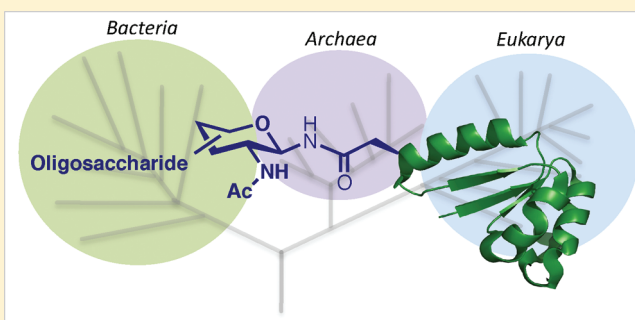


The Expanding Horizons of Asparagine-Linked Glycosylation

Angelyn Larkin[†] and Barbara Imperiali^{*,†,‡}

[†]Department of Chemistry and [‡]Department of Biology, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, Massachusetts 02139, United States

ABSTRACT: Asparagine-linked glycosylation involves the sequential assembly of an oligosaccharide onto a polyisoprenyl donor, followed by the *en bloc* transfer of the glycan to particular asparagine residues within acceptor proteins. These N-linked glycans play a critical role in a wide variety of biological processes, such as protein folding, cellular targeting and motility, and the immune response. In the past decade, research in the field of N-linked glycosylation has achieved major advances, including the discovery of new carbohydrate modifications, the biochemical characterization of the enzymes involved in glycan assembly, and the determination of the biological impact of these glycans on target proteins. It is now firmly established that this enzyme-catalyzed modification occurs in all three domains of life. However, despite similarities in the overall logic of N-linked glycoprotein biosynthesis among the three kingdoms, the structures of the appended glycans are markedly different and thus influence the functions of elaborated proteins in various ways. Though nearly all eukaryotes produce the same nascent tetradecasaccharide (Glc₃Man₉GlcNAc₂), heterogeneity is introduced into this glycan structure after it is transferred to the protein through a complex series of glycosyl trimming and addition steps. In contrast, bacteria and archaea display diversity within their N-linked glycan structures through the use of unique monosaccharide building blocks during the assembly process. In this review, recent progress toward gaining a deeper biochemical understanding of this modification across all three kingdoms will be summarized. In addition, a brief overview of the role of N-linked glycosylation in viruses will also be presented.



Glycosylation is the most abundant protein modification found in nature, occurring across all kingdoms of life.¹ The great variety of carbohydrates coupled with the vast number of possible carbohydrate linkages serve to diversify the proteome in a manner beyond what is directly encoded in the genome. The appendage of a glycan to a protein can profoundly impact structure and function, playing important roles in protein stability and rigidity, intracellular localization, cellular signaling and adhesion, and the immune response.

There are three major types of protein glycosylation that have been observed in nature: N-linked, O-linked, and GPI-anchored.² N-Linked glycosylation, the focus of this review, involves the *en bloc* transfer of an oligosaccharide onto the side chain amide nitrogen of asparagine residues within acceptor proteins. In the case of O-linked glycosylation, monosaccharides are typically added in a sequential manner onto the side chain hydroxyl oxygen atom of either serine or threonine residues, exemplified in the biosynthesis of O-linked mucin glycans,³ or may involve only a single carbohydrate, such as in dynamic O-GlcNAcylation.⁴ However, in recent years, glycosylation of other residues, including tyrosine, hydroxylysine, and hydroxyproline, has also been found.^{5,6} In the third major type of glycosylation, a complex glycosylphosphatidylinositol moiety (GPI anchor) is transferred to the C-terminus of a target protein, serving to tether it to the cell membrane.⁷ Other classes of protein glycosylation have also been described in addition to

these three main classes, including phosphoglycosylation^{8,9} and C-mannosylation,¹⁰ although the biochemical details of these modifications are still unclear and require further investigation.

While N-linked glycosylation was once believed to be limited to eukaryotes, it is now firmly established that this complex modification also occurs in bacteria and archaea.^{11,12} In the past 10 years, the field of N-linked glycosylation has witnessed enormous strides in the discovery of new and unusual carbohydrates, the elucidation of the enzymes involved in glycan assembly and processing, and understanding the biological impact that these glycan modifications have on the structure and function of target proteins. These advances have been fueled in large part by a number of technological breakthroughs, including the development of powerful mass spectrometric methods that have allowed characterization of minute quantities of low-abundance glycans and high-throughput sequencing of prokaryotic genomes, which has provided evidence that N-linked glycosylation is more widespread than initially believed.^{13,14} In this review, recent developments in the field of N-linked glycosylation across the three kingdoms of life will be discussed, with an emphasis on work conducted in the past five years. Special focus will be placed on the different means by which the glycans

Received: March 8, 2011

Revised: April 20, 2011

Published: April 20, 2011

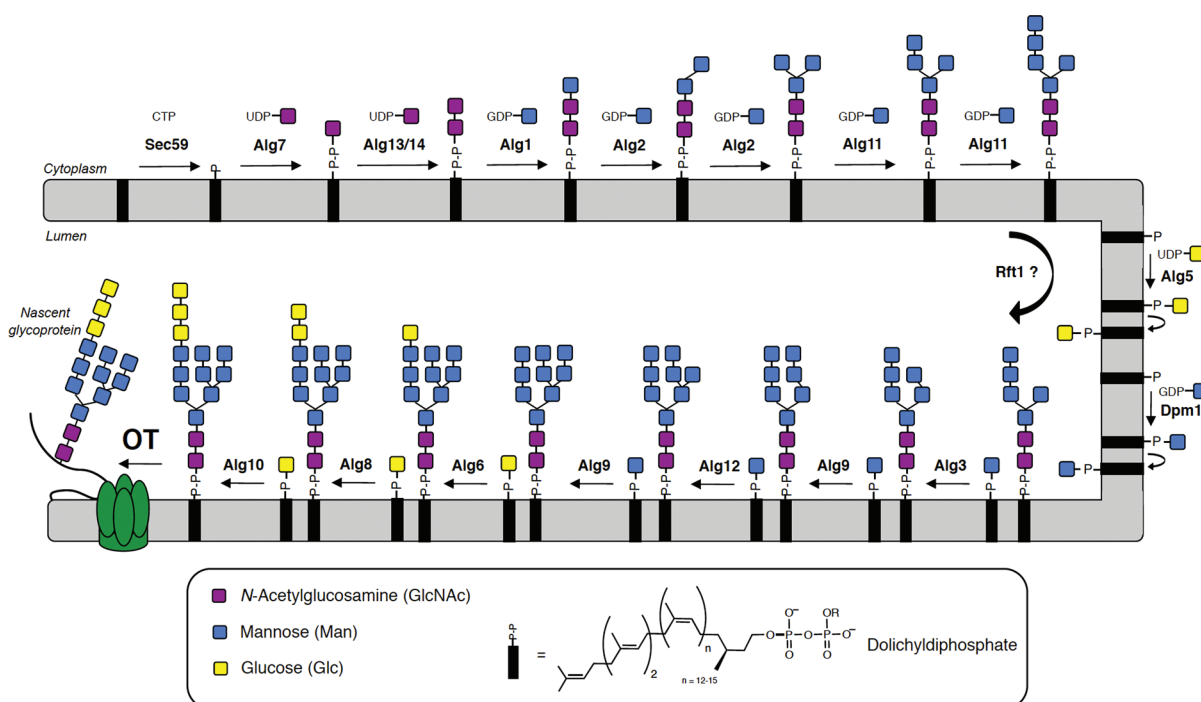


Figure 1. Pathway of N-linked glycosylation at the ER membrane in *S. cerevisiae*.

are assembled, the variety in the glycan structures, and how this structural diversity imparts function to specific glycoproteins. In addition, a brief overview of the role that N-linked glycosylation plays in viruses will be presented as well.

N-LINKED GLYCOSYLATION IN EUKARYOTES

Assembly of Eukaryotic N-Linked Glycans. In eukaryotes, N-linked glycosylation occurs at the membrane of the endoplasmic reticulum (ER) and involves the transfer of a tetradecasaccharide ($\text{Glc}_3\text{Man}_9\text{GlcNAc}_2$) from a dolichylphosphate carrier onto the amide side chain nitrogen of an acceptor protein (Figure 1). Although a majority of the genetic and biochemical characterization of this pathway has been achieved in *Saccharomyces cerevisiae*, the process is remarkably conserved in all eukaryotes, from yeast to humans.^{15,16} In *S. cerevisiae*, glycan assembly is carried out by a series of membrane-bound glycosyltransferases in the Alg family (*asparagine-linked glycosylation*), which catalyze the transfer of each monosaccharide onto a dolichylphosphate carrier. Dolichol, a long chain α -saturated polyisoprene that varies in size between 14 and 21 units depending on the cell type and species,¹⁷ is phosphorylated by a kinase (Sec59 in *S. cerevisiae*) prior to glycan assembly.^{18,19} In addition to its role as a membrane anchor, dolichylphosphate may serve other functions in the glycosylation process as well, such as promoting membrane fluidity, facilitating translocation across the ER membrane, and potentially recruiting key enzymes to the site of glycosylation.^{20–22}

The first phase of glycan assembly takes place on the cytoplasmic face of the ER membrane, where a high concentration of nucleotide sugar donors (UDP-GlcNAc and GDP-Man) act as substrates for the early glycosyltransferases. Although these enzymes have been implicated in the initial steps of glycan assembly for quite some time, only in the past five years has the biochemical activity of these glycosyltransferases been

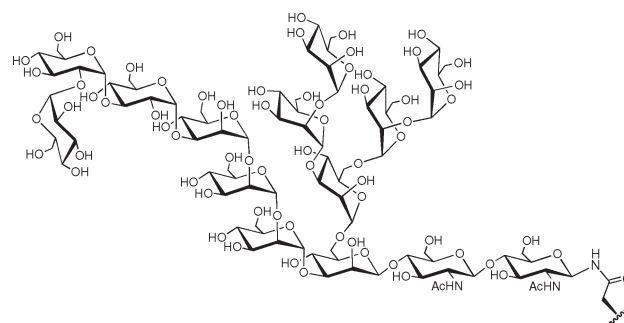


Figure 2. Structure of the eukaryotic N-linked glycan, $\text{Glc}_3\text{Man}_9\text{GlcNAc}_2$.

unequivocally established. In *S. cerevisiae*, the pathway begins with the addition of GlcNAc-1-P to the Dol-P carrier by Alg7,²³ followed by the transfer of the second GlcNAc residue by the Alg13/14 complex.^{24,25} Recent work has established that Alg13 and Alg14 interact to form a functional glycosyltransferase capable of catalyzing GlcNAc transfer.^{26,27} The initial mannosylation is performed by Alg1 and establishes the first branch point in the glycan structure.^{28,29} Alg2 and Alg11 then sequentially install the next four mannose residues, in which Alg2 conducts α -1,6 and α -1,3 mannosylations, followed by two iterative α -1,2 mannose additions by Alg11 to yield the heptasaccharide.^{15,30–32} Genetic and biochemical studies have recently determined that these early enzymes form discrete protein complexes in the ER membrane that may be important for maintaining flux through the pathway.^{33,34} After completion of the heptasaccharide, the Dol-PP-GlcNAc₂Man₅ intermediate is flipped from the cytoplasmic to the luminal face of the ER by a process that remains unclear. While genetic studies have implicated Rft1 as the required flippase,³⁵ recent biochemical evidence has contradicted this designation and suggested that Rft1 may in fact not be required for translocation of the oligosaccharide intermediate.^{36–39}

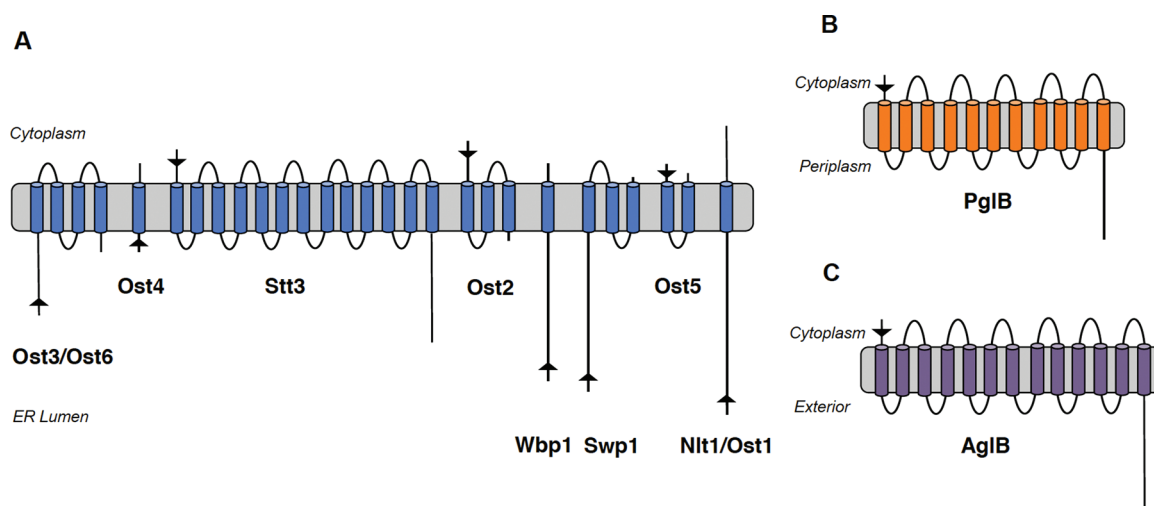


Figure 3. Comparison of the predicted membrane topology of oligosaccharyl transferases from all three domains of life: (A) OT complex from *S. cerevisiae*, (B) PglB from *C. jejuni*, and (C) AglB from *Methanococcus voltae*. The size of the membrane loops between the transmembrane domains in panels B and C is variable and has not been determined.

Further work in this area is necessary to characterize this elusive process.

Glycan assembly continues on the luminal side of the ER membrane, where the four remaining mannosylations are catalyzed by the actions of the *S. cerevisiae* enzymes Alg3, Alg9, Alg12, and Alg9;⁴⁰ Alg9 has been shown to transfer both α -1,2-linked mannoses that cap two of the three branches of the glycan.⁴¹ The three terminal glucose units are then installed onto the nascent oligosaccharide by Alg6, Alg8, and Alg10, respectively, to afford the final tetradecasaccharide Glc₃Man₉GlcNAc₂ (Figure 2).⁴⁰ Unlike the case in the first phase of the pathway, the glycosyl donors on the luminal side of the ER membrane are the dolichylphosphate-linked monosaccharides Dol-P-Man and Dol-P-Glc.⁴² These Dol-P-monosaccharides are synthesized by the enzymes Dpm1 and Alg5 on the cytoplasmic face of the ER membrane and then flipped into the lumen by an unknown mechanism.^{43–45} The underlying reasons behind why the cell switches to membrane-bound dolichylphosphate-linked glycosyl donors in the ER lumen compared with the soluble nucleotide sugars in the cytoplasm are unclear, but presumably the scarcity of nucleotide sugar transporters in the ER membrane is a critical factor.

