Glycobiology vol. 22 no. 8 pp. 1128–1139, 2012
doi:10.1093/glycob/cws083
Advance Access publication on May 11, 2012

Detailed *O*-glycomics of the Muc2 mucin from colon of wild-type, core 1- and core 3-transferase-deficient mice highlights differences compared with human MUC2

Kristina A Thomsson[1,2], Jessica M Holmén-Larsson[2], Jonas Ångström[2], Malin EV Johansson[2], Lijun Xia[3], and Gunnar C Hansson[2]

[2]Department of Medical Biochemistry, University of Gothenburg, PO Box 440, 405 30 Gothenburg, Sweden and [3]Cardiovascular Biology Research Program, Oklahoma Medical Research Foundation, Oklahoma City, OK, USA

Received on February 27, 2012; revised on April 17, 2012; accepted on May 2, 2012

The heavily *O*-glycosylated mucin MUC2 constitutes the major protein in the mucosal layer that acts as a physical barrier protecting the epithelial layer in the colon. In this study, Muc2 was purified from mucosal scrapings from the colon of wild-type (WT) mice, core 3 transferase knockout (*C3Gnt*$^{-/-}$) mice and intestinal epithelial cell-specific core 1 knockout (IEC *C1Galt1*$^{-/-}$) mice. The Muc2 *O*-glycans were released by reductive β-elimination and analyzed with liquid chromatography-mass spectrometry in the negative-ion mode. Muc2 from the distal colon of WT and *C3Gnt*$^{-/-}$ knockout mice carried a mixture of core 1- or core 2-type glycans, whereas Muc2 from IEC *C1Galt1*$^{-/-}$ mice carried highly sialylated core 3- and core 4-type glycans. A large portion of NeuAc in all mouse models was positioned on disialylated *N*-acetyllactosamine units, an epitope not reported on human colonic MUC2. Mass spectra and proton NMR spectroscopy revealed an abundant NeuAc linked to internally positioned *N*-acetylglucosamine on colonic murine Muc2, which also differs markedly from human MUC2. Our results highlight that murine colonic Muc2 *O*-glycosylation is substantially different from human MUC2, which could be one explanation for the different commensal microbiota of these two species.

Keywords: LC / mass spectrometry / mucin / murine / *O*-glycosylation

Introduction

Dense *O*-glycosylation forms a protective coat around the multimerizing MUC2 protein backbone, which constitutes the major protein of the mucosal layers that lines the colon lumen. The outer loose mucosal layer in the distal colon harbors the symbiotic commensal microbiota, whereas the inner firmly attached layer is composed of a dense MUC2 network, impenetrable for the bacteria (Johansson et al. 2008). The importance of separating the intestinal bacteria from the epithelium is illustrated by the most commonly used colitis model in mice, 3–5% dextran sodium sulfate (DSS) in the drinking water. DSS inflammation is observed after 3–5 days, but already at the first contact of DSS with the inner mucus layer of the colon, it becomes permeable to bacteria (Johansson et al. 2010). The outer and loose layer is formed from the firm layer and is continuously renewed from the inner layer. The *O*-glycans constitute ~80% of the mass of MUC2 and are linked to serine and threonine residues in PTS domains (rich in proline–serine–threonine) forming the mucin domains. The dense glycosylation makes the protein gel network of the inner layer resistant toward proteolytic degradation; however, in the outer, loosely attached layer, the glycans can act as a energy source as well as providing binding sites for the commensal bacteria, which have glycan adhesins and enzymes that can degrade glycans. The presence of the commensal bacteria flora in the colon may have several advantages also for the host as they degrade both food and mucin saccharides (Backhed et al. 2005) and may suppress the colonization of pathogens. We have previously shown that colonic MUC2 *O*-glycosylation in human sigmoideum is largely blood group independent and identical between individuals (Larsson et al. 2009). We currently aim to address whether MUC2 *O*-glycosylation could be of importance for the colonization of the commensal flora (Rawls et al. 2006).

The *O*-glycans are added to the protein in the Golgi, and they are made up by the concerted actions of glycosyltransferases adding one monosaccharide after the other, building up linear or branched sequences. On mucins, hundreds of different *O*-glycans can be found (Larsson et al. 2009). The terminal monosaccharide residues can be further modified by, for example, sulfate or acetyl groups. In order to study the impact of protein *O*-glycosylation, various mouse models have been designed by targeted deletion of glycosyltransferases.

The *O*-glycan biosynthesis is initiated by the addition of an *N*-acetylgalactosamine (GalNAc) residue to serine or threonine in the protein backbone, forming the Tn-antigen, and is then commonly modified by core transferases, where the cores 1, 2, 3 and 4 are the most abundant (Figure 1A). These

[1]To whom correspondence should be addressed: Tel: +46-31-7863707; Fax: +46-31-416108; e-mail: kristina.thomsson@medkem.gu.se

cores are formed by the core 1 β1,3-galactosyltransferase (C1galT1, also known as the T-synthase) adding a Gal forming the core 1 glycan Galβ1-3GalNAcα-Ser/Thr, the core 3 β1,3 *N*-acetylglucosaminyltransferase (C3GnT) forming GlcNAcβ1-3GalNAcα-Ser/Thr and additional core 2 and core 4 transferases acting on these precursors. The IEC *C1Galt1*$^{-/-}$ mouse developed spontaneous colitis (Fu et al. 2011). The core 2 β1,6 *N*-acetylglucosaminyltransferase 2 (*C2Gnt2*$^{-/-}$) and the core 3 (*C3Gnt*$^{-/-}$) knockout mice both displayed increased susceptibility to colitis in the DSS colitis model (An et al. 2007; Stone et al. 2009). The sulfo-transferase GlcNAc6ST2 adds sulfate to *N*-acetylglucosamine (GlcNAc) residues on MUC2 *O*-glycans, and when deleted in mice, they also showed an increased susceptibility to DSS colitis (Tobisawa et al. 2010). All these observations suggest that MUC2 *O*-glycosylation is important for homeostasis in the colon.

Characterization of murine MUC2 *O*-glycosylation in the gastrointestinal tract in wild-type (WT) and knock-out mouse models has been performed previously by us, although with older and less informative approaches (Holmen et al. 2002; Thomsson et al. 2002; Hurd et al. 2005). Among the more recent studies by other groups, where modern and sensitive mass spectrometric approaches have been applied allowing more comprehensive profiling, *O*-glycans were extracted from various tissues including colon from WT, three core 2 knock-out mouse models and from the gastric mucosa in the Fut2 null model (Magalhaes et al. 2009; Ismail et al. 2011). The approaches used were based on the analysis of permethylated derivatives of the *O*-glycans by matrix-assisted laser desorption ionization-mass spectrometry (MALDI-MS) or electrospray ionization mass spectrometry (ESI-MS) in the positive-ion mode, approaches that do not allow analysis of the sulfated glycans frequently found on mucins.

Our laboratory has over the last years applied a different methodology for screening mucin O-glycosylation using reversed-phase graphitized carbon chromatography-liquid chromatography-mass spectrometry (LC/MS) in the negative-ion mode (Andersch-Bjorkman et al. 2007; Karlsson et al. 2009; Larsson et al. 2009; Thomsson et al. 2010). Negative-ion mode MS promotes the detection of the negatively charged glycans (sialic acids, sulfate groups) which are abundantly found on mucins. LC/MS is preceded by a small-scale preparative approach of the glycans as these are released from mucins blotted to PVDF membrane after composite gel electrophoresis (Schulz et al. 2002).

In the work presented here, our aim was to specifically target the *O*-glycosylation of purified Muc2 from the gastrointestinal tract of the most commonly used WT mouse (B57/Bl6) and compare with those from the core 1 (IEC *C1Galt1*$^{-/-}$) and the core 3 (*C3Gnt*$^{-/-}$) knockout mice having the same background. This is a deeper analysis of the Muc2 *O*-glycan profiles from the colon of WT and IEC *C1Galt1*$^{-/-}$ mouse than addressed by us previously (Fu et al. 2011). The Muc2 glycans are then compared with the human MUC2 glycans and thereby generating an important source of information for studies of bacterial colonization and selection by different species.

Results

Muc2 O-glycan LC/MS profiles from colon of WT, core 1 (IEC C1Galt1$^{-/-}$*) and core 3 (C3Gnt*$^{-/-}$*) knockout mice*

Murine colonic Muc2 was purified from mucosal scrapings as outlined in Figure 1B, and the released *O*-glycans analyzed with capillary LC/MS^n operating in the negative-ion mode. The glycans were detected as deprotonated species, neutral and monosialylated or sulfated glycans as $[M–H^+]^-$ ions, and multiply charged glycans predominantly as $[M–nH^+]^{n-}$ ions, depending on an increasing number of charged groups (Supplementary data, Table S1). The glycans observed are summarized in Table 1. The base peak chromatograms from analyses of glycans from the IEC *C1Galt1*$^{-/-}$ and the WT mice are shown in Figure 2. Every peak contained several coeluting components and of which the most abundant are

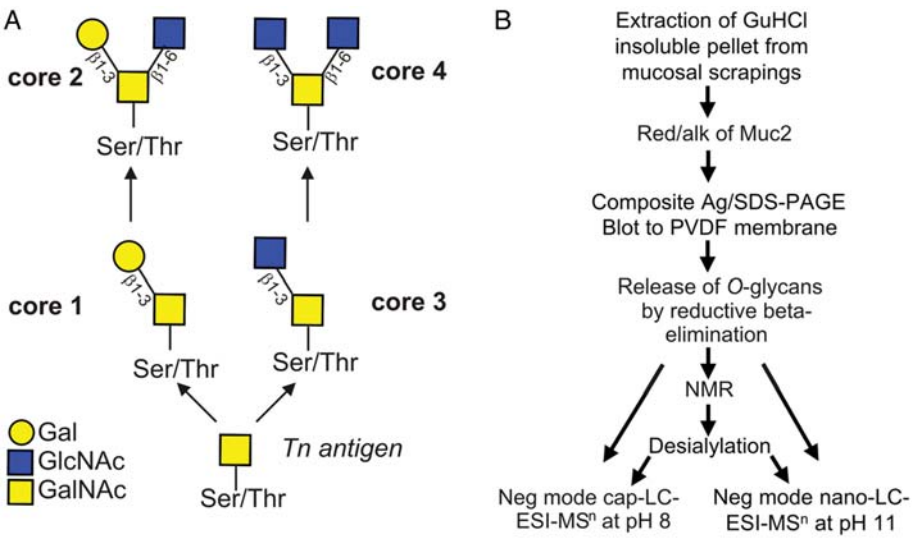

Fig. 1. Core structures commonly found on mucins (**A**). Work flow for methodology applied for structural *O*-glycan analysis from murine Muc2 from the mucosal layer of colon (**B**).

