Bacterial Hydrolysis of Host Glycoproteins - Powerful Protein Modification and Efficient Nutrient Acquisition.

Garbe, Julia; Collin, Mattias

Published in:
Journal of Innate Immunity

DOI:
10.1159/000334775

Published: 2012-01-01

Citation for published version (APA):
Bacterial Hydrolysis of Host Glycoproteins – Powerful Protein Modification and Efficient Nutrient Acquisition

Julia Garbe    Mattias Collin

Department of Clinical Sciences, Division of Infection Medicine, Lund University Biomedical Center B14, Lund, Sweden

Key Words
Glycosylation · Bacteria · Glycan · Hydrolysis · Endoglycosidase · Exoglycosidase

Abstract
Glycoproteins are ubiquitous in nature and fundamental to most biological processes, including the human immune system. The glycoprotein carbohydrate moieties, or glycans, are very diverse in their structure and composition, and have major effects on the chemical, physical and biological properties of these glycoproteins. The hydrolysis of glycoprotein glycans by bacterial glycosidases can have dramatic effects on glycoprotein function and, thereby, be beneficial for the bacteria in different ways. This review gives an introduction to the expanding field of extracellular glycosidases from bacterial pathogens with activity on host glycoproteins, describes some known and proposed consequences for the host and the bacteria and discusses some evolutionary and regulatory aspects of bacterial glycosidases.

Introduction
Covalently attached carbohydrate chains, called glycans, are abundant on proteins within most eukaryotic organisms. Glycans are classified primarily due to the nature of their linkage. N-linked glycans begin with an N-acetylglucosamine residue that is attached to an asparagine residue by the recognition of a glycosylation sequence (Asn-X-Ser/Thr), whereas O-linked glycans start with an N-acetylglactosamine that is attached to a serine or a threonine [1, 2]. The structure of these oligosaccharides is very diverse and plays an important role in the biological function of glycoproteins [2, 3]. Indeed, appropriate glycosylation is crucial for cellular functions, hence alterations in glycosylation are associated with certain medical disorders such as autoimmunity and cancer [4, 5]. N-linked glycans have a major role in protein folding, since the glycans are involved in intermolecular protein-glycan interactions and interaction with lectin-type chaperones. But the glycans are also important for correct protein sorting and transport and are therefore crucial for the glycoprotein function [6, 7]. Another essential function of N-linked glycans is to promote protein-protein interactions mediating biological responses, e.g. the binding of the immunoglobulin Fc fragment to the Fc receptor [8, 9]. Interestingly, N-linked glycans share a common core structure (\(\text{Man}_3\text{GlcNAc}_2\)), because they are derived from a block of 14 sugars that is added to the newly synthesized protein in the endoplasmic reticulum. The glycan is then further processed in the endoplasmic reticulum and Golgi complex, which results in the diverse structure of N-linked glycans [3, 6]. In contrast to the N-linked glycans, O-linked glycans do not have typical
common structural elements except for the GalNAc that is attached to the hydroxyl group of a serine or a threonine. However, the structural diversity, similar to the N-linked glycans, is given by the attachment of terminal sugars that are typically sialic acid, fucose, galactose, N-acetylgalactosamine or N-acetylgalactosamine. O-linked glycans can be further divided into 5 groups (core 1–5) based on the sugars directly attached to the first GalNAc [3]. O-linked glycans are not as well studied, but it is known that they have important functions in providing charge, water-binding properties as well as stabilizing the glycoprotein [3]. As glycans are very abundant on eukaryotic cells, it is not surprising that these carbohydrate structures are frequent targets for the binding of pathogens and toxins [10]. For instance, glycans play a major role in the adhesion of Helicobacter pylori to gastric epithelial cells [11] and for the binding of influenza A hemagglutinin to sialic acid-containing glycans [12].

Many microbial proteins target components of the immune system and could thereby contribute to immune evasion and successful colonization and/or invasion of the host. For instance, secretion of proteases that degrade proteins of the innate and acquired immune system, such as immunoglobulins or complement factors, is an effective way to interfere with the immune response and is conducted by many bacterial pathogens such as Streptococcus pyogenes, Enterococcus faecalis, Staphylococcus aureus and Pseudomonas aeruginosa [13]. Such immunomodulation can also result from glycan hydrolysis, as almost all proteins from the innate and adaptive immune system are glycosylated and these glycans play important roles in stability and in the immune response itself [14]. In addition, glycans are an important source of nutrients during growth in vivo. Many commensals, especially from the gastrointestinal tract, express glycosidases that metabolize glycans from host glycoproteins, such as mucus [15, 16]. Nutrient acquisition via deglycosylation of glycoproteins also seems to play an important role in the survival and persistence of pathogens colonizing different niches of the human body [17, 18].