Although the structure of the core Glc₃Man₉GlcNAc₂ tetradecasaccharide is widely conserved across eukaryotes, several species of protists have been found to assemble only truncated forms of the glycan.^{46,47} Recent studies have found that these truncations are due to a loss of certain sets of glycosyltransferases over the course of evolution.⁴⁸ For example, the causative agent of Chagas disease, *Trypanosoma cruzi*, produces only Dol-PP-GlcNAc₂Man₉ and is missing the glucosyltransferases Alg6, Alg8, and Alg10.⁴⁷ In contrast, *Tetrahymena pyriformis* has lost all mannosylation function in the ER lumen, producing only Dol-PP-GlcNAc₂Man₅Glc₃.⁴⁹ In an extreme case, *Plasmodium falciparum*, a causative agent of malaria in humans, is missing all the Alg glycosyltransferases except Alg7, thus generating only Dol-PP-GlcNAc and providing an explanation for the difficulty in identifying N-linked glycoproteins in this species.^{48,50}

Transfer of Eukaryotic N-Linked Glycans. After the oligosaccharide has been assembled, it is transferred to the side chain amide of acceptor proteins by the multimeric oligosaccharyl

transferase (OT) complex. In *S. cerevisiae*, OT comprises at least eight membrane-bound subunits (Ost3/Ost6, Ost4, Stt3, Ost2, Wbp1, Swp1, Ost5, and Ost1), and genetic studies have determined that at least five of these (Ost1, Ost2, Stt3, Swp1, and Wbp1) are essential for survival (Figure 3A).^{51,52} Recent evidence indicates that multiple isoforms of the OT complex containing various combinations of subunits can exist within the same cell and may serve to modulate both activity and specificity. For example, OT isoforms containing Ost1 and Swp1 are required for glycosylation of certain membrane proteins,⁵³ while the presence of either Stt3A or Stt3B impacts the kinetics and timing of glycosylation, promoting either co- or posttranslational catalysis.⁵⁴

Despite challenges associated with isolation and expression of the complex, several key studies have contributed to our understanding of the specific involvement of each subunit. For example, modification of a key cysteine residue within the Wbp1 protein disrupts catalytic activity, and incubation of OT with Dol-PP-GlcNAc₂ prior to treatment with the cysteine-modifying agent methyl methanethiosulfonate (MMTS) prevented this alkylation and preserved activity, suggesting it may play a role in substrate binding.⁵⁵ The Ost3/Ost6 proteins have been recently shown to exhibit protein-dependent oxidoreductase activity in vitro, suggesting a direct role for OT in protein folding.⁵⁶ In addition, a number of factors have emerged to suggest that Stt3 is the catalytic subunit of the OT complex. Stt3 is the most highly conserved subunit of OT, as homologues of this protein are found in all eukaryotes and throughout the archaeal and bacterial kingdoms, and the Stt3 homologue in the bacterial pathogen *Campylobacter jejuni*, PglB, is sufficient to catalyze glycan transfer.^{57,58} In addition, the protists *Leishmania major* and *Trypanosoma brucei* were found to contain only the Stt3 subunit of OT but are still capable of N-linked glycosylation.^{59–61} In a recent study, the Stt3 homologues in *L. major* were shown to complement a deletion of the *stt3* gene in *S. cerevisiae*; however, these proteins did not incorporate into the native yeast OT complex but instead formed a functional homodimer in the cell.⁵⁹ Although structural studies of the entire OT complex are currently lacking, recent progress has been made

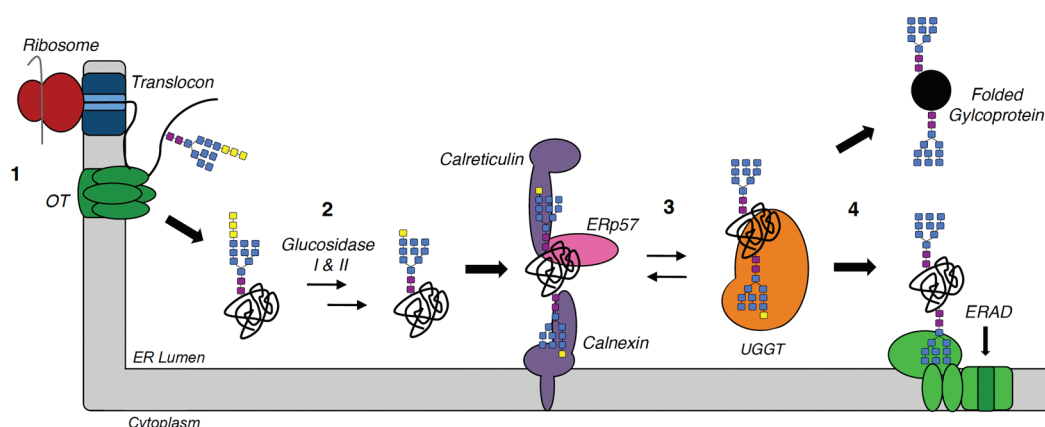


Figure 4. Calnexin—calreticulin pathway of quality control for protein folding. (1) Nascent proteins travel through the translocon and are glycosylated by OT. (2) The two terminal glucose residues are removed by glucosidases I and II. (3) The folding protein interacts with the chaperones calnexin and calreticulin and the ERp57 oxidoreductase. If misfolded, they are recognized by UGGT and cycled through the pathway again. (4) Defective proteins are demannosylated and ultimately sent for degradation through the ERAD pathway, while properly folded proteins exit the cycle.

toward understanding the proximity of the subunits to one another. To this end, the cryo-electron microscopy structure of the OT complex at 12 Å was recently reported.⁶²

The spatial requirements of the N-X-S/T glycosylation sequence within acceptor proteins have been widely analyzed and do not appear to be very stringent.⁵² However, not all asparagines found within the N-X-S/T sequon are glycosylated; in fact, it was recently estimated that roughly 35% of N-linked sequons that pass through OT are not modified.^{52,63} Early work showed that OT does not glycosylate proteins that contain a proline in position X of the N-X-S/T sequence⁶⁴ and exhibits a modest preference for N-X-T sites compared with N-X-S sites.⁶⁵ In addition, sites within 12–14 residues of the N-terminus are not glycosylated, suggesting that the active site is 30–40 Å from the ER luminal membrane.⁶⁶ The secondary structure of the glycosylation site has also been found to be a determinant for catalysis, as the N-X-S/T sequon must be able to adopt an Asx turn, which has been proposed to promote activation of the side chain amide nitrogen in the catalytic mechanism.^{67,68}

Functional Roles of Eukaryotic N-Linked Glycans. After glycosylation, the newly formed glycoproteins are processed by a series of glycosidases and glycosyltransferases in both the ER and Golgi to resculpt the core tetradecasaccharide, resulting in the vast diversity of N-linked glycan structures observed in eukaryotes. These modifications include glycan trimming, as well as the addition of carbohydrates such as sialic acid, fucose, and N-acetylgalactosamine to form complex and branched structures.^{69,70} The mature glycans are important for a number of biological functions, including intracellular targeting, cell signaling, cellular trafficking, and the immune response. Mutations in almost every step of glycan biosynthesis and processing lead to a group of debilitating and often fatal diseases that are collectively termed the congenital disorders of glycosylation (CDGs), underscoring the importance of this protein modification to human health.⁷¹ In an impressive recent study, Cantagrel et al.⁷² combined clinical pathology, molecular biology, and biochemistry to attribute a newly discovered CDG resulting in dramatic developmental problems, including severe brain malformation, to a mutation in *srd5a3*, serving to identify the encoded protein as the long-sought reductase responsible for selective reduction of the α -isoprene unit of phytylprenol to yield dolichol.

Perhaps the most well-studied role for N-linked glycans in eukaryotes involves quality control for protein folding in the ER.^{73,74} In the cotranslational pathway of N-linked glycosylation in *S. cerevisiae*, newly synthesized proteins leaving the ribosome are translocated through the ER membrane by way of the Sec61 translocation channel and then presented to the OT complex. After glycan transfer, the nascent glycoproteins enter an elaborate quality control cycle that serves to ensure proper folding and utilizes all three branches of the glycan (Figure 4). The first step in this cycle involves the removal of two of the terminal glucose units from the tetradecasaccharide by glucosidases I and II.⁷⁵ The nascent glycoprotein then encounters two ER chaperones that recognize the monoglucosylated glycan, the membrane-bound calnexin and soluble calreticulin. These chaperones sequester the newly formed protein to prevent aggregation and misfolding and also provide access to ERp57, an oxidoreductase that aids in disulfide bond formation. Once the protein is properly folded, glucosidase II removes the remaining glucose, which stimulates release of the protein from calnexin and calreticulin and progression through the remainder of the glycan-processing pathway in the ER and Golgi. However, if the protein is misfolded, it binds to the folding sensor UDP-glucose:glycoprotein glucosyltransferase (UGGT).^{75,76} UGGT contains a hydrophobic N-terminal domain that recognizes non-native protein structures and a C-terminal glucosyltransferase domain that reglucosylates misfolded proteins, causing them to rebind calnexin and calreticulin. The iterative removal and addition of glucose by glucosidase II and UGGT persists until either the protein is folded or it is recognized by the α -1,2-mannosidases present in the ER, which remove terminal mannose residues from the glycan core.⁷⁶ Progressive mannose removal lowers the affinity of the glycan for calnexin and calreticulin and ultimately serves as a folding timer to interrupt the folding cycle and target the defective glycoprotein for degradation through the ER-associated degradation pathway (ERAD).⁷⁷

In addition to their role in the quality control of protein folding, N-linked glycans perform a variety of other functions within eukaryotes. They serve as identity and localization tags to direct proteins to the proper cellular destination and regulate mobility of proteins on or within the cell membrane through interactions with lattice-forming lectins, such as galectins.⁷⁸

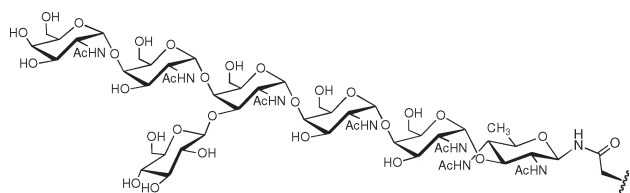


Figure 5. Structure of the N-linked GlcGalNAc₅Bac heptasaccharide identified in *C. jejuni*.

N-Linked glycans also impart structural rigidity to their targets and may provide protection from proteolysis, such as in the case of the heavily glycosylated lysosomal membrane proteins Lamp-1 and Lamp-2.^{79,80} In addition, the surface N-linked glycans of cancer cells have been found to be highly branched compared with those of healthy tissues; exploitation of this difference may present a platform for the development of clinical biomarkers of disease.^{81,82}

■ N-LINKED GLYCOSYLATION IN BACTERIA

Despite the enormous complexity and structural diversity of glycans that comprise the lipopolysaccharide (LPS) and capsular polysaccharide (CPS) of bacterial cell walls, it was believed until recently that protein glycosylation, in particular N-linked glycosylation, was absent from bacteria.^{83–85} However, in the past decade, this dogma has been shattered with discovery of both N-linked and O-linked bacterial glycoproteins. In 1999, the first evidence of a general system of N-linked glycosylation was uncovered in the human mucosal pathogen *C. jejuni*.⁸⁶ Mutations in the identified *pgl* locus (*protein glycosylation*) resulted in a loss of bacterial immunogenicity without affecting LPS or CPS biosynthesis, providing an initial link between N-linked glycosylation and host pathogenicity.^{86,87} Structural characterization of the *C. jejuni* glycoproteins PEB3 and CgpA established that the installed glycan was the heptasaccharide GlcGalNAc₅Bac, which forms a β -linkage to asparagine (Figure 5).^{88,89} To date, more than 65 glycoproteins of various biological functions have been characterized in *C. jejuni*.¹¹ In addition, the identification of N-linked glycosylation loci in newly sequenced bacterial genomes establishes that this modification occurs in other bacterial species as well.

