MMBR
Journals.ASM.org

# N-Linked Glycosylation in *Archaea*: a Structural, Functional, and Genetic Analysis

**Ken F. Jarrell,[a] Yan Ding,[a] Benjamin H. Meyer,[b] Sonja-Verena Albers,[b] Lina Kaminski,[c] Jerry Eichler[c]**

Department of Biomedical and Molecular Sciences, Queen's University, Kingston, Ontario, Canada[a]; Molecular Biology of Archaea, Max Planck Institute for Terrestrial Microbiology, Marburg, Germany[b]; Department of Life Sciences, Ben Gurion University, Beersheva, Israel[c]

SUMMARY ..... 305
INTRODUCTION ..... 305
EUKARYAL PROTEIN N-GLYCOSYLATION ..... 305
DISCOVERY OF PROKARYOTIC PROTEIN GLYCOSYLATION ..... 306
N-GLYCOSYLATION PATHWAYS IN ARCHAEAL MODEL SYSTEMS ..... 307
Pathway of N-Linked Glycosylation in *Haloferax volcanii* ..... 307
Topology of N-glycosylation in *Hfx. volcanii* ..... 310
The lipid carrier of *Hfx. volcanii* N-glycosylation ..... 310
Flippase of *Hfx. volcanii* N-glycosylation ..... 311
Current model of N-glycosylation in *Hfx volcanii* ..... 311
A Second N-Glycosylation Pathway in *Hfx. volcanii* ..... 311
Other Recent Insights into Haloarchaeal N-Glycosylation ..... 312
Pathway of N-Linked Glycosylation in Methanogens ..... 312
*Methanococcus voltae* ..... 313
*Methanococcus maripaludis* ..... 314
(i) *M. maripaludis* glycosyltransferases and the oligosaccharyltransferase ..... 314
(ii) Other *M. maripaludis* genes involved or implicated in N-glycosylation ..... 314
(iii) *In vitro* studies of purified *M. maripaludis* enzymes involved in glycan sugar biosynthesis ..... 316
(iv) Lipid carrier of *M. maripaludis* N-glycosylation ..... 317
(v) Current model of N-glycosylation in *M. maripaludis* ..... 317
Pathway of N-Linked Glycosylation in the Thermoacidophile *Sulfolobus acidocaldarius* ..... 317
Lipid carrier of *S. acidocaldarius* N-glycosylation ..... 320
Current model of N-glycosylation in *S. acidocaldarius* ..... 320
ARCHAEAL GLYCAN STRUCTURES ..... 320
Halophiles ..... 320
*Halobacterium salinarum* ..... 320
*Haloferax volcanii* ..... 323
*Haloarcula marismortui* ..... 323
Methanogens ..... 323
*Methanococcus* ..... 323
*Methanothermus fervidus* ..... 324
*Methanosaeta concilii* ..... 324
Thermoacidophiles ..... 324
*Sulfolobus* ..... 324
*Thermoplasma acidophilus* ..... 325
Hyperthermophiles ..... 325
*Pyrococcus furiosus* ..... 325
*Archaeoglobus fulgidus* ..... 326
Archaeal Virus Glycoproteins ..... 326
*N*-OLIGOSACCHARYLTRANSFERASES ..... 327
Requirement for AglB Varies ..... 328
AglB Structure ..... 328
Amino Acid Preferences in and around Archaeal N-Glycosylation Sequons ..... 329
A Variety of Linking Sugars Is Attached by AglB ..... 330
Multiple N-Glycans in a Single Species/Glycoprotein ..... 330
BIOLOGICAL EFFECTS OF N-LINKED GLYCOSYLATION PERTURBATIONS IN *ARCHAEA* ..... 331
FUTURE OUTLOOK ..... 333
ACKNOWLEDGMENTS ..... 334
REFERENCES ..... 334
AUTHOR BIOS ..... 340

Address correspondence to Ken F. Jarrell, jarrellk@queensu.ca.
Copyright © 2014, American Society for Microbiology. All Rights Reserved.
doi:10.1128/MMBR.00052-13

## SUMMARY

N-glycosylation of proteins is one of the most prevalent posttranslational modifications in nature. Accordingly, a pathway with shared commonalities is found in all three domains of life. While excellent model systems have been developed for studying N-glycosylation in both *Eukarya* and *Bacteria*, an understanding of this process in *Archaea* was hampered until recently by a lack of effective molecular tools. However, within the last decade, impressive advances in the study of the archaeal version of this important pathway have been made for halophiles, methanogens, and thermoacidophiles, combining glycan structural information obtained by mass spectrometry with bioinformatic, genetic, biochemical, and enzymatic data. These studies reveal both features shared with the eukaryal and bacterial domains and novel archaeon-specific aspects. Unique features of N-glycosylation in *Archaea* include the presence of unusual dolichol lipid carriers, the use of a variety of linking sugars that connect the glycan to proteins, the presence of novel sugars as glycan constituents, the presence of two very different N-linked glycans attached to the same protein, and the ability to vary the N-glycan composition under different growth conditions. These advances are the focus of this review, with an emphasis on N-glycosylation pathways in *Haloferax*, *Methanococcus*, and *Sulfolobus*.


## INTRODUCTION

One of the major achievements of molecular biology in the last century was the realization that DNA serves as the universal repository of genetic information. This, in turn, led to the concept of the amino acid sequence of a protein being derived from a genetic template encoded within the DNA of a cell. However, it soon became clear that the genomic content of a cell could not fully explain the protein profile seen at any given time. Today, it is clear that other cellular processes, including alternate gene splicing, differential rates of protein synthesis and/or degradation, and a variety of posttranslational protein modifications (PTMs), contribute to the expansion of the proteome beyond what is predicted by the genome.

Posttranslational modifications correspond to a broad set of protein-processing events affecting both nascent polypeptides and more mature proteins. These proteins can be modified by controlled proteolysis, such as signal peptide cleavage; by intra- or interprotein linkages, such as disulfide bridges; or by the attachment of one or multiple classes of molecules, as occurs upon ubiquitination, methylation, acetylation, or glycosylation, for example (1, 2). The covalent attachment of biochemically functional groups, either temporarily or permanently, can influence numerous properties of a protein, including its net charge, folding, stability, activity, subcellular location, and interaction partners. Furthermore, a given protein can undergo multiple PTMs.

Of the various PTMs to which a protein can be subjected, glycosylation, namely, the covalent attachment of sugar residues, is one of the more predominant and likely the most complex. Indeed, it has been postulated that more than half of all eukaryotic proteins are glycosylated (3). The sheer diversity of sugars and sugar derivatives that can be recruited for assembly into glycans, including pentoses, hexoses, hexosamines, deoxyhexoses, uronic acids, and sialic acids, as well as the range of sizes that such glycans can assume, spanning from monosaccharides to large oligosaccharides, coupled with the variety of possible linkages between the components of the glycan, lead to a seemingly endless number of possible protein-linked glycan structures. In addition, modification of the individual sugar components of a glycan by phosphorylation, sulfation, methylation, and/or O-acetylation can serve to further increase this diversity.

To date, five major classes of protein glycosylation have been described on the basis of the manner in which the glycan is linked to the modified protein (Fig. 1). Of these, N-glycosylation (4–7), O-glycosylation (8, 9), and glycosylphosphatidylinositol (GPI) anchoring (10–12) have been well studied biochemically. Less is known of phosphoglycosylation (13–16) and C-mannosylation (17–21). For the *Archaea*, only N-linked glycosylation has been well studied.

In N-glycosylation, an oligosaccharide is assembled on an isoprene-based lipid carrier and then transferred *en bloc* to an asparagine residue found within a sequon (i.e., a conserved Asn-X-Ser/Thr motif, where X is any residue but proline) in the target protein (4, 22–24). The transfer of the oligosaccharide to the protein is mediated by an oligosaccharyltransferase (OST). In higher *Eukarya*, the OST comprises a heterooligomeric complex, whereas a single-subunit OST is found in both *Bacteria* and *Archaea* (6, 7, 25–31). However, a unique N-linked glycosylation system was recently described for *Haemophilus influenzae*, where N-glycosylation occurs in the cytoplasm through the action of a novel glycosyltransferase (GT), independent of OST or a lipid-linked glycan carrier (32, 33).

## EUKARYAL PROTEIN N-GLYCOSYLATION

Work conducted over the last 40 years has provided considerable insight into the eukaryal N-glycosylation process (Fig. 2) (for a recent review, see reference 4). In eukaryotic cells, N-glycosylation begins with the assembly of a lipid-linked oligosaccharide (LLO), with the lipid carrier being identified as dolichol pyrophosphate (DolPP). The LLO is synthesized according to a bipartite assembly process that begins on the cytoplasmic side of the endoplasmic reticulum (ER) membrane with the generation of DolPP-$GlcNAc_2$-$Man_5$ heptasaccharide (M5-DLO). Assembly of the LLO-linked heptasaccharide is initiated with the transfer of *N*-acetylglucosamine (GlcNAc)-1-phosphate from UDP-GlcNAc onto a dolichol phosphate (DolP) carrier in a reaction catalyzed by the Asn-linked glycosylation phosphotransferase Alg7 to yield DolPP-GlcNAc. In the next step, the interacting enzymes Alg13 and Alg14 process a second soluble UDP-GlcNAc donor, resulting in the addition of a second GlcNAc residue to the DolPP-GlcNAc core to yield a DolPP-bound chitobiose moiety. In subsequent reactions, five GTP-activated mannoses are added in a stepwise manner via three specific GTs to build the branched heptasaccharide. In the second part of the bipartite assembly process, M5-DLO is "flipped" across the membrane to face the ER lumen in an ATP-independent manner, although the identification of the flippase remains a source of controversy (34–37). Once oriented to face the ER lumen, the flipped lipid-linked heptasaccharide is further elongated by four mannose and three glucose residues, each donated from DolP-mannose or -glucose; charged on the cytoplasmic face of the ER membrane; and translocated to face the ER lumen by unidentified flippases (38). The OST complex then transfers the fully assembled lipid-linked 14-member branched glycan onto sequons of a nascent polypeptide translocating into the ER lumen. In *Saccharomyces cerevisiae*, the OST complex contains eight different proteins, including the catalytic Stt3 subunit (27–29). The glycoprotein may then pass through the multiple