Glycoside hydrolases (glycosidases), which have the capability of hydrolyzing sugar residues from glycoproteins, are widespread among many organisms including bacteria and are classified by amino acid sequence similarities [19]. A very good overview of the 125 known glycoside hydrolase families is presented by the online web service CAZY (http://www.cazy.org/Glycoside-Hydrolases.html) initiated by Bernard Henrissat and Gideon Davies [20]. Bacterial glycosidases can be divided into exo- or endoglycosidases, depending on if they hydrolyze sugar residues from the nonreducing end (exoglycosidases) of the sugar chain or if they hydrolyze the sugar chain at defined sites within the glycan (endoglycosidases) (fig. 1, table 1). For more information about the biochemistry of bacterial glycosidases, please read Yamamoto et al. [21].

This short review describes bacterial glycosidases from pathogens that have activity on the glycans of host glycoproteins, mainly from the immune system. Effects such as immunomodulation, nutrient acquisition and improvement of adherence on host cells due to glycosidase activity are described. Finally, we discuss some evolutionary and regulatory aspects of the described glycosidases. As the focus of this review is glycosyl hydrolases, the glycosyl transferases (including sialyltransferases) are not discussed. We are aware that some of the described glycosidases have been shown to have more functions with consequences for the host and the pathogens. We cannot include all relevant information in this short review, but we hope to inspire the readership to further reading and experimentation within this evolving field.

**Immunomodulation by Streptococcus pyogenes**

*S. pyogenes* (group A Streptococcus) is a Gram-positive bacterium that causes uncomplicated infections such as pharyngitis, scarlet fever and impetigo, but can also cause very severe conditions like necrotizing fasciitis and sepsis [22]. *S. pyogenes* secretes one of the best described bacterial enzymes with glycoside hydrolase activity specific for human IgG, the secreted endo-β-N-acetylgalactosaminidase EndoS [23] (table 1, fig. 1). EndoS contains a family 18 glycoside hydrolase domain and has a size of 108 kDa. It selectively cleaves the β1–4 linkage between the two N-acetylgalactosamines leaving one N-acetylgalactosamine on the protein backbone of the γ-chain of human IgG [23]. Interestingly, EndoS requires the native fold of IgG (suggesting important protein-protein interactions) for glycan-hydrolyzing activity in contrast to the related endoglycosidases EndoF1–3 (*Elizabethkingia meningoseptica*, see below) that require or are enhanced in their activity by denaturation of the glycoprotein [24, 25].

IgG and the other immunoglobulins (IgA, IgM, IgD and IgE) are the major soluble products of the adaptive immune system and provide the host with long-term immunity against pathogens [9]. IgG is the most abundant immunoglobulin and it contains one conserved N-glycosylation site (Asn-297). The glycan of IgG is crucial for the stability and conformation of the Fc portion. Deglycosyl-
Bacterial Hydrolysis of Host Glycoproteins

Fig. 1. Schematic representation of glycan hydrolysis of N-linked and core-1 O-linked glycans. Arrows indicate cleavage sites of glycosidases (discussed in this review): with a solid line = endoglycosidases and with a dotted line = exoglycosidases, with activity only on terminal carbohydrates. Enzyme names in bold indicate cleavage sites (shown or proposed in the references). Other indicated cleavage sites are suggested cleavage sites due to similarities with other enzymes. Please note that EngPA and EngEF also have activity on core-3 O-linked glycans and that EngSP activity probably relies on the cleavage of the sialic acid by NanA. Gal = Galactose; GalNAc = N-acetylglactosamine; GlcNAc = N-acetylgalactosamine; Man = mannose; Sia = sialic acid. Symbols and coloring adhere to guidelines devised by the Consortium for Functional Glycomics (www.functionalglycomics.org/static/consortium/Nomenclature.shtml).

Sequential Deglycosylation of Glycoproteins by Streptococcus pneumoniae

*S. pneumoniae* is a Gram-positive bacterium that frequently colonizes the human nasopharynx, a major cause of pneumonia and causes other infections like otitis media, sinusitis and meningitis. *S. pneumoniae* has evolved complex mechanisms allowing evasion of opsonophagocytic killing in the human host, adherence to host epithelial surfaces and nutrient acquisition by sequential deglycosylation of human glycoproteins. The deglycosylation of IgG can have consequences such as loss of activation of the complement and loss of binding to Fc receptors [8, 9]. The activity of EndoS on opsonizing IgG has been shown to lead to a significant reduction in killing of *S. pyogenes* in blood compared to untreated IgG [26]. The effect of reduced killing of bacteria is due to the decreased ability of IgG to bind to the Fc receptor of activated monocytes after treatment with EndoS, showing that removal of the glycan from IgG leads to a reduced binding to its receptor [27]. Reduced binding to the receptor leads to a reduced phagocytosis of *S. pyogenes* by monocytes, and therefore to a better survival of the bacteria. Apart from its possible involvement in immune evasion during bacterial colonization/infection, EndoS, due to its specificity for IgG, has successfully been developed as an experimental treatment for antibody-mediated autoimmune diseases, discussed extensively elsewhere [28].