Bacterial N-Linked Glycan Assembly. The N-linked glycosylation pathway in *C. jejuni* shown in Figure 6 bears resemblance to the first half of the dolichol pathway in *S. cerevisiae* (Figure 1). The glycan is assembled on the cytoplasmic face of the periplasmic membrane and involves stepwise elaboration onto a polyisoprenyldiphosphate carrier by a series of glycosyltransferases that rely on nucleotide sugar donors. However, although the overall architecture of the pathway is similar, several differences distinguish N-linked glycosylation in eukaryotes and bacteria. For one, the polyisoprenyl carrier in bacteria is undecaprenol (Und), a long chain unsaturated polyisoprene composed of 11 units that is the most abundant isoprene in bacterial membranes and the central component of the cell wall biosynthetic pathway.¹⁷ In addition, the first sugar in the *C. jejuni* N-linked glycan is *N,N'*-diacetylbaicillosamine (Bac), rather than the GlcNAc residue found in eukaryotic glycans. In general, prokaryotes contain a variety of unique, highly modified sugars not commonly observed in eukaryotes.^{90,91} Originally discovered in *Bacillus subtilis*, Bac is found in a number of bacterial glycans,

including the LPS of *Vibrio cholerae*, the CPS of *Acinetobacter lwofii*, and the pilin of *Neisseria meningitidis*.^{92–95}

The N-linked glycosylation pathway in *C. jejuni* is initiated by the biosynthesis of UDP-Bac by the enzymes PglF, PglE, and PglD (Figure 6). Starting from UDP-GlcNAc, the dehydratase PglF generates the UDP-4-keto intermediate via an NAD⁺-promoted hydride transfer and facilitated elimination, followed by formation of the UDP-4-amino sugar by the PLP-dependent aminotransferase PglE.⁹⁶ PglD then catalyzes the acetylation of UDP-4-amino using AcCoA to afford UDP-Bac.⁹⁷ The phosphoglycosyltransferase PglC initiates carbohydrate transfer to the undecaprenylphosphate carrier by transferring Bac-1-P to Und-P, succeeded by the addition of the first and second α -1,3- and α -1,4-linked GalNAc residues by PglA and PglJ, respectively.^{98–100} Analysis of the polyisoprene specificity of these enzymes revealed that the α -unsaturation and the *cis*-double bond geometry of undecaprenol were more important than small changes in the polyisoprene length, highlighting the significance of the specific polyisoprene structure to the enzymes in the glycosylation pathway.¹⁰¹ The trisaccharide is then elaborated through the action of the polymerase PglH, which installs the next three α -1,4-linked GalNAc residues.¹⁰² Detailed kinetic examination revealed that PglH uses a single active site to conduct three galactosyl transfer reactions and exhibits an increasingly stronger affinity for its product as the glycan becomes longer, which serves to halt catalysis after the third transfer. The assembly of the heptasaccharide is completed with the addition of a branching β -1,3-linked glucose by the membrane protein PglI. The entire *pgl* locus along with two *C. jejuni* periplasmic proteins, AcrA and PEB3, has been inserted into *Escherichia coli* and shown to be capable of N-linked glycosylation, indicating that all of the necessary components of the pathway were present.⁵⁷ Interestingly, the Pgl pathway enzymes can be combined with the obligate cofactors and cosubstrates to generate the full-length Und-PP-BacGalNAc₅Glc in vitro, which represents the first complete chemoenzymatic synthesis of a polyisoprenyl-linked heptasaccharide.^{97,100}

In addition to *C. jejuni*, evidence of N-linked glycosylation pathways in other species of bacteria is slowly beginning to emerge. *Wolinella succinogenes* was shown to contain a *pgl* gene cluster, although the glycan structure as well as the identity of possible target proteins is still unclear.¹⁰³ Structural analysis of the HmcA protein from the sulfate-reducing bacterium *Desulfovibrio gigas* revealed the presence of a trisaccharide linked to asparagine, but the exact composition of this glycan remains unknown.¹⁰⁴ Several species of *Helicobacter*, including *H. canadensis* and *H. pullorum*, also possess *pgl* gene clusters. Recent biochemical studies have shown that *H. pullorum* contains two isoforms of the *pglB* OTase gene; however, only one of these genes is responsible for transfer of a linear N-linked glycan comprising five hexosamine sugars, while the role of the second isoform is still under investigation.¹⁰⁵ Bioinformatic analyses of newly sequenced bacterial genomes suggests the presence of *pgl* genes in more than 49 bacterial species, although biochemical characterization of these pathways is scant.¹¹

Interestingly, O-linked pilin glycosylation in both *N. meningitidis* and *Neisseria gonorrhoeae* is believed to strongly resemble N-linked glycosylation in *C. jejuni*.^{106,107} In the *N. gonorrhoeae* *pgl* pathway, genetic studies suggest that the enzymes PglB, PglA, and PglE are responsible for assembly of a trisaccharide onto an undecaprenyldiphosphate-linked carrier, which is then flipped into the periplasm and transferred to the hydroxyl side chain

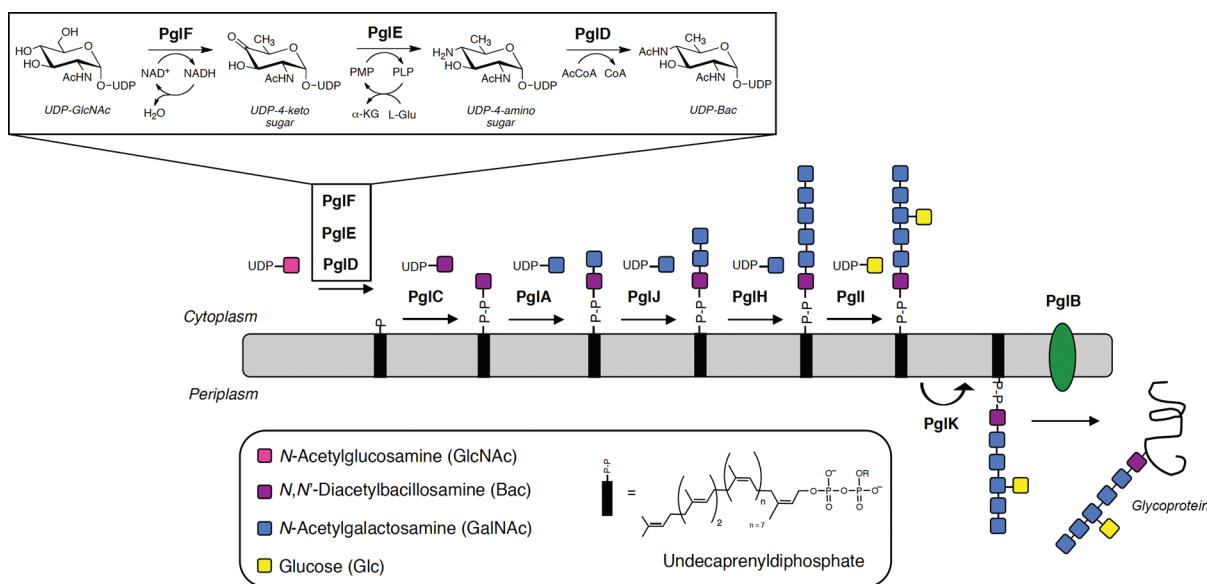


Figure 6. N-Linked glycosylation pathway in *C. jejuni*.

serine atom by the oligosaccharyl transferase PglO. This pathway differs from most identified O-linked glycosylation pathways in eukaryotes, in that the glycan is transferred to the serine residue in an *en bloc* manner, rather than through the individual transfer of each sugar directly onto the protein from a nucleotide diphosphate donor. In the case of *Neisseria*, the glycan structure was found to be Gal₂Bac,⁹⁴ although recent evidence has shown that an alternate glycosyltransferase PglH can replace PglA, resulting in the production and transfer of the disaccharide GlcBac (M. Koomey and B. Imperiali, unpublished results). The *N. gonorrhoeae* PglO OTase does not appear to require a specific glycosylation sequence in the acceptor protein and also does not bear close sequence homology to the *C. jejuni* OTase PglB.^{107,108} Detailed biochemical characterization of PglO is still required to understand the basis of glycosylation specificity.

Bacterial N-Linked Glycan Transfer. In contrast to eukaryotes, which require a multimeric OT complex, bacterial N-linked glycosylation in *C. jejuni* is conducted by a single protein, the Stt3 homologue PglB.^{57,58} PglB is predicted to comprise a large N-terminal domain consisting of 9–13 transmembrane helices, and a periplasmic C-terminal domain that includes the highly conserved WWDXXG signature sequence (Figure 3B). A preliminary crystal structure of the C-terminal soluble domain of PglB was recently reported, and although this domain is not sufficient to catalyze oligosaccharide transfer, analysis of this structure coupled with evaluation of genetic point mutations and comparison with other PglB homologues suggests the importance of the two additional protein motifs for catalysis, MXXI (residues 568–571) and XXD (residues 52–54).¹⁰⁹

Recent studies have shown that PglB is able to glycosylate folded proteins and requires the extended D/E-X₁-N-X₂-S/T glycosylation sequence, in which an acidic residue occupies the –2 position.^{110,111} Experiments to define the optimal glycosylation sequence involved testing a library of peptides in which the residues at both X positions were varied; PglB was found to glycosylate peptides containing a bulky residue at position X₁ and a smaller hydrophobic residue in position X₂ with the greatest efficiency.¹¹² The extremely high affinity that PglB exhibits for these short acceptor peptides ($K_m = 1 \mu\text{M}$ for

DQNAT) suggests that PglB does not require a specific tertiary structure for binding the fully folded protein substrate.¹¹² In addition, structural analysis of glycosylated proteins PEB3 and AcrA has established that the glycosylation sequence must be in a fairly exposed loop region of the protein for access by PglB.^{113,114} PglB has also been shown to be more promiscuous than the eukaryotic OT with respect to the glycan structures that are transferred to protein. Experiments in which PglB was substituted for the bacterial LPS O-antigen ligase demonstrated that PglB was able to transfer distinct O-antigens from both *E. coli* and *Pseudomonas aeruginosa* onto asparagine residues, though the presence of an N-acetyl group at the 2'' position of the proximal sugar is required for efficient glycosylation.^{115,116} The ability of PglB to carry out posttranslational N-linked glycosylation with decreased substrate selectivity provides an exciting platform for engineering custom glycoproteins for therapeutic use. Several groups have recently begun to explore this opportunity; for example, a recent study by Schwarz et al.¹¹⁷ showed that the *C. jejuni* pgl locus can be utilized in *E. coli* to produce homogeneous glycoproteins containing the eukaryotic Man₃GlcNAc₂–Asn linkage after an *in vitro* endoglycosidase-catalyzed transglycosylation to exchange the bacterial glycan that is installed by PglB with the eukaryotic pentasaccharide core. However, many challenges with this approach still remain, including low yields as well as the restricted placement of the glycosylation sequence to flexible, solvent-exposed loops.¹¹⁸ Further work in this area is required to address these issues before these methods can serve a practical use.

Biological Significance of Bacterial N-Linked Glycans. To date, more than 65 distinct glycoproteins have currently been identified in *C. jejuni*.¹¹ These proteins are associated with a wide variety of biological pathways within the cell, suggesting that N-linked glycans may play a role in a number of important cellular functions. Initial genetic studies in which components of the N-linked glycosylation pathway were disrupted have resulted in impaired host cell adhesion and colonization, establishing a strong tie between N-linked glycans and virulence. For example, early studies of *C. jejuni* glycoproteins indicated that they were highly immunogenic when cross-reacted with animal antisera.⁸⁶

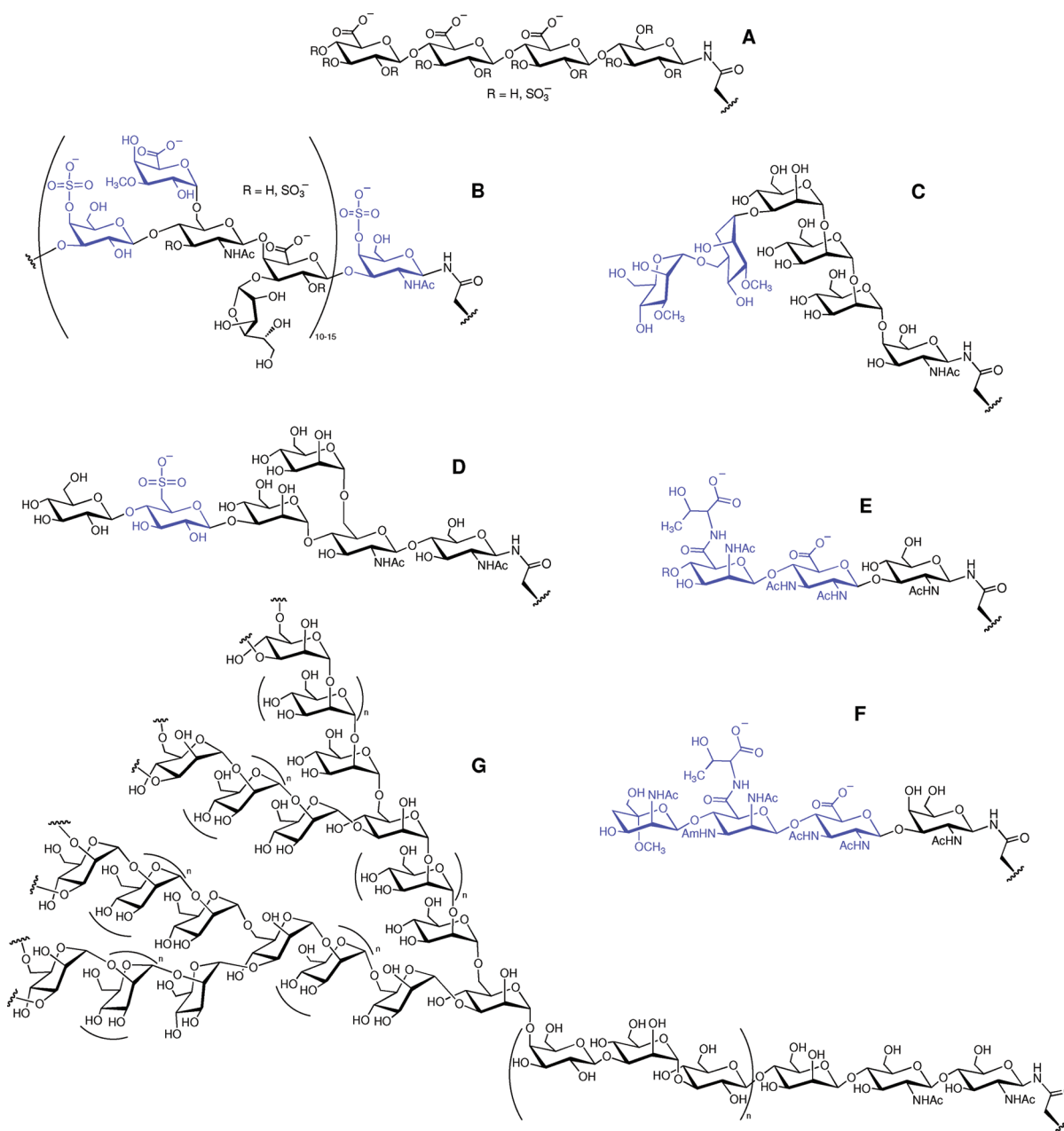


Figure 7. Structures of representative archaeal N-linked glycans: (A) *H. halobium*, shorter glycan; (B) *H. halobium*, longer glycan; (C) *M. fervidus*, in which the two residues at the nonreducing end are either Man or Glc; (D) *S. acidocaldarius*, containing the unique 6-sulfoquinovose moiety; (E) *M. voltae*, where the R indicates the presence of an unidentified hexose in some strains; (F) *M. maripaludis*; and (G) *T. acidophilum*, in which the order of the residues near the reducing end is unclear. Unusual sugar moieties are colored blue; see the text for more detail and specific references.

In addition, genetic mutants of *C. jejuni* that are impaired in glycosylation showed a reduced ability to adhere to and invade human intestinal epithelial cells, as well as decreased colonization of the intestinal tracts of mice and chicken.^{87,119} Although details about the interactions between *C. jejuni* and the mammalian immune system are largely unclear, a recent study revealed that the *C. jejuni* N-linked glycan is recognized by the human macrophage galactose-type lectin, a receptor that plays an important role in intracellular signaling.¹²⁰ When glycosylation was impaired, *C. jejuni* was found to stimulate formation of the cytokine IL-6, suggesting that glycosylation of specific proteins might help the organism evade the host immune response.

