Published in final edited form as:

ACS Chem Biol. 2018 April 20; 13(4): 833–837. doi:10.1021/acscchembio.8b00163.

Biogenesis of Asparagine-linked Glycoproteins Across Domains of Life - Similarities and Differences

Jerry Eichler¹ and Barbara Imperiali²

¹Department of Life Sciences, Ben Gurion University of the Negev, Beersheva, Israel

²Department of Biology and Department of Chemistry, Massachusetts Institute of Technology, Cambridge MA, USA

Asparagine-linked (N-linked) glycosylation is a complex protein modification reaction with diverse structural and functional implications. The myriad essential functions of N-linked glycoproteins in Eukarya are amply documented¹ and changes in glycan signatures in humans accompany diverse disease states including many cancers.^{2–4} N-glycosylation in Archaea is an almost universal modification that affects cell surface proteins, possibly providing stabilizing effects that contribute to survival in the harshest environments on the globe.^{5,6} In contrast, although an understanding of the significance of N-glycosylation in Bacteria is still emerging, it now is evident that the modification, while non-essential for survival, may play an important role in the virulence of some Gram-negative epsilon- and delta-proteobacteria.^{7,8}

The biogenesis of N-linked glycoproteins throughout evolution shares a common molecular logic.⁹ The pathways involve highly-coordinated, multistep processes, which are spatially defined on both faces of cellular membranes (Fig. 1). In Eukarya, assembly and transfer of glycans to proteins occurs at the endoplasmic reticulum (ER) membrane,¹⁰ in Archaea at the plasma membrane⁶ and in Gram-negative bacteria at the inner membrane.⁷ Membrane association is a critical feature of glycoprotein biosynthesis as it determines the targeting of newly-synthesized conjugates to the cell surface and/or for secretion from cells. Together with membrane localization, a hallmark of N-glycosylation is the use of linear polyprenol phosphate- and diphosphate-linked substrates.¹¹ These amphiphilic moieties serve both for membrane association and as activating groups for glycan assembly. Indeed, the only known dedicated function of extended linear polyprenols in cells is in membrane-associated glycan assembly for glycoconjugate biosynthesis. Still, polyprenols modulate the physical properties of biological membranes, which undoubtedly also contributes to function.^{12,13} At first glance, the polyprenols in Eukarya and Prokarya appear similar; however, there are noteworthy differences. Most intriguing is the observation that the polyprenols range in size from 7–20 isoprene units, with the longest polymers being found in Eukarya, yet with all apparently maintaining the geometric variation of E and Z isoprene units. Eukarya and Archaea feature dolichols, which are characterized by saturation of the α -isoprene unit. The polyprenols in Archaea are also saturated at the ω -isoprene unit, and may include additional saturation. Archaeal polyprenols, moreover, exhibit the greatest diversity in terms of size and

overall isoprene saturation.^{14,15} In contrast, Bacteria most commonly enlist uniformly-unsaturated C₅₅ polyprenols for N-glycosylation. These polyprenols constitute part of the same pool that is implicated in the biosynthesis of other glycoconjugates (e.g. peptidoglycan and O-antigen).¹⁶ The biological significance of the alterations in polyprenol structure is unknown, however it has been suggested that the reduced chemical stability of allylic diphosphate intermediates found in bacterial N-glycosylation may have been a factor in the emergence of α -isoprene saturation.¹³

In addition to differences in polyprenol structure, a surprising variation is found in some Archaea. This involves the action of a polyprenol phosphate glycosyl transferase, rather than a phosphoglycosyl transferase, in the first membrane-committed step of the pathways.¹⁷ Thus, these archaeal pathways feature α -linked polyprenol monophosphate intermediates¹⁸ rather than the corresponding diphosphates that are confirmed in all eukaryotic and bacterial pathways studied to date. It is unknown whether this significant difference is an adaptation to the environments in which Archaea can be found, in which case the biochemical characterization of additional pathways in this domain may reveal important trends, or indeed if this intriguing difference is a clue that may provide new insight into the evolution of the essential eukaryotic pathway (Fig. 2).

