Sucar Coating the Envelope : Glycoconjucates for Microbe-Host Crosstalk

Tytgat, Hanne L. P.

2016-11

Tytgat , H L P & de Vos , W M 2016 , ' Sucar Coating the Envelope : Glycoconjucates for Microbe-Host Crosstalk ' , Trends in Microbiology , vol. 24 , no. 11 , pp. 853-861 . https://doi.org/10.1016/j.tim.2016.0

http://hdl.handle.net/10138/228785
https://doi.org/10.1016/j.tim.2016.06.004

publishedVersion

CellPress

Opinion

Sugar Coating the Envelope: Glycoconjugates for Microbe–Host Crosstalk

Hanne L.P. Tytgat[1,2] and Willem M. de Vos[1,3,*]

Tremendous progress has been made on mapping the mainly bacterial members of the human intestinal microbiota. Knowledge on what is out there, or rather what is inside, needs to be complemented with insight on how these bacteria interact with their biotic environment. Bacterial glycoconjugates, that is, the collection of all glycan-modified molecules, are ideal modulators of such interactions. Their enormous versatility and diversity results in a species-specific glycan barcode, providing a range of ligands for host interaction. Recent reports on the functional importance of glycosylation of important bacterial ligands in beneficial and pathogenic species underpin this. Glycoconjugates, and glycoproteins in particular, are an underappreciated, potentially crucial, factor in understanding bacteria–host interactions of old friends and foes.

The Interaction Potential of Bacterial Glycoconjugates

Humans live in close contact with an enormous bacterial population tightly associated to mucosal surfaces, that is, the **microbiota** (see Glossary). Tremendous progress has been made on the identification of members of the human microbiota via vast genome-mining efforts of different research consortia, such as the EU-MetaHit and the Human Microbiome Project. The composition and coding capacity of the gut microbiota have been especially studied in-depth and revealed to consist of mainly bacteria [1–3]. These efforts have resulted in a baseline view on what is out there, or rather inside. However, how these bacteria specifically interact with their biotic environment remains an open question. Key molecules and mechanisms driving microbe–host interactions are yet to be fully unraveled. Usual suspects in this context include secreted and cell surface-associated components of bacteria. The importance of **pili** and other cell wall appendages in adhesion and interaction of bacteria with the environment has been widely shown [4]. The deep metagenomic analysis of high-level community structures, termed enterotypes, pointed towards the presence of such pili as being a major factor in survival and persistence of low-abundance species in the gut [2].

Glycoconjugates form an interesting class of molecules that can play an important role in bacteria–host interactions. Glycoconjugates comprise all glycan-modified molecules, including exo- and lipopolysaccharides, capsular polysaccharides, lipooligosaccharides, lipoglycans, peptidoglycan, teichoic acids, and **glycoproteins** [5] (Figure 1, Key Figure). A variety of studies, mainly focusing on glycoconjugates of pathogens, have shed light on the enormous diversity and versatility of these molecules in bacteria. Compared to higher organisms, bacteria are capable of producing an extraordinary amount of unique and diverse glycans, which are principally attached to the cell surface and secreted molecules. Moreover, bacteria use these glycoconjugates to establish species- and, in some cases, strain-specific barcodes on their cell surface [5,6]. These surface-glycan barcodes offer the bacteria a range of unique and specific ligands to specifically

Trends

Glycoconjugates generate a species-specific barcode on the bacterial cell surface. The extreme diversity of bacterial glycoconjugates renders them ideal ligands to establish specific interactions with the environment.

Host cells are covered with lectin receptors designed to discriminate between self and non-self glycoconjugates and signal to the immune system.

Most ground has been covered by research on glycoconjugates of species on the pathogenic side of the bacterial spectrum. Glycosylation seems to be closely intertwined with virulence. By the same token, glycosylation can be closely intertwined with symbiotic interactions of beneficial species.

Glycosylation of cell surface molecules of (beneficial) bacteria might play a crucial, yet underappreciated, role in microbiota–host interactions and offer unique insights in the understanding of these specific interactions.

[1]Laboratory of Microbiology, Wageningen University, 6708 WE Wageningen, The Netherlands
[2]Institute of Microbiology, Swiss Federal Institute of Technology, ETH Zurich, 8093 Zurich, Switzerland
[3]Faculty of Medicine, Immunobiology Research Program, Department of Bacteriology and Immunology, University of Helsinki, 00290 Helsinki, Finland

*Correspondence: willem.devos@wur.nl (W.M. de Vos).

interact with the host. Indeed, the host expresses several immune **lectin**-containing receptors to interact with bacterial glycans and thus 'sense' the presence of bacteria (Box 1). One of the best-studied lectin receptors is **DC-SIGN**, a lectin present on dendritic cells dedicated to the detection of glycoconjugates of both pathogenic and beneficial species [7,8] (Figure 1).

In view of their unique properties, we are convinced that bacterial glycoconjugates, and glycoproteins in particular, are a yet underappreciated factor governing bacteria–host interactions. Further elucidation of the functional importance and glycosylation mechanism of these structures is imperative for a more complete understanding of microbiota–host interactions and bacteria–host interactions in general. Here, some compelling cases for the importance of glycoproteins are considered, and areas for future research are highlighted.

Functional Aspects of Protein Glycosylation in Bacterial Foes

Bacterial glycoproteins constitute a recently recognized class of glycoconjugates, and most research has focused on elucidating the structure–function relations of glycosylated virulence factors of pathogens.

Research on glycoproteins of human pathogens has led to the elucidation of the main mechanisms of bacterial glycosylation and spurred the design of novel eradication strategies. Generally, all glycoconjugates are biosynthesized following two main mechanisms, namely, a sequential pathway, that is, the transfer of nucleotide-activated sugars directly to the acceptor substrate, and an *en bloc* mechanism, in which the complete glycan is synthesized on a carrier prior to transfer to its substrate. Key enzymes in both processes are **glycosyltransferases**, catalyzing the formation of glycosidic bonds. These enzymes attach glycans either to serine, threonine or in some, more rare, cases to tyrosine residues (*O*-linked glycosylation) or asparagine residues (*N*-glycosylation). For detailed information on glycoconjugate and glycoprotein biosynthesis we refer the reader to [5].

