Published in final edited form as:

Annu Rev Biochem. 2020 June 20; 89: 769–793. doi:10.1146/annurev-biochem-011520-105053.

Mucins and the Microbiome

Gunnar C Hansson

Department of Medical Biochemistry, University of Gothenburg, Gothenburg, Sweden

Abstract

Generating the barriers that protect our inner surfaces from bacteria and other challenges requires large glycoproteins called mucins. These come in two flavors, gel-forming and transmembrane, all characterized by large highly *O*-glycosylated mucin domains diversely decorated by Golgi glycosyltransferases to become extended rod-like structures. The general function of the mucins on internal epithelial surfaces is to wash away microorganisms, but even more importantly to build protective barriers. This is most evident in the large intestine where the inner mucus layer separates the numerous commensal bacteria from the epithelial cells. The host conversion of the MUC2 mucin to the outer mucus layer allow the bacteria to degrade the mucin glycans and recover its energy content that is shared with the host. The molecular nature of the mucins is complex and how these are building the extracellular complex glycocalyx and mucus is poorly understood and a future biochemical challenge.

Keywords

Mucus; Carbohydrate; Glycan; Saccharide; Intestine; Bacteria

INTRODUCTION

Maintaining the integrity of a living organism is most important. This requires the clear separation from other organisms that would like to take advantage of the other. The most efficient protection is to have an inert coat in the form of a shell or as for higher animals a skin built by dead cells. However, this principle cannot be utilized for organs that shall actively interact with the surroundings. Thus these organs require dynamic systems that still keep intruders at bay.

The inner surfaces of our bodies are coated with single or multiple layers of very active cells responsible for nutrient, liquid and gas exchange. These cells are protected by mucus made by specialized goblet cells and by surface coating by a dense glycocalyx (23, 69, 83). Glycoproteins called mucins are the main building block of both the mucus and the glycocalyx (66). The high glycan content makes the mucins water-soluble and able to bind large amounts of water reflected in the transparent nature of normal mucus and glycocalyx. The protein core of mucins is made inaccessible by the high number of attached

gunnar.hansson@medkem.gu.se .

DISCLOSURE STATEMENT

Nothing to declare.

oligosaccharides, something that is important for protecting the mucosal surfaces. This is the case in the gastrointestinal tract where the milieu is ideal for bacterial proliferation with its balanced salt and water content, richness in nutrients and ideal temperature. The number of intestinal bacteria is close to the number of cells in our body, but still the bacteria do not take over. This is largely due to the mucins and their structure and function. However, this is not the only remarkable feature of the mucins. Certain commensal bacteria are specialized in degrading the mucin glycans and utilize the obtained energy to feed not only themselves, but also their host. Some of the mysteries around the mucins and their relation to bacteria is starting to be revealed, but much more is to be learnt. The state of today's understanding will be discussed in the following.

MUCINS

Mucins are characterized by their dense coat of glycans, especially of the *O*-glycan family. A mucin was originally defined as a glycoprotein with more than 50% of its mass attributed to *O*-glycans and the name mucin used for the major components of mucus. Today these mucins are known as the gel-forming mucins. During the era of gene cloning in the 1980s, the first mucin to be sequenced was mucin-1 with the abbreviated gene name *MUC1* (43). This is a transmembrane mucin and as it later turned out very different from the classical gel-forming mucins of which the *MUC2* was the first to be cloned (48)(Fig. 1). As the naming simply followed the order of discovery, the mucin family nomenclature often appears confusing.

Mucin domains

Sequencing mucin genes revealed long stretches of, often repeated, sequences that varied in length. This type of sequences are generally named VNTR (Varying Number of Tandem Repeats) and thus these parts of the mucins were originally named this way (77, 117). The human genome contains numerous such VNTR repeats, but few are in exons. Analyzing the same mucin and its VNTR sequences show that the repeated nature is often lost during evolution, although the original repeats are sometimes possible to predict. Further studies of these sequences revealed little relation to evolution, but instead characterized by their high frequency of the hydroxy-amino acids (Ser and Thr) together with Pro (Fig. 1). This made us introduce the name **PTS** for such sequences and predicted domains (78). The hydroxyl groups are used for attaching the first *O*-glycan sugar, the GalNAc residue. The Pro amino acids ensure a non-folded structure allowing the peptidyl-GalNAc transferases of the Golgi apparatus to access the protein core (7). One can predict PTS sequences by mining genomes for long sequences rich of Ser/Thr (>25%) and Pro (>5%) in proteins with signal sequences (78).

The PTS sequences have a non-structured, random coil nature when exiting the endoplasmic reticulum (ER). Once in the Golgi apparatus the peptidyl-GalNAc transferases start to decorate the PTS sequences. Most of the Ser and Thr amino acids become glycosylated and the GalNAc is typically further modified by additional glycans. By this, the PTS sequence becomes *O*-glycosylated and a **mucin domain** is formed. The dense glycosylation pulls out the protein core and long extended rod-like structures are formed. When the length of

the mucin domain has been possible to compare with the actual amino acid composition and sequence, that the peptide backbone is maximally extended (17, 116). This is in line with observations by electron microscopy where the mucin domains are appearing as long extended rods that appear stiff (115). The properties of the mucin domains will depend on the individual mucin protein sequence and the repertoire of glycosyltransferases present in the mucin-producing cell. Less densely, glycosylated mucin domains with short non-branched glycans will likely be more flexible, whereas densely glycosylated domains with highly branched and long glycans can be predicted to be stiffer. It should be pointed out that many other proteins, not classified as mucins, have short mucin domains. This is typical for membrane proteins, like the LDL-receptor, that have a short stalk to extend the functional domain out from the cell surface membrane. Domain predicting programs have so far not included tools able to predict extracellular PTS sequences that will become *O*-glycosylated; rather they usually call these parts simply 'unstructured'. This is a misunderstanding as once *O*-glycosylated, the formed mucin domains are highly structured.

Mucin *O*-glycosylation

The predominant glycosylation of mucins is the *O*-glycans that typically make up more than 80% of the mass of a mucin. The protein core is well hidden and the glycans form the outer surface of these mucin domains. The glycans are densely packed and most of the time individual glycans will not be recognized, but rather the glycan surface are forming clustered saccharide patches (21, 131). However, due to the lack of methods there is limited understanding of the structure of such patches. Much research has been devoted to studies of bacterial binding to individual glycans, but the likely more relevant and specific bacterial binding to the glycan patches on mucins has not been studied (71).

The *O*-glycosylation of the PTS sequences start in cis-Golgi by the action of any of the peptidyl-GalNAc transferases (7, 24). These enzymes appeared early in evolution during Metazoan and at the same time as the first mucins appeared, suggesting a close relation between the two. There are 20 such transferase genes in the human genome that have different specificities, some are more promiscuous whereas others are more specific. Biotechnical and mass spectrometry developments have allowed systematic studies of individual single *O*-glycan attachment sites (120). This show that the different peptidyl-GalNAc transferases display peptide sequence specificity based both on the catalytic unit and its lectin-binding domain and thus act in concert to control mucin-type *O*-glycosylation (25, 101). Some of these transferases can act directly on the naked peptides, whereas others work only after its glycan binding domain has bound to an already attached GalNAc (7, 101). Consequentially, it is difficult to know if a certain Ser or Thr in a peptide sequence is glycosylated or not. Although Ser or Thr close to Pro is typically glycosylated, prediction of glycan attachment sites based on amino acid sequence is still complex due to the integration of target protein structure, transferase specificity, relative levels, and relative localization of peptidyl-GalNAc transferases in the Golgi apparatus of the cell.

Once the first glycan residue in the form of GalNAc is attached to the protein core, this will act as a substrate for the transferases able to add Gal, GlcNAc, GalNAc, NeuAc, and sulfate groups to this residue (3, 13, 23, 136). Studies of individual *O*-glycans show that also

this group of enzymes are dependent of the peptide sequence further contributing the spatial organization of glycans on mucins (44). Based on the coupling of Gal to GlcNAc, there are four major basic structures that make up the core of *O*-glycans, called Core1–4 (see Fig. 2a for explanation). Already at this level, there are large differences between animal species as illustrated in Fig. 2b showing that mouse colon Muc2 largely have Core2, whereas human MUC2 predominantly have Core3 structures (59, 103, 104).

The Core-structures are further elongated by the addition of the same type of sugar residues. A classical extension is one or several lactoseamine disaccharide units of type1 (Gal β 1–3GlcNAc) or type2 (Gal β 1–4GlcNAc) that in this way generate long *O*-glycans. This varies of course with tissue and species, but most of the time the mucin glycans of the intestine is shorter and have extensions on both the C-3 and C-6 of the peptide-bound GalNAc. This will generate bulky glycans that will help in extending and stiffening the mucin domains and efficiently shield the protein cores from access of proteases.

Mucin *O*-glycans are typically capped with peripheral glycan residues that cannot be further elongated. Among these are fucose (Fuc α 1–2, –3, or –4), sialic acid (NeuAc(Gc) α 2–3 or –6), GalNAc (α 1–3, β 1–4), or Gal (α 1–3). Some of these are in combination making the blood- or histo-group structures defining the A, B, H and Lewis a, b, x and y epitopes found on mucins as well as on internal cells and erythrocytes. Both the peripheral or core glycans are substituted with sulfate groups and numerous of the hydroxyl groups acetylated. This latter modification is known to be present on sialic acids, but likely abundant on other sugar residues as well (112).

Gastro-intestinal mucin *O*-glycosylation

The mucin glycan repertoire differ substantially along the gastro-intestinal tract. In general, the stomach mucins are neutral and contain few sialic acids, whereas mucins in the distal end of the intestine are rich in sialic acid and sulfate groups. The normal human stomach has an enormous glycan diversity with up to 13 residues long glycans and around 100 different species in each individual (63, 106). In 10 individuals more than 250 different glycan structures were identified.