Table I. Summary of colonic Muc2 *O*-linked oligosaccharides from WT, core 1 (IEC *C1Galt1*$^{-/-}$) and core 3 (*C3Gnt*$^{-/-}$) knockout mice

Label	Proposed oligosaccharide sequence/composition[a]	Mouse model		
		WT	Core 3 KO	Core 1 KO
Neutral				
384	Gal-GalNAcol	+	+	−
425	HexNAc-GalNAcol	−	−	+
530	Fuc-Gal-GalNAcol	+	+	−
587a	Gal-**(HexNAc-6)**GalNAcol	+	+	−
587b	Gal-GlcNAc-GalNAcol	−	−	+
733a	Fuc-Gal-GlcNAc-GalNAcol	+	−	+
733b	Fuc-Gal-**(HexNAc-6)**GalNAcol	+	+	−
749	Gal-**(Gal-4GlcNAc-6)**GalNAcol	+	+	−
790a	HexNAc-Gal-**(HexNAc-6)**GalNAcol	+	+	−
790b	HexNAc-Gal-GlcNAc-GalNAcol	−	−	+
895a	Gal-**(Fuc-Gal-GlcNAc-6**)GalNAcol	+	+	−
895b	Gal-**(Fuc-Gal-4GlcNAc-6**)GalNAcol	+	+	−
895c	Fuc-Gal-**(Gal-GlcNAc-6**)GalNAcol	+	+	−
1041a	Fuc-Gal-**(Fuc-Gal-GlcNAc-6**)GalNAcol	+	+	−
1041b	Fuc-Gal-**(Fuc-Gal-4GlcNAc-6**)GalNAcol	+	+	−
1098	HexNAc-Gal-**(Fuc-Gal-4GlcNAc-6**)GalNAcol	+	−	−
1244a	Fuc-Gal-GlcNAc-**(Fuc-Gal-GlcNAc-6)**GalNAcol	+	−	+
1244b	Fuc-Gal-GlcNAc-(Fuc-Gal-4GlcNAc-)GalNAcol	+	−	+
One acidic residue				
513	NeuAc-GalNAcol	−	−	+
667	Gal-**(SO_3^--GlcNAc-6)**GalNAcol	+	+	−
675x	$[NeuAc]_1$ Hex-GalNAcol	+	+	+
813	Fuc-Gal-**(SO_3^--GlcNAc-)**GalNAcol	+	+	−
829a	Gal-**(Gal-(SO_3^--)GlcNAc-)**GalNAcol	+	+	−
829b	Gal-**(SO_3^--Gal-GlcNAc-)**GalNAcol	+	+	−
878a	Gal-(NeuAc-)GlcNAc-GalNAcol	−	−	+
878b	NeuAc-Gal-GlcNAc-GalNAcol	−	−	+
975a	Gal-**(Fuc-Gal-(SO_3^--)GlcNAc-6)**GalNAcol	+	+	−
975b	Fuc-Gal-(SO_3^--Gal-GlcNAc-)GalNAcol	+	−	−
975c	Fuc-Gal-**(Gal-(SO_3^--)GlcNAc-6)**GalNAcol	+	+	−
1040a	Gal-(Gal-(NeuAc-)GlcNAc-)GalNAcol	+	+	−
1040b	Gal, Gal-(NeuAc-)GlcNAc-GalNAcol	−	−	+
1081	$[NeuAc]_1$ HexNAc-Gal-GlcNAc-GalNAcol	+	−	+
1121	Fuc-Gal-3**(Fuc-Gal-4(SO_3^--)GlcNAc-6)**GalNAcol	+	+	−
1186a	Fuc-Gal-(Gal-(NeuAc-)GlcNAc-)GalNAcol	+	+	−
1186b	Fuc-Gal-(NeuAc-Gal-GlcNAc-)GalNAcol	+	−	−
1243	$[NeuAc]_1$ Gal-**(HexNAc-Gal-4GlcNAc-6)**GalNAcol	+	+	−
1324	Fuc-Gal-(SO_3^--)GlcNAc-(Fuc-Gal-GlcNAc-)GalNAcol	−	−	+
1389a	Fuc-Gal-**(NeuAc, HexNAc-Gal-4GlcNAc-6)**GalNAcol	+	+	−
1389b	NeuAc-Gal-GlcNAc**(Fuc-Gal-4GlcNAc-6)**GalNAcol	−	−	+
1389c	Gal-(NeuAc-GlcNAc**(Fuc-Gal-GlcNAc-6)**GalNAcol	−	−	+
Multiple acidic residues				
1120a	Gal-(SO_3^--Gal-(NeuAc-)GlcNAc-)GalNAcol	+	+	−
1120b	SO_3^--Gal-(NeuAc-)GlcNAc-Gal-GalNAcol	−	+	−
1169	NeuAc-Gal-(NeuAc-)GlcNAc-GalNAcol	+	−	+
1266	$[NeuAc]_1$ Fuc-Gal-(Gal-(SO_3^--)GlcNAc-)GalNAcol	+	+	−
1323	$[NeuAc]_1$ Gal-(HexNAc-Gal-(SO_3^--)GlcNAc-)GalNAcol	+	+	−
1331a	NeuAc-Gal-(Gal-(NeuAc-)GlcNAc-)GalNAcol	+	+	−
1331b	NeuAc-Gal-(NeuAc-)GlcNAc-Gal-GalNAcol	−	+	−
1331c	$[NeuAc]_1$ Gal, Gal-(NeuAc-)GlcNAc-GalNAcol	−	−	+
1372	HexNAc-(NeuAc-)Gal-(NeuAc-)GlcNAc-GalNAcol	+	−	+
1469	$[NeuAc]_1$ Fuc-Gal-(HexNAc-Gal-(SO_3^--)GlcNAc-)GalNAcol	+	+	−
1477a	Fuc-Gal-(NeuAc-Gal-(NeuAc-)GlcNAc-)GalNAcol	+	+	−
1477b	$[NeuAc]_2$ $[Gal]_2$ $[Fuc]_1$ $[GlcNAc]_1$ GalNAcol	−	−	+
1526	$[SO_3^-]_1$ $[NeuAc]_1$ $[HexNAc]_3$ $[Gal]_2$ GalNAcol	+	+	−
1535a	Gal-(HexNAc-(NeuAc-)Gal-(NeuAc-)GlcNAc-)GalNAcol	+	+	−
1535b	$[NeuAc]_2$ $[HexNAc]_2$ $[Gal]_2$ GalNAcol	−	−	+
1566	$[NeuAc]_1$ Gal-(SO_3^--Gal-GlcNAc-Gal-(SO_3^--6)GlcNAc-)GalNAcol	+	+	−
1575	$[NeuAc]_1$ HexNAc-(HexNAc-Gal-(NeuAc-)GlcNAc-)GalNAcol	+	−	+
1681a	Fuc-Gal-(HexNAc-(NeuAc-)Gal-(NeuAc-)GlcNAc-)GalNAcol	+	+	−
1681b	NeuAc-Gal-(NeuAc-)GlcNAc-**(Fuc-Gal-4GlcNAc-6)**GalNAcol	−	−	+
1738a	$[NeuAc]_2$ GlcNAc-Gal-(HexNAc-Gal-GlcNAc-)GalNAcol	+	+	−
1738b	$[NeuAc]_2$ Gal-GlcNAc-(HexNAc-Gal-GlcNAc-)GalNAcol	−	−	+
1769	$[NeuAc]_1$ Gal-(HexNAc-(SO_3^-)-Gal-GlcNAc-Gal-(SO_3^--6)GlcNAc-)GalNAcol	+	−	−

Continued

Table I. (Continued)

Label	Proposed oligosaccharide sequence/composition[a]	Mouse model		
		WT	Core 3 KO	Core 1 KO
1826	NeuAc-Gal-(NeuAc-)GlcNAc-(Gal-(NeuAc-)GlcNAc)GalNAcol	−	−	+
1835	$[SO_3^-]_1$ $[NeuAc]_1$ $[Fuc]_1$ $[HexNAc]_3$ $[Gal]_3$ GalNAcol	+	+	−
1843a	$[NeuAc]_2$ Gal, Gal-GlcNAc-**(Fuc-Gal-4GlcNAc-6)**GalNAcol	−	−	+
1843b	$[NeuAc]_2$ $[Fuc]_1$ $[HexNAc]_2$ $[Gal]_3$ GalNAcol	+	+	−
1884	$[NeuAc]_1$ NeuAc,HexNAc-Gal-GlcNAc-**(Fuc-Gal-4GlcNAc-6)**GalNAcol	+	−	+
1941a	$[NeuAc]_2$ HexNAc-Gal-GlcNAc-(HexNAc-Gal-GlcNAc-)GalNAcol	+	−	+
1941b	$[NeuAc]_2$ $[HexNAc]_4$ $[Gal]_2$ GalNAcol	+	+	+
1980	$[SO_3^-]_1$ $[NeuAc]_2$ $[HexNAc]_3$ $[Gal]_3$ GalNAcol	+	+	−
2029	$[NeuAc]_3$ Gal-GlcNAc-(HexNAc-Gal-GlcNAc-)GalNAcol	+	−	+
2117	NeuAc-Gal-(NeuAc-)GlcNAc-(NeuAc-Gal-(NeuAc-)GlcNAc)GalNAcol	−	−	+
2232a	$[NeuAc]_3$ HexNAc-Gal-GlcNAc-(HexNAc-Gal-GlcNAc)GalNAcol	−	−	+
2232b	$[NeuAc]_3$ $[HexNAc]_4$ $[Gal]_2$ GalNAcol	+	−	+

[a]Assumptions: Hex residues are Gal, deoxyHex are Fuc, HexNAc within chain are GlcNAc and HexNAcol are GalNAcol. Residues attached to C6 of GalNAcol are in bold.