In contrast to the relatively well-elucidated biosynthesis, research towards the unveiling of the functional importance of glycoproteins has been lagging behind. The scarce available reports have revealed that pathogens exploit protein glycosylation in two main functional strategies. Bacteria modify their proteinaceous structures with glycans either to remain stealthy (e.g., *Helicobacter pylori* [9] and *Neisseria meningitidis* [10]) or, in contrast, to interact better with the host immune system (e.g., *Burkholderia cenocepacia* [11] and *Francisella tularensis* [12]). Other reported roles of protein glycosylation involve modulation of adhesion (e.g., *Haemophilus influenzae* [13] and *Escherichia coli* [14]), stability (e.g., *Campylobacter jejuni* [15]) and activity of the glycosylated substrate (e.g., *Porphyromonas gingivalis* [16]).

All these reports reflect the tight association between glycosylation and virulence, which offers interesting targets for applications such as vaccine design [5,17].

No Monopoly for Pathogens: Glycoproteins in Old Friends

Glycosylation of proteins offer bacteria a vast and versatile repertoire of ligands to modulate their interaction with the environment and to influence the biochemical properties of the substrate. Recent studies have supported the notion that also less- or non-pathogenic bacteria exploit the functionality of glycoproteins. Interestingly, these studies report on protein glycosylation in important gut microbiota members or so-called '**old friends**', including members of the Bacteroidetes and Firmicutes that make up the vast majority of cultured intestinal species [18].

A bacterium residing somewhere in the middle of the spectrum between pathogenic and beneficial bacteria is *Bacteroides fragilis.* Elucidation of a peculiar glycosylation system in this bacterium laid the foundation for the notion that also commensal bacteria can produce

Glossary

CAZy database: Carbohydrate-Active enZymes database (www.cazy.org). This is a database of all known enzyme families that catalyze the formation, breakdown, or modification of glycosidic bonds, such as glycosyltransferases and glycoside hydrolases [46].
DC-SIGN: DC-SIGN [Dendritic Cell-Specific ICAM-3 (intercellular adhesion molecule) grabbing non-integrin or CD209] is a human C-type lectin receptor present on immune cells, that is, dendritic cells and macrophages, specifically recognizing mannosylated and fucosylate structures on microorganisms [47].
Glycoconjugates: all molecules to which glycans, or more in particular sugar monomers or polymers, are covalently attached. Examples of glycoconjugates include exopolysaccharides, lipopolysaccharides, glycolipids, and glycoproteins.
Glycoproteins: proteins to which one or more sugar moieties are covalently linked. It was long thought that these structures were unique to Eukarya, until the discovery of the first bacterial glycoproteins in the 1970s. It is generally believed that over half of all proteins in nature are glycosylated [48].
Glycosyltransferases: enzymes which catalyze the formation of glycosidic linkages, that is, the transfer of sugar monomers from a donor to an acceptor substrate molecule. Together with glycoside hydrolases they are believed to constitute 1–3% of the genome of all living organisms. Glycosyltransferases can be classified based on different properties, including their fold and structure (CAZy database).
Lectins: carbohydrate-binding (glyco-)proteins which specifically recognize glycans. They occur in all domains of life and play important roles in interactions and immunity. Purified lectins are often used as tools to screen glycoconjugates.
Microbiota: all commensal, symbiotic, and pathogenic microorganisms sharing a defined niche. All plants and animals live in close contact with microorganisms. The terms 'microbiota' and 'microbiome' are often mistakenly used as synonyms, while the latter refers to the entire habitat, meaning the microorganisms, their genomes

glycoproteins [19–21]. Intriguingly, these bacteria incorporate host-derived glycans into their surface exposed glycoconjugates, including glycoproteins. This molecular mimicry to host glycans, that is, the incorporation of surface exposed L-fucose residues, is assumed to confer a colonization advantage [19].

Several intestinal Firmicutes were also found to produce glycoproteins. Especially in lactobacilli, several reports have provided experimental evidence for protein glycosylation and its functional impact. The glycosylated S-layer protein SlpA of *Lactobacillus acidophilus* NCFM was found to regulate dendritic and T cells via interaction with DC-SIGN [22]. In mice, it was moreover shown that this interaction could mitigate colitis and help to maintain a healthy microbiota and gut mucosal barrier function [23]. Recently, we discovered that the adhesive heterotrimeric SpaCBA pili of the model probiotic *Lactobacillus rhamnosus* GG are glycosylated and that this glycosylation is important for the interaction with DC-SIGN [24]. Other *L. rhamnosus* GG glycoproteins include the Msp1 peptidoglycan hydrolase, which can also influence survival of intestinal epithelial cells [25]. A homologue of this protein was also found to be glycosylated in *Lactobacillus plantarum* WCFS-1, in which ten more glycoproteins were identified [26,27]. In the latter species two glycosyltransferases related to protein modification could also be identified, based on homology with well-studied glycosyltransferases in pathogenic Firmicutes [28]. The recent sequencing of 213 lactobacilli type strains revealed that 22 out of 95 glycosyltransferase families of the **CAZy database** were represented [29]. Although glycosyltransferases are involved in biosynthesis of all glycoconjugates, this might hint towards a broader spectrum of protein glycosylating lactobacilli.

Only few studies have addressed the mechanisms of glycosylation in species regarded as being beneficial. These endeavors might be further hampered by the lack of genetic tools to study old friends and their resistance against transformation. Lactobacilli and other Firmicutes can provide a solution here and become model systems for glycosylation in old friends.

