Construction of a natural mucin microarray and interrogation for biologically relevant glyco-epitopes

Michelle Kilcoyne,[†] Jared Q. Gerlach,[†] Ronan Gough,[§] Mary E. Gallagher,[§] Marian Kane,[†] Stephen D. Carrington,[§] and Lokesh Joshi[†,*]

[†]Glycoscience Group, National Centre for Biomedical Engineering Science, National University of Ireland, Galway, Ireland

[§]Veterinary Science Centre, University College Dublin, Belfield, Dublin 4, Ireland

ABSTRACT

Mucins are the principal components of mucus and mucin glycosylation has important roles in defence, microbial adhesion, immunomodulation, inflammation and cancer. Mucin expression and glycosylation are dynamic, responding to changes in local environment and disease. Potentially hundreds of heterogenous glycans can substitute one mucin molecule and it is difficult to identify biologically accessible glyco-epitopes. Thirty-seven mucins, from the reproductive and gastrointestinal (GI) tracts of six species (bovine, ovine, equine, porcine, chicken and deer) and from two human-derived cell lines, were purified. Following optimisation of mucin printing to construct a novel mucin microarray, the glycoprofiles of the whole mucins were compared using a panel of lectins and one antibody. Accessible glyco-motifs of GI mucins varied according to species and localisation of mucin origin, with terminal fucose, the sialyl T-antigen and *N*-linked oligosaccharides identified as potentially important. The occurrence of T- and sialyl T-antigen varied in bovine and ovine reproductive tract mucins, and terminal *N*-acetylgalactosamine (GalNAc) and sulfated carbohydrates were detected. This study introduces natural mucin microarrays as an effective tool for profiling mucin glyco-epitopes and highlights their potential for discovery of biologically important motifs in bacterial-host interactions and fertility.

Keywords:

Glycomics

Glycosylation

High throughput

Host-pathogen interaction

Lectin

Carbohydrate

* Corresponding author. Ph.: +353 91 495768

E-mail address: Lokesh.Joshi@nuigalway.ie

INTRODUCTION

Mucins are secreted and membrane-bound glycoproteins with roles in adhesion, defence, immunomodulation, inflammation and tumorigenesis.[1, 2] Secreted mucins are the major macromolecular components of the mucus layer found on most epithelial surfaces of higher vertebrates, including the gastrointestinal (GI), reproductive and respiratory tracts,[3, 4] serving to lubricate and protect these surfaces against mechanical damage and chemical and biological insult.[4, 5] Seven of the eighteen different mucin genes (MUC) give rise to secreted mucins, which are heavily *O*-glycosylated, with 50 to 90% by weight of the molecule composed of carbohydrate.[2, 5] The bulk carbohydrate content aids in the general viscoelasticity and permeability barrier functions of the mucus layer, while the oligosaccharide structures and substitutions more specifically contribute to, or determine, the physical or biological properties of the mucin, due to their hydrophobicity, charge, identity or configuration.[4]

As the main entry point for pathogenic microbes or their toxins into the body, host mucosal surfaces are very responsive to changes in their immediate environment.[6] The expression and glycosylation of their mucin components vary depending on the location within the body and are altered by changes in hormonal status, inflammation and the presence of microbes.[7, 8] The heterogeneity of the oligosaccharide components of mucins provide multivalent binding sites, which can serve as competitive inhibitors or decoys for the interaction of pathogens or toxins with the host cell.[4] Pathogens are then removed from the body with the continuous sloughing off and replacement of the mucus layer. Mucin glycosylation also promotes microbial colonization by providing adhesion sites or a nutrient source for commensal bacteria,[5, 9] which in turn contribute to the epithelial defence against pathogens.[10, 11] By secreting enzymes that modify mature mucin oligosaccharides, pathogens can exploit a similar mechanism to subvert the mucus defence system, thus consistently

renewing adhesion targets or generating an energy source to support their growth.[12] To further our understanding of the role of mucin glycosylation in host-microbe interactions at mucosal surfaces, high throughput methods with minimal sample requirements are needed for profiling or comparing the accessible sugars or glyco-epitopes presented on mucins collected under different conditions and also for investigation of microbial interactions with these mucins.

Structural analysis of mucin oligosaccharides is normally carried out by mass spectrometric techniques.[13-17] Eight core structures have been described for mucin-type oligosaccharides (Table S-1), which are *O*-glycosidically linked to serine or threonine in the protein backbone *via N*-acetylgalactosamine (GalNAc). The oligosaccharides are further extended with GalNAc, *N*-acetylglucosamine (GlcNAc), galactose (Gal), fucose (Fuc), and sialic acids (Neu) and/or sulfate groups.[1, 5, 18] The combination of different core structures, chain elongation, branching, linkages and the variety of peripheral and terminal residues results in a high degree of mucin oligosaccharide heterogeneity,[1] with dozens to hundreds of diverse oligosaccharides densely substituting the mucin protein backbone[19] in a 'bottle-brush' configuration[20] (Figure 1A). While it is important to catalogue the different oligosaccharides which comprise the carbohydrate component of mucins, the density, distribution pattern and three-dimensional architecture of the oligosaccharides on mucins are also relevant for the formation of the glyco-epitopes that are presented to the microenvironment, and therefore impact on their biological interactions and recognition.[21-23]

There have been some reports of the integration of whole mucin molecules into high throughput platforms. Flow cytometry was used to investigate lectin interactions with porcine gastric mucin (PGM),[2] and a tissue microarray was constructed from gastric biopsy and gallbladder samples to enable histochemical examination of mucins *in situ*.[24] Synthetic glycopolymers intended to mimic the multivalency and spatial positioning of saccharides on

natural mucins have been synthesised and printed in a microarray format.[21, 25] This approach is useful for probing ligand interactions based on known structures, orientations and densities,[21] or to test new synthetic target structures. However, for discovery and interactions of accessible glyco-epitopes, only by using whole mucin molecules will the complexities of the glyco-epitopes they present and their biological interactions be understood and the relevant glyco-epitopes identified.[2, 26] Purification of mucins is a lengthy procedure and purified mucins are typically only available in very limited quantities. Hence, a microarray format will maximise the number of possible binding or analytical experiments with the scarce quantities of mucins obtainable. Here, we describe the construction and optimisation of a natural mucin microarray containing 37 mucin samples isolated from a range of animal and cell line sources and its use to profile and compare the accessible glyco-epitopes presented by each mucin.

MATERIALS AND METHODS

Materials. Nexterion® slide H microarray slides were purchased from Schott AG (Mainz, Germany). Tetramethylrhodamine isothiocyanate (TRITC)-labelled lectins were from EY Laboratories, Inc. (San Mateo, CA, USA). MECA-79 (rat IgM) antibody was purchased from Santa Cruz Biotechnology, Inc. (Heidelberg, Germany) and the TRITC-labelled polyclonal goat anti-rat IgM secondary antibody was from AbD Serotec (Oxford, UK). Glycoproteins and all other chemicals were obtained from Sigma-Aldrich Co. (Dublin, Ireland), unless otherwise noted, and were of the highest grade available.

Mucin purification. Mucus samples from most animals were obtained either post-mortem from healthy animals killed at commercial abattoirs, or after humane euthanasia by intravenous barbiturate overdose. Generally, mucosal surfaces were scraped with a scalpel blade to harvest secreted mucus and epithelial cells. Some of the bovine cervical samples

were obtained through direct aspiration from the cranial vaginas of live animals in the peri-estrus period. These experimental procedures were licensed by the Department of Health and Children, Ireland, in accordance with the Cruelty to Animals Act (Ireland 1897) and the European Community Directive 86/609/EC, and were sanctioned by the Animals Research Ethics Committee, University College Dublin, Ireland.

Mucins were isolated from the collected mucus and purified as previously described.[27, 28] In brief, mucus was solubilised with guanidine hydrochloride (4 M final concentration) and incubated on a roller for a minimum of 3 days. In some cases, the sample was briefly pre-treated with 6 M guanidine hydrochloride in phosphate buffered saline (PBS) to facilitate mucin extraction. The samples were then reduced with dithiothreitol (10 mM final concentration) and alkylated with iodoacetamide (25 mM). CsCl was added to a starting density of 1.42 g/mL, and isopycnic density gradient centrifugation was performed for 18 h at 65,000 rpm and 10 °C using a Beckman-Coulter Optima LE-80K ultracentrifuge with a 70 Ti rotor. The mucin-rich fractions were detected on slot blots using periodic acid – Schiff's (PAS) reagent and pooled. These are routinely found between densities of 1.35 and 1.45 g/mL (see Figure S-1 for an example of a typical density gradient profile). The high molecular weight components of the samples were collected by Sepharose Cl-4B chromatography and pooled. The semi-purified material was desalted by Sephadex G-25 chromatography, lyophilised and weighed. Mucin preparations were stored dry at 4 °C until used.

Mucins from the LS174T cell line were obtained from the culture media, following concentration by freeze-drying, and from the cells following extraction with 6 M guanidine hydrochloride in PBS. Mucins were harvested from HT29-MTX-E12 (E12) cells as previously described.[29] Further details of the 37 mucin samples used in this study are available in Table S-2.

Mucin microarray printing. Mucins and glycoproteins (probes) were dissolved in PBS, pH 7.4 (1.37 M NaCl, 0.027 M KCl, 0.02 M KH_2PO_4 and appropriate mixture of 0.1 M Na_2HPO_4 and NaH_2PO_4 (monobasic and dibasic) for correct pH) (see Table 1 for final print concentration and buffer composition) and printed onto Nexterion® slide H microarray slides using a SciFLEXARRAYER S3 piezoelectric printer (Scienion AG, Germany) equipped with a 90 μm uncoated glass nozzle at 62% humidity (+/-2% tolerance). Probes were printed in replicates of six, approximately 1 nL per feature, 312 features per subarray and 8 subarrays per slide. Slides were incubated in a humidity chamber overnight after printing to facilitate conjugation and the remaining functional groups were capped with 100 mM ethanolamine in 50 mM sodium borate, pH 8.0, for 1 h. The slides were then washed three times in PBS with 0.05% Tween-20 (PBS-T), once in PBS, centrifuged dry (1,500 rpm, 5 min) and stored dry at 4 °C with desiccant until use.