The normal human intestine show a specific distribution of *O*-glycans along its length (103, 104). Fucose in blood group ABH antigens is more localized to the small intestine and proximal part of the large intestine and the longitudinal glycan differences along the colon is as expected reflected in the levels of glycosyl transferases (128). That blood group status is reflected in the intestinal mucins means that there is an individual variability in glycan epitopes. Interestingly, this individual glycan variability is almost absent in the distal part of the human large intestine where the glycans are relatively uniform between individuals (57). This should likely be understood in relation to the commensal bacteria where they utilize the mucin glycans for attachment as well as a nutrient source. The glycosylation of the mucins in the distal human colon seems to be stable over time, but was found to be altered in active ulcerative colitis (56). The altered glycan repertoire was triggered by the inflammation when the disease was active and returned back to normal as the inflammation faded.

The mucin glycans show large variability between different evolutionary closely related species and this is especially evident in the respiratory and gastrointestinal tracts. Likely explanations are that different microbial pathogens has generated exogenous selection pressure to replace glycan structures that has been utilized for invading the host (42). As a consequence, there are unexplored genetic mechanisms favoring species polymorphism and a more rapid evolution of glycans than in general.

It is typical for mucins to show a large number, often more than 100, different *O*-glycan structures on the same mucin from the same organ. A major reason for this large diversity is that all biosynthetic intermediates are typically present. This is in contrast to the glycosphingolipids of the apical enterocyte cell membrane where most biosynthetic intermediates are absent (12). The differences could be due to difficulties for the glycosyltransferases to reach all potential substrates on the mucins due to their size, but is more likely of functional importance as it will allow the mucins to present an enormously rich decoy of potential bacterial binding sites.

Secreted mucins

Mucins lacking transmembrane sequences have been classified as secreted mucins (Table 1, Fig. 1). These are further separated into monomeric and gel-forming mucins. The name gel-forming is because they are able to form mucus with gel-like properties. This is not an ideal name and polymer-forming should probably better reflect their structure. There are two monomeric mucins, MUC7 and MUC20, which essentially only consist of a secreted mucin domain (10, 53).

Gel-forming mucins

The classical mucins that form polymers have a common protein domain organization shared with the von Willebrand factor (VWF), an important molecule in the vascular coagulation system. The four gel-forming mucins and VWF, all have 3.5 von Willebrand D (VWD) assemblies in their N-termini (see Fig. 3 for explanation, Table 1) named VWD1, VWD2, VWD', and VWD3. The central parts of the mucins, but not VWF, is one or several PTS sequences that after *O*-glycosylation become mucin domains. This central mucin domains are often interrupted with 100 amino acid long domains, called CysD. The C-termini of the mucins and the VWF has a Cysteine-knot domain that may be preceded by a VWD4 assembly and von Willebrand C (VWC) domains (Fig. 3).

In the VWF, the central mucin parts are replaced by von Willebrand A and B domains. Studies of the evolution of the VWD domains show that these appeared early or even before Metazoan evolution and was early during Metazoan time associated with PTS sequences as in mucins (Fig. 4) (78, 79). The VWF appeared much later during vertebrate evolution and thus the gel-forming mucins were the origin. Interestingly, the type of gel-forming mucins were lost during the evolution of insects. There is only one remaining protein called hemolectin with two functions, encapsulating bacteria and inhibiting hemolymph leakage, recapitulating the function of the mucins and the VWF, respectively (80). As the peptidyl-GalNAc transferases also appeared during Metazoan, this further argues for the early importance of mucins (7).

During evolution, the number of gel-forming mucins has changed. Most mucins are found in species having mucins not only on their internal organs, but also on their skin. Frogs have around 25 gel-forming mucins and fishes 15–20 (78, 79). In mammals, the number decreased to four or five (Fig. 4), protecting only internal organs. There are five gel-forming mucins in the human genome, but only four of these are expressed as a complete mucin protein; MUC2, MUC5AC, MUC5B, and MUC6. The *MUC19* mucin gene is not expressed in humans, but well in pigs where it is called PSM for porcine submaxillary mucin (19, 33, 109).

All mucins are assembled into covalent dimers in the ER by disulfide-bonds between two Cystein-knot domains (CK, Fig. 3) (4). The dimers are then passed into the Golgi apparatus where the PTS sequences are *O*-glycosylated and reach masses in the range of 5 MDa. The pH is gradually decreasing over the secretory pathway and when the mucin reach the trans-Golgi network (TGN), their N-termini are forming packed electron-dense assemblies that are sorted to the regulated secretory vesicles. The mucin N-termini have been blocked from forming disulfide-bonded polymers in the ER by the VWD1-VWD2 assemblies are now covalently assembled into N-terminal dimers-of-dimers or trimers-of-dimers by disulfide bonds formed within VWD3.

There are only limited detailed structural information of the different domains in the gel-forming mucins. The first related domain to be structurally determined was for the VWF as a monomeric von Willebrand D'D3 domain with two dimerization Cys replaced (32). Recently a structure of the MUC2 VWD3 domain has also been published (62). The structures show the coordination of the expected bound calcium ion. There are also some structural information for other domains in mucins from other proteins. The Cystein-knot domain is found as homodimers in several proteins and the VWC domain is found also in collagen (92).

The role of protecting the mucosal surfaces, especially the intestine, is highly demanding. The intestinal mucin MUC2 is insoluble, something that was long over-looked, but first observed by Ingemar Carlstedt in Lund (18). The insolubility, also in urea and chaotropic salts like guanidinium chloride, was linked to the appearance of non-reducible covalent bonds. Electrophoresis gels revealed higher oligomers than the expected monomer after dithiotreitol treatment (52). Such non-reducible bonds were formed during biosynthesis and passage through the later stages of the secretory pathway (5). Years later, the nature of these non-reducible bonds were shown to be isopeptide bonds between the side-chains of lysine and glutamine (100). Such bonds are stabilizing other proteins exposed to mechanical stress such as the collagen in the skin and fibrin in the blood clot (84). The bonds are formed by transglutaminases of which the TGM2 and TGM3 are found in the normal intestine and in the chronically diseased lungs (39, 105). The precise molecular details is still missing, but it can be assumed that isopeptide bonds will mechanically stabilize the mucus.

MUC2, MUC5AC, MUC5B, and MUC6 mucins

The major normal mucin in the respiratory tract and saliva is **MUC5B** (Table 1, Fig. 3) (30, 38). This mucin forms N-terminal dimers and thus linear polymers (102, 123).

The **MUC5AC** mucin has the most similar sequence to the **MUC5B** mucin (74, 82). This mucin forms N-terminal covalent dimers that is further interacting to tetrameric non-covalent polymers (S. Trillo-Muyo & G.C. Hansson, unpublished manuscript). Thus these will form net-like structures.

The **MUC2** mucin is found in the small and large intestine and have two large mucin domains interrupted with two CysD domains (49). This mucin forms N-terminal disulfide-bonded trimers that will generate stacked net-like polymers once secreted, ideal for the generation of protective filters (2).

The **MUC6** mucin is found in the stomach glands and pancreas and has a similar size as the other mucins, but a simpler organization with a single central mucin domain and short C-termini (121, 124). Little is known about its assembly and higher order polymers.

Trans-membrane mucins

All transmembrane mucins are type 1 transmembrane proteins with an N-terminal extracellular mucin domain, a transmembrane domain and a relatively short unstructured C-terminal cytoplasmic tail (Fig. 1). There are three groups of transmembrane mucins; SEA (sea urchin-enterokinase-agrin) mucins, NIDO-AMOP-VWD mucins, and a group without specific protein domains (Table 1). This last group (**MUC15**, **MUC21** and **MUC22**) has essentially only an extracellular mucin domain coupled to the transmembrane part (53, 54, 94). The NIDO-AMOP-VWD members have these domains extracellular between the N-terminal mucin domain and the transmembrane one. Interestingly there is only one member of this family in each higher species. This mucin is called **MUC4** and is not restricted to mucosal surfaces. Little is known about its normal function as it has largely been studied in relation to cancer development (29, 107, 118).

The **MUC1** mucin is part of the largest group of transmembrane mucins that all have a SEA domain between the N-terminal PTS sequence and the transmembrane domain. The SEA domain is cleaved auto-catalytically during ER folding, but still held together by strong non-covalent forces generated by four anti-parallel β -sheets (87). The **MUC1** mucin is the most studied mucin of this family and is found on many cell types outside of the mucosal surfaces, also on a number of immune cells (6). The largest known mucin have more than 20,000 amino acids and is called **MUC16**. This mucin have more than 30 repeated SEA-PTS sequences where the SEA domains are not cleaved as in the mucins with only one such domain (91, 135). This mucin is well-known because it carries the CA125 epitope, widely utilized as a marker for ovarian cancer (135).

The remaining SEA-mucins (**MUC3**, **MUC12**, **MUC13**, **MUC17**) are all found in the intestinal tract and localized to the apical surface of the polarized epithelial cells (46, 47, 132, 133). Their PTS sequences have more than 4,000 amino acids, except **MUC13** that is considerably smaller with only 150 amino acids in its PTS. The **MUC3**, 12, and 17 reach about one micrometer out from the tip of the enterocyte microvilli into the intestinal lumen. Such mucins build the densely glycosylated glycocalyx of the enterocytes (95).

MUCINS AND BACTERIA

Mucus and its main component, the mucins, as well as epithelial surface mucins have taken the role to protect inner mucosal surfaces. This requires that the molecules act with dual properties, to **protect** and to **remove**. Protection is based on building a system where the epithelium is protected from contact. Removal is based on a system that requires trapping and a possibility to move intruders to for the host less susceptible positions, in most cases colon. Protection requires a barrier that is attached to the epithelial surface. This is for example the case for the transmembrane mucins building the dense glycocalyx of intestinal enterocytes and for the gel-forming mucin in the inner mucus layer of colon. Removing require movement and is accomplished by cilia in the respiratory system and by peristalsis in the intestine.

The challenge for the host to protect its inner surfaces is enormous. Parasites have been a demanding problem during human evolution, but is smaller today where the dense human populations are more challenged by viruses. Bacteria may be most important as they are not only threatening us, but are also dividing fast and undergo rapid evolution. Mucus and mucins have an important role for protecting the organism from these challenges. We will focus our attention on the role of mucins in protecting us from bacteria in the two largest internal surfaces, the respiratory and gastrointestinal tracts.