An improved knowledge of glycosylation pathways of old friends will boost the elucidation of the functional roles of glycosylation in these species. Although much ground needs to be covered still, it can be expected that protein glycosylation is a general strategy used by a vast number of bacteria to expand the range, properties, and specificity of their ligands. Research on how these bacteria benefit from this post-translational modification will undoubtedly offer new insights in their interaction potential and in bacterial glycosylation in general. Further study of glycoproteins of old friends can reveal important new aspects governing microbiota–host interactions.

General Mechanism of O-Glycosylation in Firmicutes

Looking at the Firmicutes phylum, glycosylation of several members has been documented. Bacterial protein glycosylation was even first discovered in the Firmicute *Clostridium* [30]. But what makes this phylum of special interest is the detailed elucidation of a common mechanism involved in the modification of **serine-rich repeat proteins (SRRPs)**, important adhesins of streptococci and staphylococci (*Streptococcus parasanguinis*, *Streptococcus gordonii*, *Streptococcus pneumoniae*, *Streptococcus agalactiae*, and *Staphylococcus aureus*). Enzymes and even a dedicated secretion system for glycoproteins were discovered in these species, rendering it one of the best-studied systems. Here we highlight some recent findings of importance to the broader field of bacterial glycobiology. For a detailed overview of SRRP glycosylation we refer the reader to other publications [17,31,56].

In short, the adhesins are first targeted by an *O*-GlcNAc (*N*-acetylglucosamine) transferase consisting of two proteins GtfA, the active glycosyltransferase, and GtfB, a coactivator (in some species respectively Gtf1 and Gtf2). Recent resolution of the crystal structure of the GtfA enzyme of *S. pneumoniae* revealed a β-meander 'add-on' domain (DUF1975) crucial for

and the environmental conditions as a whole [49].

Old friends: according to a hypothesis postulated by Rook, the old friends or hygiene hypothesis, Western lifestyle has depleted the exposure of humans to microorganisms. Microbial exposure, also to pathogens, is nevertheless key in the crucial education steps of the immune system (differentiating between 'good' and 'bad' microorganisms). Rook postulates that societies with higher socioeconomic and urbanization levels are more prone to the development of allergic-like reactions to microbial contact (cf. the rise in gut-related diseases) [50].

Pili: these proteinaceous cell wall appendages can occur on the surface of both Gram-positive and Gram-negative bacteria. Their main functions comprise motility, adhesion, and colonization. The best-known and best-studied pili are the sortase-dependent pili and type IV pili [51,52].

Serine-rich repeat proteins (SRRPs): are large surface-exposed glycosylated proteins exclusively found in Gram-positive species. They have been characterized in streptococci and staphylococci and are important adhesins. The exact role of their glycosylation remains to be elucidated [31].

