ANNUAL REVIEWS

Annual Review of Biochemistry

Mucins and the Microbiome

Gunnar C. Hansson

Department of Medical Biochemistry, University of Gothenburg, SE 405 30 Gothenburg, Sweden; email: gunnar.hansson@medkem.gu.se

Annu. Rev. Biochem. 2020. 89:769–93

First published as a Review in Advance on April 3, 2020

The *Annual Review of Biochemistry* is online at biochem.annualreviews.org

https://doi.org/10.1146/annurev-biochem-011520-105053

Keywords

mucus, carbohydrate, glycan, saccharide, intestine, bacteria

Abstract

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

Contents

INTRODUCTION

Maintaining the integrity of a living organism is paramount and requires clear separation between it and other organisms. The most efficient protection is an inert coat in the form of a shell or, as for higher animals, a skin composed of dead cells. However, this principle cannot be utilized for organs that actively interact with their surroundings. Thus, such organs require protection that is dynamic but still keeps 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 of specialized goblet cells and by a surface coating consisting of a dense glycocalyx (1–3). Glycoproteins called mucins are the main building blocks of both mucus and glycocalyx (4). The high glycan content of mucins makes them water soluble and able to bind large amounts of water, reflected in the transparent nature of normal mucus and glycocalyx. Mucins' protein core is rendered inaccessible by the high number of attached oligosaccharides, which are important for protecting mucosal surfaces. For example, the gastrointestinal tract is ideal for bacterial proliferation, given its balanced salt and water content, high nutrient content, and ideal temperature. The number of intestinal bacteria is close to the number of cells in the human body, yet the bacteria do not take over—largely because of the mucins and their structure and function. However, this is not the only remarkable feature of mucins. Certain commensal bacteria specialize in degrading the mucin glycans and utilize the energy thus obtained to feed not only themselves but also their host. Some of the mysteries surrounding the mucins and their relation to bacteria are starting to be revealed, but much more has yet to be learned. Below, I discuss our current understanding of mucins.

Goblet cells: specialized cells that produce mucus and especially mucins

Glycocalyx: the glycan-rich coating of cells; intestinal enterocytes have the thickest and densest glycocalyx as the apical coating of their brush border

***O*-glycans:** glycans attached via GalNAc to the hydroxyl amino acid Ser or Thr

MUCINS

Mucins are characterized by their dense coat of glycans, especially those of the *O*-glycan family. A mucin was originally defined as a glycoprotein with more than 50% of its mass attributable to *O*-glycans, and the name mucin was used for the major components of mucus. Today, these mucins are known as the gel-forming mucins. During the early era of gene cloning, in the late 1980s, the

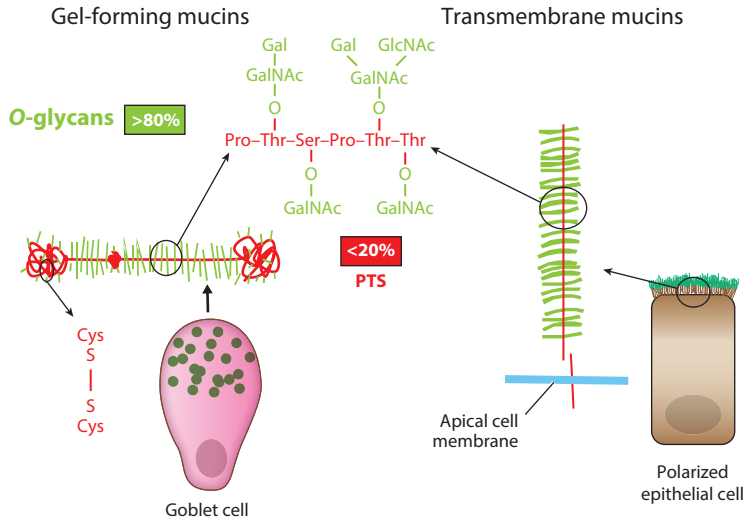

Figure 1

Gel-forming and transmembrane mucins have mucin domains encoded by PTS sequences. Abbreviations: GalNAc, *N*-acetylgalactosamine; GlcNAc, *N*-acetylglucosamine; PTS, Pro–Thr–Ser.

first mucin to be sequenced was mucin-1, with the abbreviated gene name *MUC1* (5). *MUC1* is a transmembrane mucin and, as it later turned out, very different from the classical gel-forming mucins, of which *MUC2* was the first to be cloned (6) (**Figure 1**). Because mucins were named according to their order of discovery, the nomenclature can be confusing.

Mucin Domains

Sequencing of mucin genes revealed long stretches of often repeated sequences that vary in length. Such sequences are generally referred to as VNTRs (variable number of tandem repeats); thus, these parts of the mucins were originally named VNTRs (7, 8). The human genome contains numerous such VNTR sequences, but only a few are in exons, as for mucins. Analyses of a mucin and its VNTR sequences show that the repetitive 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 showed that they are characterized by a high frequency of the hydroxy-amino acids Ser and Thr, along with Pro (**Figure 1**). As a result, we introduced the name PTS (Pro–Thr–Ser) for such sequences and predicted domains (9). The hydroxyl groups are used to attach the first *O*-glycan sugar, the *N*-acetylgalactosamine (GalNAc) residue. Pro ensures a nonfolded structure, which allows the peptidyl-GalNAc transferases of the Golgi apparatus to access the protein core (10). One can predict PTS sequences by mining genomes for long sequences rich in Ser/Thr (>25%) and Pro (>5%) in proteins with signal sequences (9).

Upon exiting the endoplasmic reticulum (ER), PTS sequences have a nonstructured, random coil nature. Once in the Golgi apparatus, 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. In this way, the PTS sequence becomes *O*-glycosylated, and a mucin domain is formed. The dense glycosylation pulls out the protein core, and long, extended rodlike structures form. When the physical length of the mucin domain has been possible to measure and when compared with the actual amino acid sequence and length, it is evident that the mucin domain peptide backbone is often maximally extended (11, 12). This is in line with

VNTR: variable number of tandem repeats

PTS: Pro–Thr–Ser

ER: endoplasmic reticulum

Mucin domain: long, densely *O*-glycosylated domain with sequences rich in Pro, Thr, and Ser, often characterized by tandem repeats

observations obtained via electron microscopy of mucin domains appearing as long, extended, stiff-looking rods (13). The properties of the mucin domains 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 nonbranched glycans will likely be more flexible, whereas more densely glycosylated domains with highly branched and long glycans are predicted to be stiffer. Note that many proteins that are not classified as mucins have short mucin domains. This is typical for membrane proteins, such as the low-density lipoprotein receptor, that have a short stalk that extends the functional domain out from the cell surface membrane.

So far, domain-predicting programs have not included tools able to predict which extracellular PTS sequences will become *O*-glycosylated; rather, they usually refer to these parts as unstructured. This is a misunderstanding, because once they are *O*-glycosylated, the mucin domains are highly structured.

Mucin *O*-Glycosylation

Mucins are glycosylated predominantly by the *O*-glycans, which 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 individual glycans usually are not recognized; rather, clustered saccharide patches form on the glycan surface (14, 15). However, due to a lack of appropriate methods there is limited understanding of the structure of such patches or the surface of mucin domains. However, a recent low-resolution cryo-electron microscopy study of mucin domains of MUC5B suggests an ordered and variable glycan surface (16). Much research has been devoted to studies of bacterial binding to individual glycans, but there are no studies on bacterial binding to glycan patches on mucins (17).

O-Glycosylation of the PTS sequences starts in the *cis*-Golgi stack by the action of any of the peptidyl-GalNAc transferases (10, 18). These enzymes appeared early in evolution, during the Metazoan era, 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 advances have allowed systematic studies of individual single-*O*-glycan attachment sites (19). These studies have shown that different peptidyl-GalNAc transferases display peptide sequence specificity based on both the catalytic unit and its lectin-binding domain, thus acting in concert to control mucin-type *O*-glycosylation (20, 21). Some of these transferases can act directly on the naked peptides, whereas others work only after their glycan-binding domain has bound to an already-attached GalNAc (10, 21). Consequently, it is difficult to determine whether a certain Ser or Thr in a peptide sequence is glycosylated or not. Although a Ser or Thr close to Pro is typically glycosylated, prediction of glycan attachment sites based on amino acid sequence is still complex because of the integration of target protein structure, transferase specificity, relative enzyme levels, and relative localization of peptidyl-GalNAc transferases in the Golgi apparatus.