Two other enzymes with family 18 glycoside hydrolase domains and similarities to EndoS were found in streptococci that do not encode the *endoS* gene. EndoS 2 was discovered in a few M serotypes of *S. pyogenes* and shares 37.7% identity on the protein level to EndoS [unpubl. results], whereas the secreted protein EndoC from human and animal isolates of group C streptococci is 67% identical to EndoS [29]. Both enzymes are able to cleave the N-linked glycan of human IgG [unpubl. results]. Based on the similarities to EndoS, EndoS 2 and EndoC most likely cleave the polysaccharide moieties on IgG in a similar way and thereby interfere with IgG-dependent immune responses. The occurrence of EndoS-like enzymes in streptococci that do not express EndoS suggests an important function of these enzymes in streptococci. However, it was recently concluded that EndoS is not a major virulence factor of a highly virulent strain of the M1 serotype in a systemic mouse model [30]. Nevertheless, EndoS hydrolysis of IgG glycans might still be an important mechanism to evade opsonophagocytosis in the human host or to gain nutrients during colonization and localized infections, but this still has to be investigated.
is accomplished by exo- and endoglycosidases specific to N- and O-linked glycans. Many glycosidases from *S. pneumoniae* have been identified and many of these have been studied in detail. Since we cannot include all known pneumococcal enzymes and their functions, we refer to an excellent review published recently by S.J. King [31].

The neuraminidase A (NanA) is a well-characterized, surface-associated virulence factor of *S. pneumoniae* that catalyzes the release of sialic acid that is α2–3 or α2–6 linked to galactose [32, 33]. BgaA (247 kDa, GH2), on the other hand, has β-galactosidase activity and is responsible for the removal of terminally linked galactose that is β1–4 linked to N-acetylglucosamine [34]. StrH is a 144-kDa N-acetylglucosaminidase with a family 20 glycoside hydrolase domain that removes terminally linked N-acetylglucosamines from human glycoproteins which are 1-linked to mannose [35] (fig. 1, table 1). The combined activity of these LPXTG cell wall-anchored exoglycosidases enables *S. pneumoniae* to sequentially deglycosylate human glycoproteins such as IgA1, human lactoferrin (hLF), human secretory component (hSC) and 1-acid glycoprotein [17, 36]. *S. pneumoniae* is also able to utilize

<table>
<thead>
<tr>
<th>Table 1. Summary of glycosidases and their characteristics described in this review</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glycosidase</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td>EndoS</td>
</tr>
<tr>
<td>EndoS49</td>
</tr>
<tr>
<td>EndoC</td>
</tr>
<tr>
<td>NanA</td>
</tr>
<tr>
<td>BgaA</td>
</tr>
<tr>
<td>StrH</td>
</tr>
<tr>
<td>EndoD</td>
</tr>
<tr>
<td>EngSP</td>
</tr>
<tr>
<td>EndoE</td>
</tr>
<tr>
<td>EngEF</td>
</tr>
<tr>
<td>PPA1560</td>
</tr>
<tr>
<td>EngPA</td>
</tr>
<tr>
<td>EndoF1</td>
</tr>
<tr>
<td>EndoF2</td>
</tr>
<tr>
<td>EndoF3</td>
</tr>
<tr>
<td>SiaC</td>
</tr>
<tr>
<td>GpdG</td>
</tr>
<tr>
<td>NanH</td>
</tr>
</tbody>
</table>

---

*a* According to the classification of CAZY (http://www.cazy.org/Glycoside-Hydrolases.html). *b* Formerly known as *Flavobacterium meningosepticum*. TF = Transferrin, hSC = human secretory component, hLF = human lactoferrin.
complex N-linked glycans from the human glycoprotein α1-acid glycoprotein as a carbon source in vitro, suggesting that it gains nutrients in vivo by glycoprotein hydrolysis [17]. The α1-acid glycoprotein is an acute-phase protein with highly sialylated complex-type N-linked glycans which increase in concentration in response to tissue injury, inflammation and infection [37]. Interestingly, α1-acid glycoprotein also has immunomodulatory functions, such as the inactivation of platelet aggregation that is dependent on the sialic acids of α1-acid glycoprotein [38]. But it is unknown if the deglycosylation of α1-acid glycoprotein diminishes its function. The activity of S. pneumoniae on the host airway defense molecules hLF, hSC and IgA1 has been suggested to contribute to the persistence of S. pneumoniae in the nasopharynx [36].