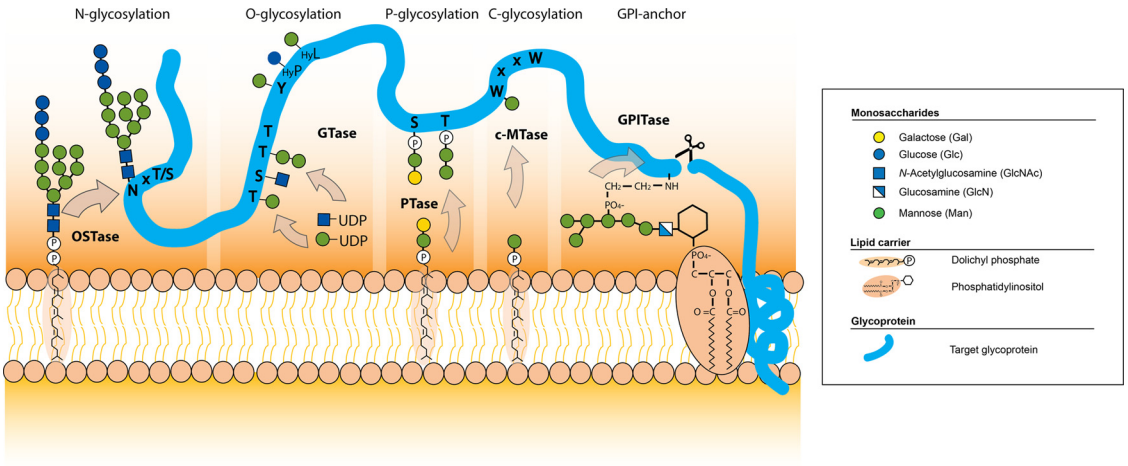



**FIG 1** Schematic overview of the different types of protein glycosylation. Acceptor residues (S, serine; T, threonine; Y, tyrosine; HyP, hydroxyproline; HyL, hydroxylysine; N, asparagine; W, tryptophan) or conserved motifs (N-X-S/T and WXXW) within a newly synthesized peptide (blue band) are indicated in boldface type. The sugar donors, shown as either nucleotide-activated or lipid-bound intermediates, are shown next to the glycosylation site. The catalyzing enzyme or enzyme complex is indicated in boldface type. GPI, glycosylphosphatidylinositol.

stacks of the Golgi apparatus before possibly being ultimately distributed from the *trans*-Golgi network to various destinations in the cell or even secreted beyond the confines of the cell. During this passage, the asparagine-linked oligosaccharide is further modified by the actions of various glycosidases and/or GTs (39). Such downstream modification is a universal feature of the eukaryotic N-glycosylation system.

## DISCOVERY OF PROKARYOTIC PROTEIN GLYCOSYLATION

Following Neuberger's demonstration that a carbohydrate group was an integral part of egg albumin (40), it became generally accepted that protein glycosylation was a trait restricted to eukaryotic cells. However, glycoproteins in the cell envelope of the halophilic archaeon *Halobacterium halobium* (renamed *Halobacterium salinarum* [41]) were detected by both carbohydrate-detecting periodic acid-Schiff (PAS) staining and concanavalin A agglutination (42). It was subsequently shown that a single 194-kDa protein staining positively for the presence of carbohydrates accounted for almost 50% of the *Hbt. salinarum* cell envelope protein content (43). Next, in 1976, Mescher and Strominger (44) demonstrated that this protein, the surface (S)-layer glycoprotein, was subject to both N- and O-glycosylations, thereby providing

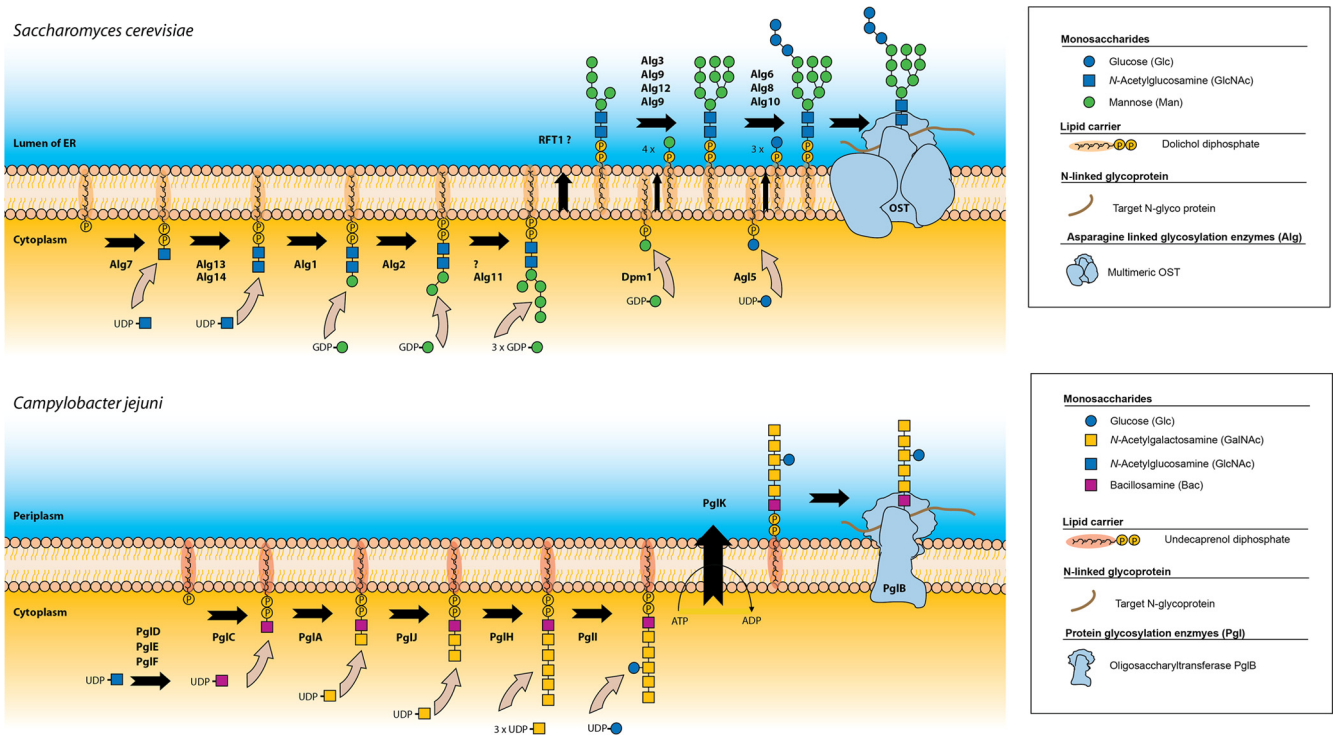



**FIG 2** The N-glycosylation pathways in *Saccharomyces cerevisiae* and *Campylobacter jejuni*.

the first example of a noneukaryotic glycoprotein. Indeed, the *Hbt. salinarum* S-layer glycoprotein consists of approximately 10 to 12% carbohydrate by weight. Analyses at that time indicated the presence of a single N-linked carbohydrate moiety composed of 1 glucose, 1 mannose, 8 to 9 galactose, and 10 to 11 glucosamine residues, along with 6 residues of an unidentified amino sugar. Those glycans linked through O-glycosidic bonds were reported to comprise 22 to 24 disaccharides of glucosylgalactose and 12 to 14 trisaccharides of hexuronic acid (HexA), glucose, and galactose. The linking sugar for both types of O-linked glycans was claimed at that time to be galactose (44). Soon after, proteins decorated by O-linked glycans containing glucose, galactose, mannose, and rhamnose (Rha) residues were detected in the cell walls of the Gram-positive bacteria *Clostridium thermosaccharolyticum* and *Clostridium thermohydrosulfuricum* (45, 46). The discovery of many other exclusively O-linked S-layer glycoproteins in *Bacteria* followed (46, 47), but it would be decades later before the first bacterial N-linked glycosylation system was described for *Campylobacter jejuni* (26, 48, 49).

Today, the N-glycosylation system of *C. jejuni*, which can be functionally transferred into *Escherichia coli*, remains the best-studied case of N-linked glycosylation in bacteria (Fig. 2) (26, 48, 50, 51). A heptasaccharide is assembled on the cytoplasmic side of the cytoplasmic membrane onto the lipid carrier undecaprenyl phosphate (UndP) through the activities of numerous protein glycosylation (*pgl*) gene products. The biosynthesis of the glycan begins with UDP-GlcNAc, which is successively modified by a dehydratase (PglF), an aminotransferase (PglE), and an acetylase (PglD) to generate UDP-2,4-diacetamido bacillosamine. Subsequently, PglC, a GT, transfers 2,4-diacetamido bacillosamine to UndP. Several other GTs then act to extend the glycan. PglA transfers *N*-acetylgalactosamine (GalNAc) from UDP-GalNAc to form an undecaprenyl pyrophosphate (UndPP)-disaccharide. PglJ adds another GalNAc to form a trisaccharide, before PglH adds the three remaining GalNAc residues, resulting in a hexasaccharide. Finally, the heptasaccharide is completed by the activity of the GT PglI, which adds the branched glucose residue (50, 52). The UndPP-heptasaccharide is then flipped across the membrane to the periplasmic side via an ATP-dependent flippase, PglK, although N-glycosylation persists in a strain lacking *pglK* (53). PglB, a monomeric OST homologous to the Stt3 subunit of the eukaryotic OST, then transfers the glycan to a wide variety of target proteins at select asparagine residues located within an expanded sequon.