Microarray incubations. Microarray slides were incubated using an eight-well gasket slide and incubation cassette system (Agilent Technologies, CA, USA) and were protected from light throughout the procedure. Briefly, 70 μL of TRITC-labelled lectin (see Table 2 for appropriate dilution of each lectin) in Tris-buffered saline (TBS; 20 mM Tris-HCl, 100 mM NaCl, 1 mM $CaCl_2$, 1 mM $MgCl_2$, pH 7.2) with 0.05% Tween-20 (TBS-T) was applied to each well of the gasket. For inhibition experiments, lectins were diluted in 100 mM of appropriate haptenic sugar or glycoprotein in TBS-T (Table 2) and resulting binding intensity compared to the same lectin incubated without inhibition on a subarray on the same slide. The microarray slide was sandwiched with the gasket, the cassette assembled and placed in a rotating incubation oven (23 °C, approximately 4 rpm) for 1 h. Slides were disassembled under TBS-T, washed 3 times in TBS-T for 2 min each with gentle agitation in a Coplin jar, with a final 3 min wash in TBS. The microarrays were dried by centrifugation and scanned immediately.

MECA-79 (rat IgM) antibody was incubated at a 1 in 100 dilution in TBS-T on the microarray slide for 1 h, washed and dried as above and then immediately incubated with TRITC-labelled polyclonal goat anti-rat IgM at 1 in 500 dilution in TBS-T. Microarray slides were washed and dried as above and scanned immediately.

Imaging, data extraction and analysis. Microarray slides were imaged in a Perkin-Elmer Scanarray Express HT (543 nm laser, 90% laser power, 70% PMT, TRITC emission filter, 5 μm resolution). Intensity values were extracted from the image files using GenePix Pro v6.1.0.4 and a proprietary *.gal file using adaptive diameter (70-130%) circular alignment based on 230 μm features and exported as text to Excel (version 2007, Microsoft) where all data calculations were performed. Local background was subtracted and background-corrected median feature intensity (F543median-B543) was used for each feature intensity value. The median of six replicate spots per subarray was handled as a single data point for graphical and statistical analysis. Data were normalised to the mean of three replicate microarray slides (subarray by subarray using subarray total intensity) and binding data was presented in histogram form of average intensity with standard deviation of three experimental replicates (n = 3, 18 data points). In certain cases, only one microarray slide was used (6 data points) and thus no error bars were depicted. The significance of inhibition data was evaluated using a standard Student's t-test (paired, two-tailed).

Monosaccharide analysis. The monosaccharide content was determined by hydrolysis of 50 μg of mucin with 2 N trifluoroacetic acid (TFA) for 4 h at 100 °C.[30, 31] Fetuin and water blanks were hydrolyzed in parallel to ensure consistency of conditions for positive control and for background correction, respectively. Hydrolysate was dried from water three times in a centrifugal evaporator and stored dry until analysed. The released sugars were taken up in 50 μL 18.2 MΩ water and identified and quantified by comparison to appropriate monosaccharide standards by high performance anion exchange chromatography

with pulsed amperometric detection (HPAEC-PAD) on an ICS3000 system (Dionex, Sunnyvale, CA). The equivalent of 0.3-2.0 μg hydrolysate (depending on the sample) was injected onto a CarboPac PA20 analytical column (3 x 150 mm) equipped with an amino trap column (3 x 30 mm) with isocratic elution in 18 mM NaOH at 0.35 mL/min over 18.5 min. The column was washed with 100 mM NaOH for 8 min and reequilibrated at 18 mM for 12 min. Each sample was injected in triplicate, the monosaccharide content was quantified by comparison to the standard curve generated and the average values are reported.

RESULTS AND DISCUSSION

Mucin purification. The buoyant densities of all mucin preparations were between 1.35 and 1.45 g/mL and all eluted from the size exclusion column immediately after the void volume, which are typical characteristics of mucins. The yield of purified mucins was generally around 1 mg/mL of mucus samples, although up to 30-fold higher yields were obtained from mucosal scrapings and direct mucosal extracts using 6M guanidine hydrochloride. The yields of mucin from LS174T and E12 cells were approximately 1 mg per 3.5×10^7 cells and 1 mg per 5.0×10^7 cells, respectively.

Optimisation of mucin printing on the microarray. Amino groups of the mucin protein backbone were targeted for conjugation to the functionalised slide surface. In general, the termini of mucins are less densely substituted by *O*-linked glycosylation (Figure 1A) and hence the protein backbone in this region is more accessible for immobilisation of the molecule. Also, this strategy for mucin conjugation would not interfere with the three-dimensional presentation of the bulk oligosaccharides of the mucin molecules (Figure 1B). Nexterion® slide H was selected as the platform for the mucin microarray as it has a hydrogel surface functionalised with amine-reactive *N*-hydroxylsuccinimide (NHS) esters, which provided a facile and robust conjugation chemistry at a physiological pH compatible

with maintaining the natural mucin conformation. In addition, the hydrogel surface has a low intrinsic background without the need for a separate blocking step.

Printing of each mucin was initially attempted at a range of concentrations (0.1 – 1 mg/mL) in PBS, and the formation of mucin features was assessed based on lectin binding. Lectins are carbohydrate-binding proteins capable of binding specifically to a range of carbohydrate structures.[13, 32] Most mucin solutions were too viscous to print at concentrations above 0.5 mg/mL. However, even at 0.5 mg/mL and below, most mucins aggregated on the slide surface and did not form extractable features when incubated with the lectins (Figure 1C, top). Inclusion of detergent in the mucin print buffer led to improved feature formation, and the detergent and mucin concentrations were individually optimised for each mucin printed (Figure 1C, bottom and Table 1). The resulting average feature size was 220 μm. The concentration of each lectin was selected so that the response of most mucin features was within the dynamic range of the microarray scanner (0 - 65,000 RFU approximately, see Table 2 for concentrations used). Standard glycoproteins were also included on the microarrays to monitor the performance of the carbohydrate recognition elements used to profile the mucin microarray (Figure S-2).

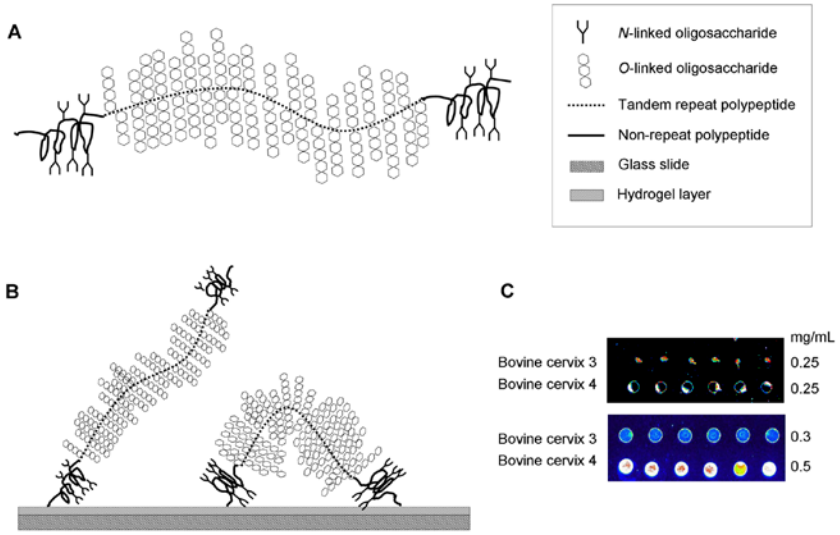

Figure 1. (A) Schematic of a mucin molecule (based in part on schematics from [5, 19]). (B) Schematic of mucin molecule immobilisation on a microarray slide surface representing the three dimensional presentation of the mucin molecule. (C) Bovine cervical mucins (samples M3 and M4) printed at concentrations indicated in PBS, pH 7.4 (top) and PBS, pH 7.4 with 0.025% Tween 20 included (bottom).

Interrogation of mucin microarray with selected carbohydrate recognition molecules. To demonstrate the utility of the mucin microarray in mucin-binding studies, the arrays were probed with a panel of 11 lectins and one antibody (Table 2), with specificities that covered the range of carbohydrate motifs known to be present on mucins.[1] Most binding experiments were carried out in triplicate and the average slide-to-slide percentage coefficients of variance (%CV) for the binding of lectins to mucins ranged from 13.4% (Jacalin) to 45.7% (SNA-I), with six out of the eleven lectins used for profiling having average %CVs less than 30% (UEA-I, MAA, WFA, VVA, PHA-E and GS-II) (Table S-3). These %CV values are in the lower range reported for typical protein-based microarrays.[33, 34] The lectin-binding profile of each mucin sample was unique, which was consistent with the reported diversity of glycosylation of mucins isolated from different sites. The resulting profiles are compared below.

Lectin binding in the presence of the appropriate haptenic carbohydrate or glycoprotein (Table 2) was determined in parallel in each case to verify that the observed lectin-mucin binding interactions were carbohydrate-mediated.[32] In all cases, binding intensity was decreased across a population of mucin features using the appropriate haptenic agent (Table S-4). Some interactions, e.g. WFA binding to bovine abomasum mucin (sample M57) in the presence of GalNAc, were almost completely abolished (98%, $p = 0.0153$). Even where lectin-mucin interactions did not produce uninhibited relative fluorescence intensity

values exceeding an accepted signal threshold for microarray studies of five times background,[35] most were still inhibited using the appropriate haptenic carbohydrate moiety, e.g. binding of Con A to deer large intestine mucin (sample M56) was 99% inhibited (p = 0.0014) in the presence of Man.