Additionally, the combined activities of the above-mentioned S. pneumoniae exoglycosidases promote resistance to opsonophagocytic killing [39]. The complement component C3 coats the pneumococcal surface in vivo, which ensures recognition by neutrophils and finally leads to opsonophagocytosis. It has been shown that the activity of NanA, BgaA and StrH on an until-now unspecified human serum glycoconjugate (not IgG) leads to a reduction of C3 deposition on the bacterial surface and therefore to a decreased clearance function of the host immune system via opsonophagocytosis [39]. Evading opsonophagocytosis most likely increases the survival and persistence of the pathogen within the human host.

S. pneumoniae also expresses an endo-β-N-acetylglucosaminidase, named EndoD, with a size of 178 kDa and 32% sequence identity to the glycosidase EndoA from the soil bacterium Arthrobacter protophormiae. EndoD has an LPXTG motif typical for cell wall–anchored proteins and shows activity on the chitobiose core of asparagine-linked oligosaccharides [40]. This endoglycosidase works synergistically with the three pneumococcal exoglycosidases mentioned above to remove complex type oligosaccharides from transferrin, fetuin and IgG [40, 41]. The activity on IgG might lead to a similar effect, as shown for EndoS activity [42]. However, it was recently shown that EndoD contains a family 85 glycoside hydrolase domain instead of the family 18 domain found in EndoS. However, GH85 enzymes have a similar configuration of catalytic amino acids as GH18 enzymes, with an asparagine instead of an aspartate residue, and therefore act on the glycosidic linkage between the two N-acetylglucosamines of N-linked glycans [43].

Based on the enzyme EngBF from the nonpathogen Bifidobacterium longum, similar endo-α-N-acetylglactosaminidases with GH101 domains were detected in bacteria like S. pneumoniae (EngSP), E. faecalis (EngEF) and Propionibacterium acnes (EngPA), as well as from other nonpathogenic bacteria with amino acid identities between 28 and 49% [44]. All these enzymes show activity on core-1 O-linked glycans. EngSP is surface-associated and acts together with the above-mentioned neuraminidase NanA to hydrolyze sialylated core-1 O-linked glycans. Moreover, it was suggested that this enzyme is common to all S. pneumoniae strains. This O-glycosidase has an important function for the pathogenesis of S. pneumoniae, as it contributes to the adherence to human airway epithelial cells and therefore to the colonization of the upper respiratory tract [45]. The mechanism is not fully understood yet, but it has been suggested that the activity of the glycosidases exposes receptors of S. pneumoniae on the epithelial surface [45]. Interestingly, it was also described for NanA that its activity on sialylated glycoconjugates on epithelial cells has a potential role in exposing a receptor for adherence, since a nanA mutant was described as showing a significantly reduced adherence to human airway epithelial cells, than the wild type [36].

**Sialidase Activity in Streptococcus oralis**

S. oralis belongs to the normal flora of the oral cavity in humans, but is also associated with systemic diseases such as endocarditis and septicemia. This Gram-positive bacterium is able to utilize the complex-type glycans of human α1-acid glycoprotein as well as mucin to sustain growth [46, 47]. Several enzymes with sialidase, N-acetylglucosaminidase and β-galactosidase activity were detected in the culture of S. oralis supplemented with α1-acid glycoprotein, suggesting that the glycans provide the bacterium with fermentable carbohydrates essential for growth during infection [46]. However, until now, only one secreted sialidase has been purified from the culture medium. It was characterized on protein level as a 144-kDa protein that releases N-acetyleneuraminic acid from different glycoconjugates like the above-mentioned α1-acid glycoprotein and bovine submaxillary mucin [48]. As described above in the chapter describing pneumococcal enzymes, α1-acid glycoprotein has an immunomodulatory function. However, whether or not the described sialidase has any effect on the immunomodulatory function of α1-acid glycoprotein is unknown.

This sialidase might also be in complex with other enzymes that have glycosidic and proteolytic activity as described for C. canimorsus [49] (see below), or it works synergistically with other secreted or surface-associated en-
zymes as does \textit{S. pneumoniae} [17, 36, 39]. Sialidases are not only involved in nutrient acquisition during growth in vivo, they are also defined as a major virulence factor involved in bacterial invasion of the host. It has been suggested that sialidase activity during an infection with \textit{S. oralis} could, for instance, effect the bacterial binding to epithelial cells for host invasion, decrease the half-life of glycoproteins and increase the formation of immune complexes [50]. There is currently no information concerning the genes encoding the described enzymes or when the proteins are expressed. This would be necessary to understand the functions of glycosidases from \textit{S. oralis} in more detail.