In 1976, namely, the year when *Hbt. salinarum* was shown to contain true glycoproteins, this organism was still considered a member of the bacterial world, albeit an unusual one. However, following Carl Woese's pioneering use of 16S/18S rRNA analysis reported the next year (54, 55), an approach that ultimately led to the redrawing of the universal tree of life to comprise three distinct domains, i.e., *Eukarya*, *Bacteria*, and *Archaea*, *Hbt. salinarum* was reassigned to the archaeal branch, and the field of archaeal protein glycosylation was founded.

## N-GLYCOSYLATION PATHWAYS IN ARCHAEAL MODEL SYSTEMS

While the structures of only a limited number of the N-linked glycans decorating archaeal glycoproteins are presently known, the variety seen in these few structures points to a degree of diversity that is unparalleled in either the eukaryal or bacterial systems (6, 56). As genome analysis predicts N-glycosylation to be a common PTM in *Archaea* (57), it would seem that archaeal N-glycosylation relies on a large number of species-specific pathways. To date, however, progress in deciphering the pathways responsible for generating such a variety of N-linked glycans has focused on a limited number of archaeal species. In the following sections, the latest insights into the mechanism of archaeal N-glycosylation are presented for the model organisms *Haloferax volcanii*, *Methanococcus voltae*, *Methanococcus maripaludis*, and *Sulfolobus acidocaldarius*. The glycan structures themselves are presented below (see Archaeal Glycan Structures). The commonalities and differences in the N-glycosylation systems of the three domains (4, 6, 7, 58–60) are summarized in Table 1. To better monitor the continued progress of these and other efforts in delineating the archaeal N-glycosylation process, the reader is directed to the *agl*genes website (http://www.bgu.ac.il/aglgenes), where an updated list of N-glycosylation pathway components is maintained. In addition, we have recently proposed an *agl* (archaeal glycosylation)-based protocol for naming relevant new genes associated with archaeal N-glycosylation that would allow the ready identification of commonalities in the pathways among diverse *Archaea* (61). The interested reader is also directed to the ProGlycProt database (http://www.proglycprot.org/), a repository for bacterial and archaeal glycoproteins with experimentally validated glycosylation sites (62).

### Pathway of N-Linked Glycosylation in *Haloferax volcanii*

Despite early structural and biochemical advances in deciphering the pathway of N-glycosylation in *Hbt. salinarum*, the haloarchaeal N-glycosylation process is currently best understood for *Hfx. volcanii* (6, 63, 64). Indeed, in the last few years, substantial progress has been made in deciphering the process of N-glycosylation in *Hfx. volcanii* with the delineation of the Agl pathway in this organism.

In *Hfx. volcanii*, *agl* genes involved in N-glycosylation were originally identified on the basis of the homology of their protein products to components of the well-defined eukaryal or bacterial N-glycosylation pathways (65). This initial approach revealed that *Hfx. volcanii* contains two genes homologous to the human *alg5* gene and four homologues of the mouse *dpm1* gene. These genes code for proteins responsible for the addition of activated glucose and mannose, respectively, onto dolichol lipid carriers in the ER membrane (66). In addition, single homologues of the fungal *alg2* and *alg11* genes were detected in the partially assembled *Hfx. volcanii* genome sequence available at that time. These genes encode proteins that add mannose residues to the di-*N*-acetylglucosamine-charged DolPP carrier found on the cytoplasmic face of the ER membrane (66, 67). Likewise, *Hfx. volcanii* was shown to include a homologue of Stt3, the catalytic subunit of the OST complex of higher eukaryotes (68). Homologues of GTs involved in N-glycosylation in the bacterium *C. jejuni* (69) were also noted. Specifically, *Hfx. volcanii* encodes homologues of *pglA*, *pglI*, and *pglJ*, the products of which are responsible for the addition of a GalNAc and a glucose branch to a UndPP-linked polysaccharide, and of *pglE*, the product of which is responsible for the transformation of *N*-acetylhexosamine into bacillosamine (26). Relying on a gene deletion approach (70), it was shown that while the absence of the *Hfx. volcanii alg11* homologue is lethal, *Hfx. volcanii* cells lacking the homologue of *stt3* and one of the *alg5* homologues, *alg5-A*, remained viable, suggesting that N-glycosylation is not an essential process in *Hfx. volcanii* (65). Still, such mutations

**TABLE 1** Comparison of N-glycosylation in the three domains

| Trait | Description | | |
|---|---|---|---|
| | *Eukarya* | *Bacteria* | *Archaea* |
| N-glycosylation | | | |
| Essential process | Yes | No | Species dependent |
| Presence in domain | Universal | Limited genera | Almost universal |
| N-glycans | | | |
| Diversity of glycans | Initially conserved 14-sugar glycan in higher species | Limited | Extensive |
| Linking sugar | GlcNAc | Diacetyl-bacillosamine, HexNAc | GlcNAc, GalNAc, glucose, other hexoses |
| Glycosyl donors | NDP-sugars, DolP-sugars | NDP-sugars | NDP-sugars, DolP-sugars |
| Multibranched N-glycan | Yes | No | Yes, species dependent |
| Repeating-unit type | No | No | Rare (*Hbt. salinarum*) |
| N-glycan modification | Yes | No | Yes |
| Oligosaccharide-lipid carrier | | | |
| Lipid carrier(s) | DolPP, DolP | UndPP | DolPP, DolP |
| Isoprene units (no. of units) | Variable (14–21) | Variable (9–12), typically 11 | Variable (8–12) |
| Degree of saturation | Monosaturated (α unit) | Not saturated | Disaturated (α and ω units) or polysaturated (α, ω, and other units) |
| Site of LLO assembly | Cytosolic face of ER membrane | Cytoplasmic face of cytoplasmic membrane | Cytoplasmic face of cytoplasmic membrane |
| Flippases | | | |
| Flippases/flippase-related proteins | Rft1 (role questioned); multiple flippases likely | PglK | AglR; other(s) likely |
| Mechanism | ATP independent | ATP dependent | Unknown |
| Oligosaccharyltransferase | | | |
| Composition | Multimeric complex, single subunit | Single subunit | Single subunit |
| Catalytic subunit | Stt3 | PglB (Stt3 homologue) | AglB (Stt3 homologue) |
| Sequon(s) recognized | N-X-S/T (X [not P]) | D/E-Z-N-X-S/T (Z and X [not P]) (*Campylobacter*), N-X-S/T (X [not P]) (others) | N-X-S/T (X [not P]), N-X-N/L/V (X [not P]) |
| Nonstandard sequons used | Yes | Yes | Yes |
| Site of glycan transfer | Lumenal face of ER membrane | External face of cytoplasmic membrane | External face of cytoplasmic membrane |
| Catalytic motifs | WWDXG, DK type, SVSE | WWDXG, MI, TIXE | WWDXG, DK (most) plus DK variant, MI, TIXE |

disrupted S-layer glycoprotein glycosylation, as visualized by a faster migration of the protein on SDS-PAGE gels. In the case of Δ*alg5-A* cells, the fact that periodic acid-Schiff glycostaining was negative pointed to a profound effect of this deletion on the glycosylation of the S-layer glycoprotein.

The roles of *Hfx. volcanii stt3* and *alg5-A*, renamed *aglB* and *aglD*, respectively, according to the nomenclature proposed by Chaban et al. (71), were further addressed biochemically (72). Matrix-assisted laser desorption ionization–time of flight mass spectrometry (MALDI-TOF MS) analysis of an S-layer glycoprotein-derived tryptic peptide containing Asn-13 and an Asn-83-containing peptide generated upon digestion of the S-layer glycoprotein with trypsin and Glu-C protease revealed peaks corresponding to these peptides modified by a pentasaccharide comprising a hexose (Hex), two 176-Da subunits, a 190-Da subunit, and a terminal hexose as well as by the mono-, di-, tri-, and tetrasaccharide precursors of this glycan. While the 176-Da subunit was determined to be hexuronic acid (HexA) based on a methyl esterification assay, the 190-Da subunit was deemed at that time to be either a dimethylated Hex or a methyl ester of HexA. In *Hfx. volcanii* cells lacking *aglD*, Asn-13 and Asn-83 were modified only by the mono-, di, tri-, and tetrasaccharides and not by the complete glycan, pointing to the involvement of AglD in the addition of the final hexose of the pentasaccharide. Indeed, AglD was bioinformatically annotated as a GT. Accordingly, a later study relying on an *in vivo* assay of AglD activity revealed that Asp-110 and Asp-112 are components of the DXD motif of AglD, a GT motif predicted to interact with divalent metal cations during the nucleophilic reaction associated with the binding of nucleotide-activated sugars (73). In that same study, Asp-201 was predicted to be the catalytic base of this GT (73). In *aglB* deletion mutants, S-layer glycoprotein-derived peptides containing Asn-13 or Asn-83 were shown by mass spectrometry to lack the pentasaccharide and its precursors, confirming that AglB is an OST. Finally, these first studies aimed at delineating the *Hfx. volcanii* N-glycosylation pathway revealed that the absence or even a disruption of N-glycosylation affected growth in 4.8 M NaCl-containing medium. Moreover, the importance of N-glycosylation for proper assembly and stability of the S-layer, composed solely of the S-layer glycoprotein (63), was shown by cryo-electron microscopy (72).

Next, the contribution of Dpm1-B, another *Hfx. volcanii* homologue of a eukaryal N-glycosylation pathway component, to S-layer glycoprotein N-glycosylation was considered. In the unannotated and partially assembled version of the *Hfx. volcanii* genome available at that time, *dpm1-B* was detected eight open reading frames upstream of the *aglB* gene, on the opposite strand (74). Mass spectrometry analysis of an S-layer glycoprotein-derived tryptic peptide containing Asn-13 from a *dpm1-B* deletion mutant revealed peaks lacking the N-linked pentasaccharide and its tetrasaccharide precursor, suggesting that Dpm1-B participates in the addition of the 190-Da moiety found at pentasaccharide position 4. Given its role in the N-glycosylation process, Dpm1-B was renamed AglE (74). In contrast to cells lacking either AglB or AglD, *aglE* deletion mutants showed no effects on cell growth or S-layer stability.