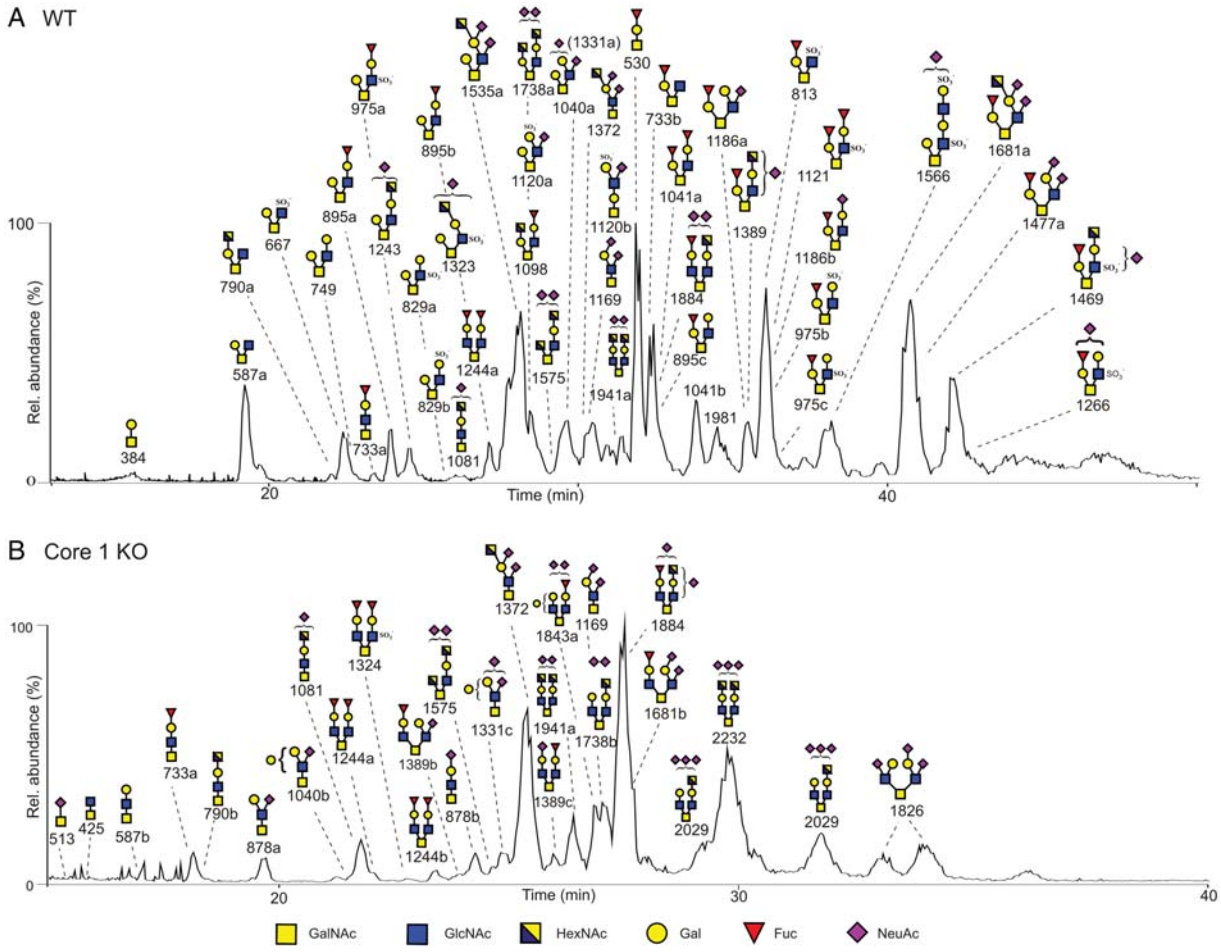

Fig. 2. Negative-ion mode capillary-LC/MS base peak chromatograms of *O*-glycan alditols from colonic Muc2 mucin from WT mouse (**A**) and from the core 1 knockout (IEC *C1Galt*$^{-/-}$) mouse (**B**). Major components are annotated with *m/z* ([M–H$^+$]$^-$) and proposed sequences interpreted from MS^2 spectra. The structural data are compiled in Table 1, with more details in Supplementary data, Table S1.

annotated with their deprotonated mass (M–H). The identification of *O*-glycan sequences was performed by the manual interpretation of tandem mass spectrometry (MS/MS) spectra, and compared with the reference spectra of *O*-glycans collected from mucins from other sources. Interpretation of diagnostic fragment ions has been investigated and discussed

elsewhere (Thomsson et al. 2000; Karlsson et al. 2004; Robbe et al. 2004a; Larsson et al. 2009). MS^2 spectra collected in this work are made available in the public UniCarb-DB database (www.unicarb-db.com). Components with identical MS/MS spectra found on Muc2 from both the IEC *C1Galt1*$^{-/-}$ and the WT mice were annotated similar. Their differing retention time in the two samples reflected that these were analyzed with different gradients and at different occasions, but they may also be linkage isomers. Complete linkage and sequence information cannot be obtained with mass spectrometry.

The base peak chromatogram obtained from the IEC *C1Galt1*$^{-/-}$ mouse was markedly different from the WT mouse, highlighting that colonic Muc2 was differently glycosylated in the two mice (Figure 2). In contrast, the base peak chromatogram from the *C3Gnt*$^{-/-}$ mouse was very similar to the WT, although the relative abundance of many glycan components varied slightly when judging from the relative MS ion intensities (Supplementary data, Figure S1). We had previously observed that colonic Muc2 from the WT and *C3Gnt*$^{-/-}$ mice migrated similarly and displayed a nearly identical Alcian Blue stained pattern when analyzed with composite gel electrophoresis (Supplementary data, Figure S1), in contrast to Muc2 from *C1Ga1t1*$^{-/-}$ mouse (not shown). This migration pattern most likely reflected that Muc2 from WT and *C3Gnt*$^{-/-}$ mice were similarly glycosylated. We know from previous studies of differently charged glycoforms of MUC5B that the glycosylation affects migration on composite gels (Thomsson et al. 2005).

Colonic Muc2 from the WT mouse carried a mixture of neutral and acidic glycans of predominantly core 1 and core 2 type. Fifteen neutral and 32 acidic (sulfated and/or sialylated) glycans were sequenced, 10 of which were low-abundant core 3 and core 4 type. Structural isomers were abundant. A complete list of characterized components is found in Table 1 (more details in Supplementary data, Table S1). Additional low-abundant components were observed in all three mouse models, where the MS^2 spectra contained fragment ions indicative of glycan alditols, but these could not always be sequenced. Nevertheless, these observations indicated that the degree of structural complexity on Muc2 from mouse colon resembled that of human colonic MUC2, where more than 100 glycans were detected (Larsson et al. 2009).

Major peaks in the LC/MS chromatogram of WT-derived glycans corresponded to the neutral glycans Fuc-Gal-GalNAcol (labeled 530) and Fuc-Gal-(HexNAc-)GalNAcol (733b), and these were flanked by sulfated and sialylated glycans (1535a, 813 and 1681a) which generated MS peak intensities of nearly equal size (Figure 2A). A common structural feature was that the core 1 type branches carried a single Gal or Fuc-Gal linked to the core *N*-acetylgalactosaminitol (GalNAcol), but no further extension was observed.

Fucose (Fuc) was found attached exclusively to Gal residues, and no Fuc linked to any *N*-acetylhexosamine (HexNAc) as in Lewis-type epitopes was detected. Sulfate was predominantly attached to the core HexNAc, but also found on terminal Gal residues. NeuAc was found on both Gal and internal HexNAc residues, the latter which was a surprising finding since this is not commonly reported on mucins. Evidence for NeuAc linked directly to the core GalNAcol (NeuAcα2-6GalNAc), which is frequent on human colonic MUC2, was found in trace amounts in the IEC *C1Galt1*$^{-/-}$ mouse. NeuAc-GalNAcol was also detected in low amounts on Muc2 from the small intestine of the WT mouse (not shown).

The glycan profile from the IEC *C1Galt1*$^{-/-}$ mouse displayed predominantly di- and triply fucosylated and sialylated glycans of core 3 and core 4 types (Figure 2B, Table 1). Six neutral, 1 sulfated and 22 sialylated glycans were sequenced, all of which were found as multiple structural isomers. All 10 core 3 and core 4 type glycans identified in the WT mouse were found in the IEC *C1Galt1*$^{-/-}$ mouse, which included the two abundant core 4-type glycans labeled 1372 and 1884 in the middle of the chromatogram (Figure 2).

A low-abundant glycan interpreted as a sialylated Hex-HexNAcol unit at *m/z* 675 was detected in the IEC *C1Galt1*$^{-/-}$ mouse, which was not expected in this mouse model as this composition would imply a core 1-type glycan. However, this particular isomer had a different MS/MS spectrum and retention time when compared with the two core 1-based glycans at *m/z* 675 found in human MUC2 (NeuAcα2-3Galβ1-3GalNAc or Galβ1-3(NeuAcα2-6)GalNAc) (Supplemental data, Figure S2). This component was also detected on colonic Muc2 from WT and the *C3Gnt*$^{-/-}$ mouse (not shown), and all three 675 isomers were found on Muc2 from WT mouse small intestine. This 675 isomer is thus not a core 1 glycan, although its precise nature cannot be established by LC/MS. A potential interpretation would be a NeuAc-HexNAc-Hexol sequence, formed from either a peeling product or, for example, the rare reducing core *O*-Man previously reported in the murine stomach tissue from the triple knockout core 2 mice (Ismail et al. 2011). A HexNAc-Hexol fragment would also generate a fragment at *m/z* 384 as we observe here. Another explanation is that it may be the rare core 8 sequence (Galα1-3GalNAc) which has been reported on mucins from human respiratory mucus (van Halbeek et al. 1994).