\textbf{N- and O-Glycosidases from \textit{Enterococcus faecalis}}

\textit{E. faecalis} is a Gram-positive bacterium and a component of the normal flora in the human gastrointestinal tract. Despite the fact that \textit{E. faecalis} is a well-known member of the microbial consortium in the human gut, it has become a dreaded multiresistant nosocomial pathogen causing infectious endocarditis, urinary tract infections and sepsis [51, 52]. \textit{E. faecalis} secretes an endo-\begin{math}\beta\end{math}-N-acetylgalactosaminidase (EndoE), identified due to similarities to EndoS. In contrast to EndoS, EndoE (88 kDa) has a unique feature: it combines two enzymatic domains with different glycoside hydrolase activities. The \begin{math}\alpha\end{math}-domain of EndoE contains a family 18 glycoside hydrolase motif and has activity on the \begin{math}N\end{math}-linked mannose-rich glycans of the high-mannose model glycoprotein RNaseB, cleaving the B1–4 linkage between the two \begin{math}N\end{math}-acetylglucosamines. The \begin{math}\beta\end{math}-domain, on the other hand, contains a family 20 glycoside hydrolase domain and is able to release the glycans from the immunoglobulin IgG in the same manner as EndoS [53] (fig. 1, table 1). This activity of EndoE most likely has the same functional consequences for IgG as described for EndoS hydrolysis of IgG. Interestingly, the GH18 \begin{math}\alpha\end{math}-domain of EndoE shows similarities to EndoS while the GH20 \begin{math}\beta\end{math}-domain with activity on IgG instead shows similarities to the \begin{math}\beta\end{math}-N-acetylgalactosaminidase StrH from \textit{S. pneumoniae} [35, 53]. However, recent studies showed that the \begin{math}\alpha\end{math}-domain of EndoE has glycosidase activity on the two biantennary \begin{math}N\end{math}-linked glycans of human lactoferrin [unpubl. data]. Lactoferrin and transferrin have multiple biological functions in blood and mucosal surfaces and are considered as components of the innate immune system with antimicrobial activity [54]. The effect of deglycosylation on the function of lactoferrin is at present unknown.

\textit{E. faecalis} also expresses the endo-\begin{math}\alpha\end{math}-N-acetylglactosaminidase EngEF with a GH101 domain. As described for EngSP from \textit{S. pneumoniae} and EngPA from \textit{P. acnes}, all enzymes of the Eng-family show activity on core-1 \begin{math}O\end{math}-linked glycans, and it was also shown that EngEF from \textit{E. faecalis} has activity on core-3 \begin{math}O\end{math}-linked glycans. There is currently no information known about the biological function of EngEF; however, it cannot be excluded that this enzyme contributes to adherence as shown for EngSP from \textit{S. pneumoniae} [45].

\textbf{Sialidase and O-Glycosidase Activity of \textit{Propionibacterium acnes}}

Sialidase activity of the Gram-positive skin commensal and opportunistic pathogen \textit{P. acnes} was described as possibly being involved in degrading host tissues. Three potential sialidases (PPA1560, PPA684 and PPA685) were identified on the genome level, as well as a sialic acid transporter (PPA686), which indicates that \textit{P. acnes} is able to import sialic acids either to gain nutrients or to evade the immune system [55, 56]. Recently, one sialidase (PPA1560) was characterized and shown to lead to the release of sialic acids from human sebocytes. Strikingly, the loss of sialic acids from the epithelial surface of the sebocytes due to PPA1560 treatment led to an increased interaction of \textit{P. acnes} with the cells, possibly increasing the adhesion of \textit{P. acnes} and thereby sensitizing these cells to \textit{P. acnes} cytotoxicity [57]. Furthermore, \textit{P. acnes} expresses the endo-\begin{math}\alpha\end{math}-N-acetylglactosaminidase EngPA with a GH101 family domain. Similar to EngEF from \textit{E. faecalis}, EngPA has activity on core-1 and core-3 \begin{math}O\end{math}-linked glycans. This enzyme could be involved in adherence, working together with a sialidase (as shown above for EngSP from \textit{S. pneumoniae}) that acts synergistically with the sialidase NanA to cleave sialylated core-1 \begin{math}O\end{math}-linked glycans [45].

\textbf{The Classic \textit{N}-Glycosidases from \textit{Elizabethkingia meningoseptica}}