In addition to pathogenicity, other roles for the N-linked glycans of *C. jejuni* have recently been described as well. The VirB10 protein, a component of the type IV secretion system that contains two glycosylation sites, was found to be incapable of properly forming a complex with other secretion system proteins when glycosylation of a key asparagine was prevented, implying that the glycan may be important in modulating protein–protein interactions.¹²¹ Analysis of the recent NMR structure of the glycoprotein AcrA, part of the *C. jejuni* multidrug efflux pump, shows that the N-linked glycan adopts a rigid, rodlike structure that appears to fold over part of the protein.¹¹⁴ The orientation and rigidity of the glycan indicate that it may serve to enhance

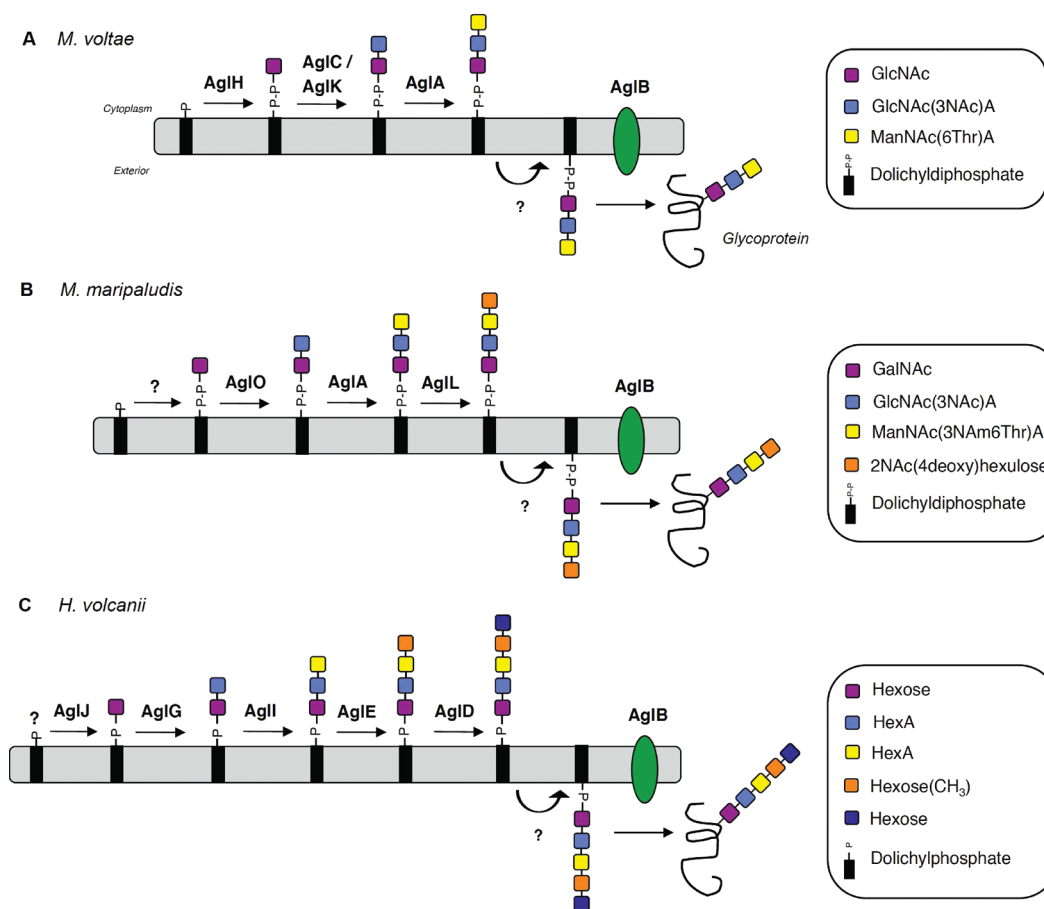


Figure 8. Current models of N-linked glycosylation in (A) *M. voltae*, (B) *M. maripaludis*, and (C) *H. volcanii* based on recent genetic studies. The exact structure of the *H. volcanii* N-linked glycan has not yet been determined, but data suggest that it is a pentasaccharide as indicated.

protein stability and provide protection from proteolysis. N-Linked glycans have also been associated with protection of *C. jejuni* from osmolytic stress. A recent study revealed that *C. jejuni* is able to produce free oligosaccharide in response to variations in the osmolarity of the external environment.¹²² While the mechanism for oligosaccharide release is still unclear, early studies clearly implicate PglB in the accumulation of these structures in the *C. jejuni* periplasm.

■ N-LINKED GLYCOSYLATION IN ARCHAEA

Structural Diversity of Archaeal N-Linked Glycans. First described by Carl Woese in 1977, archaea are single-celled organisms that share many common features with both bacteria and eukaryotes but also possess several characteristics unique to their kingdom.¹²³ The first archaeal glycoproteins were discovered in 1976 in the extreme halophile *Halobacterium salinarum*.¹²⁴ In an intriguing study, the structural analysis of the closely related species *Halobacterium halobium* revealed that the S-layer protein of the organism was modified with two very distinct types of N-linked glycans. The first was determined to be a pentasaccharide consisting of sulfated glucose and glucuronic acid (GcA) residues (Figure 7A).^{125,126} The second glycan discovered is much larger and more complex, with similarities to the glycosaminoglycans that are observed in the connective tissues of higher eukaryotes. This glycan is attached to asparagine via a β -linked sulfated GalNAc, followed by between 10 and 15 repeating units

of a branched pentasaccharide containing sulfated glucose, galacturonic acid (GalA), and galactosaminuronic acid (GalNAcA) residues, as well as a Gal in the furanose form (Figure 7B).^{127,128} These extended glycans were found to coat the outer surface of *H. halobium*, forming an acidic, two-dimensional lattice that surrounds the organism.¹²⁶

Since this discovery, many other N-linked glycoproteins have been identified across the archaeal kingdom. These glycans display an extraordinary structural variety, which may stem from the diverse habitats of the archaeal species that produce them. The S-layer protein of the hyperthermophilic methanogen *Methanothermobacter fervidus* is modified with an N-linked hexasaccharide, in which a GalNAc is attached to the Asn residue of the protein (Figure 7C).¹²⁹ The glycan includes three mannose units followed by two methylated glucose or mannose residues at the nonreducing end. Another methanogen, the mesophile *Methanococcus voltae*, was found to contain glycosylated S-layer proteins and flagella, all labeled with the trisaccharide GlcNAc- β -1,3-GlcNAc(3NAc)A- β -1,4-ManNAc(6Thr)A (Figure 7E).¹³⁰ The terminal mannosaminuronic (ManNAcA) unit of this glycan is linked to a threonine residue through an amide bond at the C6'' position. However, in some strains of *M. voltae*, the trisaccharide is further elaborated with a hexose of unknown identity in the fourth position.¹³¹ In *Methanococcus maripaludis*, the flagellin proteins are modified with a linear tetrasaccharide that is β -linked to the asparagine through a GalNAc residue (Figure 7F).¹³² GlcNAc(3NAc)A is ligated to the proximal

GalNAc, followed by a highly modified ManNAcA residue that contains an acetamido group at the C3'' position and a threonine linked via an amide to the C6'' carboxyl, similar to the *M. voltae* glycan. In addition, the terminal sugar of this oligosaccharide is the unique 2-acetamido-2,4-dideoxy-5-O-methylhexosulo-1,5-pyranose, which represents the first reported example of an aldulose in an N-linked glycan structure.¹³² Recently, the pili of *M. maripaludis* were also found to be modified with a similar N-linked branched pentasaccharide, although further analysis is required to determine the exact glycan structure.¹³³

In addition to the structural characterization of the methanogen-derived glycans described above, two other archaeal N-linked glycans have been annotated in detail. The cytochrome *b*_{558/566} protein from the thermophile *Sulfolobus acidocaldarius*, which grows optimally at 75–80 °C and pH 2–3, is modified with a hexasaccharide linked through a GlcNAc moiety that contains a 6-sulfoquinovose residue (Figure 7D).¹³⁴ *Thermoplasma acidophilum* produces a complex, highly branched N-linked glycan composed mainly of mannose that is attached to a surface membrane protein through a GlcNAc–Asn linkage (Figure 7G).¹³⁵ These large glycans encapsulate the cell membrane and are proposed to shield the organism from the harsh surroundings. Initial studies of other archaeal glycoproteins have been reported, including those of the hyperthermophile *Pyrococcus furiosus* and the halophile *Haloferax volcanii*, although detailed structures of the respective glycans have not yet been elucidated.^{12,136} For more information about proposed archaeal N-linked glycans, the reader is referred to a review by Eichler.¹³⁷

Archaeal Glycan Assembly. Unlike the analogous pathways in eukaryotes and bacteria (specifically *C. jejuni*), many of the details of N-linked glycan assembly in archaea are still unknown. A recent analysis of the nearly 60 sequenced archaeal genomes indicates that all but two (*Aeropyrum pernix* and *Methanopyrus kandleri*) contain an *stt3* homologue, which suggests that in contrast to bacteria, N-linked glycosylation in archaea may be widespread.¹³⁸ However, difficulties with genetic manipulation and large-scale growth of these organisms, together with the fact that the genes for glycoprotein biosynthesis are not clustered, have hampered genetic and biochemical studies. Despite these challenges, progress toward uncovering the biosynthetic pathways of archaeal N-linked glycans has begun to accelerate, particularly in the past five years.

Recent genetic studies in the methanogens *M. voltae* and *M. maripaludis* and the halophile *H. volcanii* have provided the first glimpses of the glycan assembly pathways in archaea; the current working models of N-linked pathways in these organisms are summarized in Figure 8. In all three models, glycan assembly is believed to take place on the cytoplasmic face of the cell membrane and is catalyzed by a series of Agl glycosyltransferases (archaeal glycosylation), after which the glycan is flipped to the exterior side of the cell membrane and transferred to asparagine residues of acceptor proteins by the Stt3 homologue AglB.¹³⁹ Early studies in *H. halobium*, *H. volcanii*, and *T. acidophilum* determined that archaeal membranes contain both the monophosphate and diphosphate forms of dolichol, where the dolichol contains only 11–12 isoprene units and is saturated at both the α and ω positions.^{140–142} In addition, dolichylphosphate-linked monosaccharides (Dol-P-sugars) have been identified within these organisms and have been implicated as glycosyl donors for individual glycosyltransferases in the pathway.^{141–144} The presence of both forms of dolichylphosphate is reminiscent of the

eukaryotic glycosylation pathway, in which Dol-P-sugars are utilized as the carbohydrate donors for the later steps of glycan biosynthesis. However, biochemical evidence is still required to define the individual roles that each of these dolichylphosphate-linked intermediates plays in N-linked glycan assembly.

As shown in Figure 8, the first step of the N-linked glycosylation pathway in these selected archaea involves the transfer of a sugar 1-phosphate onto the dolichylphosphate by the putative phosphoglycosyltransferase AglH (AglJ in *H. volcanii*).^{145,146} Strong evidence for this initial step was provided in a recent study in which the *M. voltae* *aglH* gene was able to complement a conditionally lethal mutant of the analogous phosphoglycosyltransferase gene in *S. cerevisiae*, *alg7*.¹⁴⁵ This finding suggests that in *M. voltae*, the glycosyl donor for AglH is UDP-GlcNAc and the polyisoprenyl carrier is dolichyldiphosphate. In the related methanogen *M. maripaludis*, a similar enzyme is believed to transfer GalNAc-1-P to Dol-P, though the candidate gene has not yet been identified. An alternate model for formation of the Dol-PP-GalNAc intermediate in *M. maripaludis* could involve initial formation of Dol-PP-GlcNAc by a still undiscovered AglH homologue, followed by the epimerization of Dol-PP-GlcNAc to Dol-PP-GalNAc. However, this hypothesis has yet to be supported with genetic or biochemical evidence, although an epimerase that catalyzes the conversion of Und-PP-GlcNAc to Und-PP-GalNAc in *E. coli* was recently described along with the proposal of a homologous candidate epimerase in *M. maripaludis*.^{147,148}

The remainder of the steps involved in glycan assembly remains unclear. Initial genetic deletion studies suggest two different pathways for the transfer of the second sugar GlcNAc-(3NAc)A in both *M. voltae* and *M. maripaludis*, though this finding has not been verified biochemically.^{131,149} Transfer of the third sugar is believed to be carried out by AglA or AglI; interestingly, in both the *Methanococcus* species, this sugar is a mannuronic acid derivative modified with a threonine residue at the C6'' carboxylate. After completion of the glycan, it is flipped across the cell membrane, where an OTase presumably catalyzes its transfer to acceptor proteins. Evidence for the localization of oligosaccharide transfer to the external face of the cell membrane comes from the treatment of *H. salinarium* with bacitracin, which inhibited formation of sulfated glycoproteins.¹⁵⁰ Bacitracin, a small molecule that is unable to traverse the archaeal membrane, binds to the phosphate groups of free Dol-PP that is liberated by the OTase after glycosylation, thus preventing recycling of the dolichol.¹⁵¹ However, despite this evidence, a candidate flippase gene has yet to be identified in archaea. Further work involving detailed genetic and biochemical studies is required to answer the many questions that remain about the biosynthesis of these intriguing glycans.

Archaeal Glycan Transfer. The archaeal Stt3 homologue AglB is predicted to have an overall topology similar to that of bacterial PglB, containing both a large N-terminal region composed of 11–13 transmembrane domains and an extracellular C-terminal domain that bears the signature OTase WWDXXGX sequence (Figure 3C). Recently, the X-ray crystal structure of the soluble C-terminal domain of AglB from *P. furiosus* was reported, providing the first insight into the structural organization of a segment of an archaeal Stt3 homologue.¹³⁶ Although it is unclear if this structure represents a biologically relevant form of the enzyme, as the C-terminal domain alone was unable to catalyze oligosaccharide transfer, this study revealed that the WWDXXGX motif is located in a central β -helix in the proximity of a highly

conserved DXXK motif. Initial biochemical characterization of the full-length AglB using crude mixtures of membrane-associated lipids suggests that the enzyme alone is sufficient to carry out glycosylation, and that mutations in either the WWDXXG or DXXK motif impair glycosylation activity.¹³⁶ These results are consistent with genetic studies in *M. voltae* and *M. maripaludis*, in which mutations in the *aglB* gene resulted in flagella with decreased levels of glycosylation and impaired motility.^{139,152}

In contrast to the bacterial system, the archaeal OTase AglB does not require an extended glycosylation sequence beyond the N-X-S/T sequon defined for eukaryotic glycoproteins. However, a comparative analysis of the amino acid sequence around the glycosylation site of confirmed archaeal glycoproteins indicates a slight preference for Ala and Gly at the X₊₁ position and Ser or Thr at the X position, where X₋₁-N-X-S/T-X₊₁.¹⁵³ Because AglB is thought to glycosylate proteins posttranslationally, it is proposed that the target glycosylation site should be located in a flexible, exposed region of the acceptor protein as described for the protein substrates of the *C. jejuni* PglB; however, structural analysis of glycoproteins is necessary to confirm this hypothesis.