Treatment of colonic Muc2 O-glycans with neuraminidases

O-Glycan mixtures from colonic Muc2 from the three mouse strains were subjected to general desialylation using neuraminidases from *Arthrobacter ureafaciens* and *Vibrio cholerae*. The main purpose was to aid sequence assignment, by eliminating the NeuAc residues which are the most labile residues in CID fragmentation, more fragment ions are obtained from the remaining glycan backbone. Complete desialylation was not obtained in any of the samples, despite attempts with the two different neuraminidases. This confirmed that many NeuAc residues were most likely internally positioned on the glycan core, which was reported could interfere with neuraminidase accessibility.

The base peak chromatograms from LC/MS of Muc2 *O*-glycans from WT and IEC *C1Galt1*$^{-/-}$ mice are shown in Figure 3 where some of the major sequences are annotated. As only a minor portion of the WT mouse glycans were sialylated, this glycan profile was not considerably altered (Figure 3A). This is in contrast to the highly sialylated glycan profile from the IEC *C1Galt1*$^{-/-}$ mouse, which was markedly changed (Figure 3B). A major glycan at *m/z* 952 interpreted

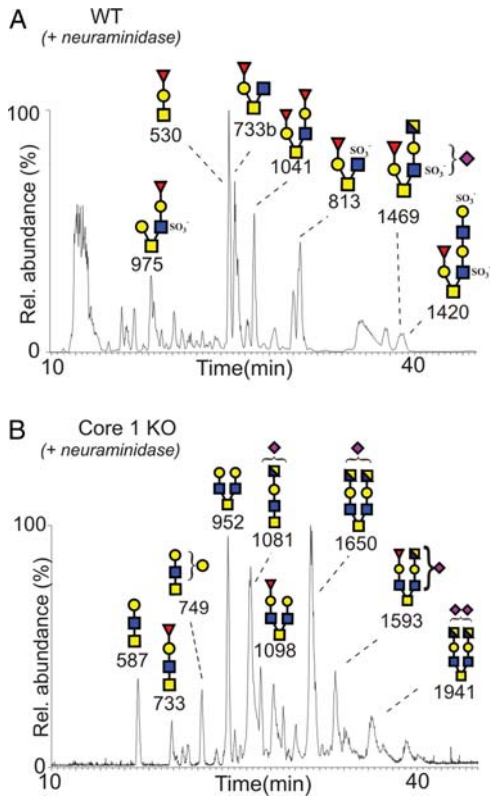

Fig. 3. Negative-ion mode capillary-LC/MS base peak chromatograms of mucin *O*-glycan alditols from colonic Muc2 from WT mouse (**A**) and from the core 1 knockout (IEC *C1Galt1*$^{-/-}$) mouse (**B**) after neuraminidase treatment. Major components are annotated with *m/z* ([M–H$^+$]) and proposed sequences interpreted from MS^2 data.

as Gal-GlcNAc-(Gal-GlcNAc-)GalNAcol was observed (Figure 3B). Sialylated "precursor" glycans in the non-treated sample with this sequence may be the triply sialylated compounds labeled 1826 eluting at 31–33 min (Figure 2B), but also mono-, di- and trisialylated isomers which originally failed to be sequenced due to too low abundance. The neuraminidase treatment thus revealed that sialylated glycans with this sequence were abundant components. Other major components in the neuraminidase-treated sample were the incompletely desialylated glycans labeled 1081 and 1650. The MS^2 spectra of the components 952 and 1650, the latter interpreted as HexNAc-Gal-GlcNAc-(HexNAc-Gal-GlcNAc)GalNAcol (+ NeuAc) are shown in Supplementary data, Figure S3.

The MS^2 spectrum of an abundant desialylated glycan at *m/z* 749 consisting of two Gal, one GlcNAc and GalNAcol from the IEC *C1Galt1*$^{-/-}$ mouse is shown in Supplementary data, Figure S3. This was interpreted as a linear (Gal-Gal-GlcNAc-) or a branched (Gal-(Gal-)GlcNAc-) glycan, based on the presence of an intense fragment ion at *m/z* 526 (b-ion), and fragment ions diagnostic of HexNAc linked to the core GalNAcol at *m/z* 407/425 indicative of a core 3-type glycan. The spectrum provided unequivocal evidence for an unusual mucin glycan epitope. It also confirmed the interpretation of three glycans from the core 1 knockout mouse (1040b, 1331c, 1843a, Table 1 and Supplementary data, Table S1) with the same epitope. The monosaccharide composition with two Hex and one HexNAc residues would otherwise have been interpreted as of core 2 type, which should not occur in this mouse model. The linear sequence (Gal-Gal-GlcNAc-) has been reported on *O*-glycans from gastrointestinal tissue from WT mouse by others (Magalhaes et al. 2009; Ismail et al. 2011).

To conclude, neuraminidase treatment of the Muc2 WT and *C1Galt1*$^{-/-}$ mouse derived glycans aided interpretation of the glycan sequence backbone and revealed a fraction of internally positioned NeuAc which could not be removed. The unique presence and interpretation of core 3- and core 4-type glycans on Muc2 from the core 1 knockout mouse (*C1Galt1*$^{-/-}$) was confirmed.

Identification of NeuAc linked to internal HexNAc residues on colonic Muc2 O-glycans from mouse

The finding that NeuAc was linked to internal HexNAc residues was surprising, since this is an unusual epitope on mucin *O*-glycans. The MS/MS spectra of two short sialylated oligosaccharides proposed to contain NeuAc linked to internal HexNAc residues are shown in Figure 4 together with the corresponding mass chromatograms. The MS^2 spectrum of an isomer labeled 878a from colonic MUC2 from the IEC *C1Galt1*$^{-/-}$ mouse consisting of NeuAc, Gal, HexNAc and GalNAcol is shown in Figure 4A. Fragment ions at *m/z* 407 and 364 were used to assign part of the sequence to Gal-HexNAc-GalNAcol. A major fragment ion was found at *m/z* 657, corresponding to the precursor ion minus 221 [M-221]. This fragment ion has been proposed to be formed by a cross ring cleavage (^{0,2}X) in a NeuAc residue and diagnostic for NeuAc linked α2-6 to the core GalNAcol, an epitope commonly found on MUC2 *O*-glycans from the human colon (Robbe et al. 2004b; Larsson et al. 2009). However, the presence of a fragment ion at *m/z* 493 that could correspond to NeuAc linked to a HexNAc, lead us to the interpretation of this sequence as Gal-(NeuAc-)GlcNAc-GalNAcol. Our interpretation was confirmed by the observation that the MS^2 spectrum of this compound differed from the MS^2 spectra of the sequences with the same composition found in the human colon interpreted as Gal-GlcNAc-(NeuAc-6)GalNAcol (see spectra database at http://www.medkem.gu.se/mucinbiology).

The MS^2 spectrum of an isomer at *m/z* 1040 from the WT mouse colon is shown in Figure 4B and was dominated by a major fragment ion at M-221 (*m/z* 819). The presence of a y-ion at *m/z* 384 and the z/y-ions at 389 and 407 supported a core 2-type glycan. This excluded the possibility of having NeuAc linked α2-6 to the core GalNAcol and the sequence was interpreted as Gal-(Gal-(NeuAc-)GlcNAc)GalNAcol.

The major sialylated glycans on Muc2 from WT mouse colon were two disialylated components labeled as 1535a and 1681a eluting at 28 and 41 min, respectively (Figure 2A). Their structures were interpreted as Fuc-Gal-(HexNAc-(NeuAc-)Gal-(NeuAc-)GlcNAc-)GalNAcol (1681a) or the non-fucosylated isomer (1535a). These were also found on Muc2 from *C3Gnt*$^{-/-}$ mice, but not from the IEC *C1Galt1*$^{-/-}$ mouse. Interpretation of the MS^n spectra obtained from the disialylated compounds 1681a and also a second compound labeled 1372 (HexNAc-(NeuAc-)Gal-(NeuAc-)GlcNAc-

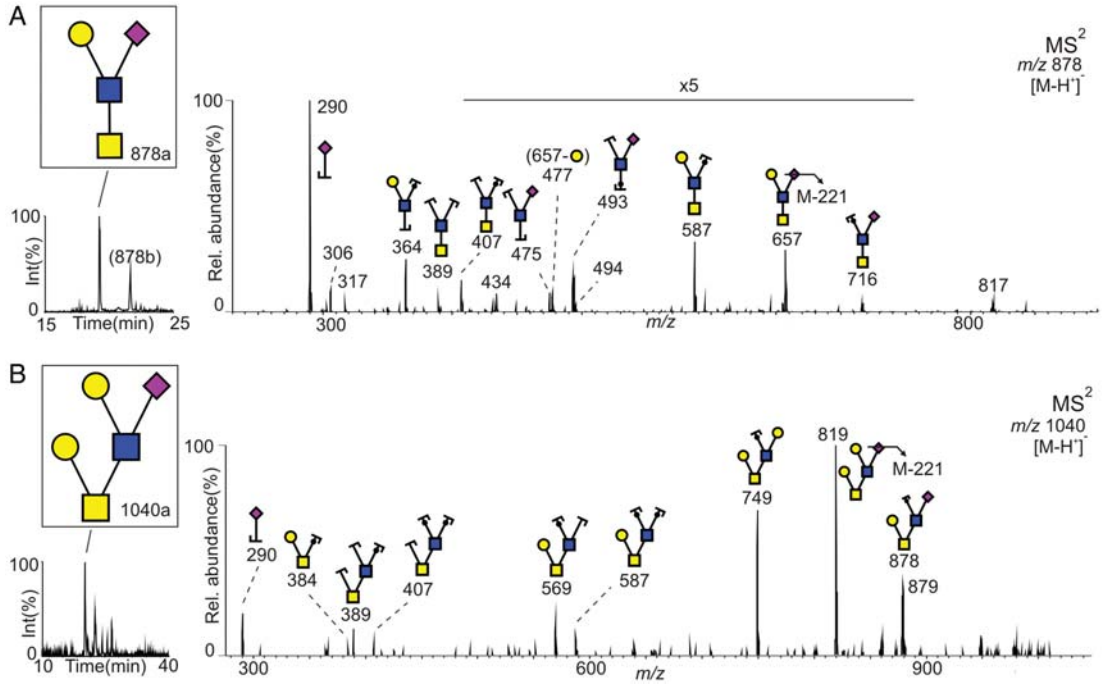

Fig. 4. MS^2 spectra from negative-ion mode capillary LC/MS of colonic Muc2 glycan alditols labeled 878a with the sequence Gal-(NeuAc-)GlcNAc-GalNAcol from the core 1 knockout (IEC *C1Galt1*$^{-/-}$) mouse (**A**) and 1040a with the sequence Gal-(Gal-(NeuAc-)GlcNAc-)GalNAcol from the WT mouse (**B**). Corresponding mass chromatograms are inserted.