THE RESPIRATORY SYSTEM

The respiratory tract is exposed to millions of particles and bacteria for every breath we take. Despite this, the normal lungs are kept 'essentially free' from bacteria (75), although DNA amplification methods show that most individuals have at least some resident bacteria (31). This is in contrast to individuals with chronic lung diseases, like Cystic Fibrosis (CF) and chronic obstructive pulmonary disease (COPD), where the lungs are heavily colonized with specific bacteria.

A hallmark of the respiratory epithelia is the ciliated cells. Their beating is continuously generating a vectorial transport cephalically of the air surface liquid (ASL). The ciliated cell membrane and cilia are coated with the transmembrane mucins MUC1, MUC4 and MUC16 where the mucin domains generate a biophysical brush (15). The high carbohydrate content of these mucins maintain the liquid around the cilia that is called periciliary liquid (PCL) and provide the low-friction required for the cilia movement (15). The coating provided by these mucins will help to protect the epithelial cells from both mechanical forces as well as act as a barrier for microorganisms.

Large animals, like pigs and humans, have numerous submucosal glands (Fig. 5a), specialized molecular machines making long and thick MUC5B bundles (36, 55). In the late secretory pathway of the MUC5B mucin producing cells where the pH is low (5.5–6), the N-terminal VWD1-VWD2 domains of the VWD3 dimer are bent inwards and by this hooks two MUC5B dimers that are facing each other at 180° together (125). Electron microscopy and 3D printed models further suggested that these double dimers are assembled side-by-side into linear structures where the remaining MUC5B including the

mucin domains are pointing outwards in all directions (Fig. 5a insert). Such non-expanded structures been observed in saliva (73). In the submucosal gland, the most peripheral cells generate a chloride- and bicarbonate-rich fluid that flushes the MUC5B mucin secreting cells. The by bicarbonate increased pH causes the MUC5B N-termini to disassemble and the flow pulls out MUC5B into linear polymers (34, 36). These polymers are gathered into thicker structures in the gland ducts and are observed at the exit of the glands as thick bundles about 25 μm in diameter (Fig. 5a). These bundles have a central core of MUC5B mucins. Calculations based on the estimated diameter of the mucin domains, suggests that each bundle contain 1,000–5,000 MUC5B polymers (34, 36). The mucin bundles must be organized by yet unidentified MUC5B intermolecular interactions and maybe also by additional molecules.

Observations of the mucin bundle movements on a tilted tracheal surface show that these are transported up-hill by the forces generated by the beating cilia driving ASL up-hill. However interestingly, the bundles did not move as expected parallel to the ASL movement, but rather perpendicular to this (Fig. 5b) (34, 36). The bundle transport velocity was also about 10-times slower than that of the ASL and not as the ASL moving evenly. Furthermore, the bundles has to be kept down on the tracheobronchial surface as they shall of course not fall off into the lumen. How efficient the bundles are cleaning the tracheobronchial surface and removing bacteria as is illustrated in Fig. 5b (34). In CF, the bundles are immobile from birth and the bacteria remain on the epithelial surface.

The mechanisms for holding down the bundles on the tracheobronchial surface and controlling their bundle movement is poorly understood today. However, the observation that the MUC5AC mucin is extending from the surface goblet cells and reach over and patchily attach and coat the MUC5B bundles suggest an interesting mechanism (34, 36, 51). In contrast to previous assumptions, gel-forming mucins seems to remain attached to the goblet cell during secretion despite their lack of transmembrane domains. This was first observed in the small and large intestine where the MUC2 is attached to the goblet cell (36, 68, 114). The molecular mechanisms for attaching the gel-forming mucins to the goblet cell is not understood, but the phenomena suggest attachment to unknown transmembrane proteins, likely present in the secretory granule membrane. From studies of the small intestine, it is known that protease activities are involved in the detachment as exemplified by the meprin β protease (114). Mice lacking this enzyme have in contrast to normal mice an attached small intestinal mucus. Observations of the bundle movement pattern, suggest that mechanical forces are involved in the perpendicular and irregular movement of the bundles. Maybe similar mechanisms as for the VWF are at place, where the blood flow can pull open an ADAM13 cleavage site in the von Willebrand A2 domain allowing a rupture of the VWF polymer. However, no VWA2 domain is present in the mucins (119, 137).

The normal organization of the mucus and mucins in the lungs differ considerably between species. Small animals, like mice, lack submucosal glands (except a few at the larynx) and have their tracheobronchial surface cleaned by small mucus clouds made by MUC5B from the surface goblet cells (127)(D. Fakih, A. Ermund & G.C. Hansson, unpublished manuscript). The reason for this is of course the lack of submucosal glands, something that is likely related to the smaller diameter of their bronchi. It is intuitively understandable

that small mucus clouds could be sufficient for sweeping and cleaning small airways, but that these should be insufficient in large airways like the centimeter wide human trachea. In human airways there are submucosal glands and mucus bundles down to about the 10th bronchial bifurcation. In the even smaller human airways the cleaning might be similar to what has been observed in the mouse. Irrespectively of bundle or cloud mucus moved by the cilia, these are highly efficient systems for keeping the lungs free from bacteria.

Upon infection or at chronic diseases like CF or COPD, the lungs can revert from cleaning to protection. In these cases, an attached mucus layer able to separate bacteria from the epithelial cells is generated (Fig. 5c) (39). The layer appear stratified and thus similar to the one that protect colon. The molecular mechanism for forming such a layer is likely the structures formed by the MUC5AC mucin, a mucin that is dramatically increased in these diseases (86, 123). The MUC5AC mucin forms a linear disulfide-bonded polymer, just like the MUC5B mucin, but in addition, the VWD3 domain dimer further interacts to form a stable dimer-dimer tetramer (S. Trillo-Muyo & G. C. Hansson, unpublished manuscript). Such polymers will by this organize net-like sheets that together with other proteins, also found in the normal colon, form a stratified mucus layer able to separate bacteria from the epithelial surface.

In chronic lung diseases like COPD and CF, the attached mucus layer is retained in the lungs (11). As a major function for mucus to bind bacteria also bacteria will be retained in the lungs. Certain bacteria, especially *Pseudomonas aeruginosa*, also have special preference for living in mucus. Retention of bacteria in the lungs will trigger inflammatory reactions, something that will eventually destroy the lung issue in these diseases.

THE DIGESTIVE SYSTEM

If the respiratory system is efficient in keeping the lungs relatively free from bacteria, the intestine is highly efficient in handling a high number of bacteria. For this, the mucins are of key importance. The main reason for the structural organization of the gel-forming mucins is likely the requirement for a digestive system for multicellular organisms. The intestine is supposed to degrade and absorb most types of food without degrading itself. This important feature is solved by limiting the host digestive system when it comes to carbohydrate degradation. The human digestive enzymes are only able to cleave the specific glycosidic linkages (Glc α 1–4Glc and Glc α 1–6Glc) found in starch and the ones in a few disaccharides (lactose, sucrose, maltose, isomaltose and trehalose). This allows the glycans on the mucins, both transmembrane and gel-forming, to be unaffected and able to protect the intestinal surface. The N- and C-termini of mucins are less glycosylated, but on the other hand rich in Cys (one every 7–10 amino acid). These mucin parts form highly compacted and stabilized structures where the pancreatic proteases cannot access their specific cleavage sequences. Together these features of MUC2, despite its more than 5,000 amino acids, means that this mucin remains intact after exposure to pancreatic proteases. Our commensal bacteria in the large intestine will on the other hand help us to degrade the MUC2 glycans and to recover its the energy content.

Stomach

The stomach is covered by an attached mucus layer built around the MUC5AC mucin where its capacity to form net-like polymers is important (37) (S. Trillo-Muyo & G. C. Hansson, unpublished manuscript). This mucus layer acts as a diffusion barrier for luminal hydrochloric acid that by this protect the surface epithelial cells (111). The stomach glands that produce the hydrochloric acid and pepsinogen are on the other hand protected in unknown ways. The MUC6 mucin is the only gel-forming mucin produced in these glands and might have an important protective role. The gland content is secreted out through the surface MUC5AC mucus layer covering the glands by opening of 'pores' by non-explored molecular mechanisms (64).

Small intestine

The small intestine has a non-attached mucus layer built around the MUC2 mucin (37, 105). The MUC2 mucin remains attached to the goblet cell after secretion and require the protease meprin β to cleave in the MUC2 to be detached (114). The detachment is well control as it is activated by the presence of bacteria and require an active bicarbonate secreting CFTR ion channel (50). The mucus is penetrable to bacteria and bacterial sized beads, but due to the abundant antibacterial peptides and proteins from especially the Paneth cells in the crypt bottom, there are no bacteria in contact with the cells (20, 98).

The enterocytes of the small intestine have the densest and thickest glycocalyx of any cell. In these, packed mucins and especially MUC17 create a barrier for bacteria, but not small digested nutrients. The thickness of the glycocalyx is similar to the length of the MUC17, MUC12, and MUC3 extended mucin domains. This group of mucins are anchored in the apical membrane, but to protect the membrane from rupture the cleaved SEA domain provide a mechanical breaking-point (96). As such mucins reach furthest away from the epithelial cells they may have yet undefined sensory functions as suggested for similar type of molecules in yeast (126).

The small intestinal mucins and their mucin domain glycan decoration is very important for the selection of commensal bacterial species. The importance of attached/detached mucus was illustrated in experiments in which germ-free animals were gavaged with normal mouse bacteria (67). The attached mucus layer in the germ-free animals retained and selected a different bacterial flora with an increased amount of Bacteroidetes and a concomitant decrease of Firmicutes (35). This altered composition was returned to normal once the mucus was normalized and detached 4–5 weeks after colonization. The CF patient problems with both intestinal obstruction and bacterial over-growth from the distal small intestine are likely related to their mucus being attached (26, 50, 60).