Functional Role of Archaeal N-Linked Glycans. As archaea populate a wide range of habitats from the extreme to the more familiar, it is not surprising that the structures of their respective N-linked glycans mimic this diversity. Archaea already possess a variety of biological adaptations to prepare them for life in chemically harsh environments. For example, archaeal membranes are composed of ether-linked membrane lipids, such as diethers, tetraethers, and macrocyclic diethers, in contrast to the ester-linked phospholipids found in both eukaryotic and bacterial membranes.¹⁵⁴ This unique lipid composition gives rise to the decreased ion permeability and planar dynamic capacity characteristic of archaeal membranes.¹⁵⁵ In addition, thermophilic and halophilic archaea employ a number of strategies for enhancing protein stability at high temperatures and salt concentrations, including an increased level of ion pairing and hydrophobic interactions, burying polar contacts in the interior of proteins, and a decreased entropy of unfolding.^{156–158}

The varied structures of N-linked glycans identified in archaea to date likely serve to equip these organisms for optimal growth and survival in their diverse surroundings. For example, the large, highly branched glycan on the surface of the hyperthermophile *T. acidophilum* is predicted to protect the organism from the heat and high acidity of its environment by forming an extensive hydrogen-bonded shell around the organism, providing external strength and rigidity and limiting access of certain ions and water molecules to the cell membrane.¹³⁵ In bacteria, these functions are provided by structures such as the LPS and CPS, but as archaea do not have a comparable structure, it is hypothesized that perhaps the surface N-linked glycans in species such as *T. acidophilum* and *H. halobium* impart similar protective features. Another important role for archaeal N-linked glycans is related to cell motility. In the methanogens *M. voltae* and *M. maripaludis*, the protein subunits that comprise the flagella are heavily modified with negatively charged N-linked glycans (Figure 7E,F). Genetic mutations that disrupt the glycosylation pathway in these organisms have resulted in impaired motility; in fact, in *M. maripaludis*, a direct relationship between glycan size and swimming capability has been established.¹⁴⁸ It is unclear exactly how glycans contribute to the proper function of flagella, although the extensive glycosylation may sheath the structure with negative charge and possibly impart stability to interactions among the individual protein subunits. Further studies are

required to gain a deeper understanding of the roles of N-linked glycans in archaea.

■ ROLE OF N-LINKED GLYCOSYLATION IN VIRUSES

In recent years, the importance of N-linked glycosylation in virology has begun to emerge, revealing the many ways in which glycans contribute to viral survival and infection. Human viruses rely on the host cell glycosylation pathway in nearly every step of the viral life cycle, from host cell recognition, viral replication, protein trafficking, and virion packaging.¹ Viruses hijack the host cell glycosylation pathways to take advantage of the many benefits available to eukaryotic proteins, including access to the host cell protein folding and quality control machinery. Nearly all of the viruses studied thus far have utilized the calnexin–calreticulin cycle to aid in protein folding and trafficking.¹⁵⁹ In addition, the rapid evolution of viral particles often results in an increased number of glycosylation sequences, introducing diversity and complexity into the viral glycoproteome that can ultimately complicate treatment of infection and alter vaccine efficiency.

The influenza virus is a common human pathogen of the upper respiratory tract and a member of the orthomyxovirus RNA virus family; in addition, it is perhaps the most well studied virus with respect to glycosylation. The virus contains two important capsular coat glycoproteins, hemagglutinin (HA) and neuraminidase (NA), that utilize N-linked glycans for a number of important functions such as receptor binding and infection. The dominant surface protein is HA, which is modified with a heterogeneous, complex glycan that depends on the viral strain but is indicative of extensive processing in the Golgi.^{160,161} HA is heavily glycosylated, containing between 5 and 11 glycosylation sites depending on the strain, all of which are clustered at the globular head of the protein and are important for interaction with a key sialic acid residue on the host cell receptor. Upon viral replication and release, NA acts to cleave the sialic acid moiety to promote diffusion of the virus. If blocked, the newly formed viral particles rebind the host cell receptor and form large aggregates. Increased glycosylation of HA to improve viral binding requires enhanced activity of NA to mediate viral release, thus demonstrating the strategic interplay of these two glycoproteins for effective influenza infection.

Another important example of viral glycosylation is in the human immunodeficiency virus (HIV), a highly mutagenic RNA virus from the retrovirus family. The virus is surrounded by the multiple copies of the envelope protein gp120, one of the most heavily glycosylated structures found in nature.¹⁶² The structure of the N-linked glycan coating gp120 was only just recently characterized and found to be a high-oligomannose structure, suggesting that this glycan is only minimally processed after transfer to gp120 by OT.¹⁶² A recent analysis of global HIV strains indicated that the gp120 protein contains between 18 and 33 possible N-linked glycosylation sites, with an average of 25 glycans per protein.¹⁶³ The vast density of carbohydrates forms a large, encompassing structure called the glycan shield, which is postulated to protect the virus from recognition and degradation from the host immune system. In particular, the glycan shield provides protection from the 2G12 neutralizing antibody, which has been shown to recognize the virus particle and render it inactive.¹⁶⁴ Rapid viral evolution introduces multiple amino acid substitutions into the gp120 sequence, changing the number and

Table 1. Comparison of N-Linked Glycosylation across the Three Kingdoms

	Eukarya	Bacteria	Archaea
site of glycan assembly	ER membrane	inner cell membrane	cell membrane
isoprene carrier	Dol-PP (14–21 isoprene units)	Und-PP (11 isoprene units)	Dol-P and Dol-PP (11 isoprene units)
glycosyl donor	NDP-sugars, Dol-P-sugars	NDP-sugars	NDP-sugars, Dol-P-sugars (?)
Asn-linked sugar	GlcNAc	Bac	GlcNAc, GalNAc, Glc
glycan diversity across kingdom	initially conserved (Glc ₃ Man ₉ GlcNAc ₂)	variable	highly variable
flippase	Rft1 (?)	PglK	unknown
oligosaccharyl transferase	OT (5–8 subunits)	PglB (1 subunit)	AgIB (1 subunit, unclear)
protein recognition sequence	N-X-S/T	D/E-X-N-X-S/T	N-X-S/T
timing of glycan transfer	cotranslational, posttranslational	posttranslational	posttranslational
glycan processing	ER, Golgi	none	unclear

organization of the glycan sites and thus allowing the virus to persist in spite of increasing pressure from host antibodies.¹⁵⁹

N-Linked glycosylation has also been found to play an important role in the infection and stability of other viruses, including the West Nile, Ebola, SARS, and Hantaan viruses.¹⁵⁹ However, further work is required to uncover the mechanisms of host recognition and viral fusion. In general, a deeper understanding of these viral glycoproteins may provide new opportunities for the development of antiviral therapeutics.

CONCLUSIONS

N-Linked glycosylation plays an important role in many biological processes across the eukaryotic, bacterial, and archaeal kingdoms. Although the overall architecture of each pathway shares many features, a comparison of the details of glycan assembly and transfer in each kingdom reveals a number of intriguing differences. Table 1 summarizes the defining characteristics of each of these pathways.

Across all three kingdoms, the biosynthesis of N-linked glycans involves the buildup of an oligosaccharide onto a polyisoprenyl carrier by a series of glycosyltransferases using a set of discrete activated glycosyl donors. In eukaryotes, these donors are both nucleotide sugars and dolichyl-linked sugars, depending on the location of the glycosyltransferases relative to the ER membrane. In contrast, studies to date have shown that bacteria exclusively rely on UDP-sugars for glycan donor biosynthesis. The glycosyl donors in archaea have not yet been clearly defined, though the presence of both dolichylphosphate- and dolichyldiphosphate-linked glycans suggests that perhaps archaea may use a combination of nucleotide- and dolichyl-activated sugars similar to eukaryotes.¹⁴⁴ In all kingdoms, the polyisoprenyl-linked oligosaccharide is flipped across the membrane prior to *en bloc* transfer to protein. The mechanism of flipping between the two faces of the membrane remains elusive. Although genetic screens originally implicated Rft1 in translocation of the heptasaccharide intermediate in *S. cerevisiae*,³⁵ recent studies have found that this protein may not be involved after all.^{36,37} The ABC transporter PglK has been annotated as the flippase in *C. jejuni*;¹⁶⁵ however, biochemical confirmation of this activity has not yet been reported. Interestingly, analysis of complete archaeal genome sequences using either Rft1 or PglK as a search model does not result in any putative flippase candidates, even though it is clear that the archaeal OTase catalyzes protein transfer on the external face of the cell membrane. Further work is needed to fully understand this interesting phenomenon.

In all three kingdoms, the action of an oligosaccharyl transferase is responsible for transferring the assembled oligosaccharide to asparagine side chains within acceptor proteins. Although the role of each of the subunits within the eukaryotic OT complex is still unresolved, recent evidence has implicated the Stt3 subunit directly in catalysis. PglB, the Stt3 homologue in *C. jejuni*, has been shown to be sufficient to conduct this reaction *in vitro*, and preliminary studies involving the archaeal Stt3 homologue in *P. furiosus* suggest that it may also function alone, though further biochemical evidence is still required. The simplicity of the prokaryotic OTases in contrast to the eukaryotic OT complex offers a valuable and tractable model system for beginning to examine this reaction in greater depth. One of the major unresolved questions in this area involves the catalytic mechanism, namely the steps by which the asparagine side chain amide is activated by OT to affect oligosaccharide transfer. Interestingly, it appears that nearly all N-linked glycan structures reported to date contain an acetylated amine at the C2'' position of the sugar proximal to the modified asparagine residue,¹¹⁶ the one exception being the N-linked pentasaccharide of *H. halobium*, in which a glucose was found to be the linking sugar (Figure 7A). This observation supports earlier predictions that the 2-acetamido group may play an important role in the enzyme mechanism.^{166,167}

As described in this review, the N-linked glycans observed in nature display a great deal of structural diversity. It appears that one of the major differences between the three kingdoms involves the origin of this diversity, specifically, whether it is introduced during the assembly process through the use of a variety of glycosyl donors, such as in bacteria, or after assembly and transfer to protein by the action of processing enzymes, as in the case of eukaryotes. Nearly all eukaryotes assemble a tetradecasaccharide (Glc₃Man₉GlcNAc₂) that is transferred to acceptor proteins by OT; however, this core structure is then modified by a series of glycosidases and glycosyltransferases in both the ER and Golgi to result in the great variety of eukaryotic N-linked glycans observed in nature. The conservation of this glycan core during the assembly process allows eukaryotes to engage the calnexin–calreticulin pathway for quality control of protein folding immediately after glycosylation, while still allowing for diversification of the glycan through the later steps. In contrast, the N-linked glycan structures identified on all *C. jejuni* glycoproteins to date are identical, indicating that the organism does not modify these oligosaccharides after protein transfer. The involvement of N-linked glycans in virulence may provide an explanation for the uniformity of bacterial glycoproteins, although the details about the roles of bacterial N-linked glycans in pathogenesis are still unresolved. It will be interesting to learn whether archaea reflect the eukaryotic

or bacterial pathways in this regard or if they derive glycan diversity in a completely unique manner.

In summary, the N-linked glycosylation pathways found in eukaryotes, bacteria, and archaea share a great deal in common but also exhibit unique differences specific to each kingdom. Future work in this field will focus on a variety of fronts, including gaining a deeper understanding of the organization of the eukaryotic OT complex, the interactions of bacterial glycoproteins with the mammalian immune system, the biochemical details of archaeal glycan assembly and transfer, and the exploitation of viral glycan profiles for the treatment of disease. Overall, it is an exciting time in glycobiology, as the next few years promise to continue to highlight the important role that N-linked glycosylation plays across all kingdoms of life.

AUTHOR INFORMATION

Corresponding Author

*Massachusetts Institute of Technology, 77 Massachusetts Ave., Cambridge, MA 02139. Phone: (617) 253-1838. Fax: (617) 452-2419. E-mail: imper@mit.edu.

Funding Sources

This work was supported by a grant from the National Institutes of Health (GM039334 to B.I.).

ABBREVIATIONS

Bac, *N,N'*-diacetylglucosamine; CPS, capsular polysaccharide; Dol, dolichol; Dol-P, dolichylphosphate; Dol-PP, dolichyldiphosphate; ER, endoplasmic reticulum; ERAD, ER-associated degradation; Gal, D-galactose; GalA, D-galacturonic acid; GalNAc, N-acetyl-D-galactose; GalNAcA, N-acetyl-D-galactosaminuronic acid; GDP, guanosine 5'-diphosphate; Glc, D-glucose; GlcA, D-glucuronic acid; GlcNAc, N-acetyl-D-glucose; GlcNAc(3NAc)A, 2,3-acetamido-2,3-dideoxy-D-glucuronic acid; GPI, glycosylphosphatidylinositol; LPS, lipopolysaccharide; Man, D-mannose; ManNAcA, D-mannosaminuronic acid; NAD⁺, nicotinamide adenine dinucleotide; OT, eukaryotic oligosaccharyl transferase complex; OTase, general oligosaccharyl transferase; PLP, pyridoxal 5'-phosphate; UDP, uridine 5'-diphosphate; Und, undecaprenol; Und-P, undecaprenylphosphate; Und-PP, undecaprenyldiphosphate.