GalNAcol) on Muc2 from WT and the *C1Galt1*$^{-/-}$ mice are discussed in Supplemental data, Figure S4.

Confirmation of NeuAc linked to internal GlcNAc residues by MS and NMR analyses

In order to confirm the fragment ions at *m/z* 493 and [M-221] interpreted as diagnostic for NeuAc linked to internal GlcNAc residues, we analyzed two reduced glycan standards containing internal NeuAcα2-6GlcNAc linkages by negative-ion mode MS^n experiments using syringe infusion (Supplemental data, Figures S5 and S6). These analyses showed that the cross-ring cleavage of NeuAc which generated the diagnostic ion at [M-221] was readily formed from NeuAc linked α2-6 to internal GlcNAc (^{0,2}X). The fragment ion at *m/z* 493 was found to be diagnostic for a NeuAc bound to an internal GlcNAc residue in the glycan backbone. MS^3 of *m/z* 493 confirmed that the fragment ion consisted of NeuAc and GlcNAc. However, this ion was only found in the MS^2 spectra of the disialylated standard DSLNT (disialyl-lacto-*N*-tetraose), but not in the monosialylated standard LSTb (LS-tetrasaccharide b). MS^2 spectra of one-third sialylated standard with the sequence NeuAcα2-6Galβ1-3Glcol revealed that the fragment ion at [M-221] could also be formed from NeuAc linked α2-6 to Gal, although at low intensity (Supplemental data, Figure S6). Together, the results show that the fragment ion at [M-221] could originate from NeuAc linked to Gal, GlcNAc or GalNAcol when fragmented by ESI-ion trap CID in the negative-ion mode, although it is much more abundantly formed when NeuAc is linked to the amino sugars. The fragment ion at *m/z* 493 can be used as a marker for NeuAc bound to internal GlcNAc, but is not always detected.

In order to obtain more information on NeuAc linkages in mouse colon, the Muc2 *O*-glycan alditol mixture from the WT mouse and five reduced glycan standards containing NeuAc-linked α2-3Gal, α2-6GlcNAc or α2-6GalNAcol were analyzed with one-dimensional and two-dimensional proton NMR spectroscopy. NeuAc residues in the Muc2 glycan mixture were easily identified in the NMR spectrum by their H3ax and H3eq signals, shown in the one-dimensional spectra in Figure 5B. Cross-peaks corresponding to NeuAc-linked α2-6 were detected at ~1.70/2.76 ppm (H3ax/H3eq), but no peaks were found in the region where NeuAc in other epitopes were expected, as in Sd^a epitopes (1.92/2.65 ppm, H3ax/H3eq) or NeuAcα2-3Gal- (1.80/2.76 ppm, H3ax/H3eq; Capon et al. 2001; Robbe et al. 2003). The lack of signals from α2-3-linked NeuAc in the Muc2 glycan mixture suggests this linkage type to be less abundant than α2-6-linked NeuAc. Interpretation of H3ax and H3eq signals from one-dimensional NMR of five sialylated glycan standards analyzed separately did support that the cross-peaks observed in the COSY spectrum corresponded to NeuAc-linked α2-6 of an internal GlcNAc (H3eq = 2.74–2.75 ppm) rather than the reducing core GalNAcol (H3eq = 2.72 ppm; Figure 5A). Although these data are strong, we cannot exclude minor amounts of GalNAc at this position. Thus, both LC-MS and NMR spectroscopy confirmed that NeuAc bound α2-6 to an internal NAc was an abundant structural motif in mouse Muc2.

A

Assignment of signals (ppm) in ^{1}H NMR analysis of NeuAc-containing glycan alditols.

NeuAcα2-3Galβ1-3GalNAcol (**A**)
NeuAcα2-3Galβ1-3(NeuAcα2-6)GalNAcol (**B**)
NeuAcα2-6Galβ1-4Glcol (**C**)
Galβ1-3(NeuAcα2-6)GlcNAcβ1-3Glcol (**D**)
NeuAcα2-3Galβ1-3(NeuAcα2-6)GlcNAcβ1-3Galβ1-4Glcol (**E**)

NeuAc		A	B	C	D	E
α2-3	H3eq	2.77	2.77			2.76
	H3ax	1.80	1.80			1.78
α2-6	H3eq		2.72	2.74	2.75	2.74
	H3ax		1.69	1.71	1.70	1.69

B

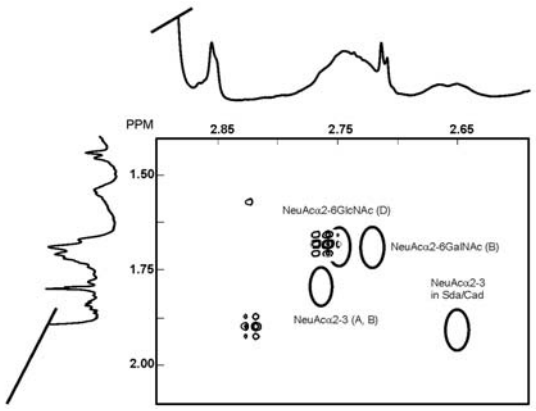

Fig. 5. Chemical shift values for NeuAc H3eq and H3ax signals for reduced sialylated standards analyzed with one-dimensional proton NMR spectroscopy at 25°C (**A**). Selected region of the two-dimensional proton NMR spectrum encompassing NeuAc H3eq/H3ax cross-peaks from the analysis of colonic Muc2 *O*-glycan alditols from WT mouse (**B**). The areas where cross-peaks from (α2-3)-linked NeuAc, (α2-6)-linked NeuAc and in Sd^a epitopes were expected according to our reference samples A–D and also published by others (Capon et al. 2001; Robbe et al. 2003) are highlighted by ellipses.

Discussion

In the work presented here, we have performed a descriptive study of the murine *O*-glycans on colonic Muc2, which constitutes the major protein component of the mucosal barrier. We have previously screened MUC2 *O*-glycosylation from the sigmoid colon of 53 individuals using the same analytical approach as used here. Our data reveal that the *O*-glycosylation of murine colonic Muc2 differs in several aspects from human colonic MUC2. Human MUC2 carries predominantly core 3- and core 4-based *O*-glycans terminated with NeuAc as illustrated in Figure 6 (Capon et al. 2001; Robbe et al. 2004b; Larsson et al. 2009). Here, we highlight that murine colonic Muc2 is instead characterized by core 1- and core 2-type glycans and also with a considerably larger percentage of neutral glycans compared with the human MUC2 (Figure 6). The predominance of these core structures and the distribution between neutral and charged glycans are in line with what we have observed in our previous, more limited studies of Muc2 derived *O*-glycans from colon from WT and the cystic fibrosis mouse model (Thomsson et al. 2002; Hurd et al. 2005). These initial interpretations were based on monosaccharide analyses of fractionated glycans, and high-temperature GC-MS of neutral and sialylated glycans of only up to five residues. The current study provided detailed structural evidence for that the sialylation of human colonic MUC2 differs from that of mouse. NeuAc terminally positioned in Sd^a epitopes (GalNAcβ1-4(NeuAcα2-3)Gal-) is observed on human MUC2 (Capon et al. 2001) and occurs in Sd^a-like epitopes in mouse as shown here. However, the frequent NeuAc linked α2-6 to the core GalNAc as found in humans was totally absent on Muc2 from the WT mouse colon. Instead NeuAc was attached to internal GlcNAc residues adjacent to the core GalNAc, most likely in a α2-6 linkage. NeuAc linked to internal GlcNAc residues is only rarely found on human mucins, but still present. We have observed very low-abundant compounds on human colonic MUC2 with NeuAc linked to internal GlcNAc (Larsson et al. 2009). Internally positioned NeuAcα2-6GlcNAcβ- epitopes are described on mammalian *N*-linked glycans (glycosuitedb.expasy.org).

There are 20 members in the mammalian sialyltransferase family identified to date. Six of these are classified as GalNAc α2,6-sialyltransferases with developmental stage- and tissue-specific expression as well as precursor specificity (Takashima 2008). ST6GalNAc V and VI have been shown to add NeuAc to sialolactotetraosylceramide on the internally positioned GlcNAc residue (NeuAcα2-3Galβ1-3GlcNAcβ1-3Galβ1-4Glcβ1-Cer). This implies that there are ST6GalNAc transferases that are candidates for sialylating internal GlcNAc residues. However, we currently do not know which of the sialyltransferases that are expressed in the mucin producing goblet cells in the human or mouse colon.

Fuc linked to GlcNAc residues, as in Lewis type sequences (Galβ1-3/4(Fucα1-3/4)GlcNAc-), found on MUC2 from human colon was not present in mouse. Blood group H-related Fucα1-2Gal- epitopes, which are found in decreasing amounts on MUC2 through the human colon and virtually absent in human sigmoid (Robbe et al. 2003; Robbe et al. 2004b; Larsson et al. 2009), were abundantly found in mouse colon.