Once the first glycan residue in the form of GalNAc is attached to the protein core, it will act as a substrate for the glycosyltransferases to add Gal, *N*-acetylglucosamine (GlcNAc), GalNAc, *N*-acetylneuraminic acid (NeuAc), and sulfate groups to this residue (1, 22–24). Studies of individual *O*-glycans show that this group of enzymes is dependent on the peptide sequence further contributing to the spatial glycan organization along the peptide core of mucins (25). On the basis of the coupling of Gal to GlcNAc, there are four major basic core structures that make up the core of *O*-glycans, called Core1–4 (**Figure 2*a***). At this level, there are already significant differences between animal species, as illustrated in **Figure 2*b***, which shows that mouse colon Muc2 has mostly Core2 structures whereas human MUC2 has predominantly Core3 structures (26–28).

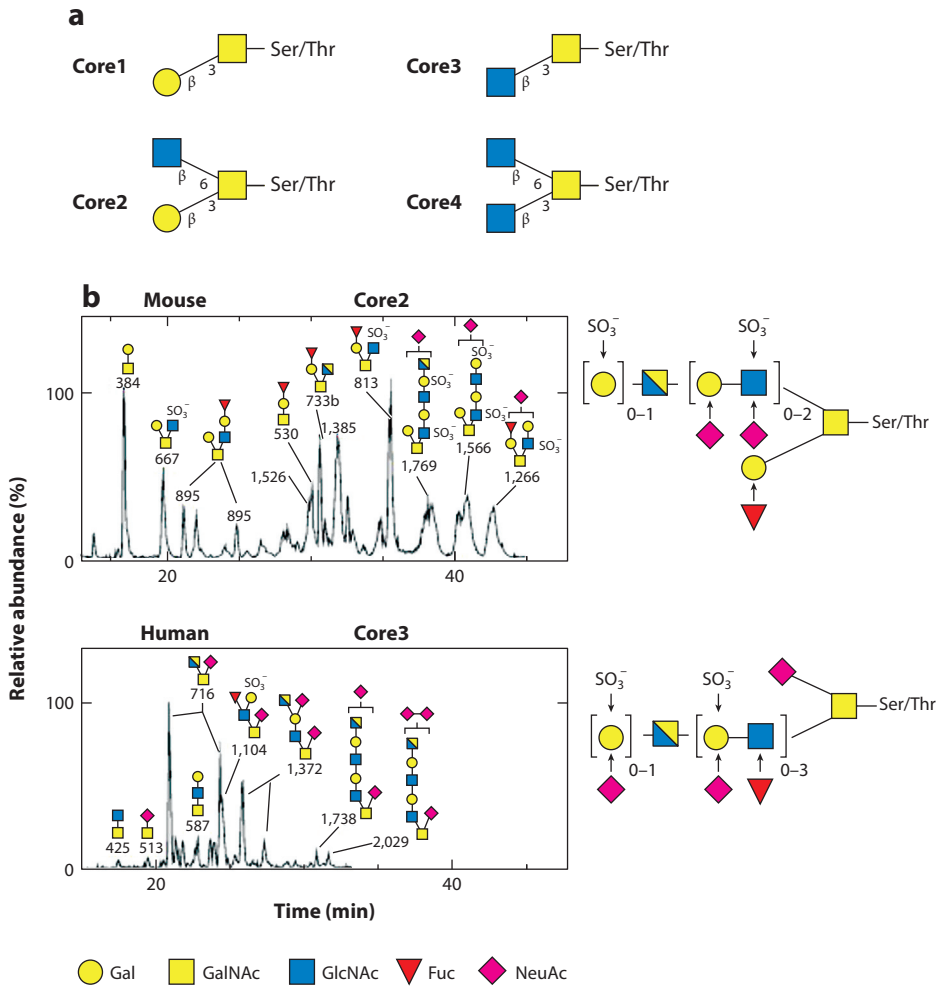

Figure 2

Mucin *O*-glycosylation. (*a*) Major *O*-glycan core structures. (*b*) Chromatograms and major structure of important *O*-glycans of mouse Muc2 and human MUC2. Abbreviations: GalNAc, *N*-acetylgalactosamine; GlcNAc, *N*-acetylglucosamine; NAc, *N*-acetylcysteine; NeuAc, *N*-acetylneuraminic acid.

The Core structures are further elongated by the addition of the same type of sugar residues. A classical extension consists of one or several lactoseamine disaccharide units of type 1 (Galβ1–3GlcNAc) or type 2 (Galβ1–4GlcNAc) that generate long *O*-glycans. The length of the extension varies, of course, by tissue and species, but usually the mucin glycans of the intestine are shorter and have extensions on both the C-3 and C-6 carbons of the peptide-bound GalNAc. This will generate bulky glycans that help extend and stiffen the mucin domains and efficiently shield the protein cores from access by proteases.

Mucin *O*-glycans are typically capped with peripheral glycan residues that cannot be further elongated. Among these are fucose (Fucα1-2, α1-3, α1-4), sialic acid [NeuAc(Gc)α2-3α2-6], GalNAc (α1-3, β1-4), and Gal (α1-3). Some of these combine, giving rise to the blood or histogroup 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 peripheral and core glycans are substituted by sulfate groups, and numerous hydroxyl groups are acetylated. The latter modification is known to be present on sialic acids, but it is likely abundant on other sugar residues as well (29).

Gastrointestinal Mucin *O*-Glycosylation

The mucin glycan repertoire differs substantially along the gastrointestinal 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 glycan diversity in the normal human stomach is enormous, with up-to-13-residue-long glycans and around 100 different species in each individual (30, 31). In 10 individuals, more than 250 different glycan structures were identified.

The normal human intestine shows a specific distribution of *O*-glycans along its length (27, 28). Fucose in blood group ABH antigens is more localized to the small intestine and proximal part of the large intestine, and longitudinal glycan differences along the colon are reflected in the levels of glycosyl transferases, as expected (32). The fact that blood group status is reflected in the intestinal mucins means that there is individual variability in glycan epitopes. Interestingly, this individual glycan variability is almost absent in the distal part of the human large intestine, where glycans are relatively uniform among individuals (33). This uniformity should be understood in relation to the commensal bacteria, which utilize the mucin glycans for attachment as well as a source of nutrients. Glycosylation of mucins in the distal human colon seems to be stable over time but was found to be altered in active ulcerative colitis (34). The altered glycan repertoire was triggered by inflammation when the disease was active and returned to normal as the inflammation faded.

The mucin glycans vary widely between different evolutionarily closely related species, and this variability is especially evident in the respiratory and gastrointestinal tracts (for a more detailed discussion of glycan variability in the intestine of humans and rodents, see the sidebar titled Intestinal Glycan Differences Among Mouse, Rat, and Human). A likely explanation is that different microbial pathogens generated exogenous selection pressure on the host to replace glycan structures, as the pathogens utilized these to invade the host (40). As a consequence, there are unexplored genetic mechanisms favoring species polymorphism and more rapid evolution of glycans than in general.

INTESTINAL GLYCAN DIFFERENCES AMONG MOUSE, RAT, AND HUMAN

The normal rodent small intestinal mucins show, similarly to humans, glycans that differ from those in the colon (26, 35). In mouse, the Fut2 enzyme that adds fucose is constituently present in the colon but is inducible in the small intestine by bacteria or parasites (36, 37). Interestingly, in rats the Fut2 enzyme is constitutively present; instead, a blood group A–type GalNAcα1-3 transferase is transiently induced upon parasite infection (38).

Human intestinal mucin glycans are characterized by the presence of Core3 and, to a lesser extent, Core4 (**Figure 2*a***), and they essentially lack Core1. Mouse glycans, by contrast, comprise predominantly Core1-based components, although they make small amounts of Core3 and Core4 (26). A major difference is that humans have an active sialyl transferase that adds NeuAcα2-6 to the peptide-bound GalNAc, an enzyme not active in mice (**Figure 2*b***). However, both mice and humans have the Sd^a-Cad blood group–like epitope [NeuAcα2-3(GalNAcβ1-4)Galβ1-3/4GlcNAc] as a major cap of their glycans in the large intestine (26, 28, 39).