Comparison of lectin-binding profiles for GI tract mucin samples. In total, 23 mucin samples from the digestive systems of the six species, chicken, equine, bovine, porcine, ovine and deer, were included on the microarray. Comparison of the glycoprofiles of these mucins shows variations progressing through the GI tract and also different patterns of expression in each species (Figures 2 and 3). These diverse profiles reflect differences in the composition and/or presentation of the mucin glyco-epitopes along the tract, as previously reported,[5, 26, 36] and/or differences in their accessibility to the relevant lectin.

Chicken GI tract mucins. Mucin samples from three regions of the chicken GI tract were included on the microarray. These showed increasing intensity of binding of UEA-I, MAA, Jacalin, PHA-E and SBA as the mucins progressed towards the terminus of the chicken digestive system (Figure 2), indicating increased presence or accessibility of α-(1,2)-linked fucose (UEA-I), α-(2,3)-linked sialylation (MAA), the T-antigen (Gal-β-(1→3)-GalNAc-α-*O*-S/T) (Jacalin), complex biantennary *N*-linked structures with outer Gal and bisecting GlcNAc (PHA-E) and terminal GalNAc epitopes (SBA), respectively (Table 3,). Indeed, as the lectins Jacalin and PNA both bind to the T-antigen but only Jacalin tolerates the presence of sialic acid, the higher intensity from Jacalin and MAA together with a lower intensity of binding from PNA suggested that the majority of the accessible T-antigen motifs were capped with α-(2,3)-linked sialic acid (sialyl T-antigen). WFA also bound strongly to the mucins of the large intestine which suggested greater accessibility of GalNAc or the presence of sulfated GalNAc epitopes.[37] A higher level of sulfation in mucins from the large

intestine could also contribute to the MAA binding intensity, as this lectin has been reported to bind to Gal-3-SO_4 as well as to α-(2,3)-linked sialic acid.[38]

Interrogation of the chicken GI tract mucins with the monoclonal antibody, MECA-79, one of very few available antibodies specific for a sulfated carbohydrate motif typical of mucins, showed no binding to any of the chicken mucin samples. This antibody recognises GlcNAc-6-SO_4, part of the 6-sulfo-sialyl Lewis x (6-sulfo-SLe^x), and 6-sulfolactose, [40] and the importance of presentation and spatial orientation of sulfate esters on carbohydrates for recognition is known. [40] However, *O*-linked oligosaccharides can be sulfated at sites other than the C-6 of GlcNAc, e.g. the C-3 of Gal.[1] Taken together with the observed binding of WFA and MAA, there still is a possibility that other sulfated motifs are present in mucins from the large intestine.

Chicken mucin preparations from the three GI tract locations studied here have been shown to attenuate the binding and internalisation of a highly invasive strain of the pathogenic microorganism, *Campylobacter jejuni*, into the human ileo-caecal adenocarcinoma cell line, HCT-8, with mucin from the large intestine showing the greatest effect (1500-fold reduction) followed by mucins from the small intestine (150-fold reduction) and the caecum (5-fold reduction).[27] This effect was carbohydrate-mediated and it was inferred that specific glycan epitopes on the mucins from the large intestine were responsible for the biological activity observed. The differences in lectin-binding profiles of chicken mucins from the three locations of the GI tract noted in this study support this hypothesis. When the relative binding intensity of each lectin to the three mucin samples were compared to their reported activity as inhibitors of *C. jejuni* internalization into HCT-8 cells, the WFA binding pattern suggested that terminal GalNAc and/or sulfated GalNAc motifs may be important (Table 3). These results demonstrate the potential of the mucin microarray platform to provide insight into the possible glycan epitopes that are involved in important biological

effects, to identify target motifs for further study and, in this case, to contribute to further understanding of the pathogenic mechanism of *C. jejuni* in humans and why this organism is not pathogenic in chickens.[41]

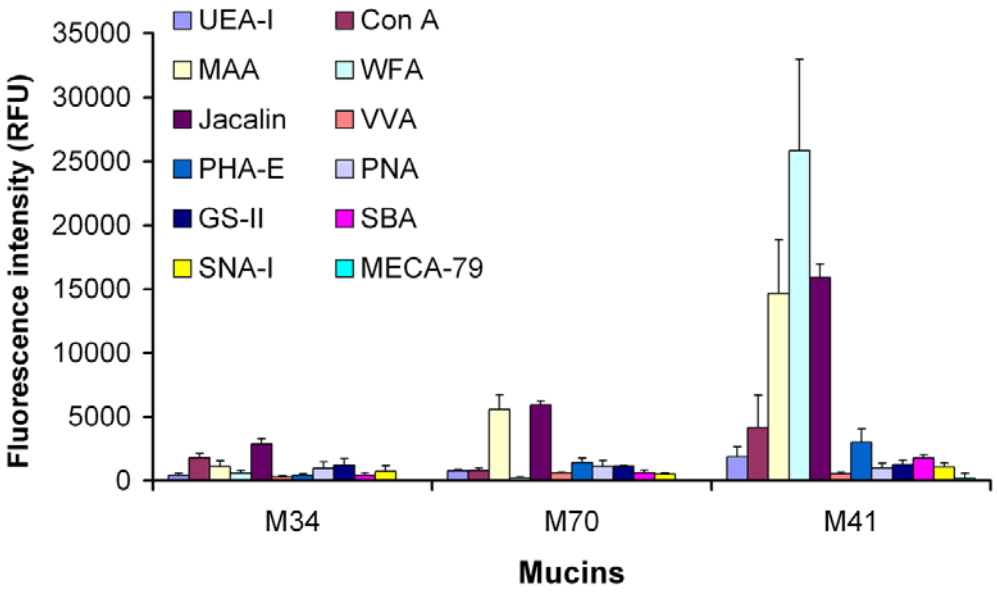

Figure 2. Histogram representing the mean fluorescence intensity from three replicate microarray slides of individual lectins binding to printed mucins from chicken digestive system, values for each microarray slide consisting of the median of six feature replicates. Error bars are the standard deviation of the mean of three microarray slides.

GI tract mucins in ruminants. Figure 3 depicts the glycoprofiles of the mucins collected from different regions of the GI tract of the four ruminant species, horse, cattle, sheep and deer. Comparing the mucins found in the true stomach (or abomasum), the α-(1→2)-linked Fuc motif as revealed by UEA-I was present in all four species, and UEA-I binding generally decreased progressing down the digestive system, except in the bovine mucins, where highest expression was in the spiral ascending colon mucin (sample M61). Mucins from the stomach of all species except the horse also showed strong binding to WFA and Jacalin, but not PNA. The latter inferred that the sialyl T-antigen was accessible on these

mucins, which was supported by the accompanying higher intensity of MAA. Of the abomasum mucins, the bovine sample (sample M57) alone had high binding of GS-II, which indicated the presence of terminal GlcNAc residues (Table 2).

With the exception of deer, there was a general trend towards increased binding of the mucins to a range of lectins as sampling point progressed down the intestine, with evidence of increased sialylation (MAA, SNA-I) and sialyl T-antigen (MAA and Jacalin, with accompanying low PNA-binding). The equine right ventral ascending colon mucin (sample M65), in particular, exhibited near saturated intensity for a range of motifs, including α-(2,3)- and α-(2,6)-linked sialic acids (MAA and SNA-I, respectively), T-antigen (Jacalin and PNA), α-(1,2)-linked fucosylation (UEA-I), high mannose *N*-linked structures (Con A) and terminal GalNAc epitopes (WFA and SBA), that may also be sulfated (WFA).[37] Also, some complex type *N*-linked oligosaccharides were accessible in a number of samples (PHA-E).

While there have been a number of lectin histochemical studies of regions of the GI tract in the ruminant species presented here,[42-44] there have been few reports characterising the glycosylation of isolated mucus or mucins. Mucus from bovine abomasum was reported to bind strongly to the GalNAc-binding lectins, RCA120, DBA, and SBA,[44] which correlated with the binding observed in this study between WFA and SBA and bovine abomasum mucin (sample M57, Figure 3). Strong binding of the mucus to WGA was also reported.[44] As WGA binds to GlcNAc and sialic acid residues, this agreed with the GS-II and MAA binding observed here, which suggested that both residues were present in bovine abomasum mucin. Weak binding of the mucus to UEA-I and PNA was also reported,[44] as was observed with the mucin microarray.

In agreement with previously reported flow cytometry data,[2] Fuc-specific UEA-I bound to PGM (sample M37, Figure S-3 and Table S-4). WGA binding to PGM was also reported, which indicated the presence of accessible sialic acid and/or GlcNAc residues. High

intensity binding of GS-II and low binding of MAA and SNA-I on our platform, however, suggested that GlcNAc residues were the dominant glyco-epitopes expressed, rather than sialic acid residues. Thus, this comparison of GI tract mucin glyco-profiles using the mucin microarray approach has confirmed literature observations and has provided further insight into the glyco-motifs present, which can be targeted for further investigation.