The participation of another homologue of eukaryal Dpm1 in *Hfx. volcanii* N-glycosylation was next demonstrated. Deletion of *dpm1-C*, harbored by *HVO_1517*, led to faster migration of the S-layer glycoprotein on SDS-PAGE gels (75). Mass spectrometry analysis of an S-layer glycoprotein-derived Asn-13-bearing peptide from *Hfx. volcanii* cells lacking *dpm1-C* revealed a profile containing a minor peak, corresponding to a monosaccharide-modified peptide, and a major peak, corresponding to the nonmodified peptide. On the basis of the demonstrated involvement of Dpm1-C (HVO_1517) in the addition of the first (or possibly the second) sugar of the N-linked pentasaccharide, it was renamed AglJ.

Given the participation of AglE in *Hfx. volcanii* N-glycosylation, additional genes in the vicinity of *aglB* were tested for roles in this PTM. Examination of genes upstream of *aglB* (*HVO_1530*) revealed two previously identified sequences corresponding to a homologue of *C. jejuni pglI* (*HVO_1528*) and to one of the *Hfx. volcanii* homologues of eukaryal *mpg1*, *mpg1-B* (*HVO_1527*) (65, 76). Another sequence found adjacent to these sequences, *HVO_1529*, was also considered, given its position in the genome, despite it being annotated as encoding ExoM, an enzyme not involved in N-glycosylation (77). Mass spectrometry analysis of an S-layer glycoprotein-derived Asn-13-containing peptide from Δ*HVO_1527* and Δ*HVO_1528* strain cells revealed peaks corresponding to the peptide modified by monosaccharide- and disaccharide-modified precursors, confirming the participation of the deleted sequences in the addition of the third subunit of the pentasaccharide N-linked to the S-layer glycoprotein, namely, a HexA. In the case of the same peptide generated from Δ*HVO_1529* cells, only monosaccharide-bearing peptide peaks were identified, pointing to the involvement of HVO_1529 in the addition of the second subunit of the pentasaccharide N-linked to the S-layer glycoprotein, also a HexA (76). Given the participation of HVO_1527, HVO_1528, and HVO_1529 in the *Hfx. volcanii* N-glycosylation process, these proteins were renamed AglF, AglI, and AglG, respectively (76). Although *aglF*, *aglI*, *aglG*, and *aglB* are found adjacent to each other in the genome, they do not form an operon, since *aglG* is found on the minus strand, while *aglF*, *aglI*, and *aglB* are found on the plus strand. Still, reverse transcription-PCR (RT-PCR) analysis revealed that *aglF* and *aglI* are cotranscribed. At the same time, quantitative PCR showed the coordinated transcription of *aglF*, *aglI*, *aglG*, and *aglB* under different growth conditions (76).

A closer examination of the region of the *Hfx. volcanii* genome carrying these genes revealed that the automatically annotated version of the genome available at that time did not recognize *aglE* as an open reading frame. Thus, the possibility that other nearby N-glycosylation-related genes may have been overlooked was considered. Accordingly, the genomic region spanning *HVO_1517* to *HVO_1530*, namely, the region beginning with the gene encoding the GT AglJ and ending with the gene encoding the OST AglB and including *aglF*, *aglI*, and *aglG*, was revisited by using various algorithms with the aim of identifying additional genes in the N-glycosylation pathway. In this manner, *HVO_1522*, *HVO_1523*, and *HVO_1524* were reannotated *aglP*, *aglQ*, and *aglR*, respectively, given that they likely code for proteins with N-glycosylation-related roles. This assumption was based on transcriptional analyses confirming that genes in this region were cotranscribed with known pathway components and showing the expression of the products of several of these genes fused to green fluorescent protein (GFP), thus confirming that they code for true proteins. Thus, except for *aglD*, all of the *agl* genes identified to this point were shown to be clustered into a single gene island (78).

Efforts next focused on confirming that the novel *agl* gene products identified upon reexamination of the relevant region of the *Hfx. volcanii* genome indeed participate in N-glycosylation. Initially, the precise function of AglP was addressed. Mass spectrometry analysis of an S-layer glycoprotein-derived Asn-13-containing peptide from cells depleted of *aglP* revealed the presence of a tetrasaccharide albeit with the fourth, 190-Da subunit being replaced by a 176-Da moiety. As the 190-kDa pentasaccharide subunit was previously identified as a dimethylated Hex or a methyl ester of HexA (72), it was predicted that the 14-Da decrease in mass in cells lacking AglP likely reflected the loss of a methyl group. This assumption was confirmed in a methyl esterification assay involving an Asn-13-containing S-layer glycoprotein-derived peptide (79). This assay revealed that the 176-Da sugar detected at position 4 of the pentasaccharide in cells where AglP is absent is a HexA, meaning that the 190-Da sugar normally found at this position is a methyl ester of HexA and that AglP is a methyltransferase. Accordingly, an *in vitro* assay demonstrated that AglP was able to transfer the [$^3$H]methyl group from [$^3$H]methyl *S*-adenosyl-L-methionine (SAM) to membrane fragments prepared from *Hfx. volcanii* cells lacking AglP. In this way, it was confirmed that AglP acts as a SAM-dependent methyltransferase that modifies the HexA found at position 4 of the pentasaccharide N-linked to the S-layer glycoprotein (79).

Upon the identification of the *agl* gene cluster spanning *HVO_1517* (*aglJ*) to *HVO_1530* (*aglB*), the possibility that this cluster could be extended to include additional *agl* genes was considered. Specifically, the involvement of *HVO_1531* in N-glycosylation was tested. Although deletion of *HVO_1531* did not affect cell viability, it did modify S-layer glycoprotein migration on SDS-PAGE gels, with the protein from the deletion strain migrating further than the protein from the parent strain. The absence of HVO_1531 also enhanced S-layer susceptibility to proteolytic digestion (80). Mass spectrometry of an S-layer glycoprotein-derived Asn-13-containing peptide from Δ*HVO_1531* cells revealed only a peak corresponding to the peptide bearing the first subunit of the N-linked pentasaccharide. As such, HVO_1531, having been confirmed as contributing to the addition of the second pentasaccharide subunit (a HexA), was renamed AglM (80).

Bioinformatics analysis predicted AglM to act as a UDP-glucose/GDP-mannose dehydrogenase. Accordingly, AglM activity was tested *in vitro* by spectrophotometrically tracking the reduc-

tion of $NAD^+$ upon the transformation of UDP-glucose, GDP-mannose, or UDP-galactose into the corresponding HexA (80). While AglM could process all of these activated sugars, it worked best with UDP-glucose, pointing to AglM as being a $NAD^+$-dependent UDP-glucose dehydrogenase involved in the formation of the HexA found at position 2 of the N-linked pentasaccharide. Further *in vitro* efforts revealed the coordinated activities of AglM and AglF, a previously reported sugar nucleotidyltransferase that works together with the GT AglI to add a second HexA to position 3 of the N-linked pentasaccharide decorating the S-layer glycoprotein (76). The formation of NADH or the generation of UDP-glucuronic acid when UDP, $NAD^+$, and glucose-1-phosphate were incubated together with purified AglM and AglF confirmed that AglM is the dehydrogenase involved in the assembly of HexA subunits of the pentasaccharide, with AglM working with AglF to generate the HexA found at pentasaccharide position 3, as follows: glucose-1-phosphate + UDP $\xrightarrow{\text{AglF}}$ UDP-glucose + $NAD^+$ $\xrightarrow{\text{AglM}}$ UDP-glucuronic acid.

Finally, quantitative PCR analysis showed the coordinated transcription of *aglI*, *aglF*, *aglG*, and *aglM* (80). This, together with a previous analysis showing that *aglB*, *aglF*, and *aglG* are transcribed in a coordinated manner (76), supports the hypothesis that the genes of the *agl* cluster work together as a single genomic unit rather than as just a set of proximal genes involved in a common process (80).

**Topology of N-glycosylation in *Hfx. volcanii*.** Having identified components involved in *Hfx. volcanii* N-glycosylation, efforts were next focused on defining the topology of the process. Initially, the topology of AglE was determined by subcellular fractionation designed to define the position of the protein fused to a *Clostridium thermocellum* cellulose-binding domain (CBD) tag. Such efforts revealed AglE to be a transmembrane (TM) protein with the N-terminal end facing the cytoplasm (74). AglP was determined to be found in the cytoplasmic fraction by both bioinformatics and subcellular localization of a CBD-AglP fusion (79). The topologies of AglJ and AglD were identified through a combination of bioinformatics topology prediction software, assays designed to assess the accessibility of CBD-tagged versions of these proteins to proteinase K, and the use of cysteine-modifying reagents. In the latter experiments, CBD-tagged versions of AglJ and AglD were mutated such that each one contained a single cysteine residue, introduced at selected positions. Labeling by membrane-permeating and impermeant cysteine-binding reagents subsequently served to define the positions of the introduced cysteines (81). In this manner, both AglJ and AglD were demonstrated to be membrane-spanning proteins presenting two and six membrane-spanning domains, respectively, with the N-terminal portion of each protein facing the cytoplasm. Given the largely cytoplasmic orientation of AglJ, AglE, AglP, and AglD (74, 79, 81), it was concluded that the assembly of the N-linked pentasaccharide occurs on the cytoplasmic leaflet of the cytoplasmic membrane.