Table 1 Human mucins

Mucins[a]	Domain structure	Chromosome	Number of amino acids	Typical localization	Reference
TM mucins					
SEA					
MUC1	PTS–SEA–TM–CT	1q22	1,255	General	5
MUC3	PTS–SEA–TM–CT	7q22.1	>2,541	Intestine	80
MUC12	PTS–SEA–TM–CT	7q22.1	5,478	Intestine	81
MUC13	PTS–SEA–TM–CT	3q21.2	512	Intestine	82
MUC16	$(PTS–SEA)_{33}$–TM–CT	19p13.2	22,152	General	77
MUC17	PTS–SEA–TM–CT	7q22.1	4,493	Intestine	79
NIDO–AMOP–VWD					
MUC4	PTS–NIDO–AMOP–VWD–TM–CT	3q29	5,284	General	73
Others					
MUC15	PTS–TM–CT	11p14.2	334	General	71
MUC21	PTS–TM–CT	6p21.33	566	Esophagus	70
MUC22	PTS–TM–CT	6p21.33	1,733	Esophagus	70
Secreted mucins					
Gel-forming					
MUC2	VWD1–VWD2–VWD3–PTS(CysD)–VWD4–CT	11p15	5,130	Intestine	66
MUC5AC	VWD1–VWD2–VWD3–PTS(CysD)–VWD4–CT	11p15	5,654	Lung, stomach	64
MUC5B	VWD1–VWD2–VWD3–PTS(CysD)–VWD4–CT	11p15	5,703	Lung, saliva	60
MUC6	VWD1–VWD2–VWD3–PTS–CT	11p15	5,534	Stomach	69
Monomeric					
MUC7	PTS	4q13.3	377	Saliva	42
MUC20	PTS	3q29	709	Kidney/urinary tract	43

[a]Not all numbers are used: MUC8 is not a mucin; MUC9 was renamed OVIOGP1; MUC10 is not found in humans; MUC11 is part of MUC12; MUC14 is called EMCM; MUC18 is not used; and MUC19 is not expressed in humans.
Abbreviations: AMOP, adhesion-associated domain; CT, cytoplasmic tail; NIDO, nidogen domain; PTS, Pro–Thr–Ser; SEA, sea urchin–enterokinase–agrin; TM, transmembrane; VWD, von Willebrand D domain.

Typically, there are many (often more than 100) different *O*-glycan structures on the same mucin from the same organ. A major reason for this wide diversity is that all biosynthetic intermediates are usually present. In contrast, the glycosphingolipids of the apical enterocyte cell membrane lack most biosynthetic intermediates (41). The reason for these differences could be that it is difficult for glycosyltransferases to reach all potential substrates on the mucins because of their size, but they are more likely of functional importance, in that they 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**; **Figure 1**). They are further divided into monomeric and gel-forming mucins. The term gel-forming refers to the ability to form mucus with gel-like properties. This is not an ideal name,

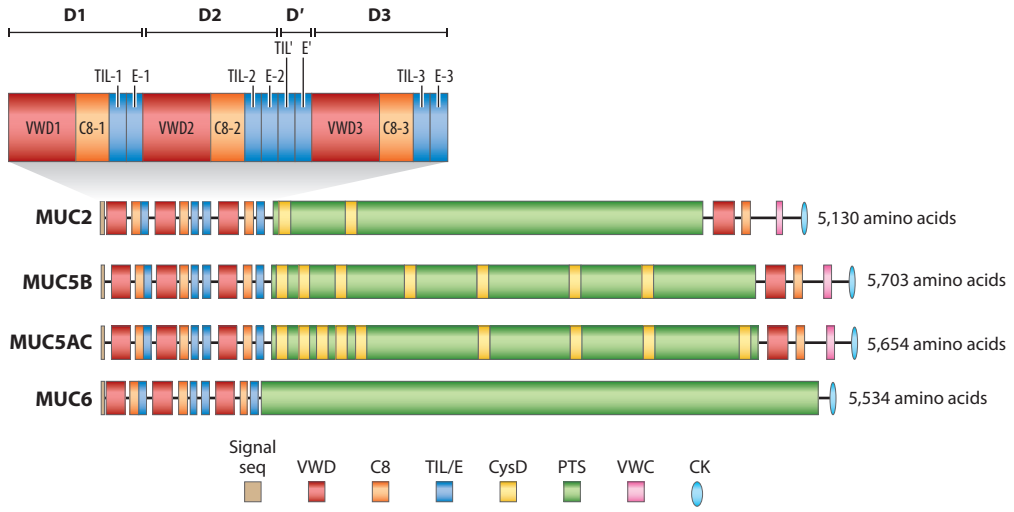

Figure 3

Domains of the four human polymer-forming mucins. Abbreviations: CK, Cys-knot domain; PTS, Pro–Thr–Ser; VWC, von Willebrand C domain; VWD, von Willebrand D domain.

and the term polymer-forming probably better reflects their structure. Two monomeric mucins, MUC7 and MUC20, essentially consist of only a secreted mucin domain (42, 43).

Polymer-Forming Mucins

The classical polymer-forming mucins have a common protein domain organization shared with von Willebrand factor (VWF), an important molecule in the vascular coagulation system. The four human gel-forming mucins and VWF all have 3.5 von Willebrand D (VWD) assemblies in their N termini (**Figure 3**; **Table 1**) named VWD1, VWD2, VWD′, and VWD3. The central part of the mucins, but not VWF, consists of one or several PTS sequences that, after *O*-glycosylation, become mucin domains. These central mucin domains are often interrupted by 100-amino-acid-long domains called CysD domains. The C termini of the mucins and the VWF contain a Cys-knot domain that may be preceded by VWD4 assembly and von Willebrand C (VWC) domains (**Figure 3**).

VWF: von Willebrand factor, an important factor of the blood coagulation system

VWD domain: von Willebrand D domain or assembly; assemblies refer to the VWD domain and associated C8, TIL, and E domains

VWC domain: von Willebrand C domain

In VWF, the central part of the mucin is replaced by von Willebrand A and B domains. Studies of the evolution of VWD domains show that these appeared early in or even before Metazoan evolution and were associated with PTS sequences, as in mucins (9, 44) (**Figure 4**). VWF appeared much later, during vertebrate evolution; thus, polymer-forming mucins were its origin. Interestingly, polymer-forming mucins were lost during the evolution of insects. There is only one remaining protein in insects called hemolectin, which has two functions: encapsulating bacteria and inhibiting hemolymph leakage, thereby recapitulating the function of the mucins and VWF, respectively (45). The fact that the peptidyl-GalNAc transferases also appeared during the Metazoan era adds further support for the early importance of mucins (10).

The number of polymer-forming mucins has changed during evolution. Most mucins are found in species that have mucins not only on their internal organs but also on their skin. Frogs have around 25 polymer-forming mucins; fishes, 15–20 (9, 44). Mammals have four or five of these mucins (**Figure 4**), which protect only internal organs. There are five polymer-forming mucins in the human genome, but only four are expressed as complete mucin proteins: MUC2, MUC5AC,

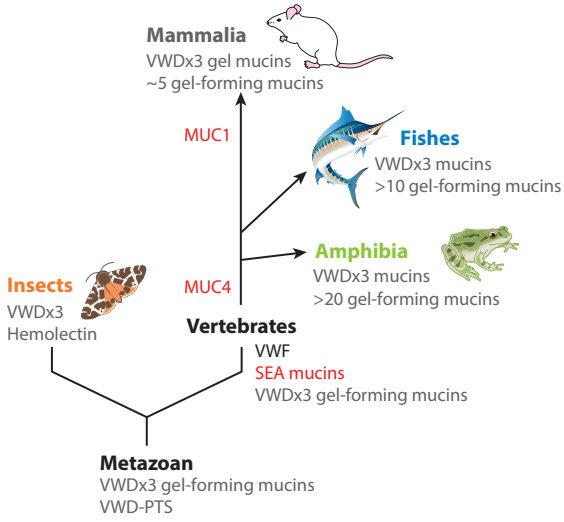

Figure 4

Evolution of mucins and their major domains. Abbreviations: PTS, Pro–Thr–Ser; SEA, sea urchin–enterokinase–agrin; VWD, von Willebrand D.

MUC5B, and MUC6. The *MUC19* gene is not expressed in humans but is highly expressed in pigs, where it is called PSM (porcine submaxillary mucin) (46–48).

All polymer-forming mucins are assembled into covalent dimers in the ER by disulfide bonds between two Cys-knot domains (49) (**Figure 3**). The dimers then pass into the Golgi apparatus, where the PTS sequences are *O*-glycosylated and reach masses in the range of 5 MDa. The pH gradually decreases over the secretory pathway, and when the mucins reach the *trans*-Golgi network, their N termini form packed electron-dense assemblies that are sorted into regulated secretory vesicles. The mucin N termini, 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.

Only limited detailed structural information on the different domains of the polymer-forming mucins is available. The first such domain to be structurally determined was a monomeric form of VWD′–VWD3 of VWF (50). A structure of the MUC2 VWD3 domain was published recently (51). These structures show the coordination of the expected bound calcium ion. Some structural information on other domains in mucins from other proteins is also available. The Cys-knot domain is found as homodimers in several proteins, and the VWC domain is found in collagen (52).

Mucins' role in protecting the mucosal surfaces, especially the intestine, is demanding. The intestinal mucin MUC2 is insoluble, a fact that was long overlooked but first observed by Carlstedt et al. (53) in Sweden. MUC2's insolubility in urea and in chaotropic salts like guanidinium chloride has been linked to the appearance of nonreducible covalent bonds. Electrophoresis gels have revealed oligomers higher than the expected monomer after dithiothreitol treatment (54). Such nonreducible bonds form during biosynthesis and passage through the later stages of the secretory pathway (55). Much later, these nonreducible bonds were shown to be isopeptide bonds between the side chains of Lys and Gln (56). Such bonds are known to stabilize other proteins exposed to mechanical stress, such as collagen in skin and fibrin in blood clots (57). These bonds are formed by transglutaminases, of which TGM2 and TGM3 are found in normal intestine and TGM2 in

chronically diseased lungs (58, 59). The precise molecular details and functions are lacking, but we can assume that isopeptide bonds mechanically stabilize mucus.