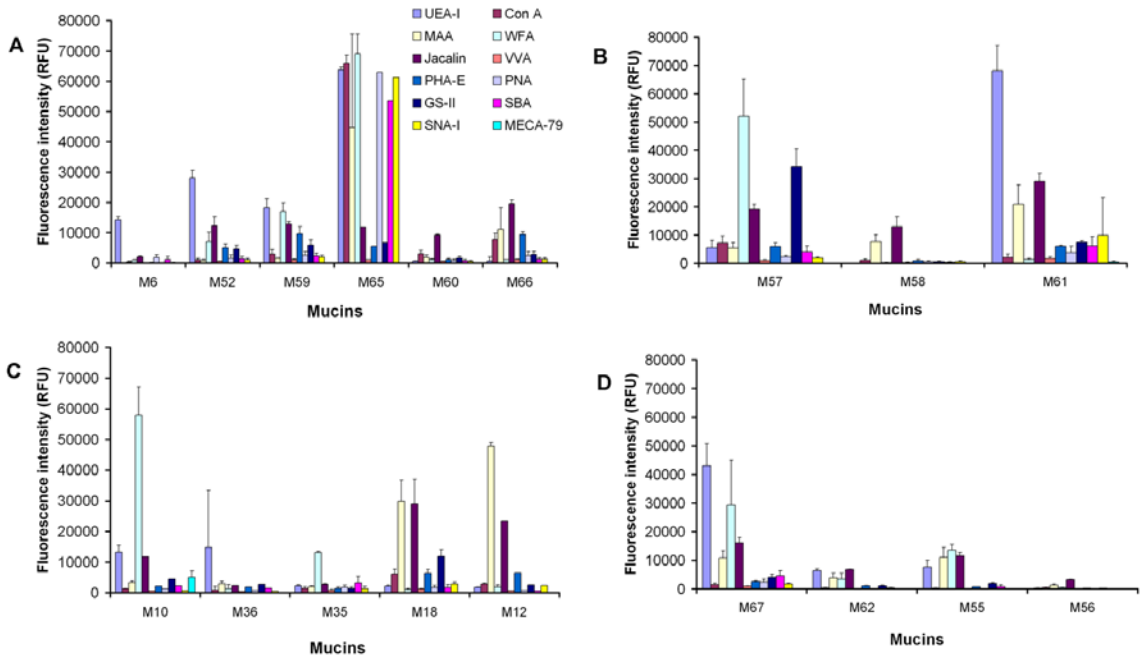

Figure 3. Histograms representing the mean fluorescence intensity from three replicate microarray slides of individual lectins binding to printed mucins from (A) equine, (B) bovine, (C) ovine and (D) deer GI tract. Mucins from left to right are presented as progressing through the GI tract. Error bars depict standard deviation of the mean of three replicate microarray slides. Histograms with no error bars represent the median of 6 data points on one slide.

Mucins from the reproductive tract. The glycoprofiles of mucins from the bovine and ovine reproductive tract showed a clear species difference between the cervical/cervico-vaginal samples (Figure 4). While WFA bound to both species, the highest intensity was

observed in bovine mucin samples, which emphasised the importance of the GalNAc epitope and possibly indicated a role for sulfated motifs in cervical mucins from the cow. By contrast, the α-(1,2)-linked fucose motif (UEA-I) was well-recognised in the cervical mucins of sheep, but the intensity of binding varied between bovine samples. There were also notable differences related to stages of the estrous cycle. In the cow, the accessibility of sialyl T-antigen during pro-estrus (sample M48) and metestrus (samples M49 and M4) was inferred from intense Jacalin and MAA binding coupled with lower intensity binding from PNA, while there seemed to be greater availability of the T-antigen (higher PNA and Jacalin intensity with lower MAA intensity) during estrus (samples M15 and M3, Figure 4A).

In contrast, the sialyl T-antigen was detected in the ovine cervical mucin samples (samples M30 to M33, Figure 4B). The ovine samples were pooled samples from two breeds of sheep, the highly fertile Belclare breed (samples M31 and M33) and less fertile Suffolk breed[46] (samples M30 and M32); samples M30 and 31 were collected 42 h post-estrus induction and samples M32 and M33, 56 h post-induction. The binding profiles for these mucins were generally similar, although the binding intensity of UEA-I and WFA was higher at the 42 h timepoint for both breeds compared to the later timepoint. The only difference that was notable between the binding profiles of the two breeds was higher binding of the MECA-79 antibody to the mucins from the Suffolk breed at the earlier time period compared to the Belclare breed. This suggests that glycan sulfation may be involved in the different fertility rates of these breeds, as the earlier time-point is the normal time for insemination. The binding of the MECA-79 antibody to bovine cervical mucins was much lower than ovine mucins in general.

Some of our observations on the glycosylation of bovine and ovine mucins have parallels in the human reproductive tract. MUC1 glycosylation in the human endometrium was reported to be altered during the menstrual cycle,[47] with expression of the T-antigen

increased during the early secretory phase when implantation occurs.[48] Also, the selectin ligands, SLe^x and SLe^a, which are known to carry the MECA-79 antigenic motif, have been reported on human uterine epithelial cells during the receptive phase where they seem to be important motifs for attachment and implantation.[47, 48] Indeed, the expression of the MECA-79 ligand has been suggested as a marker for human implantation.[49] The mucin microarray platform will facilitate more detailed monitoring of the alteration of these and other glycosylation changes during the human menstrual cycle.

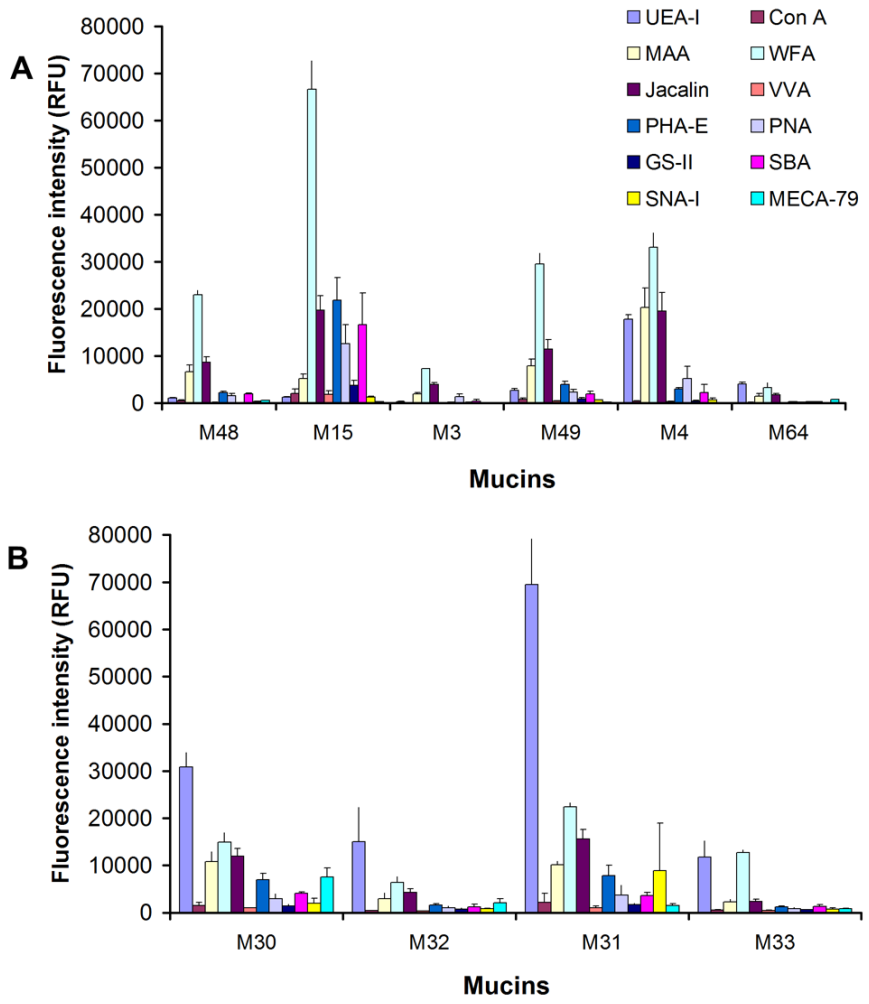

Figure 4. Histograms representing the fluorescence intensity of individual lectins bound to reproductive tract mucins from (A) bovine, in order of pro-estrus, estrus and metestrus, and

(B) ovine sources, grouped according to breed (M30 and 32, Suffolk, M31 and 33, Belclare). Error bars depict standard deviation of the mean of three microarray slides.

Comparison of mucins from human colon carcinoma-derived cell lines, LS174T and E12. The mucin microarray described here has also given a first insight into the glycosylation of mucins produced by gut epithelial cell lines. The production of mucin is an important attribute of cell lines used as *in vitro* models of the gut epithelium, because of the role played by mucins in gut physiology. Two such cell lines are LS174T, a human colon carcinoma-derived cell line that produces and secretes the mucin, MUC2,[50] and the HT29-MTX-E12 (E12) cell line, a derivative of another human colon carcinoma cell line, HT29.[51] E12 cells are goblet-like cells that secrete the mucin, MUC5AC, and uniquely develop an adherent mucus layer when grown on transwells that is similar in thickness to that found in the human small intestine (approximately 150 μm).[52]

Jacalin bound intensely to the mucins from both cell lines, which indicated the presence of the T-antigen. The accompanying high intensity of MAA and low intensity of PNA suggested that the T-antigen was mainly sialylated (Figure 5) and the presence of this motif was not surprising considering the carcinoma origin of the cells.[3, 53] The higher intensity of MAA binding to E12 mucin (sample M11) suggested the presence of Gal-3-SO_4.[38] Differential sulfation patterns for the mucins from both cell lines were also suggested by MECA-79 binding with mucin from E12 cells but not with LS174T mucin. Sulfo-mucins have been suggested to have a role in cancer progression and metastasis.[3, 24] Greater intensity for PHA-E binding to mucins from LS174T compared to E12 demonstrated that more complex *N*-linked oligosaccharides were accessible on the LS174T mucin. Thus, differences were noted in the glycosylation of mucins from mucin-producing cell lines using the mucin

microarray, which can inform the interpretation of experiments exploring bacterial adhesion and pathogenesis in the gut.[29]

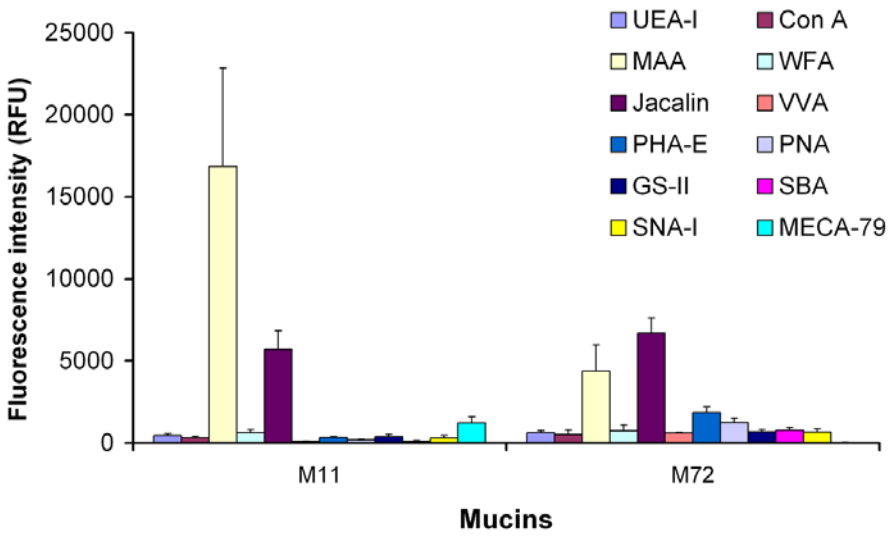

Figure 5. Histograms representing the fluorescence intensity of mucins from E12 (sample M11) and LS174T (sample M72) cell lines. Error bars depict standard deviation of the mean of three microarray slides.