The most dramatic differences between N-linked glycosylation in Eukarya and Prokarya is seen in the glycan composition (Fig. 3). In higher Eukarya, N-glycans transferred to nascent proteins most commonly include a GlcNAc₂Man₉Glc₃ branched tetradecasaccharide.¹⁰ However, unicellular protists employ specific truncated variants of this common tetradecasaccharide.¹⁹ In Prokarya, carbohydrate diversity is the norm and glycans found decorating proteins show considerable variation in both the constituent carbohydrates and glycosidic linkages.²⁰ An understanding of the functional significance of this variation is still extremely limited, although in Gram-negative pathogens such as *Campylobacter jejuni*, the glycans clearly play a specific role in host-pathogen interactions.²¹ The virulence of N-linked glycosylation-deficient strains, reflecting enzymes in the protein glycosylation (pgl) pathway, is significantly attenuated.

A key step in N-glycosylation pathways involves the action of a flippase to translocate the growing polyprenol phosphate-linked glycan across the membrane in readiness for further elaboration (in Eukarya) or for direct transfer to protein substrates (in Archaea and Bacteria). In Eukarya, the polytopic membrane protein Rft1 from *Saccharomyces cerevisiae* has been implicated in the ATP-independent translocation based on genetic studies,²² although subsequent biochemical analyses²³ have called this assignment into question. Likewise in Archaea, little is known of the flippase, although in *Haloferax volcanii*, AgIR has been implicated.²⁴ In Bacteria the *C. jejuni* PglK flippase, encoded in the *pgl* operon, is an ATP-dependent ABC-type transporter, which is well-characterized both structurally and functionally.²⁵

Following translocation, oligosaccharyl transferases (OTases) catalyze the *en bloc* transfer of glycans from the polyprenol-linked carrier to selected Asn residues of target proteins. Intriguingly, the one detail that is almost universally conserved is that the modified Asn-amide side chain is part of the -Asn-Xaa-Ser/Thr- consensus sequence where Xaa is any

residue except proline. There are some exceptions to this rule with slight extensions²⁶ or substitutions²⁷ in the consensus sequences, however, some assignments are based solely on *in vivo* analyses and would benefit from biochemical validation. In the past decade, there has been tremendous progress in the structural characterization of monomeric and oligomeric OTases. Most recently two groups have reported on cryo-EM studies of the hetero-octameric *S. cerevisiae* OTase.^{28,29} There are also a number of structures of monomeric OTases from simple protists,³⁰ as well as from Archaea^{31,32} and Bacteria,^{25,33} the latter two being designated as AglB and PglB, respectively. The monomeric OTases are homologous to STT3, the catalytic subunit of the oligomeric eukaryotic enzymes. Originally, the variation between Eukarya and Prokarya appeared to be associated with the more complex cellular compartmentalization of Eukarya, however, the discovery of monomeric protozoal OTases has introduced new subtleties that remain to be understood.^{34–36} The site and timing of N-glycosylation also varies; the most information is available on Eukarya, where the principle mode of N-glycosylation is co-translational and co-translocational,^{37,38} although occurrences of posttranslational N-glycosylation have also been reported.^{37,39} In contrast, in Prokarya, more studies are needed to establish whether there are common principles governing the timing of glycosylation.^{40,41}

Intriguing differences suggest that a deeper understanding of all of the related N-glycosylation pathways may provide insight into the evolution of the essential eukaryotic pathway. At the same time, further defining pathways of N-glycosylation in Archaea and Bacteria could help to usher in a new era of applied glycobiology.

Acknowledgments

J.E. was supported by grants from the Israel Science Foundation (ISF) (grant 109/16) and the ISF-NSFC joint research program (grant 2193/16). B.I. was supported by the NIH (grant GM- 039334).