SEA: sea urchin–enterokinase–agrin

MUC2, MUC5AC, MUC5B, and MUC6

The major normal mucin in the respiratory tract and saliva is MUC5B (60, 61) (**Figure 3**; **Table 1**). This mucin forms N-terminal dimers and thus linear polymers (62, 63).

The sequence of MUC5AC is the most similar to that of MUC5B (64, 65). This mucin forms N-terminal covalent dimers that further interact with tetrameric noncovalent polymers (S. Trillo-Muyo & G.C. Hansson, unpublished manuscript) to form netlike structures.

MUC2 is found in the small and large intestine and has two large mucin domains interspersed by two CysD domains (66). This mucin forms N-terminal disulfide-bonded trimers that, once secreted, generate stacked netlike polymers, ideal for the generation of protective filters (67).

MUC6 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 (68, 69). Little is known about its assembly and higher-order polymers.

Transmembrane 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 (**Figure 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 other than a mucin domain (**Table 1**). This last group, consisting of MUC15, MUC21, and MUC22, has essentially only an extracellular mucin domain coupled to the transmembrane part (43, 70, 71). In the NIDO–AMOP–VWD mucins, these domains are extracellular, between the N-terminal mucin domain and the transmembrane domain. Interestingly, there is only one member of this family, MUC4, in each higher species; it is not restricted to mucosal surfaces. Little is known about its normal function, as it has largely been studied in relation to cancer (72–74).

MUC1 is part of the largest group of transmembrane mucins, all of which have a SEA domain between the N-terminal PTS sequence and the transmembrane domain. The SEA domain is cleaved autocatalytically during ER folding but is still held together by strong noncovalent forces generated by four antiparallel β-sheets (75). MUC1 is the most-studied mucin of this family and is found on many cell types outside of the mucosal surfaces, as well as on a number of immune cells (76). The largest known mucin, MUC16, has more than 20,000 amino acids. This mucin has more than 30 repeated SEA–PTS sequences in which the SEA domains are not cleaved as they are in mucins with only one such domain (77, 78). This mucin is well known because it carries the CA125 epitope, which is widely utilized as a marker for ovarian cancer (78).

The remaining SEA mucins (MUC3, MUC12, MUC13, and MUC17) are all found in the intestinal tract and localize to the apical surface of polarized epithelial cells (79–82). Their PTS sequences contain more than 4,000 amino acids, with the exception of MUC13, which is considerably smaller with only 150 amino acids in its PTS. MUC3, MUC12, and MUC17 reach approximately 1 μm from the tips of the enterocyte microvilli into the intestinal lumen. Such mucins form the densely glycosylated glycocalyx of enterocytes (83).

MUCINS AND BACTERIA

Mucus and its main components, the mucins, as well as epithelial surface mucins have the role of protecting inner mucosal surfaces. This role has dual functions: to protect the epithelium and

to remove intruders. Protection is based on building a system where the epithelium is protected from contact. Removal is based on a system that requires trapping intruders and the ability to move them to less susceptible (for the host) positions, usually the colon. Protection requires a barrier that is attached to the epithelial surface. For example, transmembrane mucins build the dense glycocalyx of intestinal enterocytes, and polymer-forming mucins build the impenetrable inner mucus layer of the colon. Removal requires movement and is accomplished by cilia in the respiratory system and by peristalsis in the intestine.

The host faces a tremendous challenge in protecting its inner surfaces. Parasites have been a considerable problem throughout human evolution, but they are a less frequent threat today, as high-density human populations are more challenged by viruses. Bacteria may be the most important threat, as they can divide and evolve rapidly. Mucus and mucins play an important role in protecting the organism from these challenges. The remainder of this review focuses on the role of mucins in protecting humans from bacteria at the two largest internal surfaces: the respiratory and gastrointestinal tracts.

CF: cystic fibrosis

COPD: chronic obstructive pulmonary disease

ASL: airway surface liquid

THE RESPIRATORY SYSTEM

The respiratory tract is exposed to millions of particles and bacteria at every breath we take. Despite this exposure, healthy lungs are kept essentially free from bacteria (84), although DNA amplification methods show that most individuals have at least some resident bacteria (85). In contrast, in individuals with chronic lung diseases, such as cystic fibrosis (CF) and chronic obstructive pulmonary disease (COPD), the lungs are heavily colonized, often with specific bacteria.

A hallmark of the respiratory epithelia is the ciliated cells. The beating of the cilia continuously generates vectorial transport in the cephalic direction of the airway 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 (86). The high carbohydrate content of these mucins maintains the liquid around the cilia, called periciliary liquid, and provides the low friction required for cilia movement (86). The coating provided by these mucins helps protect the epithelial cells from mechanical forces and acts as a barrier against microorganisms.

Large animals, such as pigs and humans, have numerous submucosal glands (**Figure 5*a***), specialized molecular machines that make long and thick MUC5B bundles (87, 88). In the late secretory pathway of the MUC5B mucin-producing cells, where pH is low (5.5–6), the N-terminal VWD1 and VWD2 domains of the VWD3 dimer are bent inward, hooking together two MUC5B dimers that face each other at an angle of 180° (89). Electron microscopy and 3D-printed models further suggest that these double dimers are assembled side by side into linear structures in which the remaining MUC5B, including the mucin domains, point outward in all directions. Such nonexpanded structures have been observed in saliva (90). In the submucosal gland, the most peripheral cells generate a chloride- and bicarbonate-rich fluid that flushes the MUC5B mucin-secreting cells. Bicarbonate increases the pH, causing the MUC5B N termini complexes to disassemble, and the flow pulls MUC5B out into linear polymers (87, 91). These polymers are gathered into thicker bundles in the gland ducts and are observed at the exit of the glands as thick bundles ~25 μm in diameter (**Figure 5*a***). These bundles have a central core made of MUC5B mucins. Calculations based on the estimated diameter of the mucin domains suggest that each bundle contains 1,000–5,000 MUC5B polymers (87, 91). The mucin bundles must be organized by as-yet-unidentified MUC5B intermolecular interactions and perhaps by additional molecules.

Observations of the mucin bundle movements on a tilted tracheal surface show that they are transported uphill by the forces generated by the beating cilia that also drive the ASL uphill.

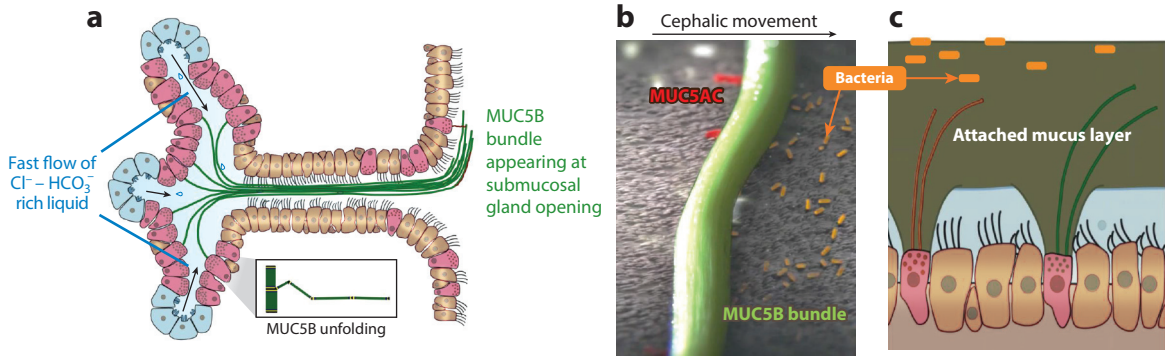

Figure 5

The mucus system of the normal and diseased respiratory tract. (*a*) The submucosal gland is a molecular machine that generates MUC5B-based bundles. (*b*) The bundles are transported cephalically on the tracheobronchial surface by the beating of cilia and movement of the airway surface liquid. Bacteria are collected by bundles sweeping over the surface. (*c*) An attached mucus layer is formed by MUC5AC and MUC5B in diseased lungs, thereby helping to separate epithelial cells from bacteria.

Interestingly, the bundles did not, as expected, move parallel to the ASL movement but rather perpendicular to it (87, 91) (**Figure 5*b***). The bundle transport velocity was also approximately a tenth as fast as that of the ASL and the bundles did not move evenly, as the ASL did. Furthermore, the bundles must be kept down on the tracheobronchial surface, as they should not fall off into the lumen. The bundles are efficient at cleaning the tracheobronchial surface and removing bacteria (91) (**Figure 5*b***). In individuals with CF, the bundles are immobile from birth, and the bacteria remain on the epithelial surface.