**The lipid carrier of *Hfx. volcanii* N-glycosylation.** In keeping with previous studies on archaeal N-glycosylation (82–84), it was predicted that the assembly of the N-linked glycan ultimately added to S-layer glycoprotein asparagine residues takes place on a lipid carrier. Hence, as a next step toward better describing the *Hfx. volcanii* N-glycosylation pathway, efforts were directed at analysis of the lipid carrier involved in the process. Previously, *Hfx. volcanii* was demonstrated to contain $C_{55}$ and $C_{60}$ DolPs charged with different sugar moieties (85); however, none of these glycans were detected on glycoproteins. More recently, a link between DolP glycosylation and N-glycosylation was provided, relying on liquid chromatography-electrospray ionization mass spectrometry (LC-ESI MS) analysis of DolP pools from *Hfx. volcanii* parent strain cells and from strains lacking the different GTs shown to be involved in the assembly of the N-linked glycan in this archaeon (75, 86). Such analysis of the glycan-charged DolP pool from parent strain cells revealed peaks corresponding to $C_{55}$ and $C_{60}$ DolPs charged with only the first four subunits of the N-linked pentasaccharide, as well as by its precursors, but not by the complete glycan bound to S-layer glycoprotein Asn-13 and Asn-83 (86). When the same analysis was also performed on lipid extracts from cells lacking AglG, AglI, or AglE, namely, those GTs involved in adding the second, third, and fourth subunits to the S-layer glycoprotein, respectively (74, 76, 87), the same effects seen at the level of the N-linked glycan were observed at the DolP level (86).

When the DolP pool from Δ*aglJ* cells was analyzed, only a barely detectable peak corresponding to hexose-charged DolP was observed, consistent with the role of AglJ in adding the first subunit of the pentasaccharide decorating S-layer glycoprotein Asn-13 (75). Because a minor amount of Hex-charged DolP was nonetheless detected in the deletion strain, the possibility that AglJ does not act alone in charging DolP with the first pentasaccharide subunit was considered. Accordingly, fractions containing Hex-charged DolP were subjected to an extended liquid chromatography run prior to a second round of mass spectrometry. In this manner, the Hex-charged DolP pool was shown to contain three distinct peaks, separated according to their retention time, representing three different species of Hex-charged DolP. Although the intensities of the first and third species were not affected by the absence of AglJ, the intensity of the second peak was reduced 7-fold in the deletion strain relative to that peak from parent strain cells. Moreover, the minor peak remaining in the absence of AglJ could not be further modified by additional subunits of the N-linked pentasaccharide. That same study identified another *Hfx. volcanii* homologue of the eukaryal Dpm1, Dpm1-D (HVO_1613), as being responsible for generating the Hex-charged DolP corresponding to the first peak described above. The absence of HVO_1613, however, had no effect on the assembly of the pentasaccharide N-linked to the S-layer glycoprotein.

Since only DolP charged with the first four subunits of the N-linked pentasaccharide decorating the *Hfx. volcanii* S-layer glycoprotein could be detected (86), the possibility that the final hexose of the pentasaccharide was derived from its own DolP carrier was proposed. This was consequently shown to be the case, since in cells lacking AglD, namely, the GT that contributes to the addition of the final sugar to the pentasaccharide, the third of the three hexose-charged DolP peaks discerned as described above was eliminated (86). Moreover, the position of this peak was consistent with the position of a mannose-charged $C_{55}$ DolP standard similarly analyzed.

Although *Hfx. volcanii* DolP serves a role similar to that of its eukaryotic counterpart, namely, acting as the lipid carrier upon which N-linked glycans are assembled, the *Hfx. volcanii* lipid can be structurally distinguished from its eukaryotic counterpart. While $C_{70}$-to-$C_{110}$ DolPP serves as the N-glycan lipid carrier in *Eukarya* (88), $C_{55}$ and $C_{60}$ DolPs assume this role in *Hfx. volcanii*. Also, whereas the eukaryal lipid is saturated solely at the α-position isoprene, the haloarchaeal molecule is saturated at both the

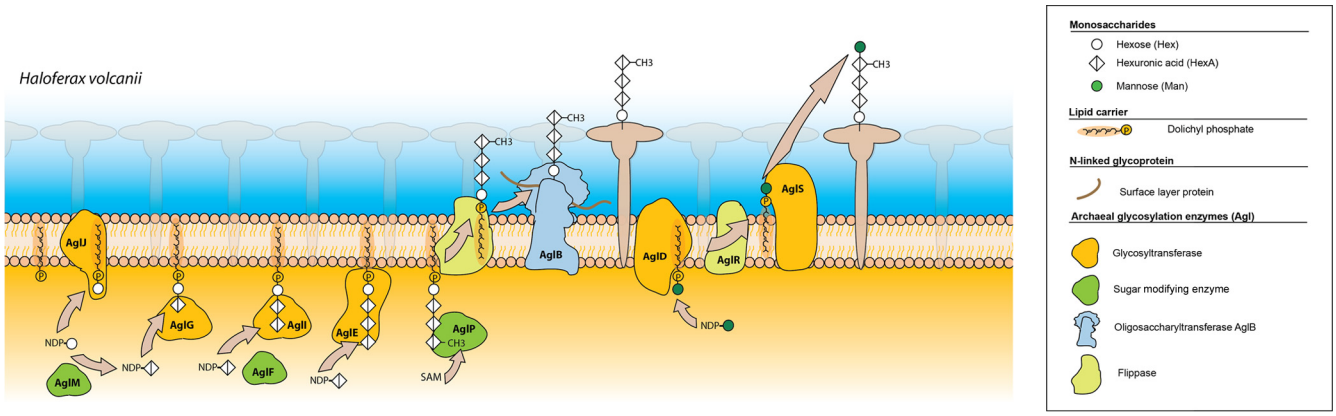



**FIG 3** Working model of the N-glycosylation pathway in *Haloferax volcanii*. Shown are the steps at which various known *agl* gene products act in the biosynthesis and assembly of the N-glycan and the steps for which Agl protein involvement remains to be identified. The actual arrangement, size, and shape of the various GTs and AglB and the mechanism by which AglB acts are not known.

α- and ω-position isoprene subunits (85, 86). Recent studies recognized the *Hfx. volcanii* geranylgeranyl reductase HVO_1799 as being able to reduce this ω-position isoprene (89). Mass spectrometry analysis revealed not only that *Hfx. volcanii* cells with a deletion of *HVO_1799* were unable to fully reduce the isoprene chains of phospholipids and glycolipids but also that the reduction of the ω-position isoprene subunit of $C_{55}$ and $C_{60}$ DolPs did not occur.

**Flippase of *Hfx. volcanii* N-glycosylation.** Having determined that the N-linked glycan decorating the *Hfx. volcanii* S-layer glycoprotein is assembled on DolP carriers on the cytoplasmic face of the cytoplasmic membrane yet is attached to the protein on the external surface of the cell, the flippase responsible for delivering tetrasaccharide- and/or monosaccharide-charged DolP carriers across the plasma membrane as well as the enzyme responsible for delivering mannose from its DolP carrier to the protein-bound tetrasaccharide were next sought. AglR was initially identified as part of the *agl* gene cluster and was shown to be cotranscribed with *aglE* (78). Deletion of *aglR* was first shown to affect S-layer glycoprotein migration on SDS-PAGE gels. When the glycan-charged DolP pool from the AglR-lacking strain was analyzed, an accumulation of tetrasaccharide-charged DolP, as well as of its trisaccharide-charged precursor, was observed. It was also noted that the level of AglD-generated DolP-mannose was higher in Δ*aglR* cells than in the parent strain (90). Moreover, the N-glycosylation profile of an S-layer glycoprotein-derived Asn-13-containing peptide from the deletion strain revealed modification by only the first four subunits of the pentasaccharide and not by the complete pentasaccharide. Furthermore, [2-$^3$H]mannose was not incorporated into the S-layer glycoprotein in the Δ*aglR* strain. Given these observations as well as the homology of AglR to Wzx, a bacterial protein thought to translocate lipid-linked O-antigen precursor oligosaccharides across the plasma membrane (91), it was concluded that AglR either serves as the DolP-mannose flippase or contributes to such activity and affects the flipping of tetrasaccharide-charged DolP (90).

AglS, also encoded by a gene within the *agl* cluster (*HVO_1526*), is a multispanning membrane protein with a major externally oriented C-terminal domain (92). Mass spectrometry analysis showed AglS to be involved in the addition of the final mannose subunit of the pentasaccharide N-linked to the S-layer glycoprotein. In seeking to define the role of AglS, bioinformatics revealed the similarity of this protein to DolP-mannose mannosyltransferases, namely, enzymes responsible for transferring mannose from DolP to a target protein. To test whether AglS is indeed responsible for delivering mannose from its DolP carrier to the S-layer glycoprotein-linked tetrasaccharide, an *in vitro* assay was developed. *Hfx. volcanii* membranes derived from either Δ*aglS* or Δ*aglD* cells were incubated with CBD-tagged AglS, and the transfer of mannose to the tetrasaccharide N-linked to the S-layer glycoprotein in these membrane fragments was visualized by mass spectrometry. In this manner, the ability of AglS to transfer mannose from DolP-mannose to the S-layer glycoprotein-bound tetrasaccharide precursor of the pentasaccharide was confirmed.

**Current model of N-glycosylation in *Hfx volcanii*.** The findings described previously led to a detailed delineation of the pathway used by *Hfx. volcanii* for N-glycosylation. The process begins with the sequential addition of nucleotide-activated versions of the first four subunits of the pentasaccharide to a common $C_{55}$ or $C_{60}$ DolP carrier via the ordered actions of the AglJ, AglG, AglI, and AglE GTs (74–76, 86). At the same time, the AglD GT adds mannose, the final subunit of the N-linked pentasaccharide, to a different DolP (72, 86). The DolP-bound tetrasaccharide and the DolP-bound mannose are flipped, the latter by AglR or in a process involving AglR, across the membrane, at which point the tetrasaccharide is transferred to target asparagine residues by AglB. Only then is the final pentasaccharide subunit, mannose, delivered from its flipped DolP carrier to the N-linked tetrasaccharide by AglS. The current model of *Hfx. volcanii* N-glycosylation is presented in Fig. 3.