Comprehensive studies of *O*-glycans extracted from total colon including both epithelia and stroma from WT mice from a mixed background have been performed by Ismail et al. (2011). They analyzed the *O*-glycans permethylated using MALDI-MS and/or direct infusion ESI-MS and MS/MS. Their results revealed predominantly core 1- and core 2-type *O*-glycans. In contrast to us, they detected glycans indicative of NeuAc linked directly to the core GalNAcol, Fuc in Lewis-type epitopes, I-branching on the core 1 Gal and the presence of NeuGc. These discrepancies are most likely explained by their *O*-glycans emanating from other proteins than Muc2 and from many different cell types, as neither the cells nor the glycoproteins were purified. Ismail et al. identified a low-abundant *O*-glycan in their colonic tissue from WT and core 2 knockout mice having a disialylated sequence similar to the glycolipid RM2 antigen (Saito et al. 2005). This epitope was now found as an abundant one on Muc2 (Figure 2; 1535a and 1681a). The RM2 antigen is composed

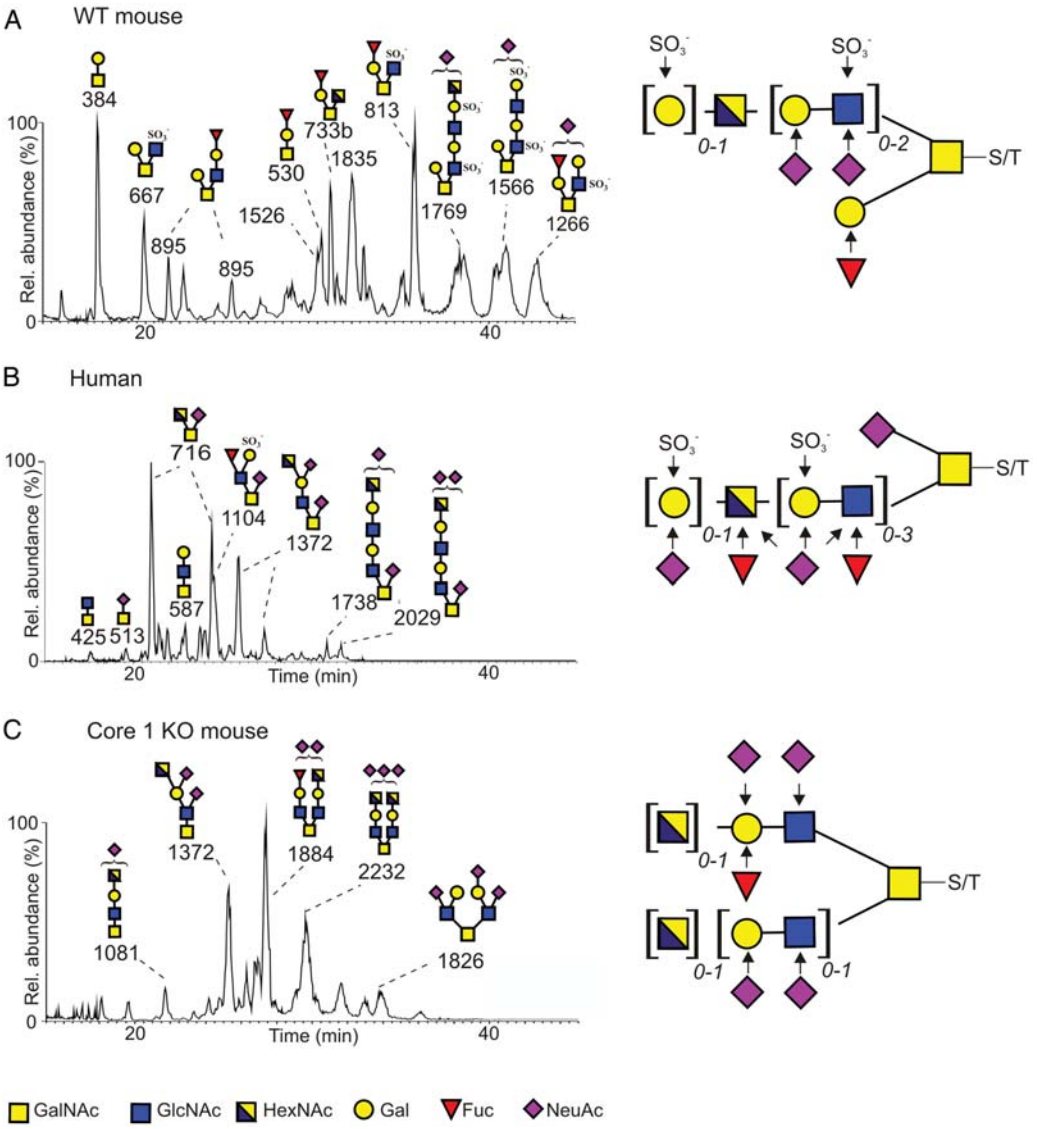

Fig. 6. Negative-ion mode capillary-LC/MS base peak chromatograms of Muc2 *O*-glycan alditols purified from a mucosal scraping of the distal colon from a WT mouse (**A**), from human biopsies from sigmoideum (**B**) and from colon of the core 1 (IEC *C1Galt1*$^{-/-}$) knockout mouse. The sequences of major components are annotated. A general structure of the oligosaccharides is depicted adjacent to each chromatogram.

of the Sd^{a} antigen adjacent to a internally positioned sialylated GlcNAc (GalNAcβ1-4(NeuAcα2-3)Gal-(NeuAcα2-6) GlcNAc-).

Sulfation of the Muc2 *O*-glycans from mouse colon is also different from the human one. Human MUC2 is mainly sulfated on Gal residues and less abundant on GlcNAc residues (Robbe et al. 2003, 2004b; Larsson et al. 2009). In mouse colon, the opposite is found as sulfation is most frequently observed C6-linked on the GlcNAc. The major Muc2 sulfation in murine colon has been shown to be mediated by the GlcNAc6ST-2 transferase (Tobisawa et al. 2010).

The glycans found on colonic Muc2 from the IEC *C1Galt1*$^{-/-}$ mouse were more sialylated than in the WT, but at the same time only poorly sulfated compared with the WT mouse. The reason for this is not obvious, but suggests that the GlcNAc6ST-2 transferase may be dependent on a C3Gal for its activity as all identified C6GlcNAc sulfated glycans also had a C3Gal substitution. Only a very minor compound in the *C1Galt1*$^{-/-}$ mouse carried a sulfate group on the C6GlcNAc. As the IEC *C1Galt1*$^{-/-}$ mouse had proportionally more sialic acids, it can be suggested that the core 3 and core 4 glycans can act as efficient precursor substrates for sialyltransferases.

Since transient glycosylation changes on MUC2 can be induced during infection/inflammation (Olson et al. 2002; Larsson et al. 2011), it could not be excluded that the ongoing inflammatory response in IEC *C1Galt1*$^{-/-}$ mouse could have been accompanied by an increased expression of sialyltransferases. However, studies of Muc2 *O*-glycosylation profiles of IEC *C1Galt1*$^{-/-}$ mice (5-week old) treated with 4

weeks of broad spectrum antibiotics revealed no major differences in the expression of sialylated or sulfated glycans (not shown). As less inflammation was observed in these mice, it is less likely that inflammation is the reason for the high sialylation of the IEC *C1Galt1*$^{-/-}$ mice.

From our previous work on the IEC *C1Galt1*$^{-/-}$ mouse, it was hypothesized that Muc2 would carry a larger portion of *O*-glycans that were not further extended on GalNAc attached on the protein core due to lack of the core 1 (C1GalT1) transferase. This was supported by an increased detection of the Tn antigen (GalNAc-Ser/Thr) in the knockout mouse compared with WT (Fu et al. 2011), but as our MS method does not allow the detection of a single GalNAc, this could not be further quantified here. However, the complex repertoire of *O*-glycans in these mice (Figure 2) still suggests a substantial amount of larger glycans.

The core 3 (C3GnT1) transferase is only expressed throughout the human gastrointestinal epithelia (Iwai et al. 2002, 2005), in contrast to the core 1 (C1GalT1) transferase, which is found in many tissues. The *C3Gnt*$^{-/-}$ mice developed normally, but were found to be more susceptible to DSS-induced colitis than the WT (An et al. 2007). Considering the *O*-glycan profile presented in this paper, we could show that core 3/4-type glycans are only expressed in low amounts in the colon of the WT mouse, and the loss of these glycans could therefore be expected to have limited effect on Muc2 and the mucus barrier. Accordingly, the animals had no inflammation unless challenged with DSS. This is in contrast to the core 1/2-type glycans that are predominant in mice, explaining why the IEC *C1Galt1*$^{-/-}$ mice develop severe colitis (Fu et al. 2011). It can be estimated that core 3/4-type glycans constitute 1% of the *O*-glycans on the WT mouse Muc2 and it was thus surprising that the *C3Gnt*$^{-/-}$ mice displayed a phenotype at all. The lack of only small amounts of core 3/4-type glycans may, however, affect the composition of the commensal microbiota and thus contribute to a higher susceptibility to DSS. In support of such an interpretation the core 3/4 glycans observed in the WT and the IEC *C1Galt1*$^{-/-}$ mice were unique. These glycans are extended and terminated differently than the core 1- and 2-type ones in the WT mice as they are on the average longer, more sialylated and less sulfated.