Monosaccharide analysis. A subset of the mucin samples were subjected to acid hydrolysis to establish the monosaccharide composition and validate the presence of the carbohydrate components of the motifs inferred by lectin binding (Table 4). All but three (M62, M55 and M70) of the 19 samples analysed had low glucose (Glc) content, which indicated there was very little glycogen present. Glycogen is a common contaminant of mucin preparations[54,55] and also gives a positive reaction with PAS stain.[56] However, as glycogen does not have any amino groups, this molecule when present was not conjugated to the microarray surface and therefore did not contribute to glycoprofiling results. The differences in overall intensity of lectin binding to various samples generally paralleled differences in carbohydrate content (e.g. bovine samples from the duodenum and spiral

colon, samples M68 and M71, respectively), but there was no clear correlation overall. The mucins from the equine right ventral colon (sample M65), which demonstrated strong binding to several lectins, had relatively low total carbohydrate content. This lends support to the hypothesis that the mucin microarray platform can reveal more about the interactions of mucins than can be ascertained from their carbohydrate content or monosaccharide composition. The extent of binding is instead determined by spatial arrangement of the glyco-epitopes, substitution or clustering effects.[22, 23]

The monosaccharide composition also provided support for conclusions drawn from comparison of the glycan profiles for different samples. For example, the carbohydrate content of the mucin from chicken large intestine (sample M41) was greater than that from the ceca (sample M70) (Table 4), which correlated with the general lower intensities of lectin responses for the cecal mucin, and the relative ratio values demonstrated that GalN was a more relatively abundant residue in the large intestine mucin compared to the cecal mucin (Table 4 and Figure 2), further reinforcing the observation of terminal GalNAc as a potentially important biological glyco-epitope. Only a trace of fucose was detected in bovine duodenum mucin (sample M58) which was supported by the absence of detectable UEA-I binding, and the higher relative proportion of Gal in the bovine spiral colon mucin (sample M61) also correlated to the observation of a higher intensity of binding to a T-antigen motif compared to sample M58 (Table 4 and Figure 3B).

Although the monosaccharide analysis validated the presence of the carbohydrate components, only the profiling of intact, accessible glyco-epitopes as presented on the whole mucin molecule, could shed light on potentially biologically important motifs.

CONCLUSION

We have successfully demonstrated the feasibility of constructing a natural mucin microarray, incorporating mucin preparations from a range of sources, including the GI and reproductive tracts from six species and from two cell lines. The advantage of this array is the retention of the three-dimensional presentation of the mucin oligosaccharides intact on their protein backbone. We have shown that this format can easily be used to provide clear and reproducible lectin and antibody binding differentials between mucin samples, suggesting differences in presented glyco-epitopes which can provide new biological insights in a much shorter time-frame and using smaller amounts of material than is possible with conventional mucin binding experiments and glycosylation analysis. With current understanding of the importance of mucins in host-microbial interactions at mucosal surfaces and the knowledge that mucin secretion and glycosylation may be influenced by 'cross-talk' between host mucosal cells and the adjacent microflora,[26, 36, 57] the natural mucin microarray platform demonstrated here has the potential to be an effective glycoprofiling and discovery tool, which can be tailored to suit the biological question under investigation and used for revealing key mucin glyco-epitopes, for interrogating biologically relevant interactions between mucins and biomolecules or microbes, and also for screening potential modulators of these interactions.

ASSOCIATED CONTENT

Supporting information. Additional information as noted in text.

ACKNOWLEDGEMENTS

RG, MEG and SDC are grateful to Dr. M. Clyne and Dr. C. Reid for the E12 and LS174T cells and Dr Anthony Corfield for advice on mucin purification. This material is

based upon work supported by the Science Foundation Ireland under grant no. 08/SRC/B1393, in support of the Alimentary Glycoscience Research Cluster (AGRC).

Table 1. Probe code, printed mucins and glycoproteins, print concentrations and detergent concentration (Tween 20). Mucins are colored according to species e.g. bovine, green; equine, blue, etc.

Code	Source	Print conc (mg/mL)	Detergent conc (%)
M3	Bovine cervix, animal 0049	0.3	0.025
M4	Bovine cervix, animal 0258	0.5	0.025
M6	Equine stomach	0.25	0.01
M10	Ovine abomasum antrum	0.25	0.01
M11	E12	0.5	0.025
M12	Ovine descending colon	0.25	0.01
M15	Bovine cervico-vaginal mucin, animal 1467	0.25	0.01
M18	Ovine spiral ascending colon	0.5	-
M30	Ovine cervix (Suffolk)	0.5	0.025
M31	Ovine cervix (Belclare)	0.5	0.025
M32	Ovine cervix (Suffolk)	0.5	0.025
M33	Ovine cervix (Belclare)	0.5	0.025
M34	Chicken proximal small intestine	0.25	-
M35	Ovine jejunum	0.5	0.01
M36	Ovine duodenum	0.25	0.01
M37	Porcine gastric mucin	0.33	0.01
M39	Equine (pregnant) cervix	0.4	0.01
M41	Chicken large intestine	0.25	0.01
M48	Bovine cervico-vaginal mucin, animal 0113	0.5	0.025
M49	Bovine cervico-vaginal mucin, animal 0278	0.25	0.01
M52	Equine duodenum	0.3	0.01
M53	Equine trachea	0.5	0.01
M55	Deer jejunum	0.25	0.025
M56	Deer spiral ascending colon	0.75	0.025
M57	Bovine abomasum	0.25	0.01
M58	Bovine duodenum	0.5	0.01
M59	Equine jejunum	0.25	-
M60	Equine left ventral colon	0.25	0.01
M61	Bovine spiral ascending colon	0.25	0.01
M62	Deer duodenum	0.5	0.025
M63	Bovine trachea	0.75	0.025
M64	Bovine endometrium (uterine horn)	0.4	0.025
M65	Equine right ventral ascending colon	0.25	0.01
M66	Equine dorsal ascending colon	0.25	0.01
M67	Deer abomasum	0.25	0.01
M70	Chicken ceca	0.5	0.01
M72	LS174T	0.5	0.01
ASF	Asialofetuin	0.25	-

RB	RNase B	0.25	-
Fetuin	Fetuin	0.25	-
Xferrin	Transferrin	0.25	-
Ovomuc	Ovomucoid	0.25	-
PBST	PBS 0.01% Tween 20		0.01

Table 2. Lectins, their specificities, concentrations used and inhibitory carbohydrates. n.a., not applicable.

Carbohydrate recognition molecule	Abbreviation	Binding specificity	Conc (μg/mL)	Inhibitory carbohydrates
Maackia amurensis agglutinin	MAA	Neu-α-(2→3)-Gal = Gal-3-SO_4 >Lac	10	100 mM lactose
Sambucus nigra agglutinin I	SNA-I	Neu-α-(2→6)-Gal(NAc) >Lac, GalNAc >Gal	15	100 mM lactose
Ulex europaeus agglutinin I	UEA-I	Fuc-α-(1→2), H type 2 antigen	10	100 mM Fuc
Concanavalin A (jack bean lectin)	Con A	Man>Glc>GlcNAc	10	100 mM Man
Wisteria floribunda agglutinin	WFA	GalNAc (GalNAc-α-(1→6)-Gal>GalNAc-α-(1→3)-GalNAc (Forsmann antigen) >GalNAc>>Lac>Gal, chondroitin sulfate	10	100 mM GalNAc
Soybean agglutinin	SBA	GalNAc>Gal	15	100 mM GalNAc
Vicia villosa agglutinin	VVA-B_4	GalNAc (Tn antigen)	5	100 mM GalNAc
Peanut agglutinin	PNA	Gal (Gal-β-(1→3)-GalNAc (T-antigen) >GalN>Lac>Gal, terminal β-Gal)	15	100 mM Gal
Artocarpus integrifolia agglutinin	Jacalin (AIA)	Gal, Gal-β-(1→3)-GalNAc (T-antigen), Gal-α-(1→6), sialylation independent	15	100 mM Gal
Griffonia simplicifolia lectin II	GS-II	GlcNAc (oligomers>monomer), terminal α- or β-GlcNAc	10	100 mM GlcNAc
Phaseolus vulgaris erythroagglutinin	PHA-E	Complex biantennary, bisecting GlcNAc, Gal-β-(1→4)-GlcNAc terminal	5	5 mg/mL bovine IgG
MECA-79 antibody	MECA-79	GlcNAc-6-SO_4 as part of the 6-sulfo-sialyl Lewis x (6-sulfo-SLe^x), 6-sulfolactose	10	n.a.

Table 3. Ratios of binding intensities of lectins to chicken intestinal mucin samples compared to the ratio of the attenuation effect of chicken mucins on internalisation of *C. jejuni* into HCT-8 cells.

		Intensity ratios for lectin binding										
Mucins	**Attenuation effect**	**UEA-I**	**Con A**	**MAA**	**WFA**	**Jacalin**	**VVA**	**PHA-E**	**PNA**	**GS-II**	**SBA**	**SNA-I**
Proximal small intestine (M34)	30	1	2.2	1	2.3	1	1	1	1	1.1	1	1.3
Cecum (M70)	1	1.7	1	5.1	1	2.1	2	3	1.1	1	1.3	1
Large intestine (M41)	300	4.2	5.1	13.3	99.3	5.5	1.8	6.4	1	1.1	3.8	2

Table 4. Quantification of monosaccharide analysis of GIT mucins. Concentration is given in pmoles per μg and the ratio relative to GalN is in italics. n.d., not detected; tr, trace.