CellPress

Key Figure

Glycoconjugates Are Important Ligands Governing Microbiota–Host Crosstalk

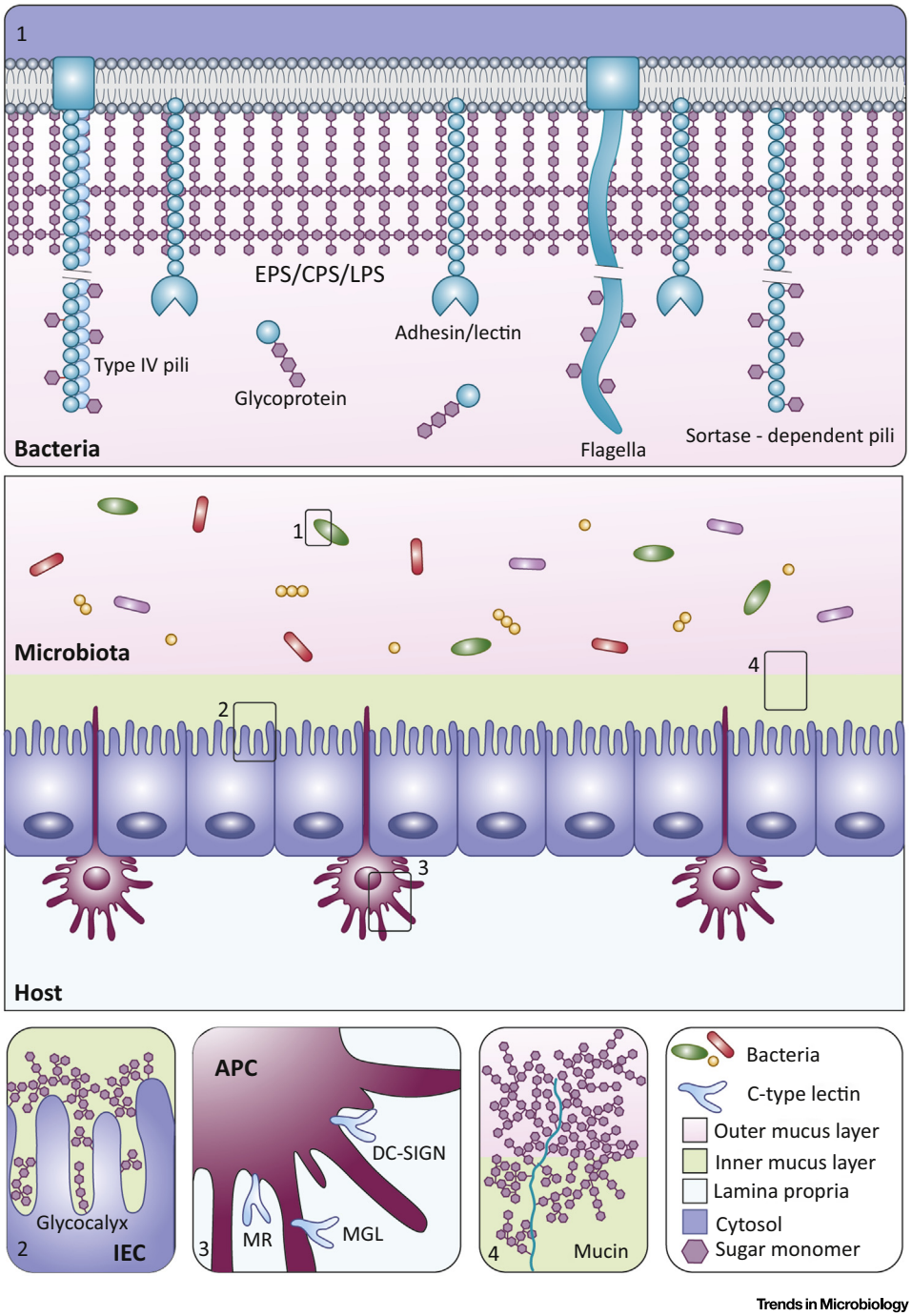

Figure 1. Bacteria are covered in glycoconjugates, forming a species- or even strain-specific barcode. These sweet ligands can include, but are not limited to, glycoproteins, pili, flagella, and capsular (CPS), exo- (EPS) and/or lipopolysaccharides

(Figure legend continued on the bottom of the next page.)

Box 1. How the Host Senses Bacterial Glycans: Immune Lectins

The host dedicates a lot of effort to the screening of glycoconjugates. From early life the immune system is trained to recognize and distinguish between self, altered self, and non-self glycosylated structures. Glycoconjugates are important pathogen-, or in general, microorganism-associated microbial patterns (MAMPs) found on the surface of invading viruses, bacteria, yeast, and parasites. The host therefore expresses several lectin receptors as part of its PRR repertoire (pattern recognition receptors), that recognize these glycan MAMPs [7,53] (see Figure 1 in main text).

C-type lectin receptors (CLR) are of special importance in humans for the recognition of bacterial glycoconjugates. These are calcium-dependent carbohydrate-binding proteins harboring a carbohydrate recognition domain (CRD) that coordinates the interaction with specific glycans. Following recognition, the glycan antigen is internalized and exposed by antigen-presenting cells. Tight binding of the bacterial glycan by the CLRs generally induces a downstream signaling pathway, resulting in the production of cytokines. CLRs recognizing mannose and fucose residues include DC-SIGN, langerin, and the mannose receptor (MR). GalNAc (*N*-acetylgalactosamine) residues are specifically bound by the MGL (macrophage galactose lectin) receptor expressed by macrophages (see Figure 1 in main text) [7,8,53].

The host relies thus on a vast repertoire of PRRs, including CLRs, for the swift recognition of invading pathogens. Most studies have focused on the elucidation of the immune response resulting from the recognition of pathogenic bacteria by CLRs, such as DC-SIGN. The detailed mechanisms related to binding of CLRs to beneficial bacteria are not yet fully understood.

complex formation with coactivator GtfB and acceptor recognition [32]. Recent insights revealed how these two subunits interact and cope with a continuously changing substrate [33]. GtfA and GtfB form a tetramer in which the conformation of GtfB, responsible for substrate recognition, is restrained by binding to GtfA, resulting in binding to unmodified substrates (Figure 2, Step I). In a second phase, GtfB also needs to recognize substrates already modified with GlcNAc residues. By breaking one of the interfaces between GtfA and GtfB, the tetramer converted to a more relaxed conformation, thus providing space to accommodate the bulkier glycosylated substrate [33].

After *O*-GlcNAc modification, a third glycosyltransferase, termed GtfC or Gtf3, adds a glucose (Figure 2, Step II). Resolution of the crystal structure of this enzyme resulted in the identification of a conserved domain for substrate recognition [34]. A glucosyltransferase, GalT1, then adds an extra glucose to the growing glycan chain (Figure 2, Step II). Study of the crystal structure of this GalT1 glucosyltransferase of *S. parasanguinis* revealed a highly conserved domain of unknown function (DUF1792). This DUF1792 domain adopts a unique fold, not found in Eukarya, and different to known glycosyltransferase folds. This fold was termed 'GT-D fold' and is characterized by binding of manganese as a cofactor and a conserved DXE and UDP-binding motif [35]. Following complete glycosylation, the proteins are secreted by a dedicated SecA2-SecY2 system [36] (Figure 2, Step III).

The findings described above illustrate the importance of the in-depth elucidation of biosynthesis mechanisms. The SRRP proteins of closely related streptococci and staphylococci have generated a unique insight in the peculiar *O*-glycosylation machinery and even revealed a unique transport mechanism for glycoproteins. These key findings have also broadened insights

(LPS) (Panel 1). The host surface is also heavily glycosylated on the glycocalyx covering intestinal epithelial cells (IECs) (Panel 2). Moreover, specialized goblet cells produce heavily glycosylated mucin proteins that are important constituents of the mucus layer, forming a barrier between the lumen and epithelial lining of the gut (Panel 4). Most microorganisms are found in the outer mucus layer, whilst the inner layer (i.e., closest to the IECs) is more sterile. The host-derived glycans are an important feeding source for microbiota, but are also recognized by specialized bacterial lectins on the cell surface. Lectins are specialized (glyco-)proteins dedicated to the recognition of specific glycan structures. These lectins enable the bacteria to prolong their residence time in the gut. The host also expresses many of these lectins as part of its immune system. Recognition of microorganism-associated glycans is key for a healthy immune system, which is reflected by the vast amount of lectin receptors expressed by the host. Antigen-presenting cells (APCs), such as macrophages, express several C-type lectin receptors, for example, DC-SIGN, MGL, and MR (Panel 3).