The mechanisms responsible for holding down the bundles on the tracheobronchial surface and controlling bundle movement are poorly understood. However, the observation that MUC5AC extends from inside the surface goblet cells, reaches over to, and patchily attaches to and coats the MUC5B bundles suggests an interesting mechanism (87, 91, 92). In contrast to previous assumptions, gel-forming mucins seem to remain attached to goblet cells during secretion, despite their lack of transmembrane domains. This was first observed in the small and large intestine, where MUC2 is attached to goblet cells (87, 93, 94). The molecular mechanisms responsible for attaching gel-forming mucins to goblet cells are not understood, but the attachment suggests anchoring to unknown transmembrane proteins that are likely present in the secretory granule membrane. Studies of the small intestine have shown that protease activities are involved in detachment, as exemplified by the meprin β protease (94). In contrast to normal mice, mice lacking this enzyme have attached small intestinal mucus. Observations of the bundle movement pattern suggest that mechanical forces are involved in organizing the perpendicular and irregular movement of the bundles. Similar mechanisms as for the VWF may be at work in which blood flow can pull open an ADAM13 cleavage site in the von Willebrand A2 (VWA2) domain, allowing rupture of the VWF polymer. However, no VWA domain is present in the mucins (95, 96).

The normal organization of the mucus and mucins in the lungs differs considerably among species. Small animals, such as mice, lack submucosal glands (except for a few at the larynx), and their tracheobronchial surface is cleaned by small mucus clouds made from MUC5B from the surface goblet cells (97, 98). The reason is, of course, the lack of submucosal glands, something that is likely related to the smaller diameter of these animals' bronchi. It is intuitively understandable that small mucus clouds could be sufficient for sweeping and cleaning small airways but insufficient

in large airways, such as the centimeter-wide human trachea. Human airways contain submucosal glands and mucus bundles down to about the tenth bronchial bifurcation. In the smallest human airways, the cleaning process might be similar to what has been observed in mice. Regardless of whether bundles or cloud mucus moved by cilia is responsible, these are highly efficient systems for keeping the lungs free from bacteria.

Upon infection, or in chronic diseases like CF or COPD, the lungs can revert from cleaning to protection. In such cases, an attached mucus layer that is able to separate bacteria from the epithelial cells is generated (58) (**Figure 5*c***). The layer appears stratified and thus similar to the one that protects the colon. The molecular mechanism responsible for forming such a layer likely involves the structures formed by MUC5AC, a mucin that is dramatically increased in these diseases (63, 99). MUC5AC forms a linear disulfide-bonded polymer, just like MUC5B, 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 thereby organize netlike sheets that, together with other proteins found in the normal colon, form a stratified mucus layer that can separate bacteria from the epithelial surface.

In chronic lung diseases such as COPD and CF, the attached mucus layer is retained in the lungs (100). As a major function of mucus is to trap bacteria, the bacteria are retained in the lungs. Certain bacteria, especially *Pseudomonas aeruginosa*, have special preferences for living in mucus. Retention of bacteria in the lungs triggers inflammatory reactions, which eventually destroy lung tissue in patients with these diseases.

THE DIGESTIVE SYSTEM

Just as the respiratory system is efficient in keeping the lungs relatively free from bacteria, the intestine is highly efficient in handling a large number of bacteria. The major intestinal mucin, MUC2, is crucial for this purpose. A major reason for the structural organization of the polymer-forming mucins is likely the requirement for a digestive system in multicellular organisms. The intestine is supposed to degrade and absorb most types of food without degrading itself. This important feature is provided by limiting the host digestive system to degrading only certain carbohydrates. Human digestive enzymes can cleave only the specific glycosidic linkages (Glcα1-4Glc and Glcα1-6Glc) found in starch and those in a few disaccharides (lactose, sucrose, maltose, isomaltose, and trehalose). This allows the glycans on the mucins, both transmembrane and polymer-forming, to be unaffected and makes them able to protect the intestinal surface. The N and C termini of the polymeric mucins are less glycosylated but are rich in Cys (1 in every 7–10 amino acids). These bonds will help the mucins to form highly compacted and stabilized structures in which pancreatic proteases cannot access their specific cleavage sequences. Together, these features of MUC2, despite its more than 5,000 amino acids, mean that this mucin remains intact after exposure to pancreatic proteases. The commensal bacteria in the human large intestine, in contrast, help us degrade MUC2 glycans and recover their energy content.

Stomach

The stomach is covered by an attached mucus layer built around MUC5AC, which, importantly, has the ability to form netlike polymers (101; S. Trillo-Muyo & G.C. Hansson, unpublished manuscript). This mucus layer acts as a diffusion barrier for luminal hydrochloric acid, thereby protecting the surface epithelial cells (102). The stomach glands that produce hydrochloric acid and pepsinogen, however, are protected in unknown ways. MUC6 is the only gel-forming mucin produced in these glands and might have an important protective role. The gland content is

secreted through the surface MUC5AC mucus layer covering the glands via opening of pores, which occurs by as-yet-unexplored molecular mechanisms (103).

Small Intestine

The small intestine has a nonattached mucus layer built around MUC2 (59, 101). MUC2 remains attached to goblet cells after secretion and requires the protease meprin β to cleave MUC2, allowing its detachment (94). The detachment is well controlled, as its activation requires the presence of bacteria as well as an active bicarbonate-secreting CF transmembrane conductance regulator (CFTR) ion channel (104). The mucus is penetrable to bacteria and bacteria-sized beads, but because of the abundance of antibacterial peptides and proteins, especially those from the Paneth cells in the crypt bottom, there are no bacteria in contact with the cells (105, 106).

The enterocytes of the small intestine have the densest and thickest glycocalyx of any cell type. In these cells, packed mucins and especially MUC17 create a barrier against 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 anchors in the apical membrane, but to protect the membrane from rupture, the cleaved SEA domain provides a mechanical breaking point (107). Because such mucins reach the farthest away from the epithelial cells, they may have as-yet-undefined sensory functions, as has been suggested for similar types of molecules in yeast (108).

The small intestinal mucins along with the glycan decoration of their mucin domains are very important for the selection of commensal bacterial species. The importance of attached/detached mucus was illustrated by experiments in which germ-free animals were first fed normal mouse bacteria (109). The attached mucus layer in the germ-free animals retained and selected different bacterial flora with an increased number of Bacteroidetes and a concomitantly lower number of Firmicutes (110). This altered composition returned to normal once the mucus was normalized and detached 4–5 weeks after colonization. In CF patients, problems with both intestinal obstruction and bacterial overgrowth from the distal small intestine are likely related to their mucus being attached (104, 111, 112).

Large Intestine

MUC2 forms the skeleton of the colon mucus system, with two major mucus layers: an outer, nonattached, less dense layer and an inner, attached, dense mucus layer (93) (**Figure 6**). The inner mucus layer has a stratified outer part, in which MUC2 is self-organized in flat netlike structures, as well as a less well organized inner part close to the epithelial cells, where mucin expands (67, 93).

Commensal bacteria. The large intestine harbors an estimated 1,000 different bacterial species, the majority of which belong to the phyla Bacteroidetes, Proteobacteria, Firmicutes, and Actinobacteria (113). Their high number and diversity have been estimated from 16S ribosomal RNA genome sequencing (114, 115). More recently, full metagenome and transcriptome sequences of the commensal bacteria and available genomic sequences for several thousand species have been obtained (116–120). Annotating these genomes and determining the functional importance of the encoded proteins are current major challenges.