Large intestine

The MUC2 mucin builds the skeleton of colon mucus system with two major mucus layers; an outer non-attached less dense and an inner attached dense mucus layer (Fig, 6) (68). The inner mucus layer has a stratified (called s) outer part where the MUC2 is organized in flat net-like structures and a less well-organized inner part close to the epithelial cells where the mucin is expanding and organized (called a) (2, 68).

Commensal Bacteria.—The large intestine harbors an estimated 1,000 different bacterial species where the majority belong to the four phyla: Bacteroidetes, Proteobacteria, Firmicutes and Actinobacteria (76). The high number and their diversity has been estimated from sequencing the genomes encoding the 16S ribosomal RNA (81, 85). Concerted actions have more recently obtained full metagenome and transcriptome sequences of the commensal bacteria and the available genomic sequences for several thousand species (1, 40, 97, 122, 134). A major challenge is now to annotate these genomes and to reveal the functional importance of the encoded proteins. One of the earliest and most studied commensal bacteria is *Bacteroides thetaiotaomicron* that has, as many other commensal bacteria, devoted a large portion of its genome to carbohydrate degradation. In this and other bacteria, carbohydrate utilization is organized in gene assemblies called polysaccharide utilization loci (PUL) including sensors, transporters and glycosidases (76, 88). Other commensal bacteria have similar or fewer genes for carbohydrate utilization, some devoted to degrading complex plant saccharides and others more specialized to mucin glycan degradation. One of the latter ones is *Akkermansia mucinifilia*, an efficient mucin glycan degrader that can live on pig stomach mucins (27, 130). The mucin O-glycans are highly diverse and there are few bacteria that can live on these glycans by themselves. Instead, several bacteria typically need to collaborate to utilize all glycans. It should be emphasized that the mucin O-glycans vary enormously along the intestinal tract and between species, urging researchers in the area to take more care in studying bacteria and mucin from the same niche.

Outer Mucus Layer and Bacteria.—The outer mucus is formed from the inner mucus layer by detachment and volume expansion, processes that are controlled by the host (Fig. 6a) (67, 68). The two processes are molecularly not fully understood and likely not linked as the detachment generate a sharp border whereas the expansion is gradual. The slow expansion of the outer mucus layer is illustrated when bacteria are aggregated by the peptidoglycan-binding protein ZG16 and by this moved further away from the inner-outer interface into the outer mucus layer (8). Several proteases are orchestrating the processes generating the outer mucus layer of which at least one has been identified; the CLCA1 protein (89, 90). This abundant mucus protein with an N-terminal metalloprotease domain, a central von Willebrand A and a C-terminal fibronectin type III domain likely have both enzymatic and structural roles in the formation of the mucus. Comparing the MUC2 mucin of the inner and outer colon mucus layer suggest cleavages in the C-terminal part of MUC2 that does not necessarily cause depolymerization as this part is heavily cross-linked by numerous disulfide bonds (68). The important consequences of the formation of the outer mucus are that bacteria can enter and that this layer is now possible to move.

That the host selects its bacteria has been nicely illustrated by elegant experiments where zebra fishes and mice were made germ-free after which bacteria were given from one to the other host (99). The mice selected out the for the mouse normal bacteria from the mixture of zebra fish bacteria.

As bacteria do not enter the inner mucus layer and the host do not secrete mucin glycan glycosidases, the outer surface of the inner mucus layer expose the intact mucin domain glycans as biosynthesized by the host (Fig. 6a) (22, 70). This is likely of fundamental

importance for host microbial selection. Studies of bacterial adhesion to glycans has largely been addressing single carbohydrate epitope-protein interactions (61, 71). This is likely not the way commensal bacteria recognize the mucin domain glycans as these glycans are densely packed. Instead, the recognition will likely be based on glycan patches where the combined surface generated by several specifically arranged glycans of the mucin domains will be recognized. Today there is limited knowledge of the structure of these patches. Such patches will be defined by both the arrangement of glycan attachment sites and the glycan structures themselves. Potential bacterial adhesins recognizing mucin glycan patches has to be carefully studied using bacteria and mucin domains from the same host, intestinal part, and bacteria localized to the inner-outer mucus layer interface.

There are of course other factors than the host glycan repertoire that is important for bacterial selection. One example is the action and specificity of antibacterial defensins (110). Important is naturally also the type of food in the host's diet and its content of plant derived complex glycans.

Once bacterial glycosidases have started to act on the mucin glycans, the host specific glycan epitopes will disappear. The bacteria will continue to degrade the glycans, and finally the protein backbone will be exposed and degraded. Only small amounts of mucins appear in the feces showing that the commensal bacteria efficiently utilize the host mucins. The released monosaccharides are used the bacterial metabolism to generate short fatty acids (SCFA); acetate, proprionate and butyrate. SCFA are diffusing back through the inner mucus layer and will feed the local epithelial cells and to some extent the body (76).

Inner Mucus Layer and Bacteria.—The inner colon mucus layer is normally devoid of bacteria as it is impenetrable to bacteria and micrometer sized particles (67, 68). This filter-like effect is generated by the nets formed by the MUC2 mucin polymers staggered on top of each other (Fig. 6a) (2, 108). The theoretical pore-sizes generated from the MUC2 ring-like polymers is considerable larger than the excluded micrometer beads, suggesting that the individual MUC2 sheets are not staggered directly on top of each other. Recent studies of how the Salmonella bacteria swim in the colon mucus is highly illustrative (41). The flagellated bacteria swim freely in the outer mucus layer and on the surface of the inner mucus layer. Here it searches for breaches in the inner mucus layer as a way to access the epithelium and be able to invade the host. Interestingly, sometimes bacteria are trapped and stopped moving at the outer surface of the inner mucus layer, supporting the concept of a net-like inner mucus layer with pores smaller than the bacteria and in this way act like a fishnet.

Interestingly, germ-free mice have an inner mucus layer that is penetrable to bacterial sized beads (67). This mucus layer has an almost identical thickness to normal animals and only a slightly lower concentration of the MUC2 mucin. Colonization of the germ-free mice with a normal bacterial mixture regenerates the normal impenetrable inner mucus layer, a process that takes up to 7 weeks. Colonization of germ-free mice with bacteria from mice with a partially penetrable inner mucus layer recapitulated the abnormal properties of the mucus layer (67). Recent observations also show that 'western type' diets, high in fat and especially low in plant polysaccharides ('fiber'), within three days renders the inner mucus layer more

penetrable to bacteria that by this come closer to the host epithelium (28, 113). Together these observations suggest that bacteria can influence not only the amount of MUC2 mucin secreted, but also its organization once it has been secreted and expanded. As normal impenetrable mucus is present in animals with the bacteria kept at a distance, this suggest that products from bacteria can influence the formation of mucus. As the inner mucus layer provide a diffusion barrier, it is likely that bacterially generated small molecules will be responsible for such a communication. Understanding the organization of a functional inner mucus layer require detail molecular understanding of the different domains of the MUC2 mucin and the other major proteins of the colon mucus and their interaction with each other. Mucin secretion and its enormous volume expansion also points to the importance of the ionic conditions and pH at mucus secretion as these likely also modulate mucus properties.

The inner mucus layer normally keeps the commensal bacteria at bay, but when this **first** defense line fails bacteria come in contact with the epithelium. This is likely the first event in the development in the inflammatory disease **ulcerative colitis** (Fig. 6b) (65, 129). When bacteria reach the crypt opening, a **second** defense line is reached as the Sentinel Goblet Cell (**SenGC**) senses the increased level of bacterial products and initiate a response by a coordinated release of a mucus plume to wash away bacteria (Fig. 6b) (9). A **third** defense line is likely mediated by the emptying of crypt goblet cells as illustrated by the event at ischemia (Fig. 6b) (45). These defense lines are easily exhausted as the regeneration of new MUC2 mucins is time consuming as their biosynthesis is demanding and prone to cause ER stress responses. Once an overwhelming amount of bacteria reach the immune cells of the lamina propria, an overt inflammatory reaction is triggered that further drive mucus secretion and cause an acute episode of ulcerative colitis.

Understanding the molecular organization of the respiratory mucus systems and the colon mucus and its interaction with the commensal bacteria is of high priority for both science and society. To make significant progress, we have to take a humble and careful approach demanding the highest scientific standards to increase our understanding of these extremely complicated systems.

ACKNOWLEDGEMENTS

I am indebted to all present and former coworkers in the Mucin Biology Groups for all their work without which this review should never have been possible and to Dr. Fredrik H. Sterky for his non-expert input to the text.

This work was supported by the European Research Council ERC (694181), 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), The Knut and Alice Wallenberg Foundation (2017.0028), Swedish Research Council (2017–00958), The Swedish Cancer Foundation, IngaBritt and Arne Lundberg Foundation, Sahlgren's University Hospital (The ALF agreement 236501), Bill and Melinda Gates Foundation, Wilhelm and Martina Lundgren's Foundation.