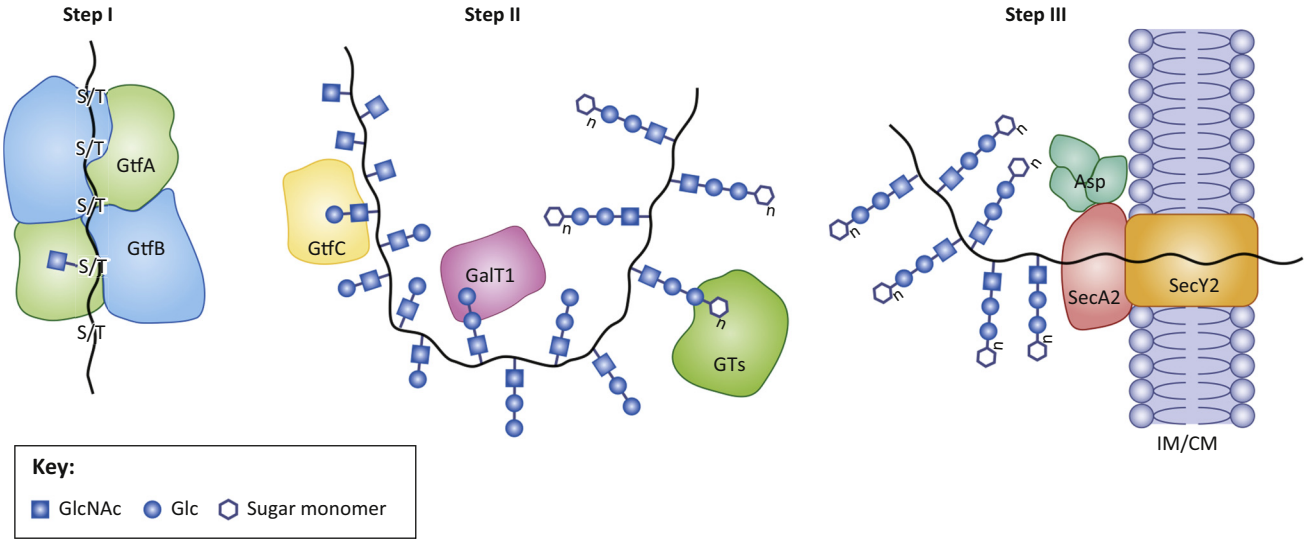

Figure 2. General Model for *O*-glycosylation in Firmicutes. The depicted three-step mechanism was found to be a common mechanism used by streptococci and staphylococci to modify serine (and threonine, S/T)-rich repeat adhesins. Identification of homologues of this mechanism in other Firmicutes might indicate that this process is a general mechanism used in Firmicutes or even Gram-positive bacteria in general for protein glycosylation. Here, the glycosylation mechanism of *Streptococcus parasanguinis* FW213 is depicted [31,35]. In Step I, a tetramer of two GtfA glycosyltransferases (GTs) and two of its coactivators involved in binding the substrate, GtfB, modify the serine-rich region of the proteins with *O*-linked *N*-acetylglucosamine (GlcNAc). In Step II, several GTs target the protein further adding two glucoses (Glc, respectively by GtfC and GalT1). The full glycan is finally synthesized by yet unknown GTs. A dedicated SecA2-SecY2 is involved in the secretion of the glycoprotein (Step III). Abbreviations: CM, cell membrane; IM, inner membrane.

in bacterial protein glycosylation beyond the Firmicutes phylum. The recent resolved crystal structures of glycosyltransferases can aid in the comprehension of catalytic mechanisms of bacterial glycosyltransferases in general, an understudied feature. However, while how glycosyltransferases recognize their target remains a major enigma, it is tempting to forecast that molecular insight in the details of *O*-linked glycan attachment and substrate selection is within reach and will promote a further understanding of these fascinating enzymes.

Homologues of these SRRP-targeting glycosyltransferases have been found in lactobacilli. A search for homologues of the GtfA and GtfB enzymes resulted in the identification of the *O*-GlcNAc transferases targeting Acm-2 in *L. plantarum* WCFS-1 [28]. Homologues of these enzymes were found in other lactobacilli, including *L. casei, L. johnsonii*, and *L. rhamnosus* GG. In some cases only one homologous protein was found, which might be the result of a fusion event. Further investigation is necessary to evaluate the target of these enzymes and the functional importance of its glycosylation. In *L. plantarum* for instance, a secreted serine- and threonine-rich peptide was found to play a role in gut homeostasis. Administration of the peptide to patients with ulcerative colitis resulted in a better maturation of dendritic cells [37,38]. The potential glycosylation status of this peptide remains to be investigated.

The identification of homologous enzymes in other species could indicate that the common mechanism for SRRP glycosylation might be a more general mechanism of *O*-glycosylation of proteins, or adhesins in particular, in Firmicutes and maybe even in Gram-positive species in general (Figure 2). More research is indispensable to evaluate these hypotheses. By contrast, caution is needed, as the strategy of relying on homologous glycosyltransferases to mine glycosylation systems in other species might result in narrow-sightedness. As

Box 2. The Gut Epithelium Is Also Covered in Glycans

The intestinal epithelial cells of the gut are covered with host glycans, consisting out of the cell surface glycocalyx and the secreted mucus layer (see Figure 1 in main text). The inner layer of the mucus layer, that is, firmly adherent to epithelial cells, is sterile, whilst the outer, less dense mucus layer is the one harboring the microbiota. Together they form an essential protective barrier between the gut lumen and intestinal epithelial cells [54,55]. But the host glycans also provide the microbiota with a glycan-rich environment. Host glycans are broken down by mucus-degrading bacteria, releasing sugars from the complex glycans for further digestion by other bacteria of the microbiota. Important mucus-degrading bacteria include *Bacteroides thetaiotaomicron, Akkermansia muciniphila* and *Ruminococcus gnavus* [56]. The availability of glycans in the gut has a large impact on the microbiota and can modulate its constitution [57].