One of the earliest-known and most-studied commensal bacteria is *Bacteroides thetaiotaomicron*, which, like many other commensal bacteria, has devoted much of its genome to carbohydrate degradation. In this and other bacteria, carbohydrate utilization is organized in gene assemblies called polysaccharide utilization loci, which include sensors, transporters, and glycosidases

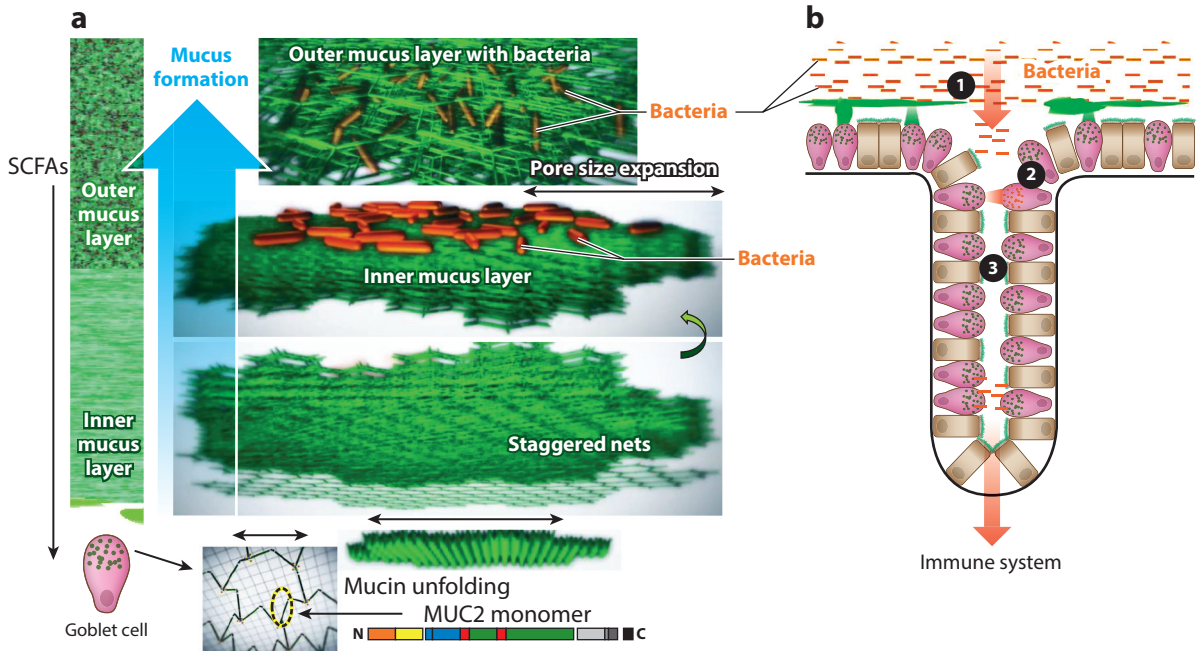

Figure 6

The mucus system and bacteria of normal and diseased colon. (*a*) Goblet cells assemble MUC2 into large polymers that, upon secretion, generate large netlike structures that are staggered from below on one another to form the attached inner mucus layer. The pore size of this inner layer is small, keeping bacteria away from the epithelial cells. At the interface with the outer mucus layer, the mucus expands and the pore size is increased, allowing bacteria to enter and degrade the mucin glycans so as to generate short-chain fatty acids (SCFAs) that feed the epithelial cells. (*b*) Ulcerative colitis is initiated when the inner mucus layer (first defense line, ❶), the sentinel goblet cells (second defense line, ❷), and a potential third defense line (❸) fail and bacteria reach the subepithelial immune system.

(113, 121). Other commensal bacteria have similar or fewer genes for carbohydrate utilization; some are devoted to degrading complex plant saccharides, and others are more specialized in mucin glycan degradation. One of the latter is *Akkermansia muciniphila*, an efficient mucin glycan degrader that can live on pig stomach mucins (122, 123). 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 of the glycans. Note that mucin *O*-glycans vary enormously along the intestinal tract and between species; researchers in this area should take great care and study bacteria and mucin from the same host and niche.

Outer mucus layer and bacteria. The outer mucus layer is formed from the inner mucus layer through detachment and volume expansion, processes that are controlled by the host (93, 109) (**Figure 6*a***). The two processes are not fully understood at the molecular level, and they likely are not linked, as detachment generates a sharp border whereas expansion is gradual. The slow expansion of the outer mucus layer is demonstrated by the aggregation of bacteria by the peptidoglycan-binding protein ZG16, which thereby moves them further away into the outer mucus layer (124). Several proteases orchestrate the processes generating the outer mucus layer, of which at least one has been identified: the CLCA1 protein (125, 126). This abundant mucus protein, which contains an N-terminal metalloproteinase domain, a central VWA domain, and a C-terminal fibronectin type III domain, likely has both enzymatic and structural roles in the

formation of mucus. A comparison between the MUC2 mucin of the inner colon mucus layer and that of the outer colon mucus layer suggests cleavages in the C-terminal part of MUC2 that do not necessarily cause depolymerization, as the C-terminal part is heavily cross-linked by numerous disulfide bonds (93). The important consequences of the formation of the outer mucus layer are that bacteria can enter and that this mucus can be transported by peristalsis.

SCFAs: short-chain fatty acids

Selection of bacteria by the host has been nicely illustrated by elegant experiments in which zebrafish and mice were made germ free, after which bacteria were transferred from one host to the other (127). The mice selected out normal (for the mouse) bacteria from the mixture of zebrafish bacteria.

As bacteria do not enter the inner mucus layer and the host does not secrete mucin glycan glycosidases, the outer surface of the inner mucus layer exposes the intact mucin domain glycans as biosynthesized by the host (128, 129) (**Figure 6*a***). This exposure is likely crucial for specific host microbial selection. Studies of bacterial adhesion to glycans have largely involved single-carbohydrate epitope–protein interactions (17, 130). This is likely not the way in which commensal bacteria recognize the mucin domain glycans, as these glycans are densely packed. Instead, recognition is probably based on glycan patches in which the combined surface generated by several specifically arranged glycans of the mucin domains is recognized. At present we have only 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 have to be carefully studied using bacteria and mucin domains from the same host, part of the intestine, and bacteria localized to the interface between the inner and outer mucus layers.

There are, of course, factors other than the host glycan repertoire that are important for bacterial selection, such as the action and specificity of antibacterial defensins (131). Also important are the host's diet and its plant-derived complex glycan content.

Once bacterial glycosidases have started to act on the mucin glycans, the host-specific glycan epitopes disappear. The bacteria continue to degrade the glycans, and finally the protein backbone is exposed and degraded. Only small amounts of mucins appear in feces, showing that the commensal bacteria efficiently utilize the host mucins. The bacterial metabolism uses the released monosaccharides to generate the short-chain fatty acids (SCFAs) acetate, propionate, and butyrate. SCFAs diffuse back through the inner mucus layer and feed the local epithelial cells and, to some extent, the body (113).

Inner mucus layer and bacteria. The inner colon mucus layer is normally devoid of bacteria because it is impenetrable to bacteria and micrometer-sized particles (93, 109). This filter-like effect is generated by the nets formed by MUC2 polymers staggered on top of one another (67, 132) (**Figure 6*a***). The theoretical pore sizes generated by the MUC2 ringlike polymers are considerably larger than bacteria and exclude micrometer-sized beads, suggesting that individual MUC2 sheets are not staggered directly on top of one another. Recent studies investigating how *Salmonella* bacteria swim in the colon mucus are illustrative (133). These flagellated bacteria swim freely in the outer mucus layer and on the surface of the inner mucus layer, where they search for breaches as a way to access the epithelium and thereby invade the host. Interestingly, bacteria are sometimes trapped and stop moving at the outer surface of the inner mucus layer, supporting the concept of a netlike inner mucus layer with pores that are smaller than the bacteria and thus act like a fishnet.

Interestingly, germ-free mice have an inner mucus layer that is penetrable to bacteria-sized beads (109). The thickness of this inner mucus layer is almost identical to that in normal animals, and it has only a slightly lower concentration of MUC2. Colonization of 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 recapitulates the abnormal properties of the mucus layer (110). Recent observations have shown that feeding of Western diets, high in fat and especially low in plant polysaccharides (fiber), within 3 days renders the inner mucus layer more penetrable to bacteria, allowing them to come closer to the host epithelium (134, 135). 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, implying that bacterial products can influence the formation of mucus. As the inner mucus layer provides a diffusion barrier, it is likely that bacterially generated small molecules will be responsible for such communication. Understanding the organization of a functional inner mucus layer requires a detailed molecular understanding of the different domains of MUC2 and the other major proteins of the colon mucus as well as their interaction with one another. Mucin secretion and its enormous volume expansion demonstrate the importance of the ionic conditions and pH during mucus secretion, as they likely also modulate mucus properties.

The inner mucus layer normally keeps the commensal bacteria at bay, but when this first line of defense fails, bacteria come in contact with the epithelium. This event is likely the first in the development of the inflammatory disease ulcerative colitis (136, 137) (**Figure 6*b***). When bacteria reach the crypt opening, they encounter a second line of defense as sentinel goblet cells sense the increased level of bacterial products and initiate a response by a coordinated release of a mucus plume to wash away bacteria (138) (**Figure 6*b***). A third line of defense is likely mediated by the emptying of crypt goblet cells (139) (**Figure 6*b***). These lines of defense are easily exhausted; the regeneration of new MUC2 mucins is time-consuming, as their biosynthesis is demanding and prone to cause ER stress responses. Once an overwhelming number of bacteria reach the immune cells of the lamina propria, an overt inflammatory reaction is triggered, increasing mucus secretion and likely initiating an acute episode of ulcerative colitis.