The Gram-negative bacterium \textit{E. meningoseptica} (formerly known as \textit{Flavobacterium meningosepticum}) secretes three of the best-characterized endo-\begin{math}\beta\end{math}-\begin{math}N\end{math}-acetylglucosaminidases (EndoF1–3) that are widely used for the characterization of glycoproteins with \begin{math}N\end{math}-linked glycans. These enzymes belong to the same family as EndoS from \textit{S. pyogenes} and EndoE from \textit{E. faecalis}, with conserved
family 18 glycosyl hydrolase motifs. Similar to EndoS and EndoE, EndoF1–3 cleaves the β1–4 linkage between the two N-acetylgalacosamines of N-linked glycans, including glycans from IgG, fetuin and α1-acid glycoprotein [58]. As mentioned above, in contrast to the specificity of for instance EndoS and EndoE to IgG, EndoF glycosidases are not specific to the proteins but rather to the glycan structure itself. The differences between EndoF1–3 are their specificities for the type of glycan structure [59]. EndoF1 is closely related to EndoH from Streptomyces plicatus and shows activity only on high-mannose and hybrid oligosaccharides, whereas EndoF2 and EndoF3 show activity on complex biantennary as well as bi- and triantennary glycans, respectively [60]. Although not well known as a human pathogen, E. meningoseptica can cause meningitis in newborns as well as other infections like pneumonia and endocarditis in immunocompromised patients [61]. The contribution of the 3 mentioned endoglycosidases in the pathogenicity of this bacterium has not been studied.

**Glycoprotein Deglycosylation Complex of Capnocytophaga canimorsus**

Another uncommon zoonotic pathogen is the Gram-negative bacterium C. canimorsus, which resides in the oral cavity of cats and dogs. Although infections of humans are rare, this bacterium can cause meningitis and fulminant septicemia with a high mortality rate [49, 62]. *C. canimorsus* is able to grow in the presence of host glycoproteins from mammalian cells including phagocytes by a mechanism that is dependent on the activity of the sialidase C (SiaC). SiaC does not supply nutrients, but it removes the terminal sialic acids and makes other carbohydrates accessible that can be metabolized by *C. canimorsus* [63]. A mouse model with a SiaC-deficient mutant strain confirmed that SiaC is a virulence factor that contributes to the persistence of this pathogen [63]. However, SiaC works together with four outer-membrane lipoproteins (GpdD–G) and one porin-like protein (GpdC) that are encoded in a so-called polysaccharide utilization locus (PUL). GpdD, E and F probably represent glycan-binding proteins. gpdG encodes an endo-β-N-acetylglucosaminidase and has activity on IgG and fetuin. The endoglycosidase GpdG cleaves the glycans between the two N-acetylgalacosamines and the sugar moiety is then transported into the cell with the help of GpdC, which is very likely a porin of the outer membrane. SiaC, located in the periplasm of this Gram-negative bacterium, cleaves the sialic acid of the glycan and it is assumed that other intracellular enzymes further process the remaining oligosaccharide [49]. The activity on human IgG, as it was also shown for EndoS from *S. pyogenes*, might contribute not only to nutrient acquisition but could also serve as an immunosuppressive mechanism. However, the molecular mechanism is not yet understood. This is the first report of a glycoprotein deglycosylation complex and the cooperation between different proteins and enzymes to deglycosylate host glycoproteins, which raises the question of whether or not other glycosidases are also part of complex systems.

**NanH from Vibrio cholerae Enhances Cholera Toxin Activity**

Another well-investigated sialidase, NanH, is secreted by the Gram-negative intestinal pathogen *V. cholerae*, which causes the water-borne disease cholera. NanH has a size of 83 kDa and is encoded on the pathogenic island of toxigenic *V. cholerae* strains [64], acts synergistically with the cholera toxin and enhances the severity of the disease [65]. NanH cleaves the sialic acids of higher-order gangliosides and thereby un masks the receptor of the cholera toxin, GM1. The binding and internalization of the toxin into the epithelial villus is therefore significantly enhanced by the activity of NanH [65]. However, the gene nanH is part of a gene cluster, which contains a series of enzymes that might be involved in the utilization of the sialic acids that could be important as a carbon and nitrogen source for *V. cholerae* during infection [64]. *V. cholerae* also produces an endo-β-N-acetylhexosaminidase that, in combination with a neuraminidase and a protease, is able to degrade mucin. It is thought that the activities of these enzymes help *V. cholerae* to colonize the intestinal epithelium which is covered with sialomucin [66]. Unfortunately, no other studies have been done on the endo-β-N-acetylhexosaminidase and it cannot be excluded that the neuraminidase activity discovered in the published study can be attributed to NanH.

**Phylogenetic Analysis of Bacterial Glycosidases**

The glycosidases described in this review are very diverse and have different functions. In order to obtain some insights into the evolution of these enzymes, we performed a phylogenetic analysis to elucidate if the en-
zymes are related and have a common evolutionary background, or if they have evolved independently.