References

- (1). Varki A (2017) Biological roles of glycans. *Glycobiology* 27, 3–49. [PubMed: 27558841]
- (2). Hoja-Lukowicz D; Przybylo M; Duda M; Pochech E; Bubka M (2017) On the trail of the glycan codes stored in cancer-related cell adhesion proteins. *Biochim Biophys Acta* 1861, 3237–3257.
- (3). Oliveira-Ferrer L; Legler K; Milde-Langosch K (2017) Role of protein glycosylation in cancer metastasis. *Semin Cancer Biol* 44, 141–152. [PubMed: 28315783]
- (4). Varki A; Kannagi R; Toole B; Stanley P In *Essentials of Glycobiology*; rd, Varki A, Cummings RD, Esko JD, Stanley P, Hart GW, Aebi M, Darvill AG, Kinoshita T, Packer NH, Prestegard JH, Schnaar RL, Seeberger PH, Eds. Cold Spring Harbor (NY), 2015, p 597–609.
- (5). Eichler J (2013) Extreme sweetness: protein glycosylation in archaea. *Nat Rev Microbiol* 11, 151–156. [PubMed: 23353769]
- (6). Jarrell KF; Ding Y; Meyer BH; Albers SV; Kaminski L; Eichler J (2014) N- linked glycosylation in Archaea: a structural, functional, and genetic analysis. *Microbiol Mol Biol Rev* 78, 304–341. [PubMed: 24847024]
- (7). Nothaft H; Szymanski CM (2010) Protein glycosylation in bacteria: sweeter than ever. *Nat Rev Microbiol* 8, 765–778. [PubMed: 20948550]
- (8). Szymanski CM; Burr DH; Guerry P (2002) *Campylobacter* protein glycosylation affects host cell interactions. *Infect Immun* 70, 2242–2244. [PubMed: 11895996]
- (9). Larkin A; Imperiali B (2011) The expanding horizons of asparagine-linked glycosylation. *Biochemistry* 50, 4411–4426. [PubMed: 21506607]

- (10). Aebi M (2013) N-linked protein glycosylation in the ER. *Biochim Biophys Acta* 1833, 2430–2437. [PubMed: 23583305]
- (11). Jones MB; Rosenberg JN; Betenbaugh MJ; Krag SS (2009) Structure and synthesis of polyisoprenoids used in N-glycosylation across the three domains of life. *Biochim Biophys Acta* 1790, 485–494. [PubMed: 19348869]
- (12). Janas T; Walinska K; Chojnacki T; Swiezewska E; Janas T (2000) Modulation of properties of phospholipid membranes by the long-chain polyprenol (C(160)). *Chem Phys Lipids* 106, 31–40. [PubMed: 10878233]
- (13). Swiezewska E; Danikiewicz W (2005) Polyisoprenoids: structure, biosynthesis and function. *Prog Lipid Res* 44, 235–258. [PubMed: 16019076]
- (14). Eichler J; Guan Z (2017) Lipid sugar carriers at the extremes: The phosphodolichols Archaea use in N-glycosylation. *Biochim Biophys Acta* 1862, 589–599.
- (15). Taguchi Y; Fujinami D; Kohda D (2016) Comparative Analysis of Archaeal Lipid-linked Oligosaccharides That Serve as Oligosaccharide Donors for Asn Glycosylation. *J Biol Chem* 291, 11042–11054. [PubMed: 27015803]
- (16). Hartley MD; Imperiali B (2012) At the membrane frontier: a prospectus on the remarkable evolutionary conservation of polyprenols and polyprenyl-phosphates. *Arch Biochem Biophys* 517, 83–97. [PubMed: 22093697]
- (17). Eichler J; Imperiali B (2018) Stereochemical Divergence of Polyprenol Phosphate Glycosyltransferases. *Trends Biochem Sci* 43, 10–17. [PubMed: 29183665]
- (18). Larkin A; Chang MM; Whitworth GE; Imperiali B (2013) Biochemical evidence for an alternate pathway in N-linked glycoprotein biosynthesis. *Nat Chem Biol* 9, 367–373. [PubMed: 23624439]
- (19). Samuelson J; Banerjee S; Magnelli P; Cui J; Kelleher DJ; Gilmore R; Robbins PW (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. [PubMed: 15665075]
- (20). Szymanski CM; Wren BW (2005) Protein glycosylation in bacterial mucosal pathogens. *Nat Rev Microbiol* 3, 225–237. [PubMed: 15738950]
- (21). Eichler J; Koomey M (2017) Sweet New Roles for Protein Glycosylation in Prokaryotes. *Trends Microbiol* 25, 662–672. [PubMed: 28341406]
- (22). Helenius J; Ng DT; Marolda CL; Walter P; Valvano MA; Aebi M (2002) Translocation of lipid-linked oligosaccharides across the ER membrane requires Rft1 protein. *Nature* 415, 447–450. [PubMed: 11807558]
- (23). Frank CG; Sanyal S; Rush JS; Waechter CJ; Menon AK (2008) Does Rft1 flip an N-glycan lipid precursor? *Nature* 454, E3–4; [PubMed: 18668045]
- (24). Kaminski L; Guan Z; Abu-Qarn M; Konrad Z; Eichler J (2012) AglR is required for addition of the final mannose residue of the N-linked glycan decorating the *Haloferax volcanii* S-layer glycoprotein. *Biochim Biophys Acta* 1820, 1664–1670. [PubMed: 22750201]
- (25). Perez C; Gerber S; Boilevin J; Bucher M; Darbre T; Aebi M; Reymond JL; Locher KP (2015) Structure and mechanism of an active lipid-linked oligosaccharide flippase. *Nature* 524, 433–438. [PubMed: 26266984]
- (26). Nita-Lazar M; Wacker M; Schegg B; Amber S; 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. [PubMed: 15574802]
- (27). Lowenthal MS; Davis KS; Formolo T; Kilpatrick LE; Phinney KW (2016) Identification of Novel N-Glycosylation Sites at Noncanonical Protein Consensus Motifs. *J Proteome Res* 15, 2087–2101. [PubMed: 27246700]
- (28). Wild R; Kowal J; Eyring J; Ngwa EM; Aebi M; Locher KP (2018) Structure of the yeast oligosaccharyltransferase complex gives insight into eukaryotic N-glycosylation. *Science* 359, 545–550. [PubMed: 29301962]
- (29). Bai L; Wang T; Zhao G; Kovach A; Li H (2018) The atomic structure of a eukaryotic oligosaccharyl transferase complex. *Nature*.
- (30). Ramirez AS; Boilevin J; Biswas R; Gan BH; Janser D; Aebi M; Darbre T; Reymond JL; Locher KP (2017) Characterization of the single-subunit oligosaccharyltransferase STT3A from