CONCLUSION

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

SUMMARY POINTS

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

FUTURE ISSUES

1. We need to reveal the molecular structure of the different domains found in mucins and how they mediate interaction between mucins and other mucus proteins to generate highly organized mucus.
2. We need to understand how mucins are unfolded upon secretion and how different ions and fluid conditions lead to different mucus properties.
3. We need to understand how gel-forming mucins detach from goblet cells in the intestine and lung.
4. We need to understand the glycan-generated surface of mucin domains and how they interact with bacteria.
5. We need to understand how different bacteria cooperate and utilize mucin glycans.
6. We need to understand how bacterial metabolites communicate with and influence the host epithelium and its mucus production.
7. We need to develop novel therapeutic approaches and agents that can enhance the protective capacity of colon mucus for the treatment of ulcerative colitis.
8. We need to develop novel therapeutic approaches and agents that can detach the attached mucus in CF and COPD.

DISCLOSURE STATEMENT

The author is not aware of any affiliations, memberships, funding, or financial holdings that might be perceived as affecting the objectivity of this review.

ACKNOWLEDGMENTS

I am indebted to all present and former coworkers in the Mucin Biology Groups for all of their work, without which this review would never have been possible, and to Dr. Fredrik H. Sterky for his input to the text. The writing of this review was supported by the European Research Council (694181), the 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 National Institutes of Health), the Knut and Alice Wallenberg Foundation (2017.0028), the Swedish Research Council (2017-00958), the Swedish Cancer Foundation, the IngaBritt and Arne Lundberg Foundation, Sahlgren's University Hospital (ALF agreement 236501), the Bill and Melinda Gates Foundation, and the Wilhelm and Martina Lundgren Science Foundation.

LITERATURE CITED

1. Corfield AP. 2015. Mucins: a biologically relevant glycan barrier in mucosal protection. *Biochim. Biophys. Acta Gen. Subj.* 1850:236–52
2. Johansson MEV, Sjovall H, Hansson GC. 2013. The gastrointestinal mucus system in health and disease. *Nat. Rev. Gastroenterol. Hepatol.* 10:352–61
3. Linden SK, Sutton P, Karlsson NG, Korolik V, McGuckin MA. 2008. Mucins in the mucosal barrier to infection. *Mucosal Immunol.* 1:183–97
4. Johansson MEV, Hansson GC. 2016. Immunological aspects of intestinal mucus and mucins. *Nat. Rev. Immunol.* 16:639–49

5. Gendler SJ, Lancaster CA, Taylor-Papadimitriou J, Duhig T, Peat N, et al. 1990. Molecular cloning and expression of human tumor-associated polymorphic epithelial mucin. *J. Biol. Chem.* 265:15286–93
6. 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
7. 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
8. 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
9. Lang T, Hansson GC, Samuelsson T. 2007. Gel-forming mucins appeared early in metazoan evolution. *PNAS* 104:16209–14
10. 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
11. 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
12. 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
13. Sheehan JK, Oates K, Carlstedt I. 1986. Electron microscopy of cervical, gastric and bronchial mucus glycoproteins. *Biochem. J.* 239:147–53
14. Cohen M, Varki A. 2014. Modulation of glycan recognition by clustered saccharide patches. *Int. Rev. Cell Mol. Biol.* 308:75–125
15. Varki A. 1994. Selectin ligands. *PNAS* 91:7390–97
16. Hughes GW, Ridley C, Collins R, Roseman A, Ford R, Thornton DJ. 2019. The MUC5B mucin polymer is dominated by repeating structural motifs and its topology is regulated by calcium and pH. *Sci. Rep.* 9:17350
17. Karlsson KA. 1989. Animal glycosphingolipids as membrane attachment sites for bacteria. *Annu. Rev. Biochem.* 58:309–50
18. de las Rivas M, Lira-Navarrete E, Gerken TA, Hurtado-Guerrero R. 2019. Polypeptide GalNAc-Ts: from redundancy to specificity. *Curr. Opin. Struct. Biol.* 56:87–96
19. 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
20. de las Rivas M, Daniel EJP, 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
21. 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
22. Arike L, Hansson GC. 2016. The densely *O*-glycosylated MUC2 mucin protects the intestine and provides food for the commensal bacteria. *J. Mol. Biol.* 428:3221–29
23. Brockhausen I, Stanley P. 2017. *O*-GalNAc glycans. In *Essentials of Glycobiology*, ed. A Varki, RD Cummings, JD Esko, pp. 113–24. Cold Spring Harbor, NY: Cold Spring Harb. Lab.
24. Zauner G, Kozak RP, Gardner RA, Fernandes DL, Deelder AM, Wuhrer M. 2012. Protein *O*-glycosylation analysis. *Biol. Chem.* 393:687–708
25. 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.* 273:26580–88

26. Holmén Larsson JM, Thomsson KA, Rodríguez-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. Gastrointest. Liver Physiol.* 305:G357–63
27. 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
28. Robbe C, Capon C, Maes E, Rousset M, Zweibaum A, et al. 2003. Evidence of regio-specific glycosylation in human intestinal mucins: presence of an acidic gradient along the gastrointestinal tract. *J. Biol. Chem.* 278:46337–48
29. Schauer R. 2009. Sialic acids as regulators of molecular and cellular interactions. *Curr. Opin. Struct. Biol.* 19:507–14
30. Jin C, Kenny DT, Skoog EC, Padra M, Adamczyk B, et al. 2017. Structural diversity of human gastric mucin glycans. *Mol. Cell Proteom.* 16:743–58
31. 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
32. van der Post S, Hansson GC. 2014. Membrane protein profiling of human colon reveals distinct regional differences. *Mol. Cell. Proteom.* 13:2277–87
33. Holmén Larsson JM, Karlsson H, Sjovall H, Hansson GC. 2009. A complex, but uniform *O*-glycosylation of the human MUC2 mucin from colonic biopsies analyzed by nanoLC/MS^n. *Glycobiology* 19:756–66
34. 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. *Inflamm. Bowel Dis.* 17:2299–307
35. 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
36. Bry L, Falk PG, Midtvedt T, Gordon JI. 1996. A model of host–microbial interactions in an open mammalian ecosystem. *Science* 273:1380–83
37. Holmén 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
38. 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
39. Capon C, Maes E, Michalski JC, Leffler H, Kim YS. 2001. Sda-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
40. Gagneux P, Varki A. 1999. Evolutionary considerations in relating oligosaccharide diversity to biological function. *Glycobiology* 9:747–55
41. 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
42. 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
43. 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
44. Lang T, Klasson S, Larsson E, Johansson MEV, Hansson GC, Samuelsson T. 2016. Searching the evolutionary origin of epithelial mucus protein components—mucins and FCGBP. *Mol. Biol. Evol.* 33:1921–36
45. Lesch C, Goto A, Lindgren M, Bidla G, Dushay MS, Theopold U. 2007. A role for hemolectin in coagulation and immunity in *Drosophila melanogaster*. *Dev. Comp. Immunol.* 31:1255–63
46. Chen Y, Zhao YH, Kalaslavadi 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. Respir. Cell Mol. Biol.* 30:155–65

47. Eckhardt AE, Timpte 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
48. 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
49. 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
50. 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
51. Javitt G, Calvo MLG, Albert L, Reznik N, Ilani T, et al. 2019. Intestinal gel-forming mucins polymerize by disulfide-mediated dimerization of D3 domains. *J. Mol. Biol.* 431:3740–52
52. 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 IIA. *J. Biol. Chem.* 279:53857–66
53. Carlstedt I, Karlsson H, Sundler F, Fransson L. 1982. An insoluble mucin complex from rat small intestine. *Adv. Exp. Med. Biol.* 144:155–57
54. 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
55. 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
56. Recktenwald CV, Hansson GC. 2016. The reduction-insensitive bonds of the MUC2 mucin are isopeptide bonds. *J. Biol. Chem.* 291:13580–90
57. Lorand L, Graham RM. 2003. Transglutaminases: crosslinking enzymes with pleiotropic functions. *Nat. Rev. Mol. Cell Biol.* 4:140–56
58. Fernández-Blanco JA, Fakih D, Arike L, Rodríguez-Pineiro A, Martinez-Abad B, et al. 2018. Attached stratified mucus separates bacteria from the epithelial cells in COPD lungs. *J. Clin. Investig. Insights* 3:e120994
59. Rodríguez-Pineiro AM, Bergström JH, Ermund A, Gustafsson JK, Schütte 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. Gastrointest. Liver Physiol.* 305:G348–56
60. Desseyn JL, Guyonnet-Dupérat 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
61. 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
62. Ridley C, Kouvatsos 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
63. Thornton DJ, Rousseau K, McGuckin MA. 2008. Structure and function of the polymeric mucins in airways mucus. *Annu. Rev. Physiol.* 70:459–86
64. Klomp LWJ, Van Rens L, Strous GJ. 1995. Cloning and analysis of human gastric mucin cDNA reveals two types of conserved cysteine-rich domains. *Biochem. J.* 308:831–38
65. 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
66. 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.* 269:2440–46
67. 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. *PNAS* 109:5645–50
68. 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. *Sci. Rep.* 8:17503