## A Second N-Glycosylation Pathway in *Hfx. volcanii*

*Hfx. volcanii* was originally reported to grow in medium containing NaCl at concentrations ranging from 1 to >4 M (93). The studies aimed at deciphering the Agl pathway responsible for the assembly and the modification of *Hfx. volcanii* glycoproteins by an N-linked pentasaccharide described above were conducted on cells grown in 3.4 M NaCl. However, recent work has shown that in cells grown in medium containing only 1.75 M NaCl, glycosylation of both DolP and the S-layer glycoprotein differs from what is seen at higher salinities (94). At higher salinities, S-layer glyco-

protein positions Asn-13 and Asn-83 are modified by a pentasaccharide, while DolP is charged both with the first four pentasaccharide sugars and by the final pentasaccharide sugar, mannose. The same S-layer glycoprotein positions are also modified by the pentasaccharide when cells are grown at a lower salt concentration (1.75 M NaCl) albeit to a lesser degree. More strikingly, cells grown in 1.75 M NaCl-containing medium present DolP charged with a distinct tetrasaccharide, comprising a sulfated hexose, two hexoses, and a rhamnose. This same "low-salt" tetrasaccharide is attached to S-layer glycoprotein Asn-498 when cells are grown under such conditions; this position is not glycosylated in cells grown in 3.4 M NaCl-containing medium (94). Thus, N-glycosylation of the *Hfx. volcanii* S-layer glycoprotein is modified as a function of environmental salinity.

To determine whether the same biosynthetic pathway is responsible for the assembly and attachment of both the N-linked pentasaccharide and the low-salt tetrasaccharide, the N-glycosylation profiles of the S-layer glycoprotein in mutants with deletions of genes belonging to the Agl pathway were considered. While the absence of *aglG*, *aglI*, and *aglE* compromised pentasaccharide assembly, as expected, such deletions had no effect on the appearance of the low-salt tetrasaccharide attached to S-layer glycoprotein Asn-498 when cells were grown in medium with reduced salinity (95). It can thus be concluded that a novel N-glycosylation pathway is involved in the assembly of the tetrasaccharide N-linked to S-layer glycoprotein Asn-498.

To identify genes implicated in the biogenesis of the low-salt tetrasaccharide, the *Hfx. volcanii* genome was scanned for genes currently annotated as serving sugar-processing-related roles. It was reasoned that because all but one of the known *agl* genes are clustered to a common genomic region (78), the same may hold true for genes of the novel N-glycosylation pathway. *HVO_2046-HVO_2061* corresponds to such a cluster. Deletion of each of these genes, followed by mass spectrometry analysis of DolP and S-layer glycoprotein-derived peptides containing Asn-498 from mutant cells grown at 1.75 M NaCl, revealed that several of these genes, now renamed *agl5* to *agl15*, indeed contribute to low-salt tetrasaccharide biogenesis.

In a current working model, Agl5 and Agl6 are involved in the addition of the linking hexose to DolP, while Agl7 is involved in the sulfation of this lipid-linked sugar. This sulfation seems to be needed for the translocation of DolP charged with the complete low-salt tetrasaccharide across the cytoplasmic membrane as well as the lipid-linked di- and trisaccharide precursors. Agl8 and Agl9 contribute to the addition of the third sugar, hexose, to disaccharide-charged DolP. Agl10 to Agl14 are involved in the appearance of the final sugar, rhamnose, on trisaccharide-charged DolP. It is also thought that Agl15 acts as a flippase or contributes to the flippase activity in this pathway, given both the results of mass spectrometry analysis and its homology to AglR (28% identity and 51% similarity), suspected of serving a similar role during N-linked pentasaccharide biogenesis (90).

Although *Hfx. volcanii* encodes components of two different N-glycosylation pathways, only a single oligosaccharyltransferase, AglB, is found in this organism (57, 96). It is known that AglB is necessary for addition of the pentasaccharide to *Hfx. volcanii* glycoproteins (72, 97). To determine whether AglB is also involved in delivering the low-salt tetrasaccharide to S-layer glycoprotein Asn-498 when cells are grown under low-salinity conditions, the N-glycosylation profile of the S-layer glycoprotein in a strain lacking AglB was assessed by mass spectrometry. Such analysis revealed peaks corresponding to the S-layer glycoprotein-derived Asn-498-containing peptide to be modified by the low-salt tetrasaccharide as well as by its precursors (95), indicating that AglB is not the OST responsible for transferring this glycan from DolP to S-layer glycoprotein Asn-498. Consequently, a currently unidentified OST seems to serve this role. Finally, interplay between the two *Hfx. volcanii* N-glycosylation pathways seems to transpire. When cells lacking GTs involved in the assembly of the pentasaccharide decorating S-layer glycoprotein Asn-13 and Asn-83 were grown in medium containing 3.4 M NaCl, a peptide containing S-layer glycoprotein Asn-498 was modified with the low-salt tetrasaccharide, despite the elevated salt concentration. This suggests that the latter pathway is recruited when the former is compromised (95).

### Other Recent Insights into Haloarchaeal N-Glycosylation

Although originally reported as being a nonmotile organism (93), it was recently shown that when grown in a defined medium, *Hfx. volcanii* displays motility (97). More recently, evidence for N-glycosylation of the *Hfx. volcanii* archaellins (formerly archaeal flagellins [98]) FlgA1 and FlgA2 was presented (99). In fact, it was reported that N-glycosylation of *Hfx. volcanii* archaellins involves the same set of Agl enzymes as those required for the assembly of the pentasaccharide N-linked to the S-layer glycoprotein and that the deletion of any gene within the *Hfx. volcanii agl* gene cluster diminished swimming motility. Indeed, mass spectrometric analysis of FlgA1 confirmed that its three predicted N-glycosylation sites are modified by the same pentasaccharide as that N-linked to the S-layer glycoprotein. As such, N-glycosylation by the Agl pathway is not restricted to the S-layer glycoprotein.

Additional insight into the process of haloarchaeal N-glycosylation has come from recent studies of other species. Like *Hfx. volcanii*, *Haloarcula marismortui* also originates from the Dead Sea and is coated with an S-layer glycoprotein (100). In both species, the S-layer glycoprotein is decorated with a similar if not the same N-linked pentasaccharide comprising a Hex, two HexAs, a methyl ester of HexA, and a mannose, and both species also employ DolP as the lipid glycan carrier (101). Moreover, it was shown that *Hfx. volcanii* strains lacking either AglJ or AglD could be functionally complemented by the introduction of rrnAC0149 and rrnAC1873, the respective *Har. marismortui* homologues of these *Hfx. volcanii* GTs (102, 103). Still, there are differences in the N-glycosylation pathways of these two haloarchaea. While the N-linked pentasaccharide of *Hfx. volcanii* is derived from a tetrasaccharide sequentially assembled onto a single DolP carrier and a final mannose residue derived from a distinct DolP carrier (86), the N-linked pentasaccharide of *Har. marismortui* is fully assembled on a single DolP carrier, from where it is transferred directly to the S-layer glycoprotein (101). Nonetheless, despite these pathway differences, the final mannose subunit of the pentasaccharide N-linked to the *Har. marismortui* S-layer glycoprotein is originally derived from a distinct DolP carrier, as in *Hfx. volcanii*.

### Pathway of N-Linked Glycosylation in Methanogens

Our current understanding of N-glycosylation in the methanogens has come from a significant body of work addressing the structure and genetics of this PTM in two species of *Methanococcus*, namely, *M. voltae* and *M. maripaludis* (104–107). In addition, *in vitro* biochemical work using purified enzymes has further con-

tributed to our understanding of the process in this group of *Archaea* (108–110).

***Methanococcus voltae.*** The initial works linking an N-glycan structure to the actions of specific gene products in *Archaea* were performed on *M. voltae* (71, 111–113). In the first of these studies, published mere days ahead of a similar study of *Hfx. volcanii* (65), the nomenclature for genes involved in archaeal N-glycosylation (i.e., *agl* genes) was initially proposed (71).

Insight into methanoarchaeal N-glycosylation began, however, with the elucidation of the structure of the glycan N-linked to the archaellins of *M. voltae* PS, namely, a novel 779-Da trisaccharide composed of sugar residues with masses of 318, 258, and 203 Da (111). Mass spectrometry analysis identified the structure as consisting of GlcNAc acting as a linking sugar, connected to a diacetylated glucuronic acid, which is in turn linked to an acetylated mannuronic acid that is further modified through the attachment of a threonine at the C-6 position [β-Man*p*NAcA6Thr-(1-4)-β-Glc*p*NAc3NAcA-(1-3)-β-Glc*p*NAc]. Other versions of *M. voltae* strain PS obtained from different laboratories presented archaellins of higher apparent molecular masses. The N-glycan attached to these archaellins was identical to the trisaccharide described above but with an additional mass of either 220 or 260 Da at the reducing end, likely representing an additional sugar (113). A logical explanation for this difference is that the original Jarrell laboratory strain of *M. voltae* strain PS, passaged over the course of >25 years, had undergone a mutation(s) in a gene(s) involved in the biosynthesis and/or attachment of the novel fourth sugar.

Having defined the structure of N-linked glycans in *M. voltae*, the extent of such modifications of target proteins was considered. The archaella of *M. voltae* consist of four related archaellins, FlaA, FlaB1, FlaB2, and FlaB3 (114, 115). The *flaA* gene is transcribed separately immediately upstream of a large operon containing *flaB1*, *flaB2*, and *flaB3*, followed by *flaCDEFGHIJ* (114, 116). The major filament proteins are FlaB1 and FlaB2, FlaA is a minor structural component, and FlaB3 is responsible for the curved, hook region of the archaella (115). All four archaellins have multiple potential sites for attachment of N-linked glycans, with two sites in FlaA, five in FlaB1, six in FlaB2, and two in FlaB3 (111). Of these 15 potential sites of N-glycosylation, the trisaccharide was detected on 14 of them by mass spectrometry. It was only the first N-linked sequon of FlaA that remained in question, since no peptides were detected in an analysis covering the region in question from this least abundant archaellin. Moreover, >95% of the examined archaellin-derived glycopeptides were fully modified. In addition to archaellins, the *M. voltae* S-layer glycoprotein also presents two potential N-linked sequons. While no studies specifically targeted the glycosylation of the *M. voltae* S-layer glycoprotein, 59% coverage of the S-layer glycoprotein was obtained during the course of studies performed on the archaellins. From this work, the tryptic peptide $T^{75–91}$, containing one of the potential N-linked sequons, was shown to indeed bear the 779-Da glycan (111).