GIT mucins	Code	Monosaccharides			Conc (pmol/ug) *Ratio relative to GalN*		
		Fuc	**GalN**	**GlcN**	**Gal**	**Glc**	**Man**
Chicken							
Ceca	M70	48.8	9.4	57.2	64.2	25.3	1.3
		5.12	*1*	*6.08*	*6.83*	*2.69*	*0.14*
Large intestine	M41	225.9	76.9	268.8	249.5	38.4	41.9
		2.94	*1*	*3.495*	*3.83*	*0.5*	*0.54*
Equine							
Stomach	M6	172	63.3	102.4	429.5	6	n.d.
		2.72	*1*	*1.62*	*6.79*	*0.1*	
Duodenum	M52	99.1	37.6	106.8	128.6	11.9	13.5
		2.64	*1*	*2.84*	*3.42*	*0.32*	*0.36*
Jejunum	M59	129.1	70.7	118.2	219.8	51.7	tr
		1.83	*1*	*1.67*	*3.11*	*0.73*	
Right ventral	M65	47.05	57.3	84.3	91.8	21.6	24.7
ascending colon		*0.82*	*1*	*1.47*	*1.60*	*0.38*	*0.43*
Left ventral	M60	28	36.3	54	55.1	11.2	19
ascending colon		*0.77*	*1*	*1.49*	*1.52*	*0.31*	*0.52*
Dorsal ascending	M66	30.9	60	96.5	71.4	18.25	30.3
colon		*0.515*	*1*	*1.61*	*1.19*	*0.30*	*0.505*
Bovine							
Duodenum	M58	tr	18.7	25.4	21.6	2.6	n.d.
			1	*1.36*	*1.16*	*0.14*	
Spiral colon	M61	98.1	53.3	89	166.1	39.4	24.9
		1.84	*1*	*1.67*	*3.11*	*0.74*	*0.47*
Ovine							
Abomasum antrum	M10	125	126.6	174.5	232.1	17.45	n.d.

		0.99	*1*	*1.38*	*1.83*	*0.14*	
Duodenum	M36	63	34.9	63.9	67.1	13.2	19.3
		1.81	*1*	*1.83*	*1.92*	*0.38*	*0.55*
Jejunum	M35	55	38	58.5	64.2	21.3	5.7
		1.46	*1*	*1.54*	*1.69*	*0.56*	*0.15*
Descending colon	M12	73.7	91.4	365.6	332.1	41.3	17
		0.81	*1*	*4*	*3.63*	*0.45*	*0.19*
Deer							
Abomasum	M67	128.2	103.7	191.9	284	11.7	tr
		1.24	*1*	*1.85*	*2.74*	*0.11*	
Duodenum	M62	9.1	4.8	18.75	19.1	5.95	2.5
		1.9	*1*	*3.9*	*3.98*	*1.24*	*0.52*
Jejunum	M55	52.6	68.9	170.2	170.6	211.7	49.8
		0.76	*1*	*2.47*	*2.47*	*3.07*	*0.72*
Spiral ascending colon	M56	19	11.2	31.3	33.5	3.7	tr
		1.70	*1*	*2.79*	*2.99*	*0.33*	
Porcine							
Gastric	M37	392.5	167.3	593.9	578.8	n.d.	n.d.
		2.35	*1*	*3.55*	*3.46*		

REFERENCES

(1) Brockhausen, I. In *Comprehensive Glycosciences*; Kamerling, J. P., Ed.; Elsevier: Burlington, MA, 2007; Vol. 3, pp 33-59.
(2) Jeffers, F.; Fuell, C.; Tailford, L. E.; MacKenzie, D. A.; Bongaerts, R. J.; Juge, N. *Carbohydr. Res.* **2010**, *345*, 1486-1491.
(3) Conze, T.; Carvalho, A. S.; Landegren, U.; Almeida, R.; Reis, C. A.; David, L.; Söderberg, O. *Glycobiol.* **2010**, *20*, 199-206.
(4) Bhavanandan, V.; Gupta, D.; Woitach, J.; Guo, X.; Jiang, W. *Biosci. Reports* **1999**, *19*, 209-217.
(5) Derrien, M.; van Passel, M. W. J.; van de Bovenkamp, J. H. B.; Schipper, R.; de Vos, W.; Dekker, J. *Gut Microbes* **2010**, *1*, 254-268.
(6) Lebeer, S.; Vanderleyden, J.; De Keersmaecker, S. C. J. *Nat. Rev. Micro.* **2010**, *8*, 171-184.
(7) Corfield, A. P.; Wiggins, R.; Edwards, C.; Myerscough, N.; Warren, B. F.; Soothill, P.; Millar, M. R.; Horner, P. *Adv. Exp. Med. Biol.* **2003**, *535*, 3-15.
(8) Singh, P. K.; Hollingsworth, M. A. *Trends Cell Biol.* **2006**, *16*, 467-476.
(9) Carrington, S. D.; Clyne, M.; Reid, C. J.; Fitzpatrick, E.; Corfield, A. P. In *Microbial Glycobiology: Structures, Relevance and Applications*; Moran, A. P., Holst, O., Brennan, P. J., von Itzstein, M., Eds.; Academic Press: London, 2009, pp 655-672.
(10) Artis, D. *Nat. Rev. Immunol.* **2008**, *8*, 411-420.
(11) Liévin-Le Moal, V.; Servin, A. L. *Clin. Microbiol. Rev.* **2006**, *19*, 315-337.
(12) McGuckin, M. A.; Lindén, S. K.; Sutton, P.; Florin, T. H. *Nat. Rev. Micro.* **2011**, *9*, 265-278.
(13) Jensen, P. H.; Kolarich, D.; Packer, N. H. *FEBS J.* **2010**, *277*, 81-94.
(14) Premaratne, P.; Welén, K.; Damber, J.-E.; Hansson, G.; Bäckström, M. *Tumor Biol.* **2011**, *32*, 203-213.
(15) Karlsson, N. G.; Nordman, H.; Karlsson, H.; Carlstedt, I.; Hansson, G. C. *Biochem. J.* **1997**, *326*, 911-917.
(16) Robbe-Masselot, C.; Maes, E.; Rousset, M.; Michalski, J.-C.; Capon, C. *Glycoconj. J.* **2009**, *26*, 397-413.
(17) Robbe-Masselot, C.; Herrmann, A.; Carlstedt, I.; Michalski, J.-C.; Capon, C. *Glycoconj. J.* **2008**, *25*, 213-224.
(18) Thomsson, K. A.; Bäckström, M.; Holmén Larsson, J. M.; Hansson, G. C.; Karlsson, H. *Anal. Chem.* **2010**, *82*, 1470-1477.
(19) Rose, M. C.; Voynow, J. A. *Physiol. Revs.* **2006**, *86*, 245-278.
(20) Miner-Williams, W.; Moughan, P. J.; Fuller, M. F. *J. Agric. Food Chem.* **2009**, *57*, 6029-6035.
(21) Godula, K.; Rabuka, D.; Nam, K. T.; Bertozzi, C. R. *Angew. Chem. Int. Ed.* **2009**, *48*, 4973-4976.
(22) Kato, K.; Takeuchi, H.; Ohki, T.; Waki, M.; Usami, K.; Hassan, H.; Clausen, H.; Irimura, T. *Biochem. Biophys. Res. Comm.* **2008**, *371*, 698-701.
(23) Dam, T. K.; Brewer, C. F. *Glycobiol.* **2010**, *20*, 270-279.
(24) Ganesh, I. M.; Subramani, D.; Halagowder, D. *J. Carcinog.* **2007**, *6*, 10-14.
(25) Godula, K.; Bertozzi, C. R. *J. Am. Chem. Soc.* **2010**, *132*, 9963-9965.
(26) Kim, Y. S.; Ho, S. B. *Curr. Gastroenterol. Rep.* **2010**, *12*, 319-330.