The presence of the microbiota also influences intestinal glycosylation. Recent work indicates that *R. gnavus* influences the glycosylation pattern and production of intestinal mucus by goblet cells [58]. *Lactobacillus casei* and *B. thetaiotaomicron* were shown to influence galactosylation on the cell surface by influencing gene expression of glycosyltransferases and mucins [59]. The latter species also was found to regulate fucosylation in the distal gut of mice [60].

A recent study by Pickard *et al.* [41] probably provides the most striking example of the importance of host glycosylation in microbiota–host crosstalk. They showed that a prolonged exposure to Toll-like receptor ligands, that is, a systemic pathogenic load, induces rapid fucosylation of intestinal epithelial cells in the small intestine of mice [41]. This is peculiar, as this is the only part of the gut that is not constitutively covered in fucose residues. Metabolic activity of members of the microbiota then liberates the fucose residues for further digestion by other members of the microbiota. In this way the host wants to support and protect its microbiota during periods of prolonged pathogenic stress. Normally these free fucose residues are rapidly scavenged by beneficial members of the microbiota. This is important as also some important pathogens harbor fucose catabolism pathways [39,41].

Next to the important role of host glycosylation in the bacterial metabolism, the host glycans are also important docking sites for bacteria. Also, bacteria express several lectins, enabling their close association with the host. Indeed, several successful microbiota members and probiotics express surface-exposed appendages harboring mucus-binding domains (e.g., *L. rhamnosus* GG).

glycosyltransferases are highly promiscuous and diverse enzymes, it is imperative to also refrain from relying too much on the assumption of conservation. For all we know, *N*-glycosylation systems could still be discovered in Gram-positive species.

Fucosylation in Bacterial Glycoproteins: From Curiosity to Ubiquity?

The sugar L-fucose is abundantly present in the gut and can be seen as a mediator of host–microbe symbiosis (Box 2). Indeed, many commensal bacteria release fucose from the mucus, for example, *Bacteroides thetaiotaomicron*, which can then be used by other bacteria as a source of energy [39,40]. The host can even induce fucosylation in the small intestine to support the commensal gut microbiota during pathogen-induced stress [41]. But also several pathogens harbor fucose catabolism enzymes enabling them to benefit from the fucose-rich environment of the gut [40,42].

Interestingly, several bacteria have been reported to attach L-fucose-rich glycoconjugates to their cell surface. *H. pylori* uses surface fucosylation in order to remain stealth from the immune system [9]. But also commensal species, such as *B. fragilis*, build in fucose residues in their surface glycans to obtain a competitive colonization advantage in the gut [19,20]. Lactobacilli also seem to produce fucosylated glycoconjugates: the SpaCBA pili of *L. rhamnosus* GG were found to be glycosylated [24], and fucosyltransferases were found in the genomes of *L. gasseri* and *L. delbrueckii* [29].

It will be extremely interesting to see whether more peculiar examples of surface fucosylation in old friends and foes turn up in the coming years. The principle of host mimicry to enhance host interaction via adhesion and/or immune interaction provides bacteria with an elegant way to reside in the competitive environment of the gut. This notion is underlined by examples of sialylated and mannosylated bacterial glycoproteins [43–45].

Concluding Remarks

Glycoconjugates–in particular, the glycoproteins described here – provide bacteria on both the beneficial side and the more pathogenic side of the spectrum with a diverse and versatile range of ligands for close interaction with the biotic environment (Figure 1). The field of bacterial glycobiology is expanding at a fast pace, beyond the first breakthrough studies. Although long neglected, it is now generally accepted that glycoproteins are widespread, and are also found in beneficial species. Much work is currently focusing on linking genomic data to the glycosylation potential of model organisms. Once established, this will enable the mining of the vast amount of available omics data for glycosylation systems, even in novel isolates.

The scarce reports on protein glycosylation of commensals points towards primordial roles of these molecules in host interaction, that is, colonization and immune interaction (Figure 1). Future studies will undoubtedly further strengthen the notion that glycoproteins are pivotal for several gut microbiota species as to establish a tight and specific interaction with the host. Other promising research lines are the further elucidation of the glycosylation machinery in Gram-positive species and the exploration of the functional role and distribution of surface fucosylation in beneficial species (see Outstanding Questions).

In conclusion, we can state that the established field of protein glycosylation in pathogens can be used as an excellent starting point for the exploration of glycoproteins in beneficial species. In both old friends and foes it will be exciting to further research the functional role of these structures. The future of glycome and glycoproteome research of beneficial and pathogenic members of the microbiota looks promising and exciting and may uncover unique features governing microbiota–host interactions and influencing host health.

Outstanding Questions

What is the importance of glycoconjugates, and of glycoproteins in particular, in microbiota–host interactions? Can glycoconjugates indeed be considered as a missing key feature contributing to bacteria–host and microbiota–host interactions?

Are glycoconjugates underappreciated biomarkers of host health and disease? Can the glycosylation status of surface molecules of specific microbiota members be connected to host health?

How widespread are glycoproteins in beneficial bacteria? To what ends are bacteria considered to be beneficial exploiting the properties of glycoconjugates to their own advantage? Have these bacteria evolved similar functional uses as more pathogenic bacteria? Or have they developed unique applications?

Are adhesins, and by extension, proteins in general, of Firmicutes, and potentially of all Gram-positive species, glycosylated by a common conserved mechanism?

What role can omics data play in the prediction of the glycosylation potential of strains?

Several important general questions on bacterial glycosylation processes remain: for example, how do glycosyltransferases select the residues to modify? What are the consequences of culturing conditions on this post-translational modification, and how relevant are these studies for the *in vivo* situation?

Acknowledgments

This research was supported by an ERC grant Microbes Inside (250172) and the Soehngen Institute of Anaerobic Microbiology (SIAM Gravity Grant 024.002.002), funded by the Netherlands Organization for Scientific Research (NWO).