Next, studies targeting genes encoding potential enzymes involved in the biosynthesis or assembly of the N-linked glycan in *M. voltae* resulted in the identification of the OST, AglB, and a GT, termed AglA, which was responsible for the transfer of the third sugar of the glycan (71). Insertional inactivation of these genes resulted in archaellins that migrated faster on SDS-PAGE than the wild-type protein, suggesting an effect on the attached glycan. A faster-migrating S-layer glycoprotein was also observed in these mutants. Although both mutants were unstable, mass spectrometry analysis of purified archaella isolated from the *aglA* mutant clearly indicated the loss of the terminal modified mannuronic acid in the glycan in the majority of the peptides carrying N-linked sequons. Furthermore, inactivation of either *aglA* or *aglB* had consequences for archaellum assembly and function. Inactivation of *aglB* led to nonarchaellated (and, hence, nonmotile) cells, while inactivation of *aglA* resulted in poorly motile cells that had fewer and seemingly shorter archaella. This was thus the first genetic evidence demonstrating that interference with the N-glycosylation system had severe effects on archaellation. Direct biochemical evidence for the transfer of DolP-linked glycans to model peptides by a polyhistidine-tagged version of *M. voltae* AglB expressed in *E. coli* was subsequently provided by Larkin et al. (110) (see below).

Attempts to identify the first GT in the pathway, namely, the enzyme responsible for attachment of GlcNAc to the dolichol carrier, led to the identification of a gene (*Mv1751*), subsequently designated *aglH* (112). AglH is the only protein in the *M. voltae* genome that belongs to Pfam PF00953 (GT family 4), which also includes the eukaryotic enzyme *N*-acetylglucosamine-1-transferase (Alg7). Alg7 is known to catalyze the first committed step in the N-glycosylation pathway of eukaryotic cells, namely, the conversion of UDP-*N*-acetyl-D-glucosamine and DolP to UMP and DolPP-*N*-acetyl-D-glucosamine (66). AglH shares 25% identity and 41% similarity with *S. cerevisiae* Alg7 as well as motifs conserved in *N*-acetylglucosamine-1-phosphate (GlcNAc-1-P) transferase (GPT) proteins, including the motif believed to be the GPT active-site nucleophile region (VFVGDT in Alg7 and VFPGDT in AglH) (112, 117). Efforts to show the direct involvement of AglH in the *M. voltae* N-linked glycosylation pathway proved unsuccessful, since all attempts to inactivate the encoding gene were unsuccessful (71). Because of the many features that AglH shares with eukaryotic Alg7, it was proposed that the archaeal enzyme could serve as a functional homologue of Agl7. Experiments showed that *M. voltae aglH* was indeed able to complement a conditional lethal mutation in the *alg7* gene of *S. cerevisiae*, demonstrating that AglH was a GlcNAc-1-P transferase and likely the GT that catalyzes the first step of the *M. voltae* N-glycosylation pathway (112). Very similar results were recently reported for *S. acidocaldarius* (118) (see below).

A subsequent report identified two putative GTs whose inactivation resulted in faster-migrating archaellins on SDS-PAGE gels, typical of glycosylation-defective mutants (113). With the GTs for the first and third sugars identified, these two new GTs, designated AglC and AglK, were proposed to be involved in the biosynthesis or transfer of the second sugar. These studies also revealed that the inactivation of either gene led to the appearance of nonarchaellated cells. More recently, *in vitro* studies of the *M. voltae* N-linked glycosylation process were conducted by using synthetic and semisynthetic substrates together with a polyhistidine-tagged version of the relevant enzymes purified after expression in *E. coli*. In these studies, the involvement of AglC in the formation of the second sugar predicted from previous genetic work was substantiated. Purified AglC was thus directly shown to be a Glc-2,3-diNAcA (2,3-diacetamido-2,3-dideoxy-D-glucuronic acid) GT, converting UDP-Glc-2,3-diNAcA and DolP-GlcNAc to DolP-GlcNAc-Glc-2,3-diNAcA. The specificity of the reaction was for DolP-GlcNAc, with no conversion being detected when either DolP or DolPP-GlcNAc was utilized as a potential substrate (110). Further *in vitro* examination of the roles of AglH and AglK in the

*M. voltae* N-glycosylation process revealed other interesting findings (110). The first of these concerned AglH, the GT proposed to act in the first step of the process (112). When AglH was overexpressed in *E. coli* and membranes from this strain served as the enzyme source, no transfer of GlcNAc from UDP-GlcNAc to either short ($C_{55}$ and $C_{60}$) or long ($C_{85}$ to $C_{105}$) DolP was observed, despite previous findings that *aglH* genes from both *M. voltae* and *S. acidocaldarius* were able to functionally complement a conditional lethal *alg7* mutant of *S. cerevisiae* (112, 118). The failure of the *in vitro* enzymatic assays to match the *in vivo* complementation results for AglH remains unexplained. Nonetheless, studies by Larkin et al. (110) instead point to AglK, rather than AglH, as being the GT needed for the first step of the N-glycosylation process. AglK shows high sequence similarity to eukaryotic Agl5, a DolP-Glc transferase. Purified polyhistidine-tagged AglK was shown to catalyze the formation of DolP-GlcNAc when incubated in the presence of $C_{55}$ and $C_{60}$ DolPs and UDP-GlcNAc, with the product bearing an anomeric α-linkage. No transfer to the lipid was observed in the presence of UDP-Glc, UDP-GalNAc, UDP-Gal, or GDP-Man. Moreover, the transfer of GlcNAc to $C_{55}$ to $C_{60}$ DolPs was much preferred over transfer to the longer ($C_{85}$ to $C_{105}$) DolP tested. The nucleotide product of the *in vitro* reaction was shown to be UDP rather than UMP, demonstrating that AglK is a DolP-GlcNAc synthase, catalyzing the transfer of GlcNAc and not GlcNAc-1-P to DolP, resulting in the formation of DolP-GlcNAc and not DolPP-GlcNAc. Finally, AglK shares high sequence similarity (43%) to *Hfx. volcanii* AglJ (75, 110), which performs the first glycosyltransfer step in this halophile, although the sugar transferred here is different, in this case being a hexose rather than a GlcNAc (75).

These *in vitro* studies continued with analyses of purified polyhistidine-tagged AglB (110). AglB was shown to efficiently transfer DolP-GlcNAc-[$^3$H]Glc-2,3-diNAcA to an acceptor peptide, Ac(YKYNESSYK*p*NF)$NH_2$ (based on the FlaB2 sequence [114]) in reactions dependent on divalent metal ions. Interestingly, DolP-GlcNAc was not a substrate in these *in vitro* reactions, although *in vivo*, archaellins are efficiently modified with a single-sugar glycan in *aglO* mutants of *M. maripaludis*, where transfer of the second sugar to the lipid carrier is prevented (119) (see below).

***Methanococcus maripaludis.*** The development of genetic techniques for *M. maripaludis* that were not available for *M. voltae* led to a shift in the genetic and structural research on N-glycosylation in methanogens to this species. The structure of the glycan N-linked to the archaellins of *M. maripaludis* was determined to be a tetrasaccharide with similarities to the structure of its *M. voltae* counterpart (120). Interestingly, the linking sugar in *M. maripaludis* was GalNAc, however, and not the GlcNAc found in *M. voltae*. The second sugar in the *M. maripaludis* glycan was the same as that in *M. voltae*, namely, a diacetylated glucuronic acid. The third sugar was only slightly different from that in *M. voltae*. Both are modified mannuronic acid moieties with a threonine attached at the C-6 position, but in the case of *M. maripaludis*, there is an additional acetamidino group added at position C-3. The terminal (fourth) sugar of the *M. maripaludis* glycan is a novel sugar, (5S)-2-acetamido-2,4-dideoxy-5-*O*-methyl-α-L-erythro-hexos-5-ulo-1,5-pyranose. It was later reported that the major pilin of *M. maripaludis* (MMP1685; EpdE) is also modified with an N-linked glycan, but unexpectedly, this glycan is not identical to the tetrasaccharide decorating archaellin Asn residues (121). Instead, the N-linked glycan found on the pilin is a pentasaccharide corresponding to the tetrasaccharide attached to archaellin yet bearing an extra hexose attached as a branch to the linking GalNAc subunit. It is also possible that the single sequon present in the *M. maripaludis* S-layer protein (MMP0383) (122) is also N-glycosylated. To date, this has not been addressed experimentally.