(27) Alemka, A.; Whelan, S.; Gough, R.; Clyne, M.; Gallagher, M. E.; Carrington, S. D.; Bourke, B. *J. Med. Microbiol.* **2010**, *59*, 898-903.
(28) Davies, J. R.; Carlstedt, I. In *Glycoprotein Methods and Protocols: The Mucins*; Corfield, A., Ed.; Humana Press Inc.: Totowa, NJ, 2000; Vol. 125, pp 3-13.
(29) Alemka, A.; Clyne, M.; Shanahan, F.; Tompkins, T.; Corcionivoschi, N.; Bourke, B. *Infect. Immun.* **2010**, *78*, 2812-2822.
(30) Kilcoyne, M.; Shah, M.; Gerlach, J. Q.; Bhavanandan, V.; Nagaraj, V.; Smith, A. D.; Fujiyama, K.; Sommer, U.; Costello, C. E.; Olszewski, N.; Joshi, L. *J. Plant Physiol.* **2009**, *166*, 219-232.
(31) Xia, B.; Royall, J. A.; Damera, G.; Sachdev, G. P.; Cummings, R. D. *Glycobiol.* **2005**, *15*, 747-775.
(32) Gerlach, J. Q.; Kilcoyne, M.; Eaton, S.; Bhavanandan, V.; Joshi, L. In *The Molecular Immunology of Complex Carbohydrates-3*; Wu, A. M., Ed.; Springer: New York, 2011; Vol. 705, pp 257-269.
(33) Daly, D. S.; Anderson, K. K.; Seurynck-Servoss, S. L.; Gonzalez, R. M.; White, A. M.; Zangar, R. C. *Stat. Appl. Genet. Mol. Biol.* **2010**, *9*, Article 14.
(34) Guilleaume, B.; Buneß, A.; Schmidt, C.; Klimek, F.; Moldenhauer, G.; Huber, W.; Arlt, D.; Korf, U.; Wiemann, S.; Poustka, A. *PROTEOMICS* **2005**, *5*, 4705-4712.
(35) Manimala, J. C.; Roach, T. A.; Li, Z.; Gildersleeve, J. C. *Glycobiol.* **2007**, *17*, 17C-23C.
(36) Chow, J.; Lee, S. M.; Shen, Y.; Khosravi, A.; Mazmanian, S. K. In *Advances in Immunology*; Fagarasan, S., Cerutti, A., Eds.; Academic Press: San Diego, 2010; Vol. 107, pp 243-274.
(37) Miyata, S.; Nishimura, Y.; Hayashi, N.; Oohira, A. *Neurosci.* **2005**, *136*, 95-104.
(38) Bai, X.; Brown, J. R.; Varki, A.; Esko, J. D. *Glycobiol.* **2001**, *11*, 621-632.
(39) Kawasaki, N.; Haishima, Y.; Ohta, M.; Itoh, S.; Hyuga, M.; Hyuga, S.; Hayakawa, T. *Glycobiol.* **2001**, *11*, 1043-1049.
(40) Bruehl, R. E.; Bertozzi, C. R.; Rosen, S. D. *J. Biol. Chem.* **2000**, *275*, 32642-32648.
(41) Hendrixson, D. R.; DiRita, V. J. *Mol. Microbiol.* **2004**, *52*, 471-484.
(42) Schumacher, U.; Duku, M.; Katoh, M.; Jörns, J.; Krause, W. J. *Anat. Rec. A Discov. Mol. Cell. Evol. Biol.* **2004**, *278A*, 540-550.
(43) Hoang, V. C.; Williams, M. A. K.; Simpson, H. V. *Vet. Parasitol.* **2010**, *171*, 354-360.
(44) Rinaldi, M.; Dreesen, L.; Hoorens, P.; Li , R.; Claerebout, E.; Goddeeris, B.; Vercruysse, J.; Van Den Broek, W.; Geldhof, P. *Vet. Res.* **2011**, *42*, 61.
(45) Bhavanandan, V. P.; Katlic, A. W. *J. Biol. Chem.* **1979**, *254*, 4000-4008.
(46) Fair, S.; Hanrahan, J. P.; O'Meara, C. M.; Duffy, P.; Rizos, D.; Wade, M.; Donovan, A.; Boland, M. P.; Lonergan, P.; Evans, A. C. O. *Theriogenology* **2005**, *63*, 1995-2005.
(47) Lagow, E.; DeSouza, M.; Carson, D. *Human Reprod. Update* **1999**, *5*, 280-292.
(48) Jeschke, U.; Walzel, H.; Mylonas, I.; Papadopoulos, P.; Shabani, N.; Kuhn, C.; Schulze, S.; Friese, K.; Karsten, U.; Anz, D.; Kupka, M. S. *J. Histochem. Cytochem.* **2009**, *57*, 871-881.
(49) Foulk, R.; Zdravkovic, T.; Genbacev, O.; Prakobphol, A. *J. Assist. Reprod. Genet.* **2007**, *24*, 316-321.
(50) Klinken, B. J.-W.; Oussoren, E.; Weenink, J.-J.; Strous, G. J.; Büller, H. A.; Dekker, J.; Einerhand, A. W. C. *Glycoconj. J.* **1996**, *13*, 757-768.
(51) Jintapattanakit, A.; Junyaprasert, V. B.; Kissel, T. *J. Pharm. Sci.* **2009**, *98*, 4818-4830.

(52) Keely, S.; Rullay, A.; Wilson, C.; Carmichael, A.; Carrington, S.; Corfield, A.; Haddleton, D. M.; Brayden, D. J. *Pharm. Res.* **2005**, *22*, 38-49.
(53) Irimura, T.; Denda, K.; Iida, S.; Takeuchi, H.; Kato, K. *J. Biochem.* **1999**, *126*, 975-986.
(54) Delmotte, P.; Degroote, S.; Merten, M. D.; Van Seuningen, I.; Bernigaud, A.; Figarella, C.; Roussel, P.; Périni, J.-M. *Glycoconj. J.* **2001**, *18*, 487-497.
(55) Ichikawa, T.; Ishihara, K.; Kusakabe, T.; Kurihara, M.; Kawakami, T.; Takenaka, T.; Saigenji, K.; Hotta, K. *Am. J. Physiol.* **1998**, *274*, G138-G146.
(56) Kilcoyne, M.; Gerlach, J. Q.; Farrell, M. P.; Bhavanandan, V. P.; Joshi, L. *Anal. Biochem.* **2011**, *416*, 18-26.
(57) Freitas, M.; Axelsson, L.-G.; Cayuela, C.; Midtvedt, T.; Trugnan, G. *Histochem. Cell Biol.* **2005**, *124*, 423-433.

Supporting Information

Construction of a natural mucin microarray and interrogation for biologically relevant presented glyco-epitopes

Michelle Kilcoyne,[†] Jared Q. Gerlach,[†] Ronan Gough,[§] Mary E. Gallagher,[§] Marian Kane,[†] Stephen D. Carrington,[§] and Lokesh Joshi[†,*]

[†]Glycoscience Group, National Centre for Biomedical Engineering Science, National University of Ireland, Galway, Ireland

[§]Veterinary Science Centre, University College Dublin, Belfield, Dublin 4, Ireland

Table of contents

Tables

Figures

Figure S-1. Analysis of the fractions produced by a density gradient of ovine ileum. The presence of mucins is implied in the fractions that have a strong response to PAS staining. The strongest PAS bands occur between a buoyant density of 1.35 and 1.45 g/ml which corresponds to the buoyant density range of mucins.

Figure S-2. Histogram of fluorescently labelled-lectins binding to glycoprotein standards printed at 0.25 mg/mL. Printed PBS and PBS-T (not shown) did not give signal intensity above background and no autofluorescence of the printed mucins or glycoproteins was noted. Error bars depict standard deviation of the mean of three microarray slides.

Figure S-3. Histogram of fluorescently labelled-lectins binding to porcine gastric mucin (M37). Error bars depict standard deviation of the mean of three microarray slides.

Table S-1. Mucin-type *O*-linked core structures

Core type	Structure
1	Gal-β-(1→3)-GalNAc-α-
2	GlcNAc-β-(1→6)-(Gal-β-(1→3))-GalNAc-α-
3	GlcNAc-β-(1→3)-GalNAc-α-
4	GlcNAc-β-(1→6)-(GlcNAc-β-(1→3))-GalNAc-α-
5	GalNAc-β-(1→3)-GalNAc-α-
6	GlcNAc-β-(1→6)-GalNAc-α-
7	GalNAc-α-(1→6)-GalNAc-α-
8	Gal-α-(1→3)-GalNAc-α-

Table S-2. Code, mucin, mucin source and purification summary.

Code	Source	Mucus production and harvesting conditions
3	Bovine cervix, animal 0049	Animal in estrus Native tissue, harvested by scraping and washing with 6M GHCl
4	Bovine cervix, animal 0258	Animal in metoestrus Native tissue, harvested by scraping and washing with 6M GHCl
6	Equine stomach	Native tissue, harvested by washing with 8M GHCl
10	Ovine abomasum antrum	Native tissue, harvested by scraping
11	E12	Cells were grown on transwell filters (44cm^2). Extracellular mucus was harvested 21 days after seeding through treatment with 10mM *N*-acetylcysteine which solubilised the mucus such that it could be removed by aspiration.
12	Ovine descending colon	Native tissue, harvested by scraping
15	Bovine cervico-vaginal mucin, animal 1467	Animal in oestrus Mucus secretions aspirated from a live animal
18	Ovine spiral colon	Native tissue, harvested by scraping
30	Ovine cervix (breed: Suffolk)	Animal in early estrus Native tissue, harvested by washing with 8M GHCl
31	Ovine cervix (breed: Belclare)	Animal in early estrus Native tissue, harvested by washing with 8M GHCl
32	Ovine cervix (breed: Suffolk)	Animal in late estrus Native tissue, harvested by washing with 8M GHCl
33	Ovine cervix (breed: Belclare)	Animal in late estrus Native tissue, harvested by washing with 8M GHCl.
34	Chicken proximal small intestine	Native tissue, harvested by scraping
35	Ovine jejunum	Native tissue, harvested by scraping
36	Ovine duodenum	Native tissue, harvested by scraping
37	Porcine gastric mucin (PGM)	This is commercially sourced and prepared mucin that we have purified further
39	Equine (pregnant) cervix, animal AP026	Native tissue, harvested by washing with 8M GHCl
41	Chicken large intestine	Native tissue, harvested by scraping
48	Bovine cervico-vaginal mucin, animal 0113	Animal in pro-estrus Mucus secretions aspirated from a live animal
49	Bovine cervico-vaginal mucin, animal 278	Animal in metoestrus Mucus secretions aspirated from a live animal

52	Equine duodenum	Native tissue, harvested by scraping
53	Equine trachea	Native tissue, harvested by scraping
55	Deer jejunum	Native tissue, harvested by scraping
56	Deer spiral ascending colon	Native tissue, harvested by scraping
57	Bovine abomasum	Native tissue, harvested by scraping
58	Bovine duodenum	Native tissue, harvested by scraping
59	Equine jejunum	Native tissue, harvested by scraping
60	Equine left ventral ascending colon	Native tissue, harvested by scraping
61	Bovine spiral colon	Native tissue, harvested by scraping
62	Deer duodenum	Native tissue, harvested by scraping
63	Bovine trachea	Native tissue, harvested by scraping
64	Bovine endometrium (uterine horn)	Native tissue, harvested from the uterine horn by washing with 8M GHCl
65	Equine right ventral ascending colon	Native tissue, harvested by scraping
66	Equine dorsal ascending colon	Native tissue, harvested by scraping
67	Deer abomasum	Native tissue, harvested by scraping
70	Chicken ceca	Native tissue, harvested by scraping
72	LS174T	Cells were grown in T-175s. The media was removed by aspiration and condensed by freeze drying. The cells were lifting using a cell scraper and the flasks were washed out using 6MGHCl in PBS. All these components were then combined.