Trypanosoma brucei using synthetic peptides and lipid-linked oligosaccharide analogs. *Glycobiology* 27, 525–535. [PubMed: 28204532]

- (31). Matsumoto S; Shimada A; Nyirenda J; Igura M; Kawano Y; Kohda D (2013) Crystal structures of an archaeal oligosaccharyltransferase provide insights into the catalytic cycle of N-linked protein glycosylation. *Proc Natl Acad Sci U S A* 110, 17868–17873. [PubMed: 24127570]
- (32). Igura M; Maita N; Kamishikiryo J; Yamada M; Obita T; Maenaka K; Kohda D (2008) Structure-guided identification of a new catalytic motif of oligosaccharyltransferase. *EMBO J* 27, 234–243. [PubMed: 18046457]
- (33). Lizak C; Gerber S; Numao S; Aeby M; Locher KP (2011) X-ray structure of a bacterial oligosaccharyltransferase. *Nature* 474, 350–355. [PubMed: 21677752]
- (34). Hese K; Otto C; Routier FH; Lehle L (2009) The yeast oligosaccharyltransferase complex can be replaced by STT3 from *Leishmania major*. *Glycobiology* 19, 160–171. [PubMed: 18955371]
- (35). Jinnelov A; Ali L; Tinti M; Guthrie MLS; Ferguson MAJ (2017) Single-subunit oligosaccharyltransferases of *Trypanosoma brucei* display different and predictable peptide acceptor specificities. *J Biol Chem* 292, 20328–20341. [PubMed: 28928222]
- (36). Poljak K; Breitling J; Gauss R; Rugarabamu G; Pellanda M; Aeby M (2017) Analysis of substrate specificity of *Trypanosoma brucei* oligosaccharyltransferases (OSTs) by functional expression of domain-swapped chimeras in yeast. *J Biol Chem* 292, 20342–20352. [PubMed: 29042445]
- (37). Ruiz-Canada C; Kelleher DJ; Gilmore R (2009) Cotranslational and posttranslational N-glycosylation of polypeptides by distinct mammalian OST isoforms. *Cell* 136, 272–283. [PubMed: 19167329]
- (38). Nilsson I; Kelleher DJ; Miao Y; Shao Y; Kreibich G; Gilmore R; von Heijne G; Johnson AE (2003) Photocross-linking of nascent chains to the STT3 subunit of the oligosaccharyltransferase complex. *J Cell Biol* 161, 715–725. [PubMed: 12756234]
- (39). Sato T; Sako Y; Sho M; Momohara M; Suico MA; Shuto T; Nishitoh H; Okiyonedo T; Kokame K; Kaneko M; Taura M; Miyata M; Chosa K; Koga T; Morino-Koga S; Wada I; Kai H (2012) STT3B-dependent posttranslational N-glycosylation as a surveillance system for secretory protein. *Mol Cell* 47, 99–110. [PubMed: 22607976]
- (40). Kowarik M; Numao S; Feldman MF; Schulz BL; Callewaert N; Kiermaier E; Catrein I; Aeby M (2006) N-linked glycosylation of folded proteins by the bacterial oligosaccharyltransferase. *Science* 314, 1148–1150. [PubMed: 17110579]
- (41). Silverman JM; Imperiali B (2016) Bacterial N-Glycosylation Efficiency Is Dependent on the Structural Context of Target Sequences. *J Biol Chem* 291, 22001–22010. [PubMed: 27573243]