References

1. Qin, J. *et al.* (2010) A human gut microbial gene catalogue established by metagenomic sequencing. *Nature* 464, 59–65
2. Arumugam, M. *et al.* (2011) Enterotypes of the human gut microbiome. *Nature* 473, 174–180
3. The Human Microbiome Project Consortium (2012) Structure, function and diversity of the healthy human microbiome. *Nature* 486, 207–214
4. Danne, C. and Dramsi, S. (2012) Pili of Gram-positive bacteria: roles in host colonization. *Res. Microbiol.* 163, 645–658
5. Tytgat, H.L. and Lebeer, S. (2014) The sweet tooth of bacteria: common themes in bacterial glycoconjugates. *MMBR* 78, 372–417
6. Zhu, F. and Wu, H. (2015) Insights into bacterial protein glycosylation in human microbiota. *Sci. China Life Sci.* 59, 11–18
7. Rabinovich, G.A. *et al.* (2012) Glycobiology of immune responses. *Ann. N. York Acad. Sci.* 1253, 1–15
8. van Kooyk, Y. (2008) C-type lectins on dendritic cells: key modulators for the induction of immune responses. *Biochem. Soc. Trans.* 36, 1478–1481
9. Pohl, M.A. *et al.* (2009) Host-dependent Lewis (Le) antigen expression in *Helicobacter pylori* cells recovered from Leb-transgenic mice. *J. Exp Med.* 206, 3061–3072
10. Borud, B. *et al.* (2011) Genetic and molecular analyses reveal an evolutionary trajectory for glycan synthesis in a bacterial protein glycosylation system. *Proc. Natl. Acad. Sci. U.S.A.* 108, 9643–9648
11. Hanuszkiewicz, A. *et al.* (2014) Identification of the flagellin glycosylation system in *Burkholderia cenocepacia* and the contribution of glycosylated flagellin to evasion of human innate immune responses. *J. Biol. Chem.* 289, 19231–19244
12. Balonova, L. *et al.* (2012) Characterization of protein glycosylation in *Francisella tularensis* subsp. *holarctica*: identification of a novel glycosylated lipoprotein required for virulence. *Mol. Cell. Proteom.* 11, M111 015016
13. Grass, S. *et al.* (2003) The *Haemophilus influenzae* HMW1 adhesin is glycosylated in a process that requires HMW1 C and phosphoglucomutase, an enzyme involved in lipooligosaccharide biosynthesis. *Mol. Microbiol.* 48, 737–751
14. Charbonneau, M.E. *et al.* (2007) O-linked glycosylation ensures the normal conformation of the autotransporter adhesin involved in diffuse adherence. *J. Bacteriol.* 189, 8880–8889
15. Alemka, A. *et al.* (2013) (2013) N-glycosylation of Campylobacter jejuni surface proteins promotes bacterial fitness. *Infect. Immun.* 81, 1674–1682
16. Vanterpool, E. *et al.* (2005) Inactivation of *vimF*, a putative glycosyltransferase gene downstream of *vimE*, alters glycosylation and activation of the gingipains in *Porphyromonas gingivalis* W83. *Infect. Immun.* 73, 3971–3982
17. Lu, Q. *et al.* (2015) Sweet talk: Protein glycosylation in bacterial interaction with the host. *Trends Microbiol.* 23, 630–641
18. Rajilic-Stojanovic, M. and de Vos, W.M. (2014) The first 1000 cultured species of the human gastrointestinal microbiota. *FEMS Microbiol. Rev.* 38, 996–1047
19. Coyne, M.J. *et al.* (2005) Human symbionts use a host-like pathway for surface fucosylation. *Science* 307, 1778–1781
20. Fletcher, C.M. *et al.* (2011) Theoretical and experimental characterization of the scope of protein O-glycosylation in *Bacteroides fragilis*. *J. Biol. Chem.* 286, 3219–3226
21. Fletcher, C.M. *et al.* (2009) A general O-glycosylation system important to the physiology of a major human intestinal symbiont. *Cell* 137, 321–331