Table S-3. Comparison of average slide-to-slide percentage coefficient of variance (%CV) across mucins incubated with different fluorescently-labelled lectins. Mucins are grouped together and grouped according to species for ease of comparison.

	Average %CV										
Probes	**UEA-I**	**Con A**	**MAA**	**WFA**	**Jacalin**	**VVA**	**PHA-E**	**PNA**	**GS-II**	**SBA**	**SNA-I**
All mucins	26.6	41.8	26.4	25.0	13.4	24.1	20.6	41.3	27.6	42.1	45.7
Glycoproteins (-ovomucoid)	31.9 (34.7)	16.7 (17.2)	27.7 (28.5)	72.5 (57.6)	32.8 (25.3)	49.5 (37.1)	34.2 (39.2)	123.0 (41.0)	39.4 (32.8)	42.7 (24.8)	124.7 (48.0)
Bovine mucins	13.5	42.8	25.1	19.4	15.3	32.4	16.9	43.0	32.6	47.8	46.5
Ovine mucins	30.2	44.2	20.4	24.1	16.2[a]	21.8[a]	22.7[a]	34.7[b]	28.5[a]	35.9[a]	49.2[a]
Cell line mucins	24.1	41.8	35.6	40.6	16.9	11.7	14.7	19.4	28.7	31.3	40.3
Chicken mucins	31.3	34.3	30.6	30.7	8.8	19.4	26.1	44.7	24.9	29.2	37.3
Equine mucins	35.8	39.7	37.6	21.3[c]	11.3[d]	21.2[d]	28.4[d]	45.6[d]	27.8[d]	45.5[d]	50.9[d]
Deer mucins	21.8	35.5	31.5	37.7	7.1	18.6	6.8	48.7	19.1	41.5	35.7

[a]Less M10, M36, M12, [b]less M36, M12, [c]less M66, [d]less M65.

Table S-4. Complete lectin inhibition data for all mucins presented in **Excel file: Lectin Inhibition All Mucins.xml.**

UIRFU values in parentheses are below signal threshold of 5 times background. Positive percentage values indicate true inhibition, values in parentheses indicate enhanced RFU values in presence of hapten (as in the case of PHA-L). *, $p = 0.02$ to 0.05; **, $p \leq 0.01$, all figures rounded to hundredth digit.

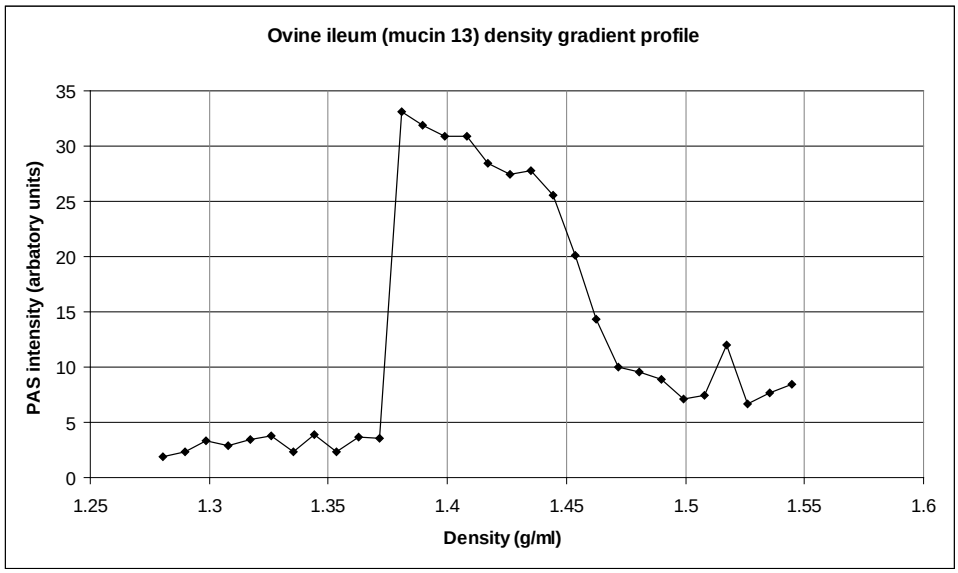

Figure S-1. Analysis of the fractions produced by a density gradient of ovine ileum. The presence of mucins is implied in the fractions that have a strong response to PAS staining. The strongest PAS bands occur between a buoyant density of 1.35 and 1.45 g/ml which corresponds to the buoyant density range of mucins.

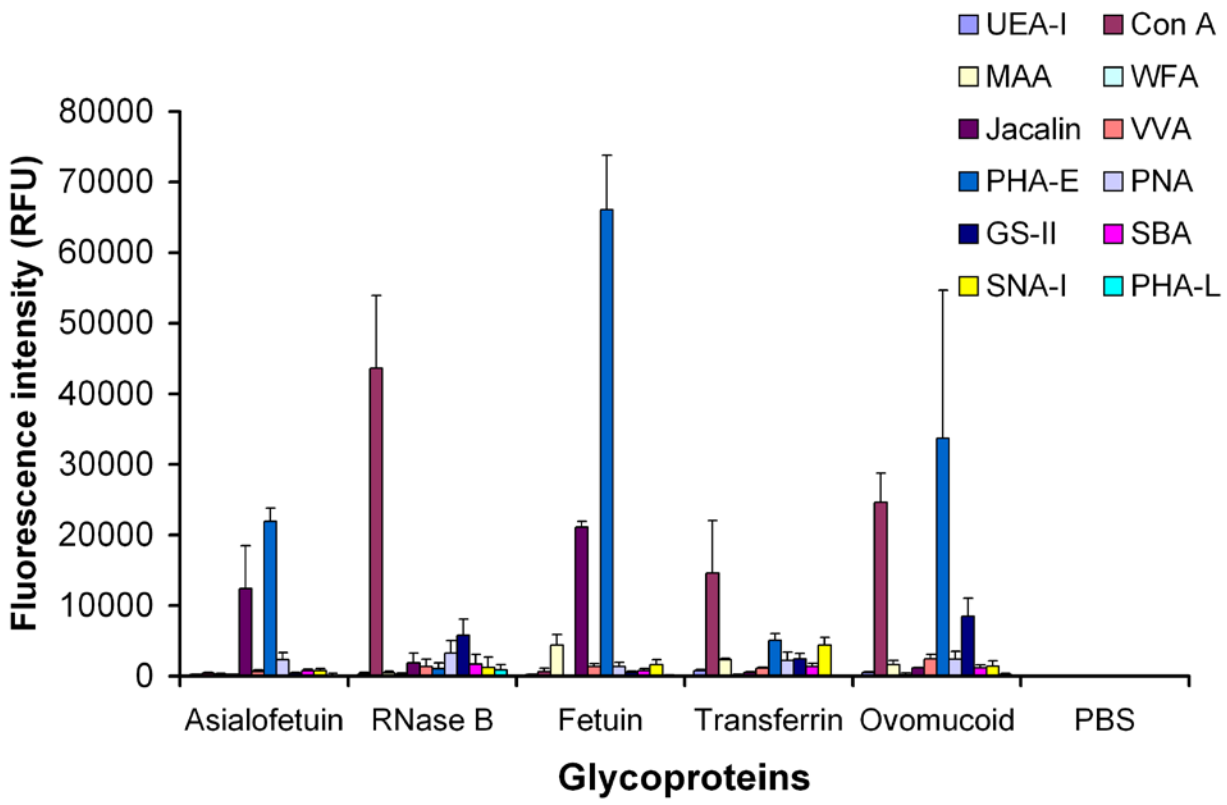

Figure S-2. Histogram of fluorescently labelled-lectins binding to glycoprotein standards printed at 0.25 mg/mL. Printed PBS and PBS-T (not shown) did not give signal intensity above background and no autofluorescence of the printed mucins and glycoproteins was noted. Error bars depict standard deviation of the mean of three microarray slides.

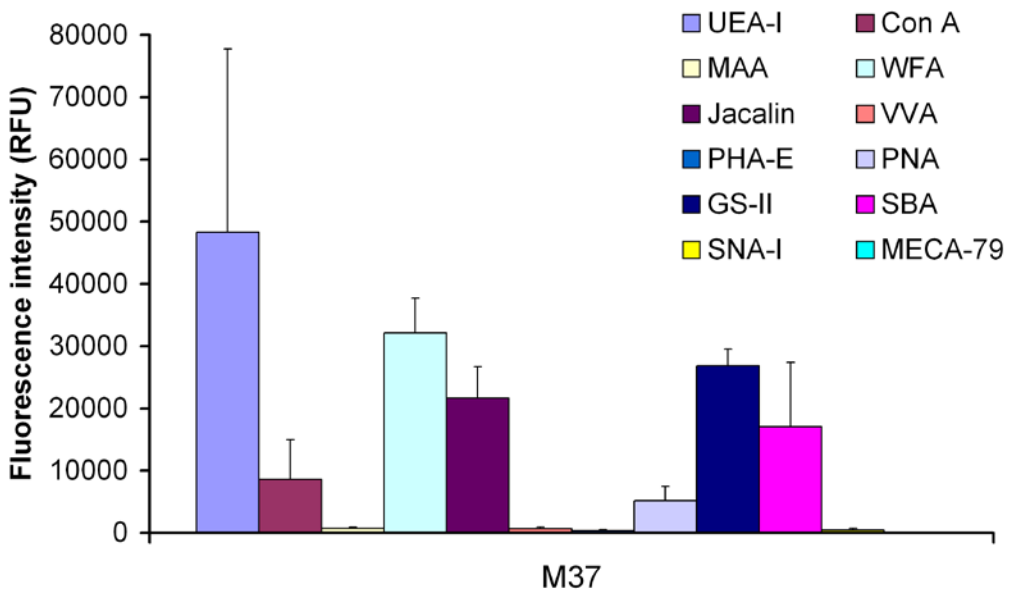

Figure S-3. Histogram of fluorescently labelled-lectins binding to porcine gastric mucin (M37). Error bars depict standard deviation of the mean of three microarray slides.