REFERENCES

- (1) Varki, A., Cummings, R. D., Freeze, H. H., Stanley, P., Bertozzi, C. R., Hart, G. W., and Etzler, M. E. (2009) *Essentials of Glycobiology*, 2nd ed., Cold Spring Harbor Laboratory Press, Plainview, NY.
- (2) Walsh, C. T. (2006) *Posttranslational modification of proteins: Expanding nature's inventory*, Roberts and Company Publishers, Greenwood Village, CO.
- (3) Jensen, P. H., Kolarich, D., and Packer, N. H. (2010) Mucin-type O-glycosylation: Putting the pieces together. *FEBS J.* 277, 81–94.
- (4) Hurtado-Guerrero, R., Dorfmueller, H. C., and van Aalten, D. M. F. (2008) Molecular mechanisms of O-GlcNAcylation. *Curr. Opin. Struct. Biol.* 18, 551–557.
- (5) Spiro, R. G. (2002) Protein glycosylation: Nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. *Glycobiology* 12, 43R–56R.
- (6) Zarschler, K., Janesch, B., Pabst, M., Altmann, F., Messner, P., and Schäffer, C. (2010) Protein tyrosine O-glycosylation: A rather unexplored prokaryotic glycosylation system. *Glycobiology* 20, 787–798.
- (7) Paulick, M. G., and Bertozzi, C. R. (2008) The glycosylphosphatidylinositol anchor: A complex membrane-anchoring structure for proteins. *Biochemistry* 47, 6991–7000.

- (8) Mehta, D. P., Etchinson, J. R., Wu, R., and Freeze, H. H. (1997) UDP-GlcNAc:Ser-protein N-acetylglucosamine-1-phosphotransferase from *Dictyostelium discoideum* recognizes serine-containing peptides and eukaryotic cysteine proteinases. *J. Biol. Chem.* 272, 28638–28645.
- (9) Haynes, P. A. (1998) Phosphoglycosylation: A new structural class of glycosylation? *Glycobiology* 8, 1–5.
- (10) Doucey, M.-A., Hess, D., Cacan, R., and Hofsteenge, J. (1998) Protein C-mannosylation is enzyme-catalyzed and uses dolichyl-phosphate-mannose as a precursor. *Mol. Biol. Cell* 9, 291–300.
- (11) Nothaft, H., and Szymanski, C. M. (2010) Protein glycosylation in bacteria: Sweeter than ever. *Nat. Rev. Microbiol.* 8, 765–778.
- (12) Calo, D., Kaminski, L., and Eichler, J. (2010) Protein glycosylation in archaea: Sweet and extreme. *Glycobiology* 20, 1065–1076.
- (13) North, S. J., Hitchen, P. G., Haslam, S. M., and Dell, A. (2009) Mass spectrometry in the analysis of N-linked and O-linked glycans. *Curr. Opin. Struct. Biol.* 19, 498–506.
- (14) Mariño, K., Bones, J., Kattla, J. J., and Rudd, P. M. (2010) A systematic approach to protein glycosylation analysis: A path through the maze. *Nat. Chem. Biol.* 6, 713–723.
- (15) Huffaker, T. C., and Robbins, P. W. (1983) Yeast mutants deficient in protein glycosylation. *Proc. Natl. Acad. Sci. U.S.A.* 80, 7466–7470.
- (16) Lehle, L., Strahl, S., and Tanner, W. (2006) Protein glycosylation, conserved from yeast to man: A model organism helps elucidate congenital human diseases. *Angew. Chem., Int. Ed.* 45, 6802–6818.
- (17) Jones, M. B., Rosenberg, J. N., Betenbaugh, M. J., and Krag, S. S. (2009) Structure and synthesis of polyisoprenoids used in N-glycosylation across the three domains of life. *Biochim. Biophys. Acta* 1790, 485–494.
- (18) Heller, L., Orlean, P., and Adair, W. L., Jr. (1992) *Saccharomyces cerevisiae* sec59 cells are deficient in dolichol kinase activity. *Proc. Natl. Acad. Sci. U.S.A.* 89, 7013–7016.
- (19) Shiridas, P., and Waechter, C. J. (2006) Human dolichol kinase, a polytopic endoplasmic reticulum membrane protein with a cytoplasmically oriented CTP-binding site. *J. Biol. Chem.* 281, 31696–31704.
- (20) Schutzbach, J. S., and Jensen, J. W. (1989) Bilayer membrane destabilization induced by dolichylphosphate. *Chem. Phys. Lipids* 51, 213–218.
- (21) Chojnacki, T., and Dallner, G. (1988) The biological role of dolichol. *Biochem. J.* 251, 1–9.
- (22) Schenk, B., Fernandez, F., and Waechter, C. J. (2001) The in(side) and out(side) of dolichyl phosphate biosynthesis and recycling in the endoplasmic reticulum. *Glycobiology* 11, 61R–70R.
- (23) Lehrman, M. A. (1991) Biosynthesis of N-acetylglucosamine-P-P-dolichol, the committed step of asparagine-linked oligosaccharide assembly. *Glycobiology* 1, 553–562.
- (24) Bickel, T., Lehle, L., Schwarz, M., Aebi, M., and Jakob, C. A. (2005) Biosynthesis of lipid-linked oligosaccharides in *Saccharomyces cerevisiae*: Alg13p and Alg14p form a complex required for the formation of GlcNAc2-PP-Dolichol. *J. Biol. Chem.* 280, 34500–34506.
- (25) Gao, X.-D., Tachikawa, H., Sato, T., Jigami, Y., and Dean, N. (2005) Alg14 recruits Alg13 to the cytoplasmic face of the endoplasmic reticulum to form a novel bipartite UDP-N-acetylglucosamine transferase required for the second step of N-linked glycosylation. *J. Biol. Chem.* 280, 36254–36262.
- (26) Gao, X.-D., Moriyama, S., Miura, N., Dean, N., and Nishimura, S.-I. (2008) Interactions between the C-Termini of Alg13 and Alg14 mediates formation of the active UDP-N-acetylglucosamine transferase complex. *J. Biol. Chem.* 283, 32534–32541.
- (27) Wang, X., Weldeghorghis, T., Zhang, G., Imperiali, B., and Prestegard, J. H. (2008) Solution structure of Alg13: The sugar donor subunit of a yeast N-acetylglucosamine transferase. *Structure* 16, 965–975.
- (28) Couto, J. R., Huffaker, T. C., and Robbins, P. W. (1984) Cloning and expression in *Escherichia coli* of a yeast mannosyltransferase from the asparagine-linked glycosylation pathway. *J. Biol. Chem.* 259, 378–382.
- (29) Watt, G. M., Revers, L., Webberley, M. C., Wilson, I. B. H., and Flitsch, S. L. (1997) Efficient enzymatic synthesis of the core

trisaccharide of N-glycans with a recombinant β -mannosyltransferase. *Angew. Chem., Int. Ed.* 36, 2354–2356.

(30) Yamakazi, H., Shiraishi, N., Takauchi, K., Ohnishi, Y., and Horinouchi, S. (1998) Characterization of ALG2 encoding a mannosyltransferase in the zygomycete fungus *Rhizomucor pusillus*. *Gene* 221, 179–184.

(31) Cipollo, J. F., Trimble, R. B., Chi, J. H., and Dean, N. (2001) The yeast ALG11 gene specifies addition of the terminal α 1,2-Man to the Man₅GlcNAc₂-PP-dolichol N-glycosylation intermediate formed on the cytosolic side of the endoplasmic reticulum. *J. Biol. Chem.* 276, 21828–21840.

(32) O'Reilly, M. K., Zhang, G., and Imperiali, B. (2006) In vitro evidence for the dual function of Alg2 and Alg11: Essential mannosyltransferases in N-linked glycoprotein biosynthesis. *Biochemistry* 45, 9593–9603.

(33) Gao, X.-D., Nishikawa, A., and Dean, N. (2004) Physical interactions between the Alg1, Alg2, and Alg11 mannosyltransferases of the endoplasmic reticulum. *Glycobiology* 14, 559–570.

(34) Noffz, C., Keppler-Ross, S., and Dean, N. (2009) Hetero-oligomeric interactions between early glycosyltransferases of the dolichol cycle. *Glycobiology* 19, 472–478.

(35) Helenius, J., Ng, D. T. W., Marolda, C., Walter, P., Valvano, M. A., and Aebi, M. (2002) Translocation of lipid-linked oligosaccharides across the ER membrane requires Rft1 protein. *Nature* 415, 447–450.

(36) Frank, C. G., Sanyal, S., Rush, J. S., Waechter, C. J., and Menon, A. K. (2008) Does Rft1 flip an N-linked glycan precursor? *Nature* 454, E3–E5.

(37) Sanyal, S., and Menon, A. K. (2008) Specific transbilayer translocation of dolichol-linked oligosaccharides by an endoplasmic reticulum flippase. *Proc. Natl. Acad. Sci. U.S.A.* 106, 767–772.

(38) Sanyal, S., and Menon, A. K. (2009) Flipping lipids: Why an' what's the reason for? *ACS Chem. Biol.* 4, 895–909.

(39) Rush, J. S., Gao, N., Lehrman, M. A., Matveev, S., and Waechter, C. J. (2009) Suppression of Rft1 expression does not impair the transbilayer movement of Man₅GlcNAc₂-PP-Dolichol in sealed microsomes from yeast. *J. Biol. Chem.* 284, 19835–19842.

(40) Burda, P., and Aebi, M. (1999) The dolichol pathway of N-linked glycosylation. *Biochim. Biophys. Acta* 1426, 239–257.

(41) Frank, C. G., and Aebi, M. (2005) ALG9 mannosyltransferase is involved in two different steps of lipid-linked oligosaccharide biosynthesis. *Glycobiology* 15, 1156–1163.

(42) Maeda, Y., and Kinoshita, T. (2008) Dolichol-phosphate mannose synthase: Structure, function and regulation. *Biochim. Biophys. Acta* 1780, 861–868.

(43) Schutzbach, J. S., Zimmerman, J. W., and Forsee, W. T. (1993) The purification and characterization of recombinant yeast dolichyl-phosphate-mannose synthase. Site-directed mutagenesis of the putative dolichol recognition sequence. *J. Biol. Chem.* 268, 24190–24196.

(44) Sanyal, S., and Menon, A. K. (2010) Stereoselective transbilayer translocation of mannosyl phosphoryl dolichol by an endoplasmic reticulum flippase. *Proc. Natl. Acad. Sci. U.S.A.* 107, 11289–11294.

(45) Heesen, S. T., Lehle, L., Weissmann, A., and Aebi, M. (1994) Isolation of the ALG5 locus encoding the UDP-glucose:dolichyl-phosphate glucosyltransferase from *Saccharomyces cerevisiae*. *Eur. J. Biochem.* 224, 71–79.

(46) Guha-Niyogi, A., Sullivan, D. R., and Turco, S. J. (2001) Glycoconjugate structures of parasitic protozoa. *Glycobiology* 11, 45R–59R.

(47) Parodi, A. J. (1993) N-Glycosylation in trypanosomatid protozoa. *Glycobiology* 3, 193–199.

(48) Samuelson, J., Banerjee, S., Magnelli, P., Cui, J., Kelleher, D. J., Gilmore, R., and Robbins, P. W. (2005) The diversity of dolichol-linked precursors to Asn-linked glycans likely results from secondary loss of sets of glycosyltransferases. *Proc. Natl. Acad. Sci. U.S.A.* 102, 1548–1553.

(49) Yagodnik, C., De la Canal, L., and Parodi, A. J. (1987) *Tetrahymena pyriformis* cells are deficient in all mannose-P-dolichol-dependent mannosyltransferases but not in mannose-P-dolichol synthesis. *Biochemistry* 26, 5937–5943.

(50) Berhe, S., Gerold, P., Kedees, M. H., Holder, A. A., and Schwarz, R. T. (2000) *Plasmodium falciparum*: Merozoite surface proteins 1 and 2 are not posttranslationally modified by classical N- or O-glycans. *Exp. Parasitol.* 94, 194–197.

(51) Dempsey, R. E., Jr., and Imperiali, B. (2002) Oligosaccharyl transferase: Gatekeeper to the secretory pathway. *Curr. Opin. Chem. Biol.* 6, 844–850.

(52) Kelleher, D. J., and Gilmore, R. (2006) An evolving view of the eukaryotic oligosaccharyltransferase. *Glycobiology* 16, 47R–62R.

(53) Wilson, C. M., Roebuck, Q., and High, S. (2008) Ribophorin I regulates substrate delivery to the oligosaccharyltransferase core. *Proc. Natl. Acad. Sci. U.S.A.* 105, 9534–9539.

(54) Ruiz-Canada, C., Kelleher, D. J., and Gilmore, R. (2009) Cotranslational and posttranslational N-glycosylation of polypeptides by distinct mammalian OST isoforms. *Cell* 136, 272–283.

(55) Pathak, R., Hendrickson, T. L., and Imperiali, B. (1995) Sulfhydryl modification of the yeast Wbp1p inhibits oligosaccharyl transferase activity. *Biochemistry* 34, 4179–4185.

(56) Schulz, B. L., Stirnimann, C. U., Grimshaw, J. P. A., Brozzo, M. S., Frisch, F., Mohorko, E., Capitani, G., Glockshuber, R., Grütter, M. G., and Aebi, M. (2009) Oxidoreductase activity of oligosaccharyltransferase subunits Ost3p and Ost6p defines site-specific glycosylation efficiency. *Proc. Natl. Acad. Sci. U.S.A.* 106, 11061–11066.

(57) Wacker, M., Linton, D., Hitchen, P. G., Nita-Lazar, M., Haslam, S. M., North, S. J., Panico, M., Morris, H. R., Dell, A., Wren, B. W., and Aebi, M. (2002) N-linked glycosylation in *Campylobacter jejuni* and its functional transfer into *E. coli*. *Science* 298, 1790–1793.

(58) Glover, K. J., Weerapana, E., Numao, S., and Imperiali, B. (2005) Chemoenzymatic synthesis of glycopeptides with PglB, a bacterial oligosaccharyl transferase from *Campylobacter jejuni*. *Chem. Biol.* 12, 1311–1315.

(59) Parsaie Nasab, F., Schulz, B. L., Gamarro, F., Parodi, A. J., and Aebi, M. (2008) All in one: *Leishmania major* STT3 proteins substitute for the whole oligosaccharyltransferase complex in *Saccharomyces cerevisiae*. *Mol. Biol. Cell* 19, 3758–3768.

(60) Hese, K., Otto, C., Routier, F. O. H., and Lehle, L. (2009) The yeast oligosaccharyltransferase complex can be replaced by STT3 from *Leishmania major*. *Glycobiology* 19, 160–171.

(61) Manthri, S., Güther, M. L. S., Izquierdo, L., Acosta-Serrano, A., and Ferguson, M. A. J. (2008) Deletion of the *Tb*ALG3 gene demonstrates site-specific N-glycosylation and N-glycan processing in *Trypanosoma brucei*. *Glycobiology* 18, 367–383.