As shown in the phylogenetic analysis in figure 2, several clusters of proteins with similarities can be detected and, not surprisingly, these similarities are mainly based on the glycoside hydrolase domains or functions, respectively. In particular, the enzymes with a GH18 domain like EndoS, EndoC, EndoS₂, EndoF₂, the α-domain of EndoE and EndoF₃ cluster together, suggesting that they are closely related. The only two enzymes with a GH18 domain not included in this cluster are EndoF₁ and GpdG, suggesting that these enzymes have an origin different from the EndoS group. Another enzyme that hydrolyzes the N-linked glycans between the two GlcNAc similar to EndoF₁ and the EndoS group is EndoD, which contains a GH85 domain. As described above, this enzyme has a similar configuration of catalytic amino acids and catalyzes the glycan hydrolysis in a similar way, but probably evolved from a different origin. Another striking feature of all these enzymes is their ability to cleave the N-linked glycans of IgG (see also fig. 1, table 1), and it seems very likely that immunomodulation of IgG, the most abundant immunoglobulin, is a very effective way to evade the immune system or even to gain nutrients. However, as the enzymes of different GH families show activity on IgG, it may be, from the evolutionary point of view, an old mechanism of immunomodulation that evolved more than once. The study of the different catalytic mechanisms of these enzymes could help the understanding of their evolutionary background.

Fig. 2. Phylogenetic analysis of bacterial glycosidases. They were first aligned using the ClustalW algorithm of the MacVector™ 11.0.2 sequence analysis software suite. Phylogenetic trees were generated using the neighbor-joining method, gaps were distributed proportionally with uncorrected p values and were validated by bootstrap analysis (32,000 repetitions). The GenBank accession numbers of the enzymes are indicated in parentheses together with the organism name.

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Accession Number</th>
<th>Domain</th>
</tr>
</thead>
<tbody>
<tr>
<td>EngSP (S. pneumoniae)</td>
<td>YP_873926.1</td>
<td>GH101</td>
</tr>
<tr>
<td>EngPA (P. acnes)</td>
<td>YP_056270.1</td>
<td>GH33</td>
</tr>
<tr>
<td>EngEF (Enterococcus faecalis)</td>
<td>NP_815498.1</td>
<td>GH18</td>
</tr>
<tr>
<td>Sialidase PPA1560 (P. acnes)</td>
<td>YP_056262</td>
<td>GH33</td>
</tr>
<tr>
<td>EndoF₁ (E. meningoseptica)</td>
<td>P36911</td>
<td>GH18</td>
</tr>
<tr>
<td>NanA (S. pneumoniae)</td>
<td>ACJ76926</td>
<td>GH33</td>
</tr>
<tr>
<td>SiaC (C. canimorsus)</td>
<td>YP_004739708</td>
<td>GH33</td>
</tr>
<tr>
<td>NanH (V. cholerae)</td>
<td>AAW31751</td>
<td>GH33</td>
</tr>
<tr>
<td>EndoD (S. pneumoniae)</td>
<td>BAB62042</td>
<td>GH85</td>
</tr>
<tr>
<td>BgaA (S. pneumoniae)</td>
<td>NP_345155</td>
<td>GH2</td>
</tr>
<tr>
<td>EndoF₂ (E. meningoseptica)</td>
<td>P36913</td>
<td>GH18</td>
</tr>
<tr>
<td>EndoE alpha (E. faecalis)</td>
<td>AAR20477</td>
<td>GH18</td>
</tr>
<tr>
<td>EndoF₃ (E. meningoseptica)</td>
<td>P36912</td>
<td>GH85</td>
</tr>
<tr>
<td>EndoS49 (S. pyogenes)</td>
<td>YP_002286383</td>
<td>GH2</td>
</tr>
<tr>
<td>EndoS (S. pyogenes)</td>
<td>AF296340</td>
<td>GH85</td>
</tr>
<tr>
<td>EndoC (C. streptococci)</td>
<td>ADC53484</td>
<td>GH85</td>
</tr>
<tr>
<td>HexB (H. sapiens)</td>
<td>NP_000512</td>
<td>GH20</td>
</tr>
<tr>
<td>HexA (H. sapiens)</td>
<td>AAD13932</td>
<td>GH20</td>
</tr>
<tr>
<td>StrH (S. pneumoniae)</td>
<td>YP_815512</td>
<td>GH20</td>
</tr>
<tr>
<td>EndoE beta (E. faecalis)</td>
<td>AAR20477</td>
<td>GH18</td>
</tr>
<tr>
<td>GpdG (C. canimorsus)</td>
<td>YP_004740097.1</td>
<td>GH18</td>
</tr>
</tbody>
</table>

As shown in the phylogenetic analysis in figure 2, several clusters of proteins with similarities can be detected and, not surprisingly, these similarities are mainly based on the glycoside hydrolase domains or functions, respectively. In particular, the enzymes with a GH18 domain like EndoS, EndoC, EndoS₂, EndoF₂, the α-domain of EndoE and EndoF₃ cluster together, suggesting that they are closely related. The only two enzymes with a GH18 domain not included in this cluster are EndoF₁ and GpdG, suggesting that these enzymes have an origin different from the EndoS group. Another enzyme that hydrolyzes the N-linked glycans between the two GlcNAc similar to EndoF₁ and the EndoS group is EndoD, which contains a GH85 domain. As described above, this enzyme has a similar configuration of catalytic amino acids and catalyzes the glycan hydrolysis in a similar way, but probably evolved from a different origin. Another striking feature of all these enzymes is their ability to cleave the N-linked glycans of IgG (see also fig. 1, table 1), and it seems very likely that immunomodulation of IgG, the most abundant immunoglobulin, is a very effective way to evade the immune system or even to gain nutrients. However, as the enzymes of different GH families show activity on IgG, it may be, from the evolutionary point of view, an old mechanism of immunomodulation that evolved more than once. The study of the different catalytic mechanisms of these enzymes could help the understanding of their evolutionary background.