TERMS AND DEFINITIONS

AMOP	adhesion-associated domain
ASL	air surface liquid
COPD	chronic obstructive pulmonary disease

CF	cystic fibrosis
ER	endoplasmic reticulum
Glycocalyx	the glycan-rich coating of cells. Intestinal enterocytes have the thickest and densest glycocalyx as apical coating of their brush-border
Goblet cell	group of specialized cell producing mucus and especially mucins
Mucin domain	long densely <i>O</i> -glycosylated domains with sequences rich in Pro, Thr and Ser (PTS-), often characterized by tandem repeats
NIDO	nidogen domain
<i>O</i>-glycans	glycans attached via GalNAc to the hydroxyl amino acids Ser or Thr
PTS	proline-threonine-serine domain
SEA	sea urchin-enterokinase-agrin
SCFA	short fatty acids
TM	transmembrane
TGN	trans-Golgi network
VNTR	variable number of tandem repeats
VWF	von Willebrand factor, an important factor of the blood coagulation system
VWC	von Willebrand C domain
VWD	von Willebrand D domain or assemblies. Assemblies refer to VWD domain and associated C8, TIL, and E domains

REFERENCES

1. Almeida A, Mitchell AL, Boland M, Forster SC, Gloor GB et al. 2019. A new genomic blueprint of the human gut microbiota *Nature* 568:499–504 [PubMed: 30745586]
2. Ambort D, Johansson MEV, Gustafsson JK, Nilsson H, Ermund A et al. 2012. Calcium and pH-dependent Packing and Release of the Gel-forming MUC2 Mucin *Proc. Natl. Acad. Sci. USA* 109:5645–50 [PubMed: 22451922]
3. Arike L, Hansson GC. 2016. The Densely *O*-glycosylated MUC2 Mucin Protects the Intestine and Provides Food for the Commensal Bacteria *J. Mol. Biol.* 428:3221–9 [PubMed: 26880333]
4. Asker N, Axelsson MAB, Olofsson SO, Hansson GC. 1998. Dimerization of the human MUC2 mucin in the endoplasmic reticulum is followed by a *N*-glycosylation-dependent transfer of the mono- and dimers to the Golgi apparatus *J. Biol. Chem.* 273:18857–63 [PubMed: 9668061]
5. Axelsson MAB, 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.* 273:18864–70 [PubMed: 9668062]

6. Beatson R, Tajadura-Ortega V, Achkova D, Picco G, Tsurouktsoglou TD et al. 2016. The mucin MUC1 modulates the tumor immunological microenvironment through engagement of the lectin Siglec-9 *Nature Immunol.* 17:1273–81 [PubMed: 27595232]
7. Bennett EP, Mandel U, Clausen H, Gerken TA, Fritz TA, Tabak LA. 2012. Control of mucin-type O-glycosylation: A classification of the polypeptide GalNAc-transferase gene family *Glycobiology* 22:736–56 [PubMed: 22183981]
8. Bergström JH, Birchenough GMH, Katona G, Schroeder BO, Schutte A et al. 2016. Gram-positive bacteria are held at a distance in the colon mucus by the lectin-like protein ZG16 *Proc. Natl. Acad. Sci. USA* 113:13833–8 [PubMed: 27849619]
9. Birchenough GMH, Nystrom ELN, Johansson MEV, Hansson GC. 2016. A sentinel goblet cell guards the colonic crypt by triggering Nlrp6-dependent Muc2 secretion *Science* 352:1535–42 [PubMed: 27339979]
10. Bobek LA, Liu JH, Sait SNJ, Shows TB, Bobek YA, Levine MJ. 1996. Structure and chromosomal localization of the human salivary mucin gene, MUC7 *Genomics* 31:277–82 [PubMed: 8838308]
11. Boucher RC. 2019. Muco-Obstructive Lung Diseases *New England Journal of Medicine* 380:1941–53
12. Breimer ME, Hansson GC, Karlsson KA, Leffler H. 1982. Glycosphingolipids of rat tissues. Different composition of epithelial and nonepithelial cells of small intestine *J. Biol. Chem.* 257:557–68 [PubMed: 7053385]
13. Brockhausen I, Stanley P. 2017. O-GalNAc Glycans. In *Essentials of Glycobiology*, ed. Varki A, Cummings RD, and Esko JD, NY: Cold Spring Harbor Laboratory Press
14. Bry L, Falk PG, Midtvedt T, Gordon JL. 1996. A model of host-microbial interactions in an open mammalian ecosystem *Science* 273:1380–3 [PubMed: 8703071]
15. Button B, Cai LH, Ehre C, Kesimer M, Hill DB et al. 2012. A Periciliary Brush Promotes the Lung Health by Separating the Mucus Layer from Airway Epithelia *Science* 337:937–41 [PubMed: 22923574]
16. 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.* 358:657–64 [PubMed: 11577689]
17. Carlstedt I, Herrmann A, Karlsson H, Sheehan JK, Fransson L, Hansson GC. 1993. Characterization of two different glycosylated domains from the insoluble mucin complex of rat small intestine *J. Biol. Chem.* 268:18771–81 [PubMed: 8360170]
18. Carlstedt I, Karlsson H, Sundler F, Fransson L. 1982. An insoluble mucin complex from rat small intestine *Adv. Exp. Med. Biol.* 144:155–7 [PubMed: 7044058]
19. Chen Y, Zhao YH, Kalasavadi TB, Hamati E, Nehrke K et al. 2004. Genome-Wide Search and Identification of a Novel Gel-Forming Mucin MUC19/Muc19 in Glandular Tissues *Am. J. Resp. Cell Mol. Biol.* 30:155–65
20. Clevers HC, Bevins CL. 2013. Paneth cells: maestros of the small intestinal crypts *Annu. Rev. Physiol.* 75:289–311 [PubMed: 23398152]
21. Cohen M, Varki A. 2014. Chapter Three - Modulation of Glycan Recognition by Clustered Saccharide Patches. In *International Review of Cell and Molecular Biology*, ed. Jeon KW, 75–125. Academic Press
22. Corfield A. 2018. The Interaction of the Gut Microbiota with the Mucus Barrier in Health and Disease in *Human Microorganisms* 6:78
23. Corfield AP. 2015. Mucins: A biologically relevant glycan barrier in mucosal protection *Biochim. Biophys. Acta* 1850:236–52 [PubMed: 24821013]
24. de las Rivas M, Lira-Navarrete E, Gerken TA, Hurtado-Guerrero R. 2019. Polypeptide GalNAc-Ts: from redundancy to specificity *Current Opinion Struct. Biol.* 56:87–96
25. de las Rivas M, Paul Daniel EJ, Coelho H, Lira-Navarrete E, Raich L et al. 2018. Structural and Mechanistic Insights into the Catalytic-Domain-Mediated Short-Range Glycosylation Preferences of GalNAc-T4 *ACS Cent. Sci.* 4:1274–90 [PubMed: 30276263]

26. De Lisle RC. 2007. Altered transit and bacterial overgrowth in the cystic fibrosis mouse small intestine *Am. J. Physiol. Gastroint. Liver Physiol.* 293:G104–G111
27. Derrien M, Vaughan EE FAU - Plugge C, Plugge CM FAU - de Vos W, de Vos WM. 2004. *Akkermansia muciniphila* gen. nov., sp. nov., a human intestinal mucin-degrading bacterium *Int. J. Syst. Evol. Biol.* 54:1469–76
28. Desai MS, Seekatz AM, Koropatkin NM, Kamada N, Hickey CA et al. 2016. A Dietary Fiber-Deprived Gut Microbiota Degrades the Colonic Mucus Barrier and Enhances Pathogen Susceptibility *Cell* 167:1339–53 [PubMed: 27863247]
29. Desseyn JL, Clavereau I, Laine A. 2002. Cloning, chromosomal localization and characterization of the murine mucin gene orthologous to human MUC4 *Eur. J. Biochem.* 269:3150–9 [PubMed: 12084055]
30. Desseyn JL, Guyonnet-Dep, rat V, Porchet N, Aubert JP, Laine A. 1997. Human mucin gene MUC5B, the 10.7-kb large central exon encodes various alternate subdomains resulting in a super-repeat *J. Biol. Chem.* 272:3168–78 [PubMed: 9013550]
31. Dickson RP, Erb-Downward JR, Martinez FJ, Huffnagle GB. 2016. The Microbiome and the Respiratory Tract *Ann. Rev. Physiol.* 78:481–504
32. Dong X, Leksa NC, Chhabra ES, Arndt JW, Lu Q et al. 2019. The von Willebrand factor D'D3 assembly and structural principles for factor VIII binding and concatemer biogenesis *Blood* 133:1523–33 [PubMed: 30642920]
33. Eckhardt AE, Timpl CS, DeLuca AW, Hill RL. 1997. The complete cDNA sequence and structural polymorphism of the polypeptide chain of porcine submaxillary mucin *J. Biol. Chem.* 272:33204–10 [PubMed: 9407109]
34. Ermund A, Meiss LN, Dolan B, Bahr A, Klymiuk N, Hansson GC. 2018. The mucin bundles responsible for airway cleaning are retained in cystic fibrosis and by cholinergic stimulation *Eur. Resp. J.* 52:1800457
35. Ermund A, Meiss LN, Gustafsson JK, Hansson GC. 2015. Hyper-osmolarity and calcium chelation: Effects on cystic fibrosis mucus *Eur. J. Pharmacol.* 764:109–17 [PubMed: 26134505]
36. Ermund A, Meiss LN, Rodríguez-Pineiro AM, Bähr A, Nilsson HE et al. 2017. The normal trachea is cleaned by MUC5B mucin bundles from the submucosal glands coated with the MUC5AC mucin *Biochem. Biophys. Res. Commun.* 492:331–7 [PubMed: 28859985]
37. Ermund A, Schutte A, Johansson MEV, Gustafsson JK, Hansson GC. 2013. Studies of mucus in mouse stomach, small intestine, and colon. I. Gastrointestinal mucus layers have different properties depending on location as well as over the Peyer's patches *Am. J. Physiol. Gastroint. Liver Physiol.* 305:G341–G347
38. Escande F, Porchet N, Aubert JP, Buisine MP. 2002. The mouse *Muc5b* mucin gene: cDNA and genomic structures, chromosomal localization and expression *Biochem J* 363:589–98 [PubMed: 11964160]
39. Fernández-Blanco JA, Fakih D, Arike L, Rodríguez-Pineiro A, Martínez-Abad Bet et al. 2018. Attached stratified mucus separates bacteria from the epithelial cells in COPD lungs *J. Clin. Invest. Insights* 3:e120994
40. Forster SC, Kumar N, Anonye BO, Almeida A, Viciani E et al. 2019. A human gut bacterial genome and culture collection for improved metagenomic analyses *Nature Biotechnology* 37:186–92
41. Furter M, Sellin ME, Hansson GC, Hardt WD. 2019. Mucus Architecture and Near-Surface Swimming Affect Distinct *Salmonella Typhimurium* Infection Patterns along the Murine Intestinal Tract *Cell Reports* 27:2665–78 [PubMed: 31141690]
42. Gagneux P, Varki A. 1999. Evolutionary considerations in relating oligosaccharide diversity to biological function *Glycobiology* 9:747–55 [PubMed: 10406840]
43. Gendler SJ, Lancaster CA, Taylor-Papadimitriou J, Duhig T, Peat Net et al. 1990. Molecular cloning and expression of human tumor-associated polymorphic epithelial mucin *J. Biol. Chem.* 265:15286–93 [PubMed: 1697589]
44. Gerken TA, Owens CL, Pasumarthy M. 1998. Site-specific core 1 O-glycosylation pattern of the porcine submaxillary gland mucin tandem repeat - Evidence for the modulation of glycan length by peptide sequence *J. Biol. Chem.* 273:26580–8 [PubMed: 9756896]

