. U.S.A. **100**, 12045–12050 (2003). doi:10.1073/pnas.2131556100
54. Szulman, A.E.: The histological distribution of the blood group substances in man as disclosed by immunofluorescence. III. The A, B, and H antigens in embryos and fetuses from 18 mm in length. J. Exp. Med. **119**, 503–515 (1964). doi:10.1084/jem.119.4.503
55. Yuan, M., Itzkowitz, S.H., Palekar, A., Shamsuddin, A.M., Phelps, P.C., Trump, B.F., *et al.*: Distribution of blood group antigens A, B, H, Lewis a, and Lewis b in human normal, fetal, and malignant colonic tissue. Cancer Res. **45**, 4499–4511 (1985)
56. Dall'Olio, F., Malagolini, N., Di Stefano, G., Ciambella, M., Serafini-Cessi, F.: Postnatal development of rat colon epithelial cells is associated with changes in the expression of the beta 1,4-N-acetylgalactosaminyltransferase involved in the synthesis of Sda antigen of alpha 2,6-sialyltransferase activity towards N-acetyl-lactosamine. Biochem. J. **270**, 519–524 (1990)
57. Malagolini, N., Santini, D., Chiricolo, M., Dall'Olio, F.: Biosynthesis and expression of the Sda and sialyl Lewis x antigens in normal and cancer colon. Glycobiology **17**, 688–697 (2007). doi:10.1093/glycob/cwm040
58. Kawamura, Y.I., Kawashima, R., Fukunaga, R., Hirai, K., Toyama-Sorimachi, N., Tokuhara, M., *et al.*: Introduction of Sd (a) carbohydrate antigen in gastrointestinal cancer cells eliminates selectin ligands and inhibits metastasis. Cancer Res. **65**, 6220–6227 (2005). doi:10.1158/0008-5472.CAN-05-0639
59. Malagolini, N., Dall'Ollio, F., Serafini-Cessi, F.: UDP-GalNAc: NeuAc alpha 2,3Gal beta-R (GalNAc to Gal) beta 1,4-N-acetylgalactosaminyltransferase responsible for the Sda specificity in human colon carcinoma CaCo-2 cell line. Biochem. Biophys. Res. Commun. **180**, 681–686 (1991). doi:10.1016/S0006-291X(05)81119-2