(62) Li, H., Chavan, M., Schindelin, H., Lennarz, W. J., and Li, H. (2008) Structure of the oligosaccharyl transferase complex at 12 Å resolution. *Structure* 16, 432–440.

(63) Petrescu, A.-J., Millac, A.-L., Petrescu, S. M., Dwek, R. A., and Wormald, M. R. (2004) Statistical analysis of the protein environment of N-glycosylation sites: Implications for occupancy, structure, and folding. *Glycobiology* 14, 103–114.

(64) Bause, E. (1983) Structural requirements of N-glycosylation of proteins. Studies with proline peptides as conformational probes. *Biochem. J.* 209, 331–336.

(65) Gavel, Y., and von Heijne, G. (1990) Sequence differences between glycosylated and non-glycosylated Asn-X-Thr/Ser acceptor sites: Implications for protein engineering. *Protein Eng.* 3, 433–442.

(66) Nilsson, I. M., and von Heijne, G. (1993) Determination of the distance between the oligosaccharyltransferase active site and the endoplasmic reticulum membrane. *J. Biol. Chem.* 268, 5798–5801.

(67) O'Connor, S. E., and Imperiali, B. (1997) Conformational switching by asparagine-linked glycosylation. *J. Am. Chem. Soc.* 119, 2295–2296.

(68) Imperiali, B., Shannon, K. L., Unno, M., and Rickert, K. W. (1992) A mechanistic proposal for asparagine-linked glycosylation. *J. Am. Chem. Soc.* 114, 7944–7945.

(69) Herscovics, A. (1999) Importance of glycosidases in mammalian glycoprotein biosynthesis. *Biochim. Biophys. Acta* 1473, 96–107.

(70) Roth, J. (2002) Protein N-glycosylation along the secretory pathway: Relationship to organelle topography and function, protein quality control, and cell interactions. *Chem. Rev.* 102, 285–304.

- (71) Freeze, H. H. (2006) Genetic defects in the human glycome. *Nat. Rev. Genet.* 7, 537–551.
- (72) Cantagrel, V., Lefeber, D. J., Ng, B. G., Guan, Z., Silhavy, J. L., Bielas, S. L., Lehle, L., Hombauer, H., Adamowicz, M., Swiezewska, E., De Brouwer, A. P., Blümel, P., Sykut-Cegielska, J., Houliston, R. S., Swiston, D., Ali, B. R., Dobyns, W. B., Babovic-Vuksanovic, D., van Bokhoven, H., Wevers, R. A., Raetz, C. R. H., Freeze, H. H., Morava, É., Al-Gazali, L., and Gleeson, J. G. (2010) SRD5A3 is required for converting polyprenol to dolichol and is mutated in a congenital glycosylation disorder. *Cell* 142, 203–217.
- (73) Moreman, K. M., and Molinari, M. (2006) N-Linked glycan recognition and processing: The molecular basis of endoplasmic reticulum quality control. *Curr. Opin. Struct. Biol.* 16, 592–599.
- (74) Aebi, M., Bernasconi, R., Clerc, S., and Molinari, M. (2009) N-Glycan structures: Recognition and processing in the ER. *Trends Biochem. Sci.* 35, 74–82.
- (75) Parodi, A. J. (2000) Protein glucosylation and its role in protein folding. *Annu. Rev. Biochem.* 69, 69–93.
- (76) Molinari, M. (2007) N-Glycan structure dictates extension of protein folding or onset of disposal. *Nat. Chem. Biol.* 3, 313–320.
- (77) Vembar, S. S., and Brodsky, J. L. (2008) One step at a time: Endoplasmic reticulum-associated degradation. *Nat. Rev. Mol. Cell Biol.* 9, 944–957.
- (78) Dennis, J. W., Lau, K. S., Demetriou, M., and Nabi, I. R. (2009) Adaptive regulation at the cell surface by N-glycosylation. *Traffic* 10, 1569–1578.
- (79) Wyss, D., and Wagner, G. (1996) The structural role of sugars in glycoproteins. *Curr. Opin. Biotechnol.* 7, 409–416.
- (80) Kundra, R., and Kornfeld, S. (1999) Asparagine-linked oligosaccharides protect Lamp-1 and Lamp-2 from intracellular proteolysis. *J. Biol. Chem.* 274, 31039–31046.
- (81) Lau, K. S., and Dennis, J. W. (2008) N-Glycans in cancer progression. *Glycobiology* 18, 750–760.
- (82) Arnold, J. N., Saldova, R., Hamid, U. M. A., and Rudd, P. M. (2008) Evaluation of the serum N-linked glycome for the diagnosis of cancer and chronic inflammation. *Proteomics* 8, 3284–3293.
- (83) Schmidt, M. A., Riley, L. W., and Benz, I. (2003) Sweet new world: Glycoproteins in bacterial pathogens. *Trends Microbiol.* 11, 554–561.
- (84) Messner, P. (2004) Prokaryotic glycoproteins: Unexplored but important. *J. Bacteriol.* 186, 2517–2519.
- (85) Hitchen, P. G., and Dell, A. (2006) Bacterial glycoproteomics. *Microbiology* 152, 1575–1580.
- (86) Szymanski, C. M., Yao, R., Ewing, C. P., Trust, T. J., and Guerry, P. (1999) Evidence for a system of general protein glycosylation in *Campylobacter jejuni*. *Mol. Microbiol.* 32, 1022–1030.
- (87) Szymanski, C. M., Burr, D. H., and Guerry, P. (2002) *Campylobacter* protein glycosylation affects host cell interactions. *Infect. Immun.* 70, 2242–2244.
- (88) Linton, D., Allan, E., Karlyshev, A. V., Cronshaw, A. D., and Wren, B. W. (2002) Identification of N-acetylgalactosamine-containing glycoproteins PEB3 and CgpA in *Campylobacter jejuni*. *Mol. Microbiol.* 43, 497–508.
- (89) Young, N. M., Brisson, J.-R., Kelly, J., Watson, D. C., Tessier, L., Lanthier, P. H., Jarrell, H. C., Cadotte, N., St. Michael, F., Aberg, E., and Szymanski, C. M. (2002) Structure of the N-linked glycan present on multiple glycoproteins in the Gram-negative bacterium, *Campylobacter jejuni*. *J. Biol. Chem.* 277, 42530–42539.
- (90) Thibodeaux, C. J., Melançon, C. E., and Liu, H.-w. (2007) Unusual sugar biosynthesis and natural product glycodiversification. *Nature* 446, 1008–1016.
- (91) Holden, H. M., Cook, P. D., and Thoden, J. B. (2010) Biosynthetic enzymes of unusual microbial sugars. *Curr. Opin. Struct. Biol.* 20, 543–550.
- (92) Sharon, N. (2007) Celebrating the golden anniversary of the discovery of bacillosamine, the diamino sugar of a *Bacillus*. *Glycobiology* 17, 1150–1155.
- (93) Kocharova, N. A., Perepelov, A. V., Zatonsky, G. V., Shashkov, A. S., Knirel, Y. A., Jansson, P.-E., and Weintraub, A. (2001) Structural studies of the O-specific polysaccharide of *Vibrio cholerae* O8 using solvolysis with triflic acid. *Carbohydr. Res.* 330, 83–92.
- (94) Stimson, E., Virji, M., Makepeace, K., Dell, A., Morris, H. R., Payne, G., Saunders, J. R., Jennings, M. P., Barker, S., Panico, M., Blench, I., and Moxon, E. R. (1995) Meningococcal pilin: A glycoprotein substituted with digalactosyl 2,4-diacetamido-2,4,6-trideoxyhexose. *Mol. Microbiol.* 17, 1201–1214.
- (95) Hanuszkiewicz, A., Kaczyński, Z., Lindner, B., Goldmann, T., Vollmer, E., Debarry, J., Heine, H., and Holst, O. (2008) Structural analysis of the capsular polysaccharide from *Acinetobacter lwoffii* F78. *Eur. J. Org. Chem.* 2008, 6183–6188.
- (96) Schoenhofen, I. C., McNally, D. J., Vinogradov, E., Whitfield, D., Young, N. M., Dick, S., Wakarchuk, W. W., Brisson, J.-R., and Logan, S. M. (2006) Functional characterization of dehydratase/aminotransferase pairs from *Helicobacter* and *Campylobacter*: Enzymes distinguishing the pseudaminic acid and bacillosamine biosynthetic pathways. *J. Biol. Chem.* 281, 723–732.
- (97) Olivier, N. B., Chen, M. M., Behr, J. R., and Imperiali, B. (2006) In vitro biosynthesis of UDP-N,N'-diacetyl bacillosamine by enzymes of the *Campylobacter jejuni* general protein glycosylation system. *Biochemistry* 45, 13659–13669.
- (98) Glover, K. J., Weerapana, E., Chen, M. M., and Imperiali, B. (2006) Direct biochemical evidence for the utilization of UDP-bacillosamine by PglC, an essential glycosyl-1-phosphate transferase in the *Campylobacter jejuni* N-linked glycosylation pathway. *Biochemistry* 45, 5343–5350.
- (99) Weerapana, E., Glover, K. J., Chen, M. M., and Imperiali, B. (2005) Investigating bacterial N-linked glycosylation: Synthesis and glycosyl acceptor activity of the undecaprenyl pyrophosphate-linked bacillosamine. *J. Am. Chem. Soc.* 127, 13766–13767.
- (100) Glover, K. J., Weerapana, E., and Imperiali, B. (2005) In vitro assembly of the undecaprenylpyrophosphate-linked heptasaccharide for prokaryotic N-linked glycosylation. *Proc. Natl. Acad. Sci. U.S.A.* 102, 14255–14259.
- (101) Chen, M. M., Weerapana, E., Ciepihal, E., Stupak, J., Reid, C. W., Swiezewska, E., and Imperiali, B. (2007) Polyisoprenol specificity in the *Campylobacter jejuni* N-linked glycosylation pathway. *Biochemistry* 46, 14342–14348.
- (102) Troutman, J. M., and Imperiali, B. (2009) *Campylobacter jejuni* PglH is a single active site processive polymerase that utilizes product inhibition to limit sequential glycosyl transfer reactions. *Biochemistry* 48, 2807–2816.
- (103) Baar, C., Eppinger, M., Raddatz, G., Simon, J. R., Lanz, C., Klimmek, O., Nandakumar, R., Gross, R., Rosinus, A., Keller, H., Jagtap, P., Linke, B., Meyer, F., Lederer, H., and Schuster, S. C. (2003) Complete genome sequence and analysis of *Wolinetella succinogenes*. *Proc. Natl. Acad. Sci. U.S.A.* 100, 11690–11695.
- (104) Santos-Silva, T., Dias, J. M., Dolla, A., Durand, M.-C., Gonçalves, L. L., Lampreia, J., Moura, I., and Romão, M. J. o. (2007) Crystal structure of the 16 heme cytochrome from *Desulfovibrio gigas*: A glycosylated protein in a sulphate-reducing bacterium. *J. Mol. Biol.* 370, 659–673.
- (105) Jervis, A. J., Langdon, R., Hitchen, P., Lawson, A. J., Wood, A., Fothergill, J. L., Morris, H. R., Dell, A., Wren, B., and Linton, D. (2010) Characterization of N-linked protein glycosylation in *Helicobacter pullorum*. *J. Bacteriol.* 192, 5228–5236.
- (106) Power, P. M., Roddam, L. F., Rutter, K., Fitzpatrick, S. Z., Srikantha, Y. N., and Jennings, M. P. (2003) Genetic characterization of pilin glycosylation and phase variation in *Neisseria meningitidis*. *Mol. Microbiol.* 49, 833–847.
- (107) Vik, Å., Aas, F. E., Anonsen, J. H., Bilsborough, S., Schneider, A., Egge-Jacobsen, W., and Koomey, M. (2009) Broad spectrum O-linked protein glycosylation in the human pathogen *Neisseria gonorrhoeae*. *Proc. Natl. Acad. Sci. U.S.A.* 106, 4447–4452.
- (108) Faridmoayer, A., Fentabil, M. A., Mills, D. C., Klassen, J. S., and Feldman, M. F. (2007) Functional characterization of bacterial oligosaccharyltransferases involved in O-linked protein glycosylation. *J. Bacteriol.* 189, 8088–8098.