45. Grootjans J, Hundscheidt IH, Lenaerts K, Boonen B, Renes IB et al. 2013. Ischemia-induced mucus barrier loss and bacterial penetration are rapidly counteracted by increased goblet cell secretory activity in human and rat colon. *Gut* 62:250–8 [PubMed: 22637697]
46. Gum JR, Crawley SC, Hicks JW, Szymkowski DE, Kim YS. 2002. MUC17, a novel membrane-tethered mucin. *Biochem. Biophys. Res. Commun.* 291:466–75 [PubMed: 11855812]
47. Gum JR, Hicks JW, Lagace RE, Byrd JC, Toribara NW et al. 1991. Molecular cloning of rat intestinal mucin. Lack of conservation between mammalian species. *J. Biol. Chem.* 266:22733–8 [PubMed: 1939280]
48. Gum JR, Hicks JW, Toribara NW, Rothe EM, Lagace RE, Kim YS. 1992. The human MUC2 intestinal mucin has cysteine-rich subdomains located both upstream and downstream of its central repetitive region. *J. Biol. Chem.* 267:21375–83 [PubMed: 1400449]
49. Gum JR, Hicks JW, Toribara NW, Siddiki B, Kim YS. 1994. Molecular cloning of human intestinal mucin (MUC2) cDNA. Identification of the amino terminus and overall sequence similarity to prepro-von Willebrand factor. *J. Biol. Chem.* 269:2440–6 [PubMed: 8300571]
50. Gustafsson JK, Ermund A, Ambort D, Johansson MEV, Nilsson HE et al. 2012. Bicarbonate and functional CFTR channel is required for proper mucin secretion and link Cystic Fibrosis with its mucus phenotype. *J. Exp. Med.* 209:1263–72 [PubMed: 22711878]
51. Hansson GC. 2019. Mucus and mucins in diseases of the intestinal and respiratory tracts. *J. Internal Medicine* 285:479–90
52. Herrmann A, Davies JR, Lindell G, Martensson S, Packer NH et al. 1999. Studies on the “Insoluble” glycoprotein complex from human colon. *J. Biol. Chem.* 274:15828–36 [PubMed: 10336486]
53. Higuchi T, Orita T, Nakanishi S, Katsuya K, Watanabe H et al. 2004. Molecular cloning, genomic structure, and expression analysis of MUC20, a novel mucin protein, up-regulated in injured kidney. *J. Biol. Chem.* 279:1968–79 [PubMed: 14565953]
54. Hijikata M, Matsushita I, Tanaka G, Tsuchiya T, Ito H et al. 2011. Molecular cloning of two novel mucin-like genes in the disease-susceptibility locus for diffuse panbronchiolitis. *Human Genetics* 129:117–28 [PubMed: 20981447]
55. Hoegger MJ, Fischer AJ, McMenimen JD, Ostedgaard LS, Tucker AJ et al. 2014. Impaired mucus detachment disrupts mucociliary transport in a piglet model of cystic fibrosis. *Science* 345:818–22 [PubMed: 25124441]
56. Holmén Larsson JM, Karlsson H, Gråberg Crespo J, Johansson MEV, Eklund L et al. 2011. An Altered O-glycosylation Profile of the MUC2 Mucin Occurs in Active Ulcerative Colitis and is Associated with Increased Inflammation. *Infl. Bowel Dis.* 17:2299–307
57. Holmen Larsson JM, Karlsson H, Sjövall H, Hansson GC. 2009. A complex, but uniform O-glycosylation of the human MUC2 mucin from colonic biopsies analyzed by nanoLC/MSn. *Glycobiology* 19:756–66 [PubMed: 19321523]
58. 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 [PubMed: 12652082]
59. Holmen-Larsson JM, Thomsson KA, Rodriguez-Pineiro AM, Karlsson H, Hansson GC. 2013. Studies of mucus in mouse stomach, small intestine, and colon. III. Gastrointestinal Muc5ac and Muc2 mucin O-glycan patterns reveal a regiospecific distribution. *Am. J. Physiol. Gastroint. Liver Physiol.* 305:G357–G363
60. Houwen RH, van der Doef HP, Sermet I, Munck A, Hauser B et al. 2010. Defining DIOS and Constipation in Cystic Fibrosis With a Multicentre Study on the Incidence, Characteristics, and Treatment of DIOS. *J. Ped. Gastroent. Nutrition* 50:38–42
61. Hultgren SJ, Abraham S, Caparon M, Falk P, Stgeme JW, Normark S. 1993. Pilus and Nonpilus Bacterial Adhesins - Assembly and Function in Cell Recognition. *Cell* 73:887–901 [PubMed: 8098994]
62. Javitt G, Calvo M, LGM, Albert L, Reznik N, Ilani T et al. 2019. Intestinal Gel-Forming Mucins Polymerize by Disulfide-Mediated Dimerization of D3 Domains. *J. Mol. Biol.*
63. Jin C, Kenny DT, Skoog EC, Padra M, Adamczyk B et al. 2017. Structural Diversity of Human Gastric Mucin Glycans. *Mol. Cell Proteomics* 16:743–58 [PubMed: 28461410]

64. Johansson M, Synnerstad I, Holm L. 2000. Acid transport through channels in the mucous layer of rat stomach *Gastroenterology* 119:1297–304 [PubMed: 11054387]
65. Johansson MEV, Gustafsson JK, Holmen-Larsson J, Jabbar KS, Xia L et al. 2014. Bacteria penetrate the normally impenetrable inner colon mucus layer in both murine colitis models and in patients with ulcerative colitis *Gut* 213:281–91
66. Johansson MEV, Hansson GC. 2016. Immunological aspects of intestinal mucus and mucins *Nature Rev. Immunol.* 16:639–49 [PubMed: 27498766]
67. Johansson MEV, Jacobsson HE, Holmén-Larsson J, Schütte A, Ermund A et al. 2015. Normalization of the host intestinal mucus systems requires long-term colonization *Cell Host Microbe* 18:582–92 [PubMed: 26526499]
68. Johansson MEV, Phillipson M, Petersson J, Holm L, Velcich A, Hansson GC. 2008. The inner of the two Muc2 mucin dependent mucus layers in colon is devoid of bacteria. *Proc. Natl. Acad. Sci. USA* 105:15064–9 [PubMed: 18806221]
69. Johansson MEV, Sjövall H, Hansson GC. 2013. The gastrointestinal mucus system in health and disease *Nature Rev. Gastroenterol. Hepatol.* 10:352–61 [PubMed: 23478383]
70. Juge N. 2019. Special Issue: Gut Bacteria-Mucus Interaction *Microorganisms* 7:6
71. Karlsson KA. 1989. Animal glycosphingolipids as membrane attachment sites for bacteria *Ann. Rev. Biochem.* 58:309–50 [PubMed: 2673013]
72. Karlsson NG, Herrmann A, Karlsson H, Johansson MEV, Carlstedt I, Hansson GC. 1997. The glycosylation of rat intestinal Muc2 mucin varies between rat strains and the small and large intestine. A study of O-linked oligosaccharides by a mass spectrometric approach *J. Biol. Chem.* 272:27025–34 [PubMed: 9341141]
73. Kesimer M, Makhov AM, Griffith JD, Verdugo P, Sheehan JK. 2010. Unpacking a gel-forming mucin: a view of MUC5B organization after granular release *Am. J. Physiol.* 298:L15–L22
74. Klomp LWJ, Vanrens L, Strous GJ. 1995. Cloning and analysis of human gastric mucin cDNA reveals two types of conserved cysteine-rich domains *Biochem. J.* 308:831–8 [PubMed: 8948439]
75. Knowles MR, Boucher RC. 2002. Mucus clearance as a primary innate defense mechanism for mammalian airways *J. Clin. Invest.* 109:571–7 [PubMed: 11877463]
76. Koropatkin NM, Cameron EA, Martens EC. 2012. How glycan metabolism shapes the human gut microbiota *Nature Rev. Microbiol.* 10:323–35 [PubMed: 22491358]
77. Lancaster CA, Peat N, Duhig T, Wilson D, Taylor-Papadimitriou J, Gendler SJ. 1990. Structure and expression of the human polymorphic epithelial mucin gene: an expressed VNTR unit *Biochem. Biophys. Res. Commun.* 173:1019–29 [PubMed: 2268309]
78. Lang T, Hansson GC, Samuelsson T. 2007. Gel-forming mucins appeared early in metazoan evolution. *Proc. Natl. Acad. Sci. USA* 104:16209–14 [PubMed: 17911254]
79. Lang T, Klasson S, Larsson E, Johansson MEV, Hansson GC, Samuelsson T. 2016. Searching the Evolutionary Origin of Epithelial Mucus Protein Components *FCöMucins and FCGBPMol. Biol. Evolution* 33:1921–36
80. Lesch C, Goto A, Lindgren M, Bidla G, Dushay MS, Theopold U. 2007. A role for Hemolectin in coagulation and immunity in *Drosophila melanogaster* *Developmental Comp. Immunol.* 31:1255–63
81. Ley RE, Backhed F, Turnbaugh P, Lozupone CA, Knight RD, Gordon JI. 2005. Obesity alters gut microbial ecology *Proc. Natl. Acad. Sci. U. S. A.* 102:11070–5 [PubMed: 16033867]
82. Li DZ, Gallup M, Fan N, Szymkowski DE, Basbaum CB. 1998. Cloning of the amino-terminal and 5'-flanking region of the human MUC5AC mucin gene and transcriptional up-regulation by bacterial exoproducts *J. Biol. Chem.* 273:6812–20 [PubMed: 9506983]
83. Linden SK, Sutton P, Karlsson NG, Korolik V, McGuckin MA. 2008. Mucins in the mucosal barrier to infection *Mucosal Immunol.* 1:183–97 [PubMed: 19079178]
84. Lorand L, Graham RM. 2003. Transglutaminases: crosslinking enzymes with pleiotropic functions *Nature Reviews Molecular Cell Biology* 4:140–56 [PubMed: 12563291]
85. Lozupone C, Knight R. 2005. UniFrac: a new phylogenetic method for comparing microbial communities *Appl Environ. Microbiol.* 71:8228–35 [PubMed: 16332807]