Steps of N-glycosylation: 1. Stepwise assembly of lipid-linked glycan; 2. Translocation of lipid-linked glycan across the membrane; 3. Transfer of glycan to protein

Fig. 1 - An overview of the components of N-glycosylation across evolution.

Across evolution, N-linked glycans are initially assembled in a stepwise manner on a lipid carrier (step 1). Once assembled, the lipid-linked glycan is translocated across the membrane (step 2) and the glycan moiety is transferred to selected Asn residues of a target protein by the actions of an oligosaccharyl transferase (step 3). At the same time, each component of the N-glycosylation process described above presents domain-specific traits, as described in the different text boxes. For further details on steps 1–3, see Fig. 3. The purple lipid corresponds to phosphorylated dolichol or polyprenol, while the yellow rectangle corresponds to the glycan. The blue entity corresponds to the oligosaccharyl transferase. DolP, dolichol phosphate; DolPP, dolichol diphosphate; OTase, oligosaccharyl transferase; PrenPP, polyprenol diphosphate.

Fig. 2 - The initial and final steps of N-linked protein glycosylation.

In Eukarya, Bacteria and some Archaea, the N-glycosylation pathway is initiated with polyprenol phosphate phosphoglycosyltransferases (PGTs), which catalyze synthesis of α -linked polyprenol diphosphate-linked sugars. In some Archaea, retaining polyprenol phosphate glycosyltransferases (GTs) mediate synthesis of α -linked polyprenol phosphate-linked sugars. Following glycan assembly, both mono- and diphosphate-linked intermediates serve as substrates for protein N-glycosylation catalyzed by OTases. The sugar drawn in blue corresponds to a hexose or a hexosamine and the mono- and diphosphate linkages are highlighted in yellow.

Fig. 3 - An overview of the steps involved in N-glycosylation across evolution.

Step 1. Stepwise assembly of lipid-linked glycan. Examples of lipid-linked glycan assembly in each domain are presented. In *Homo sapiens* (Eukarya), a heptasaccharide is sequentially assembled on DolPP in the ER membrane facing the cytoplasm. Once assembled, the DolPP-linked heptasaccharide is flipped to face the ER lumen, at which point mannose and glucose moieties are added from the corresponding DolP carriers, which are charged on the cytoplasmic face of the ER membrane and flipped to face the ER lumen, to yield the DolPP-charged tetradecasaccharide. In other Eukarya, simpler DolPP-linked glycans are assembled.

In *Campylobacter jejuni* (Bacteria), a heptasaccharide is assembled onto a polyprenol-PP carrier on the cytoplasmic face of the plasma membrane. In Archaea, considerable variability is seen in terms of glycan composition and the lipid carrier, as exemplified by the N-glycosylation pathways of *Haloferax volcanii* (top), *Methanococcus voltae* (middle) and *Sulfolobus acidocaldarius* (bottom). Where known, the enzymes involved in each step are listed. The identities of the sugars in the three panels are listed in the legend in the bottom right-hand corner.

Step 2. Translocation of lipid-linked glycans. Current knowledge of the components and mechanisms involved in translocation of lipid-linked glycans across a membrane in each domain is provided.

Step 3. Transfer of glycan to protein. Current knowledge of the transfer of the glycan to target protein Asn residues in each domain is provided.