The β-domain of EndoE from E. faecalis and StrH from S. pneumoniae share a GH20 domain and cluster together. It has previously been shown that enzymes with a GH20 domain are highly conserved and that they have a common evolutionary origin [67]. We included two human hexosaminidases (HexA and HexB) with GH20 domains to our phylogenetic analysis and we could show that the β-domain of EndoE and StrH are related to the human enzymes, underscoring their common evolutionary background. Furthermore, this might indicate that host genes could have at some point during coevolution been horizontally transferred to the colonizing bacteria.
The Eng-family of glycosidases (GH101) that has activity on O-linked glycans and which can be found in different pathogenic as well as nonpathogenic bacteria, cluster together because they share a conserved GH101 domain and have similar activities on O-linked glycans. Nevertheless, all identified enzymes with a GH101 domain are present only in bacteria, indicating that they might be unique to bacteria and play an important role in the bacterial lifestyle. O-glycosidases could be important not only to contribute to the pathogenesis of the bacteria, but also to gain nutrients from glycoproteins such as mucins. Interestingly, the sialidases that share a GH33 domain do not cluster together, suggesting that these enzymes are very diverse and might have a different evolutionary origin. A possible explanation could be that sialic acids are present on the terminal ends of glycans that are very diverse and could therefore have led to the evolutionary diversification of sialidases.

**Regulation of Glycosidase Expression**

Very little is known about the regulation of extracellular glycosidases in bacteria. More knowledge could help us to understand the importance of glycosidases for the bacterium during colonization and infection. For instance, it was shown that the genes encoding EndoS and EndoE have so-called catabolite response elements in their promoters, indicating that they are regulated via carbon catabolite repression (CCR) [68, 69]. CCR is a major regulatory system in bacteria and regulates the expression of genes involved in the use of carbon sources. Recently, it was suggested that there is a close link between basic metabolic processes and the pathogenesis of bacteria because several known virulence factors in bacteria are under the control of the CCR system [69]. Regulation via a major regulatory system like CCR indicates that the enzymes are key molecules in the life of bacteria.

Moreover, there is evidence that the glycan itself can trigger the expression of genes that encode proteins involved in glycan metabolism. It has been shown for the human gut bacterium, *Bacteroides thetaiotaomicron*, that O- and N-linked glycans from mucin induce the expression of several PULs [70]. As described for *C. canimorsus*, PULs encode a complex system of enzymes involved in the deglycosylation of host glycoproteins [49]. Many PULs also contain a sensory regulator, and it is assumed that this regulator can sense glycans and induce the expression of genes involved in the metabolism of glycans [70].

**Concluding Remarks**

A large number of known or putative glycosyl hydrolase genes can be found in the genomes of both bacterial pathogens and commensals. The high abundance and, in many cases, high level of conservation indicate that these enzymes are important molecules for colonizing and infecting the human host. The immunomodulation of host proteins by single enzymes or even by protein complexes is widely distributed among bacteria and is very diverse. Enzymes with activity on the glycans of key proteins from the immune system could contribute to the pathogenesis of these bacteria either by evading the host response, or in a more subtle way, by gaining nutrients and/or supporting the survival in the human body.

The field of glycosidases from pathogens with activity on human glycoproteins is developing but should be expanded in relation to the enzymes from nonpathogenic bacteria colonizing the host. It could be very interesting to investigate if the glycosidases of commensals also have related activities and to elucidate how similar they are to the enzymes from pathogens, structurally and functionally. Such comparisons could potentially reveal how these enzymes evolved with respect to selection pressure originating from the immune system and/or from nutrient limitation. Bacterial enzymes with activity on O-linked glycans, represented by the Eng-family of glycosidases, are most likely just the tip of the iceberg with many more to be discovered in both pathogens and commensals.

**References**


Bacterial Hydrolysis of Host Glycoproteins


60 Trimble RB, Tarentino AL: Identification of distinct endoglycosidase (endo) activities in Flavobacterium meningosepticum: endo F1, endo F2, and endo F3. Endo F1 and endo H hydrolyze only high-mannose and hybrid glycans. J Biol Chem 1991;266:1646–1651.