86. Ma J, Rubin BK, Voynow JA. 2018. Mucins, Mucus, and Goblet Cells *Chest* 154:169–76 [PubMed: 29170036]
87. Macao B, Johansson DGA, Hansson GC, Härd T. 2006. Auto-proteolysis coupled to protein folding in the SEA domain of the membrane-bound MUC1 mucin *Nature Struct. Mol. Biol.* 13:71–6 [PubMed: 16369486]
88. Martens EC, Chiang HC, Gordon JI. 2008. Mucosal Glycan Foraging Enhances Fitness and Transmission of a Saccharolytic Human Gut Bacterial Symbiont *Cell Host & Microbe* 4:447–57 [PubMed: 18996345]
89. Nystrom EEL, Arike L, Ehrencrona E, Hansson GC, Johansson MEV. 2019. CLCA1 forms non-covalent oligomers in colonic mucus and has MUC2-processing properties *J. Biol. Chem.*
90. Nystrom EEL, Birchenough GMH, van der Post S, Arike L, Gruber AD et al. 2018. Calcium-activated Chloride Channel Regulator 1 (CLCA1) Controls Mucus Expansion in Colon by Proteolytic Activity *EBioMed* 33:134–43
91. O'Brien TJ, Beard JB, Underwood LJ, Shigemasa K. 2004. The CA125 gene: a newly discovered extension of the glycosylated N-terminal domain doubles the size of this extracellular superstructure *Tumor Biol* 23:154–69
92. O'Leary JM, Hamilton JM, Deane CM, Valeyev NV, Sandell LJ, Downing AK. 2004. Solution Structure and Dynamics of a Prototypical Chordin-like Cysteine-rich Repeat (von Willebrand Factor Type C Module) from Collagen IIAJ. *Biol. Chem.* 279:53857–66 [PubMed: 15466413]
93. Olson FJ, Johansson MEV, Klinga-Levan K, Bouhours D, Enerback L et al. 2002. Blood group A glycosyltransferase occurring as alleles with high sequence difference is transiently induced during *Nippostrongylus brasiliensis* parasite infection *J. Biol. Chem.* 277:15044–52 [PubMed: 11842091]
94. Pallesen LT, Berglund L, Rasmussen LK, Petersen TE, Rasmussen JT. 2002. Isolation and characterization of MUC15, a novel cell membrane-associated mucin *Eur. J. Biochem.* 269:2755–63 [PubMed: 12047385]
95. Pelaseyed T, Gustafsson JK, Gustafsson IJ, Ermund A, Hansson GC. 2013. Carbachol-induced MUC17 endocytosis is concomitant with NHE3 internalization and CFTR membrane recruitment in enterocytes *Am. J. Physiol. Cell Physiol.* 305:C457–C467 [PubMed: 23784542]
96. Pelaseyed T, Zäch M, Petersson AC, Svensson F, Johansson DG, Hansson GC. 2013. Unfolding dynamics of the mucin SEA domain probed by force spectroscopy suggest that it acts as a cell protective device *FEBS J.* 280:1491–501 [PubMed: 23331320]
97. Proctor LM, Creasy HH, Fettweis JM, Lloyd-Price J, Mahurkar A et al. 2019. The Integrative Human Microbiome Project *Nature* 569:641–8 [PubMed: 31142853]
98. Propheter DC, Chara AL, Harris TA, Ruhn KA, Hooper LV. 2017. Resistin-like molecule +1 is a bactericidal protein that promotes spatial segregation of the microbiota and the colonic epithelium *Proc. Natl. Acad. Sci. USA* 114:11027–33 [PubMed: 28973871]
99. 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* 127:423–33 [PubMed: 17055441]
100. Recktenwald CV, Hansson GC. 2016. The reduction-insensitive bonds of the MUC2 mucin are isopeptide bonds *J. Biol. Chem.* 291:13580–90 [PubMed: 27129250]
101. Revoredo L, Wang S, Bennett EP, Clausen H, Moremen KW et al. 2016. Mucin-type O-glycosylation is controlled by short- and long-range glycopeptide substrate recognition that varies among members of the polypeptide GalNAc transferase family *Glycobiology* 26:360–76 [PubMed: 26610890]
102. Ridley C, Kouvatso N, Raynal BD, Howard M, Collins RF et al. 2014. Assembly of the Respiratory Mucin MUC5B: a new model for gel-forming mucin *J. Biol. Chem.* 289:16409–20 [PubMed: 24778189]
103. Robbe C, Capon C, Coddeville B, Michalski JC. 2004. Structural diversity and specific distribution of O-glycans in normal human mucins along the intestinal tract *Biochem. J.* 384:307–16 [PubMed: 15361072]

104. Robbe C, Capon C, Maes E, Rousset M, Zweibaum A et al. 2003. Evidence of Region-specific Glycosylation in Human Intestinal Mucins: Presence of an acidic gradient along the gastrointestinal tract. *J. Biol. Chem.* 278:46337–48 [PubMed: 12952970]
105. Rodriguez-Pineiro AM, Bergstrom JH, Ermund A, Gustafsson JK, Schutte A et al. 2013. Studies of mucus in mouse stomach, small intestine, and colon. II. Gastrointestinal mucus proteome reveals Muc2 and Muc5ac accompanied by a set of core proteins. *Am. J. Physiol. Gastroint. Liver Physiol.* 305:G348–G356
106. Rossez Y, Maes E, Lefebvre DT, Gosset P, Ecobichon C et al. 2012. Almost all human gastric mucin O-glycans harbor blood group A, B or H antigens and are potential binding sites for *Helicobacter pylori*. *Glycobiology* 22:1193–206 [PubMed: 22522599]
107. Rossi EA, McNeer RR, Price-Schiavi SA, van den Brande JMH, Komatsu M et al. 1996. Sialomucin complex, a heterodimeric glycoprotein complex. *J. Biol. Chem.* 271:33476–85 [PubMed: 8969211]
108. Round AN, Rigby NM, Garcia de la Torre A, Macierzanka A, Mills ENC, Mackie AR. 2012. Lamellar Structures of MUC2-Rich Mucin: A Potential Role in Governing the Barrier and Lubricating Functions of Intestinal Mucus. *Biomacromolecules* 13:3253–61 [PubMed: 22978827]
109. Rousseau K, Kirkham S, Johnson L, Fitzpatrick B, Howard M et al. 2008. Proteomic analysis of polymeric salivary mucins: no evidence for MUC19 in human saliva. *Biochem. J.* 413:545–52 [PubMed: 18426393]
110. Salzman NH, Ghosh D, Huttner KM, Paterson Y, Bevins CL. 2003. Protection against enteric salmonellosis in transgenic mice expressing a human intestinal defensin. *Nature* 422:522–6 [PubMed: 12660734]
111. Schade C, Flemstrom G, Holm L. 1994. Hydrogen ion concentration in the mucus layer on top of acid-stimulated and -inhibited rat gastric mucosa. *Gastroenterology* 107:180–8 [PubMed: 8020660]
112. Schauer R. 2009. Sialic acids as regulators of molecular and cellular interactions. *Current Opinion Struct. Biol.* 19:507–14
113. Schroeder BO, Birchenough GMH, Stahlman M, Arike L, Johansson MEV et al. 2018. Bifidobacterium or fiber protect against diet-induced deterioration of the inner colonic mucus layer. *Cell Host & Microbe* 23:27–40 [PubMed: 29276171]
114. Schütte A, Ermund A, Becker-Pauly C, Johansson MEV, Rodriguez-Pineiro AM et al. 2014. Microbial Induced Mepripin beta Cleavage in MUC2 Mucin and Functional CFTR Channel are Required to Release Anchored Small Intestinal Mucus. *Proc. Natl. Acad. Sci. USA* 111:12396–401 [PubMed: 25114233]
115. Sheehan JK, Oates K, Carlstedt I. 1986. Electron microscopy of cervical, gastric and bronchial mucus glycoproteins. *Biochem. J.* 239:147–53 [PubMed: 3800974]
116. Shogren R, Gerken TA, Jentoft N. 1989. Role of glycosylation on the conformation and chain dimensions of o-linked glycoproteins: Light-scattering studies of ovine submaxillary mucin. *Biochemistry* 28:5525–36 [PubMed: 2775721]
117. Silverman HS, Sutton-Smith M, McDermott K, Heal P, Leir SH et al. 2003. The contribution of tandem repeat number to the O-glycosylation of mucins. *Glycobiology* 13:265–77 [PubMed: 12626424]
118. Singh AP, Moniaux N, Chauhan SC, Meza JL, Batra SK. 2004. Inhibition of MUC4 expression suppresses pancreatic tumor cell growth and metastasis. *Cancer Res.* 64:622–30 [PubMed: 14744777]
119. Springer TA. 2014. von Willebrand factor, Jedi knight of the bloodstream. *Blood* 124:1412–25 [PubMed: 24928861]
120. Steentoft C, Vakhrushev SY, Joshi HJ, Kong Y, Vester-Christensen MB et al. 2013. Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology. *EMBO J.* 32:1478–88 [PubMed: 23584533]
121. Svensson F, Lang T, Johansson MEV, Hansson GC. 2018. The central exons of the human MUC2 and MUC6 mucins are highly repetitive and variable in sequence between individuals. *Scientific Reports* 8:17503 [PubMed: 30504806]
122. The human microbiome project consortium. 2012. Structure, function and diversity of the healthy human microbiome. *Nature* 486:207–14 [PubMed: 22699609]