Characterization of the colon microbiota using modern DNA-sequencing methodology has provided evidence that each species carries unique core microbiota. 85% of the 16S rRNA sequences are not shared between mouse and human (Ley et al. 2005), but both are still dominated by the Firmicutes and Bacteroidites. Little is known on how this relatively stable colon microbiota is maintained, but cross-species bacterial transplantation between mouse and zebra fish suggest that the host select bacterial species (Rawls et al. 2006). Diet can also influence the commensal flora as illustrated in mouse models of obesity (Ley et al. 2005). Considering that the microbiota have their habitat in the outer mucus layer of colon may suggest that the mucus and its main component, the MUC2 mucin, could be an important reason for this selection (Johansson et al. 2008). The observations that the colon MUC2 is so differently *O*-glycosylated in different species and a uniform glycosylation in all human sigmoid colons (Larsson et al. 2009) may suggest MUC2 glycans as an important factor. It is well known that commensal bacteria carry surface adhesins with specificity for glycan epitopes (Karlsson 1989; Juge 2012). Thus, it could be hypothesized that bacterial adhesion to host mucin glycans is a selection mechanism. The results presented highlight the differences in MUC2 glycans between humans and mice as potential explanation for their different microbiota. The mechanism for this selection can be explored in more detail using the two glycan-deficient mouse models characterized here.

Material and methods

Animals and mucin sample preparation

The generation of IEC *C1Galt1*$^{-/-}$ and the *C3Gnt*$^{-/-}$ mice in the C57Bl/6 background is described elsewhere (An et al. 2007; Fu et al. 2011). Mucosal intestinal scrapings from *C1Galt1*$^{-/-}$ mice, *C3Gnt*$^{-/-}$ mice and WT mice (C57Bl/6) were collected and Muc2 was purified as described previously for total mucus (Johansson et al. 2009). Briefly, mucosal scrapings from 0.5–4 colons were extracted with guanidinium chloride and in the presence of protease inhibitor, followed by reduction and alkylation. Mucin oligosaccharides from a human biopsy (sigmoid colon) were purified from healthy individuals as described (Larsson et al. 2009). Ethical approval for the animal experiments was granted by the Animal Ethics committee and for humans by the Human Research Ethical Committee, both in Gothenburg.

Glycans

Reference glycans from MUC7 from human saliva and from MUC2 produced in LS174T cells were prepared as described (Axelsson et al. 1998; Thomsson et al. 2005). Sialylated glycan standards DSLNT, LSTb and 6′-sialyllactose were obtained from Dextra Laboratories (Reading, UK). *O*-glycan standards NeuAcα2-3Galβ1-3GalNAcol and NeuAcα2-3Galβ1-3(NeuAcα2-6)GalNAcol were purified as described (Olson et al. 2005).

Chemicals

HPLC grade acetonitrile and analytical grade acetic acid, methanol and ammonium hydroxide (25% solution) were from Merck (Darmstadt, Germany). All chemicals unless otherwise stated, were purchased from Sigma-Aldrich (St Louis, MO).

Mucin glycan release

Aliquots of purified Muc2 mucins were analyzed by agarose-SDS composite gel electrophoresis (Schulz et al. 2002). Mucins purified from one mouse colon were dissolved in 100-μL sample buffer and 20–25 μL loaded in each well. The gels were developed for 3 h at 30 mA or overnight at 10 mA on ice at 4°C. The gels were wet blotted to Immobilon PVDF P^{SQ} membrane (Millipore, Billerica, MA) at 40W for 3 h in a glycine buffer (25 mM Tris, 192 mM glycine, 0.04% SDS, 20% MeOH) on ice at 4°C and stained with Alcian Blue. The glycans were released from Muc2 by reductive beta elimination (0.5 M $NaBH_4$ in 50 mM KOH, 50°C over night) as glycan alditols and desalted (Schulz et al. 2002).

Syringe infusion of sialylated glycan alditol standards

DSLNT, LSTb and 6′-sialyllactose were reduced with 0.5 M $NaBH_4$ in 20 mM KOH (2 h, 50°C), desalted with AG50-8W cation exchange media (Bio-Rad, Hercules, CA) and excess borate evaporated with repeated extractions with 1% HAc in MeOH and lyophilized in a speedivac. The glycan alditols were dissolved in 40% acetonitrile (ACN) at 10 µg/mL and analyzed by syringe infusion at 10 µL/min into the LTQ mass spectrometer run in the negative-ion mode (Thomsson et al. 2010).

Desialylation of Muc2 O-glycan alditols

Released *O*-glycan alditol mixtures from Muc2 from WT, IEC *C1Galt1*$^{-/-}$ and *C3Gnt*$^{-/-}$ mice were subjected to neuraminidases reported to release sialic acids linked in a variety of linkages (α2-3,6,8). Glycans from Muc2 from WT mice (one colon) were dissolved in 75 µL of 50 mM sodium phosphate buffer, pH 6, and 7 µL of neuraminidase from *A. ureafaciens* (0.5 U/µL in reaction buffer; Sigma-Aldrich) at 37°C overnight, then desalted with 100 µL graphitized carbon (Hypercarb SPE, Hypersep, Hypersil, UK) in a centrifugal filter unit (0.5 mL, PVDF with pore size 0.65 µm, Millipore). The glycans were eluted from the column with 2 mL of 25% ACN + 0.05% trifluoro acetic acid (TFA), lyophilized and analyzed with MS. The glycans were further treated twice with neuraminidase from *V. cholerae* (1 U/mL, 100 µL, Roche) and 200 µL of 50 mM triethylammonium acetate buffer at pH 6.5 with 4 mM $CaCl_2$ at 37°C overnight followed by desalting as above. Glycans from Muc2 (half a colon) from IEC *C1Galt1*$^{-/-}$ and *C3Gnt*$^{-/-}$ mice were incubated with 3 µL of neuraminidase from *A. ureafaciens* at 37°C in reaction buffer over night. The glycans were purified with a C18 Zip Tip (Millipore), and the flow through fraction containing glycans was desalted using Supel-tips carbon pipette tips (10 µL; Supelco, Sigma-Aldrich). The glycans were eluted with 25% ACN with 0.05% TFA.

MS analysis of O-glycan alditols with capillary- or nano-ESI LC/MS^n

Glycans from Muc2 from ~7% of a murine colon were injected per analysis. Samples were dissolved in water. The HPLC columns were of graphitized carbon (Hypercarb) and packed in-house for capillary-LC (10 cm × 250 µm i.d.) or nano-LC (20–25 cm × 0.075 mm i.d.) and used with flow rates of 3 µL/min and 400 nL/min, respectively. Different solvent gradients with 8 mM NH_4HCO_3 (A) and ACN (B) for capillary LC were applied, usually 3–45% B or 5–33% B in 46 min after loading the samples in 0% B. Nano LC/MS^n was performed using 0.04% NH_3 (A) and ACN (B) at a gradient of 5–50% B in 48 min. The instrument was an LTQ linear quadrupole ion trap mass spectrometer (Thermo, San José, CA) operating in the negative ion mode. Capillary and nano-LC/MS^n (n = 2,3,4) using graphitized carbon columns of *O*-glycans has been described in detail elsewhere (Andersch-Bjorkman et al. 2007; Karlsson et al. 2009; Thomsson et al. 2010). Obtained MS^n spectra were interpreted manually by aid of the GlycoWorkbench software (Ceroni et al. 2008), the UniCarb-DB database, compared with available published reference spectra (Karlsson et al. 2004; Robbe et al. 2004a) and also to an in-house database containing MS^n spectra on MUC2 *O*-glycans from human colon. (http://www.medkem.gu.se/mucinbiology/) (Larsson et al. 2009). Spectra collected in the current work are available in the public UniCarb-DB database (www.unicarb-db.com).

Proton NMR spectroscopy

Reduced *O*-glycans from colonic Muc2 from seven WT mice and five reduced sialylated glycan standards (~50 µg of each) were subjected to deuterium exchange with 2 × 150 µL of D_2O (99.990%), then 2 × 150 µL of D_2O (99.996%) for 30 min, with lyophilization between each exchange. ^{1}H NMR spectra were acquired on a Varian 600 MHz instrument at 25°C. Two-dimensional double-quantum-filtered correlated spectroscopy (DQF-COSY) spectra were recorded by the standard pulse sequence (Marion and Wuthrich 1983). The residual water resonance (HDO) was used as internal standard at 4.798 ppm.

Supplementary data

Supplementary data for this article is available online at http://glycob.oxfordjournals.org/.

Funding

This work was supported by the Swedish Research Council (7461, 21027 and 342-2004-4434), The Swedish Cancer Foundation, The Knut and Alice Wallenberg Foundation, IngaBritt and Arne Lundberg Foundation, Sahlgren's University Hospital (LUA-ALF), Wilhelm and Martina Lundgren's Foundation, Torsten och Ragnar Söderbergs Stiftelser, The Sahlgrenska Academy, National Institute of Allergy And Infectious Diseases (U01AI095473, the content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH), NIDDK (R01DK085691 to L.X.) and The Swedish Foundation for Strategic Research—The Mucus-Bacteria-Colitis Center (MBC) of the Innate Immunity Program.

Acknowledgements

Hasse Karlsson is acknowledged for his help with the nano-LC/MS experiments and for comments to the MS data interpretation. We are grateful to Niclas G Karlsson for valuable input to the discussion.

Conflict of interest

None declared.

Abbreviations

ACN, acetonitrile; C1galT1, core 1 β1,3-galactosyltransferase; *C3Gnt*$^{-/-}$, core 3 β1,3-*N*-acetylglucosaminyltransferase; DSLNT, disialyl-lacto-*N*-tetraose; ESI-MS, electrospray ionization mass spectrometry; Fuc, fucose; GalNAc, *N*-acetylgalactosamine; GalNAcol, *N*-acetylgalactosaminitol; GlcNAc, *N*-acetylglucosamine; Hex, hexose; HexNAc, *N*-acetylhexosamine; IEC *C1Galt1*$^{-/-}$, epithelial cell-specific core 1 β1,3-galactosyltransferase; LC/MS, liquid chromatography-mass spectrometry; LSTb, LS-tetrasaccharide b; MALDI-MS, matrix-assisted laser desorption ionization-mass spectrometry;

MS/MS, tandem mass spectrometry; PTS, proline–serine–threonine; TFA, trifluoro acetic acid; WT, wild-type.

References

An G, Wei B, Xia B, McDaniel JM, Ju T, Cummings RD, Braun J, Xia L. 2007. Increased susceptibility to colitis and colorectal tumors in mice lacking core 3-derived O-glycans. *J Exp Med*. 204:1417–1429.