69. Toribara NW, Roberton AM, Ho SB, Kuo WL, Gum E, et al. 1993. Human gastric mucin. *J. Biol. Chem.* 268:5879–85
70. 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. *Hum. Genet.* 129:117–28
71. 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
72. Desseyn JL, Clavereau I, Laine A. 2002. Cloning, chromosomal localization and characterization of the murine mucin gene orthologous to human *MUC4*. *Eur. J. Biochem*. 269:3150–59
73. 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
74. 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
75. 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. *Nat. Struct. Mol. Biol.* 13:71–76
76. Beatson R, Tajadura-Ortega V, Achkova D, Picco G, Tsourouktsoglou TD, et al. 2016. The mucin MUC1 modulates the tumor immunological microenvironment through engagement of the lectin Siglec-9. *Nat. Immunol.* 17:1273–81
77. 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
78. Yin BWT, Dnistrian A, Lloyd KO. 2002. Ovarian cancer antigen CA125 is encoded by the *MUC16* mucin gene. *Int. J. Cancer* 98:737–40
79. Gum JR Jr., Crawley SC, Hicks JW, Szymkowski DE, Kim YS. 2002. MUC17, a novel membrane-tethered mucin. *Biochem. Biophys. Res. Commun.* 291:466–75
80. Gum JR 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*. 266:22733–38
81. Williams SJ. 1999. The novel mucin genes down-regulated in colorectal cancer identified by differential display. *Cancer Res*. 59:4083–89
82. 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
83. 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:C457–67
84. Knowles MR, Boucher RC. 2002. Mucus clearance as a primary innate defense mechanism for mammalian airways. *J. Clin. Investig.* 109:571–77
85. Dickson RP, Erb-Downward JR, Martinez FJ, Huffnagle GB. 2016. The microbiome and the respiratory tract. *Annu. Rev. Physiol.* 78:481–504
86. 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
87. 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.* 492:331–37
88. 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
89. Trillo-Muyo S, Nilsson HE, Recktenwald CV, Ermund A, Ridley C, et 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
90. 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. Lung Cell. Mol. Physiol.* 298:L15–22
91. 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. Respir. J.* 52:1800457

92. Hansson GC. 2019. Mucus and mucins in diseases of the intestinal and respiratory tracts. *J. Intern. Med.* 285:479–90
93. 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. *PNAS* 105:15064–69
94. Schütte A, Ermund A, Becker-Pauly C, Johansson MEV, Rodríguez-Pineiro AM, et al. 2014. Microbial-induced meprin β cleavage in MUC2 mucin and functional CFTR channel are required to release anchored small intestinal mucus. *PNAS* 111:12396–401
95. Springer TA. 2014. von Willebrand factor, Jedi knight of the bloodstream. *Blood* 124:1412–25
96. Zhang X, Halvorsen K, Zhang CZ, Wong WP, Springer TA. 2009. Mechanoenzymatic cleavage of the ultralarge vascular protein von Willebrand factor. *Science* 324:1330–34
97. Van As A. 1977. Pulmonary airway clearance mechanisms: a reappraisal. *Am. Rev. Respir. Dis.* 115:721–26
98. Fakih D, Rodríguez-Pineiro AM, Trillo-Muyo S, Evans CM, Ermund E, Hansson GC. 2020. Normal murine respiratory tract has its mucus concentrated in clouds based on the Muc5b mucin. *Am. J. Physiol. Lung Cell. Mol. Physiol.* **https://doi.org/10.1152/ajplung.00485.2019**
99. Ma J, Rubin BK, Voynow JA. 2018. Mucins, mucus, and goblet cells. *Chest* 154:169–76
100. Boucher RC. 2019. Muco-obstructive lung diseases. *N. Engl. J. Med.* 380:1941–53
101. Ermund A, Schütte 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. Gastrointest. Liver Physiol.* 305:G341–47
102. 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* 107:180–88
103. Johansson M, Synnerstad I, Holm L. 2000. Acid transport through channels in the mucous layer of rat stomach. *Gastroenterology* 119:1297–304
104. Gustafsson JK, Ermund A, Ambort D, Johansson MEV, Nilsson HE, et al. 2012. Bicarbonate and functional CFTR channel are required for proper mucin secretion and link cystic fibrosis with its mucus phenotype. *J. Exp. Med.* 209:1263–72
105. Clevers HC, Bevins CL. 2013. Paneth cells: maestros of the small intestinal crypts. *Annu. Rev. Physiol.* 75:289–311
106. Propheter DC, Chara AL, Harris TA, Ruhn KA, Hooper LV. 2017. Resistin-like molecule β is a bactericidal protein that promotes spatial segregation of the microbiota and the colonic epithelium. *PNAS* 114:11027–33
107. 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
108. 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
109. 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
110. Jakobsson HE, Rodríguez-Pineiro AM, Schütte A, Ermund A, Boysen P, et al. 2015. The gut microbiota composition impairs the colon inner mucus layer barrier. *EMBO Rep.* 16:164–77
111. De Lisle RC. 2007. Altered transit and bacterial overgrowth in the cystic fibrosis mouse small intestine. *Am. J. Physiol. Gastrointest. Liver Physiol.* 293:G104–11
112. 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. Pediatr. Gastroenterol. Nutr.* 50:38–42
113. Koropatkin NM, Cameron EA, Martens EC. 2012. How glycan metabolism shapes the human gut microbiota. *Nat. Rev. Microbiol.* 10:323–35
114. Ley RE, Backhed F, Turnbaugh P, Lozupone CA, Knight RD, Gordon JI. 2005. Obesity alters gut microbial ecology. *PNAS* 102:11070–75

115. Lozupone C, Knight R. 2005. UniFrac: a new phylogenetic method for comparing microbial communities. *Appl. Environ. Microbiol.* 71:8228–35
116. 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
117. 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. *Nat. Biotechnol.* 37:186–92
118. Proctor LM, Creasy HH, Fettweis JM, Lloyd-Price J, Mahurkar A, et al. 2019. The integrative human microbiome project. *Nature* 569:641–48
119. Human Microbiome Project Consortium. 2012. Structure, function and diversity of the healthy human microbiome. *Nature* 486:207–14
120. Xiao L, Feng Q, Liang S, Sonne SB, Xia Z, et al. 2015. A catalog of the mouse gut metagenome. *Nat. Biotechnol.* 33:1103–8
121. 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
122. Derrien M, Vaughan EE, Plugge CM, de Vos WM. 2004. *Akkermansia muciniphila* gen. nov., sp. nov., a human intestinal mucin–degrading bacterium. *Int. J. Syst. Evol. Biol.* 54:1469–76
123. van Passel MWJ, Kant R, Zoetendal EG, Plugge CM, Derrien M, et al. 2011. The genome of *Akkermansia muciniphila*, a dedicated intestinal mucin degrader, and its use in exploring intestinal metagenomes. *PLOS ONE* 6:e16876
124. Bergström JH, Birchenough GMH, Katona G, Schroeder BO, Schütte A, et al. 2016. Gram-positive bacteria are held at a distance in the colon mucus by the lectin-like protein ZG16. *PNAS* 113:13833–38
125. Nyström EEL, Arike L, Ehrencrona E, Hansson GC, Johansson MEV. 2019. Calcium-activated chloride channel regulator 1 (CLCA1) forms non-covalent oligomers in colonic mucus and has mucin 2–processing properties. *J. Biol. Chem.* 294:17075–89
126. Nyström 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
127. 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
128. Corfield A. 2018. The interaction of the gut microbiota with the mucus barrier in health and disease in human. *Microorganisms* 6:78
129. Juge N. 2019. Special issue: Gut bacteria–mucus interaction. *Microorganisms* 7:6
130. Hultgren SJ, Abraham S, Caparon M, Falk P, St. Geme JW 3rd, Normark S. 1993. Pilus and nonpilus bacterial adhesins: assembly and function in cell recognition. *Cell* 73:887–901
131. Salzman NH, Ghosh D, Huttner KM, Paterson Y, Bevins CL. 2003. Protection against enteric salmonellosis in transgenic mice expressing a human intestinal defensin. *Nature* 422:522–26
132. 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
133. 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 Rep.* 27:2665–78
134. 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
135. 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
136. Johansson MEV, Gustafsson JK, Holmén 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

137. van der Post S, Jabbar KS, Birchenough GMH, Arike L, Akhtar N, et al. 2019. Structural weakening of the colonic mucus barrier is an early event in ulcerative colitis pathogenesis. *Gut* 68:2142–51
138. Birchenough GMH, Nyström ELN, Johansson MEV, Hansson GC. 2016. A sentinel goblet cell guards the colonic crypt by triggering Nlrp6-dependent Muc2 secretion. *Science* 352:1535–42
139. Grootjans J, Hundscheild 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* 62:250–58