123. Thornton DJ, Rousseau K, McGuckin MA. 2008. Structure and Function of the Polymeric Mucins in Airways Mucus *Ann. Rev. Physiol.* 70:459–86 [PubMed: 17850213]
124. Toribara NW, Robertson AM, HO SB, Kuo WL, Gum Eet al. 1993. Human gastric mucin *J. Biol. Chem.* 268:5879–85 [PubMed: 7680650]
125. Trillo-Muyo S, Nilsson HE, Recktenwald CV, Ermund A, Ridley Cet al. 2018. Granule-stored MUC5B mucins are packed by the non-covalent formation of N-terminal head-to-head tetramers *J. Biol. Chem.* 293:5746–54 [PubMed: 29440393]
126. Vadaie N, Dionne H, Akajagbor DS, Nickerson SR, Krysan DJ, Cullen PJ. 2008. Cleavage of the signaling mucin Msb2 by the aspartyl protease Yps1 is required for MAPK activation in yeast *J. Cell Biol.* 181:1073–81 [PubMed: 18591427]
127. van As A. 1977. Pulmonary Airway Clearance Mechanisms: A Reappraisal *Am. Rev. Resp. Dis.* 115:721–6 [PubMed: 855956]
128. van der Post S, Hansson GC. 2014. Membrane protein profiling of human colon reveals distinct regional differences *Mol. Cell. Proteomics* 13:2277–87 [PubMed: 24889196]
129. van der Post S, Jabbar KS, Birchenough GMH, Arike L, Akhtar Net al. 2019. Structural weakening of the colonic mucus barrier is an early event in ulcerative colitis pathogenesis *Gut* doi:10.1136/gutjnl-2018-317571:
130. van Passel MWJ, Kant R, Zoetendal EG, Plugge CM, Derrien Met al. 2011. The Genome of *Akkermansia muciniphila*, a Dedicated Intestinal Mucin Degradar, and Its Use in Exploring Intestinal Metagenomes *PLoS ONE* 6:e16876 [PubMed: 21390229]
131. Varki A. 1994. Selectin ligands *Proc. Natl. Acad. Sci. USA* 91:7390–7 [PubMed: 7519775]
132. Williams SJ. 1999. The novel mucin genes down-regulated in colorectal cancer identified by differential display *Cancer Res.* 59:4083–9 [PubMed: 10463611]
133. Williams SJ, Wreschner DH, Tran M, Eyre HJ, Sutherland GR, McGuckin MA. 2001. MUC13, a novel human cell surface mucin expressed by epithelial and hemopoietic cells *J. Biol. Chem.* 276:18327–36 [PubMed: 11278439]
134. Xiao L, Feng Q, Liang S, Sonne SB, Xia Zet al. 2015. A catalog of the mouse gut metagenome *Nature Biotechnol.* 33:1103–8 [PubMed: 26414350]
135. Yin BWT, Dnistrian A, Lloyd KO. 2002. Ovarian cancer antigen CA125 is encoded by the MUC16 mucin gene *Int. J. Cancer* 98:737–40 [PubMed: 11920644]
136. Zauner G, Kozak RP, Gardner RA, Fernandes DL, Deelder AM, Wuhrer M. 2012. Protein O-glycosylation analysis *Biol. Chem.* 393:687–708 [PubMed: 22944673]
137. Zhang X, Halvorsen K, Zhang CZ, Wong WP, Springer TA. 2009. Mechanoenzymatic Cleavage of the Ultralarge Vascular Protein von Willebrand Factor *Science* 324:1330–4 [PubMed: 19498171]

INTESTINAL GLYCAN DIFFERENCES BETWEEN MOUSE, RAT AND HUMAN

The normal rodent small intestinal mucins show, similarly to humans, glycans that differ from the colon (59, 72). In mouse, the Fut2 enzyme adding fucose is constitutively present in colon, but inducible in the small intestine by bacteria or parasites (14, 58). Interestingly, in rats the Fut2 enzyme is constitutively present and instead a blood group A-type GalNAc α 1–3 transferase is transiently induced upon parasite infection (93).

Human intestinal mucin glycans are characterized by having Core3, to a lesser extent Core4 (Fig. 2A) and are essentially lacking Core1. Mouse on the other hand have predominantly Core1-based components, although they make minor amounts of Core3 and Core4 (59). A major difference is that humans have an active sialyl transferase adding NeuAc α 2–6 to the peptide-bound GalNAc, an enzyme not active in mouse (Fig. 2B). On the contrary, both mouse and humans have the Sd^a-Cad blood group-like epitope (NeuAc α 2–3(GalNAc β 1–4)Gal β 1–3/4GlcNAc) as a major capping of their glycans in the large intestine (16, 59, 104).

SUMMARY POINTS

1. Mucins are characterized by long, extended rod-like mucin domains generated by dense *O*-glycosylation.
2. Intestine and lungs are protected from bacteria by transmembrane mucin generating an epithelial cell apical glycocalyx and by gel-forming mucins forming mucus of different shapes.
3. Mucus can form a protective layer that physically separate bacteria from the surface epithelial cells in the large intestine and diseased lung.
4. Mucins normally anchored to goblet cells can be detached in controlled ways and by this trap and remove bacteria.
5. The mucin *O*-glycans of the colon mucus layers act as attachment sites for commensal bacteria and as nutrient source for bacteria and host.

FUTURE ASPECTS

It is very important to:

1. To reveal the molecular structure of the different domains found in mucins and how these mediate interaction between mucins and other mucus proteins to generate highly organized mucus.
2. To understand how the mucins are unfolded upon secretion and how different ions and fluid conditions give different mucus properties.
3. To understand how the gel-forming mucins are detached from their attachment to the goblet cells in the intestine and lung.
4. To understand the glycan-generated surface of mucin domains and how these interact with bacteria.

5. To understand how different bacteria cooperate and utilize the mucin glycans.
6. To understand how bacterial metabolites communicate with and influence the host epithelium and its mucus production.
7. To develop novel therapeutic approaches and agents that can enhance the protective capacity of the colon mucus for the treatment of ulcerative colitis.
8. To develop novel therapeutic approaches and agents that can detach the attached mucus at CF and COPD.

Figure 1.
Gel-forming and transmembrane mucins have mucin domains encoded by PTS sequences.

Figure 2.

Mucin O-glycosylation. **a)** Major O-glycan core structures. **b)** Chromatograms and major structure of major O-glycans of mouse and human MUC2 mucin.

Figure 3.
Domains of the four human gel-forming mucins.

Figure 4.
Evolution of mucins and their major domains

Figure 5.

The mucus system of the normal and diseased respiratory tract. **a)** The submucosal gland is a molecular machine for the generation of mucin MUC5B based bundles. **b)** The bundles are transported on the tracheobronchial surface by the beating cilia and moving ASL cephalically. Bacteria are collected by the bundles sweeping over the surface. **c)** An attached mucus layer is formed by the MUC5AC and MUC5B mucins in diseased lungs and by this help to protect the epithelial cells from bacteria.

Figure 6.

The mucus system and bacteria of the normal colon. **a)** Goblet cells are assembling MUC2 into large polymers that upon secretion generate large net-like structures that are staggered on below each other to form the attached inner mucus layer. The pore sizes of this inner layer is small and consequently keep bacteria away from the epithelial cells. At the interphase to the outer mucus layer, the pore sizes are increased allowing bacteria to enter and degrade the mucin glycan to generate short fatty acids (SCFA) feeding the epithelial cells. **b)** Ulcerative colitis is initiated when the inner mucus layer (first defense line, 1), the sentinel Goblet Cell (second defense line, 2), and a potential third defense line (3) is failing and bacteria reach the subepithelial immune system.

Table 1.

The mucins in humans

HUMAN			Number of	Typical	References
MUCINS^a	Domain Structure^b	Chromosome	amino acids	localization	
TM -Mucins					
SEA					
MUC1	PTS-SEA-TM-CT	1q22	1255	General	(43)
MUC3	PTS-SEA-TM-CT	7q22.1	>2541	Intestine	(47)
MUC12	PTS-SEA-TM-CT	7q22.1	5478	Intestine	(132)
MUC13	PTS-SEA-TM-CT	3q21.2	512	Intestine	(133)
MUC16	(PTS-SEA) ₃₃ -TM-CT	19p13.2	22,152	General	(53)
MUC17	PTS-SEA-TM-CT	7q22.1	4493	Intestine	(46)
NIDO-AMOP-VWD					
MUC4	PTS-NIDO-AMOP-VWD-TM-CT	3q29	5284	General	(107)
Others					
MUC15	PTS-TM-CT	11p14.2	334	General	(94)
MUC21	PTS-TM-CT	6p21.33	566	Esophagus	(53)
MUC22	PTS-TM-CT	6p21.33	1733	Esophagus	(54)
Secreted Mucins					
Gel-forming					
MUC2	VWD1-VWD2-VWD3-PTS(CysD)-VWD4-CT	11p15	5130	Intestine	(49)
MUC5AC	VWD1-VWD2-VWD3-PTS(CysD)-VWD4-CT	11p15	5654	Lung, stomach	(74)
MUC5B	VWD1-VWD2-VWD3-PTS(CysD)-VWD4-CT	11p15	5703	Lung, saliva	(30)
MUC6	VWD1-VWD2-VWD3-PTS-CT	11p15	5534	Stomach	(124)
Monomeric					
MUC7	PTS	4q13.3	377	Saliva	(10)
MUC20	PTS	3q29	709	Kidney-urinary	(53)

^a All the numbers are not used: MUC8, not a mucin; MUC9, renamed to OVIOGP1; MUC10, not found in humans; MUC11, part of MUC12; MUC14, called ECM; MUC18, not used; MUC19, not expressed in humans.

^b PTS, proline-threonine-serine domain; SEA, sea urchin-enterokinase-agrin; TM, transmembrane; CT, cytoplasmic tail; VWD, von Willebrand D domain; NIDO, nidogen domain; AMOP, adhesion-associated domain