Andersch-Bjorkman Y, Thomsson KA, Holmen Larsson JM, Ekerhovd E, Hansson GC. 2007. Large scale identification of proteins, mucins, and their O-glycosylation in the endocervical mucus during the menstrual cycle. *Mol Cell Proteomics*. 6:708–716.

Axelsson MA, Asker N, Hansson GC. 1998. O-glycosylated MUC2 monomer and dimer from LS 174T cells are water-soluble, whereas larger MUC2 species formed early during biosynthesis are insoluble and contain nonreducible intermolecular bonds. *J Biol Chem*. 273:18864–18870.

Backhed F, Ley RE, Sonnenburg JL, Peterson DA, Gordon JI. 2005. Host-bacterial mutualism in the human intestine. *Science*. 307:1915–1920.

Capon C, Maes E, Michalski JC, Leffler H, Kim YS. 2001. 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.

Ceroni A, Maass K, Geyer H, Geyer R, Dell A, Haslam SM. 2008. GlycoWorkbench: A tool for the computer-assisted annotation of mass spectra of glycans. *J Proteome Res*. 7:1650–1659.

Fu J, Wei B, Wen T, Johansson ME, Liu X, Bradford E, Thomsson KA, McGee S, Mansour L, Tong M, et al. 2011. Loss of intestinal core 1-derived O-glycans causes spontaneous colitis in mice. *J Clin Invest*. 121:1657–1666.

Holmen JM, Olson FJ, Karlsson H, Hansson GC. 2002. Two glycosylation alterations of mouse intestinal mucins due to infection caused by the parasite *Nippostrongylus brasiliensis*. *Glycoconj J*. 19:67–75.

Hurd EA, Holmen JM, Hansson GC, Domino SE. 2005. Gastrointestinal mucins of Fut2-null mice lack terminal fucosylation without affecting colonization by *Candida albicans*. *Glycobiology*. 15:1002–1007.

Ismail MN, Stone EL, Panico M, Lee SH, Luu Y, Ramirez K, Ho SB, Fukuda M, Marth JD, Haslam SM, et al. 2011. High-sensitivity O-glycomic analysis of mice deficient in core 2 β1,6-*N*-acetylglucosaminyltransferases. *Glycobiology*. 21:82–98.

Iwai T, Inaba N, Naundorf A, Zhang Y, Gotoh M, Iwasaki H, Kudo T, Togayachi A, Ishizuka Y, Nakanishi H, et al. 2002. Molecular cloning and characterization of a novel UDP-GlcNAc:GalNAc-peptide beta1,3-N-acetylglucosaminyltransferase (beta 3Gn-T6), an enzyme synthesizing the core 3 structure of O-glycans. *J Biol Chem*. 277:12802–12809.

Iwai T, Kudo T, Kawamoto R, Kubota T, Togayachi A, Hiruma T, Okada T, Kawamoto T, Morozumi K, Narimatsu H. 2005. Core 3 synthase is down-regulated in colon carcinoma and profoundly suppresses the metastatic potential of carcinoma cells. *Proc Natl Acad Sci USA*. 102:4572–4577.

Johansson ME, Gustafsson JK, Sjoberg KE, Petersson J, Holm L, Sjovall H, Hansson GC. 2010. Bacteria penetrate the inner mucus layer before inflammation in the dextran sulfate colitis model. *PLoS One*. 5:e12238.

Johansson ME, Phillipson M, Petersson J, Velcich A, Holm L, Hansson GC. 2008. The inner of the two Muc2 mucin-dependent mucus layers in colon is devoid of bacteria. *Proc Natl Acad Sci USA*. 105:15064–15069.

Johansson ME, Thomsson KA, Hansson GC. 2009. Proteomic analyses of the two mucus layers of the colon barrier reveal that their main component, the Muc2 mucin, is strongly bound to the Fcgbp protein. *J Proteome Res*. 8:3549–3557.

Juge N. 2012. Microbial adhesins to gastrointestinal mucus. *Trends Microbiol*. 20:30–39.

Karlsson H, Larsson JM, Thomsson KA, Hard I, Backstrom M, Hansson GC. 2009. High-throughput and high-sensitivity nano-LC/MS and MS/MS for O-glycan profiling. *Methods Mol Biol*. 534:117–131.

Karlsson KA. 1989. Animal glycosphingolipids as membrane attachment sites for bacteria. *Annu Rev Biochem*. 58:309–350.

Karlsson NG, Schulz BL, Packer NH. 2004. Structural determination of neutral O-linked oligosaccharide alditols by negative ion LC-electrospray-MSn. *J Am Soc Mass Spectrom*. 15:659–672.

Larsson JM, Karlsson H, Crespo JG, Johansson ME, Eklund L, Sjovall H, Hansson GC. 2011. Altered O-glycosylation profile of MUC2 mucin occurs in active ulcerative colitis and is associated with increased inflammation. *Inflamm Bowel Dis*. 17:2299–2307.

Larsson JM, Karlsson H, Sjovall H, Hansson GC. 2009. A complex, but uniform O-glycosylation of the human MUC2 mucin from colonic biopsies analyzed by nanoLC/MSn. *Glycobiology*. 19:756–766.

Ley RE, Backhed F, Turnbaugh P, Lozupone CA, Knight RD, Gordon JI. 2005. Obesity alters gut microbial ecology. *Proc Natl Acad Sci USA*. 102:11070–11075.

Magalhaes A, Gomes J, Ismail MN, Haslam SM, Mendes N, Osorio H, David L, Le PJ, Haas R, Dell A, et al. 2009. Fut2-null mice display an altered glycosylation profile and impaired BabA-mediated *Helicobacter pylori* adhesion to gastric mucosa. *Glycobiology*. 19:1525–1536.

Marion D, Wuthrich K. 1983. Application of phase sensitive two-dimensional correlated spectroscopy (COSY) for measurements of ^{1}H-^{1}H spin-spin coupling constants in proteins. *Biochem Biophys Res Commun*. 113:967–974.

Olson FJ, Backstrom M, Karlsson H, Burchell J, Hansson GC. 2005. A MUC1 tandem repeat reporter protein produced in CHO-K1 cells has sialylated core 1 *O*-glycans and becomes more densely glycosylated if coexpressed with polypeptide-GalNAc-T4 transferase. *Glycobiology*. 15:177–191.

Olson FJ, Johansson ME, Klinga-Levan K, Bouhours D, Enerback L, Hansson GC, Karlsson NG. 2002. Blood group A glycosyltransferase occurring as alleles with high sequence difference is transiently induced during a *Nippostrongylus brasiliensis* parasite infection. *J Biol Chem*. 277:15044–15052.

Rawls JF, Mahowald MA, Ley RE, Gordon JI. 2006. Reciprocal gut microbiota transplants from zebrafish and mice to germ-free recipients reveal host habitat selection. *Cell*. 127:423–433.

Robbe C, Capon C, Coddeville B, Michalski JC. 2004a. 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.

Robbe C, Capon C, Coddeville B, Michalski JC. 2004b. Structural diversity and specific distribution of O-glycans in normal human mucins along the intestinal tract. *Biochem J*. 384:307–316.

Robbe C, Capon C, Maes E, Rousset M, Zweibaum A, Zanetta JP, Michalski JC. 2003. Evidence of regio-specific glycosylation in human intestinal mucins: Presence of an acidic gradient along the intestinal tract. *J Biol Chem*. 278:46337–46348.

Saito S, Egawa S, Endoh M, Ueno S, Ito A, Numahata K, Satoh M, Kuwao S, Baba S, Hakomori S, et al. 2005. RM2 antigen (beta1,4-GalNAc-disialyl-Lc4) as a new marker for prostate cancer. *Int J Cancer*. 115:105–113.

Schulz BL, Packer NH, Karlsson NG. 2002. Small-scale analysis of O-linked oligosaccharides from glycoproteins and mucins separated by gel electrophoresis. *Anal Chem*. 74:6088–6097.

Stone EL, Ismail MN, Lee SH, Luu Y, Ramirez K, Haslam SM, Ho SB, Dell A, Fukuda M, Marth JD. 2009. Glycosyltransferase function in core 2-type protein O glycosylation. *Mol Cell Biol*. 29:3770–3782.

Takashima S. 2008. Characterization of mouse sialyltransferase genes: Their evolution and diversity. *Biosci Biotechnol Biochem*. 72:1155–1167.

Thomsson KA, Backstrom M, Holmen Larsson JM, Hansson GC, Karlsson H. 2010. Enhanced detection of sialylated and sulfated glycans with negative ion mode nanoliquid chromatography/mass spectrometry at high pH. *Anal Chem*. 82:1470–1477.

Thomsson KA, Hinojosa-Kurtzberg M, Axelsson KA, Domino SE, Lowe JB, Gendler SJ, Hansson GC. 2002. Intestinal mucins from cystic fibrosis mice show increased fucosylation due to an induced Fucα1–2 glycosyltransferase. *Biochem J*. 367:609–616.

Thomsson KA, Karlsson H, Hansson GC. 2000. Sequencing of sulfated oligosaccharides from mucins by liquid chromatography and electrospray ionization tandem mass spectrometry. *Anal Chem*. 72:4543–4549.

Thomsson KA, Schulz BL, Packer NH, Karlsson NG. 2005. MUC5B glycosylation in human saliva reflects blood group and secretor status. *Glycobiology*. 15:791–804.

Tobisawa Y, Imai Y, Fukuda M, Kawashima H. 2010. Sulfation of colonic mucins by N-acetylglucosamine 6-O-sulfotransferase-2 and its protective function in experimental colitis in mice. *J Biol Chem*. 285:6750–6760.

van Halbeek H, Strang AM, Lhermitte M, Rahmoune H, Lamblin G, Roussel P. 1994. Structures of monosialyl oligosaccharides isolated from the respiratory mucins of a non-secretor (O, Lea+b-) patient suffering from chronic bronchitis. Characterization of a novel type of mucin carbohydrate core structure. *Glycobiology*. 4:203–219.