22. Konstantinov, S.R. *et al.* (2008) S layer protein A of *Lactobacillus acidophilus* NCFM regulates immature dendritic cell and T cell functions. *Proc. Natl. Acad. Sci. U.S.A.* 105, 19474–19479
23. Lightfoot, Y.L. *et al.* (2015) SIGNR3-dependent immune regulation by *Lactobacillus acidophilus* surface layer protein A in colitis. *EMBO J.* 34, 881–895
24. Tytgat, H.L. *et al.* (2016) Probiotic gut microbiota isolate interacts with dendritic cells via glycosylated heterotrimeric pili. *PLoS ONE* 11, e0151824
25. Lebeer, S. *et al.* (2012) The major secreted protein Msp1/p75 is O-glycosylated in *Lactobacillus rhamnosus* GG. *Microb. Cell Fact.* 11, 15
26. Fredriksen, L. *et al.* (2012) The major autolysin Acm2 from *Lactobacillus plantarum* undergoes cytoplasmic O-glycosylation. *J. Bacteriol.* 194, 325–333
27. Fredriksen, L. *et al.* (2013) *Lactobacillus plantarum* WCFS1 O-linked protein glycosylation: An extended spectrum of target proteins and modification sites detected by mass spectrometry. *Glycobiology* 23, 1439–1451
28. Lee, I.C. *et al.* (2014) GtfA and GtfB are both required for protein O-glycosylation in *Lactobacillus plantarum*. *J. Bacteriol.* 196, 1671–1682
29. Sun, Z. *et al.* (2015) Expanding the biotechnology potential of lactobacilli through comparative genomics of 213 strains and associated genera. *Nat. Commun.* 6, 8322
30. Sleytr, U.B. (1975) Heterologous reattachment of regular arrays of glycoproteins on bacterial surfaces. *Nature* 257, 400–402
31. Zhu, F. *et al.* (2015) Glycosyltransferase-mediated sweet modification in oral streptococci. *J. Dental Res.* 94, 659–665
32. Shi, W.W. *et al.* (2014) Structure of a novel O-linked *N*-acetyl-D-glucosamine (O-GlcNAc) transferase, GtfA, reveals insights into the glycosylation of pneumococcal serine-rich repeat adhesins. *J. Biol. Chem.* 289, 20898–20907
33. Chen, Y. *et al.* (2016) Mechanism of a cytosolic O-glycosyltransferase essential for the synthesis of a bacterial adhesion protein. *Proc. Natl. Acad. Sci. U.S.A.* 113, E1190–E1199
34. Zhu, F. *et al.* (2015) A conserved domain is crucial for acceptor substrate binding in a family of glucosyltransferases. *J. Bacteriol.* 197, 510–517
35. Zhang, H. *et al.* (2014) The highly conserved domain of unknown function 1792 has a distinct glycosyltransferase fold. *Nat. Commun.* 5, 4339
36. Feltcher, M.E. and Braunstein, M. (2012) Emerging themes in SecA2-mediated protein export. *Nat. Rev. Microbiol.* 10, 779–789
37. Al-Hassi, H.O. *et al.* (2014) Altered human gut dendritic cell properties in ulcerative colitis are reversed by *Lactobacillus plantarum* extracellular encrypted peptide STp. *Mol. Nutr. Food Res.* 58, 1132–1143
38. Bernardo, D. *et al.* (2012) Microbiota/host crosstalk biomarkers: regulatory response of human intestinal dendritic cells exposed to *Lactobacillus* extracellular encrypted peptide. *PLoS ONE* 7, e36262
39. Pickard, J.M. and Chervonsky, A.V. (2015) Intestinal fucose as a mediator of host–microbe symbiosis. *J. Immunol.* 194, 5588–5593
40. Pacheco, A.R. and Sperandio, V. (2015) Enteric pathogens exploit the microbiota-generated nutritional environment of the gut. *Microbiol. Spectr.* Published online June 11, 2015. http://dx.doi.org/10.1128/microbiolspec.MBP-0001-2014
41. Pickard, J.M. *et al.* (2014) Rapid fucosylation of intestinal epithelium sustains host-commensal symbiosis in sickness. *Nature* 514, 638–641
42. Stahl, M. *et al.* (2011) L-Fucose utilization provides *Campylobacter jejuni* with a competitive advantage. *Proc. Natl. Acad. Sci. U.S.A.* 108, 7194–7199
43. Morello, E. *et al.* (2015) Evidence for the sialylation of PilA, the PI-2a pilus-associated adhesin of *Streptococcus agalactiae* strain NEM316. *PLoS ONE* 10, e0138103
44. Liu, C.F. *et al.* (2013) Bacterial protein-O-mannosylating enzyme is crucial for virulence of *Mycobacterium tuberculosis*. *Proc. Natl. Acad. Sci. U.S.A.* 110, 6560–6565
45. Mistou, M.Y. *et al.* (2009) Molecular dissection of the *secA2* locus of group B *Streptococcus* reveals that glycosylation of the Srr1 LPXTG protein is required for full virulence. *J. Bacteriol.* 191, 4195–4206
46. Cantarel, B.L. *et al.* (2009) The Carbohydrate-Active EnZymes database (CAZy): an expert resource for glycogenomics. *Nucleic Acids Res.* 37, D233–D238
47. Geijtenbeek, T.B. *et al.* (2000) Identification of DC-SIGN, a novel dendritic cell-specific ICAM-3 receptor that supports primary immune responses. *Cell* 100, 575–585
48. Apweiler, R. *et al.* (1999) On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. *Biochim. Biophys. Acta* 1473, 4–8
49. Marchesi, J.R. and Ravel, J. (2015) The vocabulary of microbiome research: a proposal. *Microbiome* 3, 31
50. Rook, G.A. *et al.* (2014) Microbial 'old friends', immunoregulation and socioeconomic status. *Clin. Exp. Immunol.* 177, 1–12
51. Imam, S. *et al.* (2011) Identification of surprisingly diverse type IV pili, across a broad range of Gram-positive bacteria. *PLoS ONE* 6, e28919
52. Ton-That, H. and Schneewind, O. (2004) Assembly of pili in Gram-positive bacteria. *Trends Microbiol.* 12, 228–234
53. van Kooyk, Y. and Rabinovich, G.A. (2008) Protein–glycan interactions in the control of innate and adaptive immune responses. *Nat. Immunol.* 9, 593–601
54. Johansson, M.E. *et al.* (2015) Normalization of host intestinal mucus layers requires long-term microbial colonization. *Cell Host Microbe* 18, 582–592
55. Johansson, M.E. *et al.* (2011) The two mucus layers of colon are organized by the MUC2 mucin, whereas the outer layer is a legislator of host-microbial interactions. *Proc. Natl. Acad. Sci. U.S.A.* 108 (Suppl. 1), 4659–4665
56. Ouwerkerk, J.P. *et al.* (2013) Glycobiome: bacteria and mucus at the epithelial interface. *Best Pract. Res. Clin. Gastroenterol.* 27, 25–38
57. Koropatkin, N.M. *et al.* (2012) How glycan metabolism shapes the human gut microbiota. *Nat. Rev. Microbiol.* 10, 323–335
58. Graziani, F. *et al.* (2016) *Ruminococcus gnavus* E1 modulates mucin expression and intestinal glycosylation. *J. Appl. Microbiol.* 120, 1403–1417
59. Varyukhina, S. *et al.* (2012) Glycan-modifying bacteria-derived soluble factors from *Bacteroides thetaiotaomicron* and *Lactobacillus casei* inhibit rotavirus infection in human intestinal cells. *Microb. Infect.* 14, 273–278
60. Bry, L. *et al.* (1996) A model of host-microbial interactions in an open mammalian ecosystem. *Science* 273, 1380–1383