- (109) Maita, N., Nyirenda, J., Igura, M., Kamishikiryo, J., and Kohda, D. (2010) Comparative structural biology of eubacterial and archaeal oligosaccharyltransferases. *J. Biol. Chem.* 285, 4941–4950.
- (110) Kowarik, M., Numao, S., Feldman, M. F., Schulz, B. L., Callewaert, N., Kiermaier, E., Catrein, I., and Aebi, M. (2006) N-Linked glycosylation of folded proteins by the bacterial oligosaccharyltransferase. *Science* 314, 1148–1150.
- (111) Nita-Lazar, M., Wacker, M., Schegg, B., Amber, S., and Aebi, M. (2005) The N-X-S/T consensus sequence is required but not sufficient for bacterial N-linked protein glycosylation. *Glycobiology* 15, 361–367.
- (112) Chen, M. M., Glover, K. J., and Imperiali, B. (2007) From peptide to protein: Comparative analysis of the substrate specificity of N-linked glycosylation in *C. jejuni*. *Biochemistry* 46, 5579–5585.
- (113) Rangarajan, E. S., Bhatia, S., Watson, D. C., Munger, C., Cygler, M., Matte, A., and Young, N. M. (2007) Structural context for protein N-glycosylation in bacteria: The structure of PEB3, an adhesin from *Campylobacter jejuni*. *Protein Sci.* 16, 990–995.
- (114) Slynko, V., Schubert, M., Numao, S., Kowarik, M., Aebi, M., and Allain, F. H. T. (2009) NMR structure determination of a segmentally labeled glycoprotein using in vitro glycosylation. *J. Am. Chem. Soc.* 131, 1274–1281.
- (115) Feldman, M. F., Wacker, M., Hernandez, M., Hitchen, P. G., Marolda, C. L., Kowarik, M., Morris, H. R., Dell, A., Valvano, M. A., and Aebi, M. (2005) Engineering N-linked protein glycosylation with diverse O-antigen lipopolysaccharide structures in *Escherichia coli*. *Proc. Natl. Acad. Sci. U.S.A.* 102, 3016–3021.
- (116) Wacker, M., Feldman, M. F., Callewaert, N., Kowarik, M., Clarke, B. R., Pohl, N., Hernandez, M., Vines, E. D., Valvano, M., Whitfield, C., and Aebi, M. (2006) Substrate specificity of bacterial oligosaccharyltransferase suggests a common transfer mechanism for the bacterial and eukaryotic systems. *Proc. Natl. Acad. Sci. U.S.A.* 103, 7088–7903.
- (117) Schwarz, F., Huang, W., Li, C., Schulz, B. L., Lizak, C., Palumbo, A., Numao, S., Neri, D., Aebi, M., and Wang, L.-X. (2010) A combined method for producing homogeneous glycoproteins with eukaryotic N-glycosylation. *Nat. Chem. Biol.* 6, 264–266.
- (118) Fisher, A. C., Haitjema, C. H., Guarino, C., Celik, E., Endicott, C. E., Reading, C. A., Merritt, J. H., Ptak, A. C., Zhang, S., and DeLisa, M. P. (2011) Production of secretory and extracellular N-Linked Glycoproteins in *Escherichia coli*. *Appl. Environ. Microbiol.* 77, 871–881.
- (119) Hendrixson, D. R., and DiRita, V. J. (2004) Identification of *Campylobacter jejuni* genes involved in commensal colonization of the chick gastrointestinal tract. *Mol. Microbiol.* 52, 471–484.
- (120) Van Sorge, N. M., Bleumink, N. M. C., Van Vliet, S. J., Saeland, E., Van Der Pol, W. L., Van Kooyk, Y., and Van Putten, J. P. M. (2009) N-Glycosylated proteins and distinct lipooligosaccharide glycoforms of *Campylobacter jejuni* target the human C-type lectin receptor MGL. *Cell. Microbiol.* 11, 1768–1781.
- (121) Larsen, J. C., Szymanski, C., and Guerry, P. (2004) N-Linked protein glycosylation is required for full competence in *Campylobacter jejuni* 81-176. *J. Bacteriol.* 186, 6508–6514.
- (122) Nothaft, H., Liu, X., McNally, D. J., Li, J., and Szymanski, C. M. (2009) Study of free oligosaccharides derived from the bacterial N-glycosylation pathway. *Proc. Natl. Acad. Sci. U.S.A.* 106, 15019–15024.
- (123) Woese, C. R., and Fox, G. E. (1977) Phylogenetic structure of the prokaryotic domain: The primary kingdoms. *Proc. Natl. Acad. Sci. U.S.A.* 74, 5088–5090.
- (124) Mescher, M. F., and Strominger, J. L. (1976) Purification and characterization of a prokaryotic glucoprotein from the cell envelope of *Halobacterium salinarum*. *J. Biol. Chem.* 251, 2005–2014.
- (125) Wieland, F., Heitzer, R., and Schaefer, W. (1983) Asparaginylglucose: Novel type of carbohydrate linkage. *Proc. Natl. Acad. Sci. U.S.A.* 80, 5470–5474.
- (126) Lechner, J., and Wieland, F. (1989) Structure and biosynthesis of prokaryotic glycoproteins. *Annu. Rev. Biochem.* 58, 173–194.
- (127) Wieland, F., Paul, G., and Sumper, M. (1985) Halobacterial flagellins are sulfated glycoproteins. *J. Biol. Chem.* 260, 15180–15185.
- (128) Paul, G., Lottspeich, F., and Wieland, F. (1986) Asparaginyl-N-acetylgalactosamine. Linkage unit of halobacterial glycosaminoglycan. *J. Biol. Chem.* 261, 1020–1024.
- (129) Kärcher, U., Schröder, H., Haslinger, E., Allmaier, G., Schreiner, R., Wieland, F., Haselbeck, A., and König, H. (1993) Primary structure of the heterosaccharide of the surface glycoprotein of *Methanothermus fervidus*. *J. Biol. Chem.* 268, 26821–26826.
- (130) Voisin, S., Houlston, R. S., Kelly, J., Brisson, J.-R., Watson, D., Bardy, S. L., Jarrell, K. F., and Logan, S. M. (2005) Identification and characterization of the unique N-linked glycan common to the flagellins and S-layer glycoprotein of *Methanococcus voltae*. *J. Biol. Chem.* 280, 16586–16593.
- (131) Chaban, B., Logan, S. M., Kelly, J. F., and Jarrell, K. F. (2009) AglC and AglK are involved in biosynthesis and attachment of diacetylated glucuronic acid to the N-glycan in *Methanococcus voltae*. *J. Bacteriol.* 191, 187–195.
- (132) Kelly, J., Logan, S. M., Jarrell, K. F., VanDyke, D. J., and Vinogradov, E. V. (2009) A novel N-linked flagellar glycan from *Methanococcus maripaludis*. *Carbohydr. Res.* 344, 648–653.
- (133) Ng, S. Y. M., Wu, J., Nair, D. B., Logan, S. M., Robotham, A., Tessier, L., Kelly, J. F., Uchida, K., Aizawa, S.-I., and Jarrell, K. F. (2011) Genetic and mass spectrometry analyses of the unusual type IV-like pili of the archaeon *Methanococcus maripaludis*. *J. Bacteriol.* 193, 804–814.
- (134) Zähringer, U., Moll, H., Hettmann, T., Knirel, Y. A., and Schäfer, G. (2000) Cytochrome b_{558/566} from the archaeon *Sulfolobus acidocaldarius* has a unique Asn-linked highly branched hexasaccharide chain containing 6-sulfoquinovose. *Eur. J. Biochem.* 267, 4144–4149.
- (135) Yang, L. L., and Haug, A. (1979) Purification and partial characterization of a procaryotic glycoprotein from the plasma membrane of *Thermoplasma acidophilum*. *Biochim. Biophys. Acta* 556, 265–277.
- (136) Igura, M., Maita, N., Kamishikiryo, J., Yamada, M., Obita, T., Maenaka, K., and Kohda, D. (2008) Structure-guided identification of a new catalytic motif of oligosaccharyltransferase. *EMBO J.* 27, 234–243.
- (137) Eichler, J., and Adams, M. W. W. (2005) Posttranslational protein modification in archaea. *Microbiol. Mol. Biol. Rev.* 69, 393–425.
- (138) Magidovich, H., and Eichler, J. (2009) Glycosyltransferases and oligosaccharyltransferases in archaea: Putative components of the N-glycosylation pathway in the third domain of life. *FEMS Microbiol. Lett.* 300, 2005–2014.
- (139) Chaban, B., Voisin, S., Kelly, J., Logan, S. M., and Jarrell, K. F. (2006) Identification of genes involved in the biosynthesis and attachment of *Methanococcus voltae* N-linked glycans: Insight into N-linked glycosylation pathways in Archaea. *Mol. Microbiol.* 61, 259–268.
- (140) Mescher, M. F., Hansen, U., and Strominger, J. L. (1976) Formation of lipid-linked sugar compounds in *Halobacterium salinarum*. Presumed intermediates in glycoprotein synthesis. *J. Biol. Chem.* 251, 7289–7294.
- (141) Zhu, B. C. R., Drake, R. R., Schweingruber, H., and Laine, R. A. (1995) Inhibition of glycosylation by amphomycin and sugar nucleotide analogs PP36 and PP55 indicates that *Haloflex volcanii* β -glucosylates both glycoproteins and glycolipids through lipid-linked sugar intermediates: Evidence for three novel glycoproteins and a novel sulfated dihexosyl-archaeol glycolipid. *Arch. Biochem. Biophys.* 319, 355–364.
- (142) Lechner, J., Wieland, F., and Sumper, M. (1985) Biosynthesis of sulfated saccharides N-glycosidically linked to the protein via glucose. Purification and identification of sulfated dolichyl monophosphoryl tetrasaccharides from halobacteria. *J. Biol. Chem.* 260, 860–866.
- (143) Lechner, J., Wieland, F., and Sumper, M. (1985) Transient methylation of dolichyl oligosaccharides is an obligatory step in halobacterial sulfated glycoprotein biosynthesis. *J. Biol. Chem.* 260, 8984–8989.
- (144) Guan, Z., Naparstek, S., Kaminski, L., Konrad, Z., and Eichler, J. (2010) Distinct glycan-charged phosphodolichol carriers are required for the assembly of the pentasaccharide N-linked to the *Haloflex volcanii* S-layer glycoprotein. *Mol. Microbiol.* 78, 1294–1303.
- (145) Shams-Eldin, H., Chaban, B., Niehus, S., Schwarz, R. T., and Jarrell, K. F. (2008) Identification of the archaeal *alg7* gene homolog (encoding N-acetylglucosamine-1-phosphate transferase) of the N-linked

glycosylation system by cross-domain complementation in *Saccharomyces cerevisiae*. *J. Bacteriol.* 190, 2217–2220.

(146) Kaminski, L., Abu-Qarn, M., Guan, Z., Naparstek, S., Ventura, V. V., Raetz, C. R. H., Hitchen, P. G., Dell, A., and Eichler, J. (2010) AglJ adds the first sugar of the N-linked pentasaccharide decorating the *Haloflex volcanii* S-layer glycoprotein. *J. Bacteriol.* 192, 5572–5579.

(147) Rush, J. S., Alaimo, C., Robbiani, R., Wacker, M., and Waechter, C. J. (2010) A novel epimerase that converts GlcNAc-P-P-undecaprenol to GalNAc-P-P-undecaprenol in *Escherichia coli* O1570. *J. Biol. Chem.* 285, 1671–1680.

(148) Jarrell, K. F., Jones, G. M., and Nair, D. B. (2010) Biosynthesis and role of N-linked glycosylation in cell surface structures of archaea with a focus on flagella and S-layers. *Int. J. Microbiol.* 2010, 470138.

(149) VanDyke, D. J., Wu, J., Logan, S. M., Kelly, J. F., Mizuno, S., Aizawa, S.-I., and Jarrell, K. F. (2009) Identification of genes involved in the assembly and attachment of a novel flagellin N-linked tetrasaccharide important for motility in the archaeon *Methanococcus maripaludis*. *Mol. Microbiol.* 72, 633–644.

(150) Mescher, M. F., and Strominger, J. L. (1978) Glycosylation of the surface glycoprotein of *Halobacterium salinarum* via a cyclic pathway of lipid-linked intermediates. *FEBS Lett.* 89, 37–41.

(151) Storm, D. R., and Strominger, J. L. (1973) Complex formation between bacitracin peptides and isoprenyl pyrophosphates. *J. Biol. Chem.* 248, 3940–3945.

(152) VanDyke, D. J., Wu, J., Logan, S. M., Kelly, J. F., Mizuno, S., Aizawa, S.-I., and Jarrell, K. F. (2009) Identification of genes involved in the assembly and attachment of a novel flagellin N-linked tetrasaccharide important for motility in the archaeon *Methanococcus maripaludis*. *Mol. Microbiol.* 72, 633–644.

(153) Abu-Qarn, M., and Eichler, J. (2007) An analysis of amino acid sequences surrounding archaeal glycoprotein sequons. *Archaea* 2, 73–81.

(154) Sprott, G. D. (1992) Structures of archaeobacterial membrane lipids. *J. Bioenerg. Biomembr.* 24, 555–566.

(155) Valentine, D. L. (2007) Adaptations to energy stress dictate the ecology and evolution of the Archaea. *Nat. Rev. Microbiol.* 5, 316–323.

(156) Madern, D., Ebel, C., and Zaccai, G. (2000) Halophilic adaptation of enzymes. *Extremophiles* 4, 91–98.

(157) Vieille, C., and Zeikus, G. J. (2001) Hyperthermophilic enzymes: Sources, uses, and molecular mechanisms for thermostability. *Microbiol. Mol. Biol. Rev.* 65, 1–43.

(158) Luke, K. A., Higgins, C. L., and Wittung-Stafshede, P. (2007) Thermodynamic stability and folding of proteins from hyperthermophilic organisms. *FEBS J.* 274, 4023–4033.

(159) Vigerust, D. J., and Shepherd, V. L. (2007) Virus glycosylation: Role in virulence and immune interactions. *Trends Microbiol.* 15, 211–218.

(160) Keil, W., Geyer, R., Dabrowski, U., Niemann, H., Stirn, S., and Klenk, H. D. (1985) Carbohydrates of influenza virus. Structural elucidation of the individual glycans of the FPV hemagglutinin by two-dimensional ¹H n.m.r. and methylation analysis. *EMBO J.* 4, 2711–2720.

(161) Wang, C.-C., Chen, J.-R., Tseng, Y.-C., Hsu, C.-H., Hung, Y.-F., Chen, S.-W., Chen, C.-M., Khoo, K.-H., Cheng, T.-J., Cheng, Y.-S. E., Jan, J.-T., Wu, C.-Y., Ma, C., and Wong, C.-H. (2009) Glycans on influenza hemagglutinin affect receptor binding and immune response. *Proc. Natl. Acad. Sci. U.S.A.* 106, 18137–18142.

(162) Doores, K. J., Bonomelli, C., Harvey, D. J., Vasiljevic, S., Dwek, R. A., Burton, D. R., Crispin, M., and Scanlan, C. N. (2010) Envelope glycans of immunodeficiency virions are almost entirely oligomannose antigens. *Proc. Natl. Acad. Sci. U.S.A.* 107, 13800–13805.

(163) Korber, B., Gaschen, B., Yusim, K., Thakallapally, R., Kesmir, C., and Detours, V. (2001) Evolutionary and immunological implications of contemporary HIV-1 variation. *Br. Med. Bull.* 58, 19–42.

(164) Calarese, D. A., Lee, H.-K., Huang, C.-Y., Best, M. D., Astronomo, R. D., Stanfield, R. L., Kattinger, H., Burton, D. R., Wong, C.-H., and Wilson, I. A. (2005) Dissection of the carbohydrate specificity

of the broadly neutralizing anti-HIV-1 antibody 2G12. *Proc. Natl. Acad. Sci. U.S.A.* 102, 13372–13377.

(165) Kelly, J., Jarrell, H., Millar, L., Tessier, L., Fiori, L. M., Lau, P. C., Allan, B., and Szymanski, C. M. (2006) Biosynthesis of the N-linked glycan in *Campylobacter jejuni* and addition onto protein through block transfer. *J. Bacteriol.* 188, 2427–2434.

(166) Brameld, K. A., Shrader, W. D., Imperiali, B., and Goddard, W. A., III (1998) Substrate assistance in the mechanism of family 18 chitinases: Theoretical studies of potential intermediates and inhibitors. *J. Mol. Biol.* 280, 913–923.

(167) Tai, V. W. F., and Imperiali, B. (2001) Substrate specificity of the glycosyl donor for oligosaccharyl transferase. *J. Org. Chem.* 66, 6217–6228.