**(i) *M. maripaludis* glycosyltransferases and the oligosaccharyltransferase.** The first genes shown to be involved in the *M. maripaludis* archaellin N-glycosylation process were those encoding the GTs AglL, AglA, and AglO, involved in the addition of the second, third, and fourth tetrasaccharide subunits, respectively, as well as that encoding the OST AglB (119). The genes encoding the three GTs are clustered close together in the genome (*mmp1088* [*aglL*], *mmp1080* [*aglA*], and *mmp1079* [*aglO*]), but unlike the case for *Hfx. volcanii* (78), they are not clustered with *aglB* (*mmp1424*). The deletion of each gene causes a stepwise decrease in the size of the archaellins detected by Western blotting, as the resulting glycan is truncated from four sugars to three, two, one, and, finally, the nonglycosylated version found in the *aglB* mutant. One of the GTs, AglA, shows significant sequence similarity to the homologous enzyme in *M. voltae* (35% identity and 55% similarity), reflecting the fact that AglA transfers an almost identical sugar (a modified mannuronic acid) to the diacetylated glucuronic acid that is the second glycan subunit in both organisms. Interestingly, deletion of *aglL* results in the loss of not only the terminal sugar but also the threonine attached to the third glycan subunit, suggesting that the addition of the threonine does not occur until the complete tetrasaccharide has been assembled. Thus far, the identity of the GT that mediates the transfer of the first sugar, GalNAc, to the lipid carrier is unknown. However, the *M. maripaludis* genome contains a limited number of predicted GTs that could mediate this first step. To date, most of these annotated GT-encoding genes (as well as genes encoding other potential GTs not yet annotated as such) have been targeted for deletion (119; our unpublished data). Many of these genes have been successfully deleted without any apparent effect on the size of the archaellins, indicating that they do not code for the first GT (119; our unpublished data). For other candidates, repeated attempts to delete the genes have thus far proven unsuccessful. This category includes *mmp1423*, located immediately adjacent to *aglB*, and *mmp1170* (119). MMP1170 shares very high similarity (79%) to *M. voltae* AglK, shown by *in vitro* experiments to catalyze the first GT activity in the N-glycosylation pathway of this organism (110), although the *M. voltae* enzyme transfers GlcNAc and not GalNAc, as occurs in the *M. maripaludis* pathway. The failure to isolate a mutant in the first GT might indicate that it serves a dual role, acting in both N-glycosylation and a second essential pathway, such as that used for the biosynthesis of glycolipids or coenzyme M. The structure of coenzyme M, an essential cofactor for the terminal step in methanogenesis (123), includes a UDP-GlcNAc-ManNAcA disaccharide head group (124).

**(ii) Other *M. maripaludis* genes involved or implicated in N-glycosylation.** After the identification of the four gene products involved in the attachment of sugar subunits to the lipid carrier and the transfer of the tetrasaccharide onto asparagines within target sequons, other genes involved in the biosynthesis or modification of the individual glycan subunits were identified. These genes are organized into groups found at two major locations on the *M. maripaludis* chromosome. One group is clustered together with genes encoding the GTs identified by genetic and mass spec-

trometry work as being involved in N-glycosylation and with genes whose roles in acetamido sugar biosynthesis have been demonstrated *in vitro* following expression and purification of the protein products in *E. coli*. Presently, this region encompasses *mmp1076-mmp1094* (108, 119, 125). The second cluster is located in the region bounded by at least *mmp0350-mmp0359* (S. Siu, J. Kelly, S. Logan, and K. F. Jarrell, unpublished data). Interestingly, homologues of the genes in both these *M. maripaludis* clusters are also found in *M. voltae* A3, although they all appear at a single locus in *M. voltae*. Moreover, this stretch of *M. voltae* genes includes an additional gene not found in the *M. maripaludis* genome, namely, *Mvol0233*, predicted to be a D-glucuronyl C5 epimerase, the enzyme responsible for the conversion of glucuronic acid into iduronic acid.

*(a) Biosynthetic genes in M. maripaludis gene cluster 1.* In the first cluster of *M. maripaludis* N-glycosylation genes, immediately adjacent to *aglA*, yet transcribed from the opposite strand, is an operon comprising three genes, *mmp1081-mmp1083*, subsequently renamed *aglXYZ* (125). These three genes show high sequence similarity to *wbuXYZ*, *E. coli* genes implicated in the biosynthesis of 2,6-dideoxy-2-acetamidino-L-galactose, found in the O-antigen (126). In the bacterial system, the three gene products are thought to act in concert, with the glutaminidase WbuY generating ammonia that is delivered to the amidotransferase WbuX via an ammonia tunnel formed by WbuZ. An *M. maripaludis* mutant from which *aglX* was deleted presented an N-linked glycan missing the terminal sugar as well as the threonine and acetamidino group of the third sugar (125). In the *aglXYZ* system, AglY, AglZ, and AglX were proposed to correspond to the glutaminidase, ammonia tunnel, and amidotransferase, respectively. Evidence supporting these roles included the effects of the gene deletions observed when cells were grown with excess or limited ammonia. Under either condition, deletion of the amidotransferase *aglX* resulted in the appearance of smaller archaellins. The effects of the deletion of either *aglY* or *aglZ*, however, were observed only when cells were grown in ammonia-limited medium (125). Under these growth conditions, *aglY* and *aglZ* mutants presented a glycan that was identical to that observed in cells carrying a deletion of *aglX*, while a wild-type glycan was observed in *aglY* and *aglZ* mutants when the mutants were grown in medium containing sufficient ammonia. Presumably, if ammonia is plentiful, AglX does not require the production and specific tunneling of ammonia from AglY and AglZ. As mentioned above, deletion of *aglX* resulted in several losses to the glycan; the fourth sugar was absent, as were the threonine attachment and acetamidino modification of the third sugar. One explanation for this is that AglL cannot transfer the fourth sugar onto the third unless the acetamidino group is already present. Without the fourth sugar, the threonine modification of the third sugar does not occur, as observed previously for the *aglL* mutant. Finally, the *aglX*, *aglY*, and *aglZ* deletion strains were all archaellated and motile.

Located immediately adjacent to *aglXYZ* is an operon of four genes transcribed from the opposite strand, *mmp1084-mmp1087*. Two of these genes, *mmp1084* and *mmp1085*, demonstrated effects on the N-linked glycan and were designated *aglU* and *aglV*, respectively (127). Deletion of either gene resulted in a minor decrease in the apparent molecular mass of the archaellin. Assessment of the role of AglV was straightforward. It shows high sequence similarity to SAM-dependent methyltransferases, possessing motifs found in proteins with such activity. Deletion of the gene resulted in a single small change in the attached glycan, namely, a loss of the methyl group normally attached to the terminal sugar. Hence, this was the first gene identified to be involved in the biosynthesis of the unique terminal sugar of the *M. maripaludis* N-linked glycan. The role of AglU is more intriguing. Deletion of *aglU* resulted in a very specific single change in the glycan structure, namely, the loss of the threonine group attached to the third sugar. Accordingly, AglU could participate in threonine biosynthesis or in the actual transfer of the threonine to the third sugar. Deletion of *aglU* did not, however, result in threonine auxotrophy, as these mutant cells grew as well as the wild type in a minimal medium lacking all amino acids, suggesting that AglU is not involved in threonine biosynthesis. Genes encoding proteins with high sequence similarity to AglU are found only in *M. maripaludis* and *M. voltae*, both of which are known to contain threonine-modified mannuronic acid in their N-linked glycans (111, 120). Interestingly, homologues of AglU are not found in all sequenced *M. maripaludis* strains, suggesting that some strains do not carry this unusual modification. Genes involved in the attachment of amino acids to sugars in capsular polysaccharides and O-antigens are rarely known. One example of a gene known to serve this role is *orf3*, which is responsible for the addition of the serine onto the capsule polysaccharide glycan backbone in *E. coli* O8:K40 (128). AglU shares 22% identity and 39% similarity with the ORF3 serine transferase. Both proteins belong to the Gn_AT_II superfamily and share a conserved glutaminase domain in the region of the N terminus. Based on the data described above, it was concluded that AglU is the enzyme responsible for the transfer of threonine onto the third sugar.

The last two genes in the operon (*mmp1086* and *mmp1087*) share 56% sequence identity. Both genes have a predicted FeS/SAM-binding site at the N-terminal region and a DUF4008 domain in the C-terminal region. Deletion of either gene alone did not lead to any change in glycan composition, as determined by mass spectrometry (127). Moreover, the possibility that one gene product could compensate for the loss of the other, given their sequence similarity, was discarded, as the glycan structure of a mutant carrying deletions of both *mmp1086* and *mmp1087* was identical to that of the parent strain. It remains possible, considering that *mmp1086* and *mmp1087* are cotranscribed with *aglU* and *aglV*, that one or both of these genes contributes to N-glycosylation under specific growth conditions, as was found to be true for *aglY* and *aglZ* (125).

Immediately adjacent to *mmp1087* is the gene that codes for AglL, the GT responsible for the addition of the terminal sugar discussed above. The next gene, *mmp1089*, part of an operon extending from *mmp1089* to *mmp1094*, presents an interesting dilemma. MMP1089 is annotated as a polysaccharide biosynthesis protein but has conserved domains associated with the export of O-antigen, colonic acid, and teichoic acids (COG2244). Indeed, MMP1089 was among the initial set of genes targeted in *M. maripaludis* as the flippase required for translocating the lipid-linked glycan from the cytoplasmic side of the membrane to the external face, where it is acted upon by AglB and transferred to target proteins (119). Unfortunately, repeated attempts to delete this gene all proved unsuccessful. To date, the only identified archaeal protein with suspected flippase or flippase-related function is *Hfx. volcanii* AglR, which appears to deliver DolP-mannose across the membrane (90) (see above). It is of note that like AglR, MMP1089 also contains the conserved domain associated with export of O-

antigen and teichoic acid. Since *aglB* deletions are viable in *M. maripaludis*, deletion of the flippase gene should be possible, unless the flippase for N-glycosylation also acts in another essential process. Recently, it was shown that the photoreceptor opsin also acts as an ATP-independent phospholipid flippase (129). Thus, opsin is a member of the so-called moonlighting proteins in which two distinct functions are found within a single polypeptide chain (130). It was suggested that the difficulty in identifying flippases involved in N-linked glycosylation, such as the eukaryotic M5-DLO flippase, could be due to the fact that the protein is a moonlighting protein that serves another defined function that has allowed its role as a flippase to be overlooked. Such a possibility may hamper the identification of flippases in *Archaea* as well.

MMP1090 is annotated as a glucose 4-epimerase. *In vitro* work on MMP1090, expressed and purified as a C-terminally polyhistidine-tagged version in *E. coli*, confirmed its 4-epimerase activity (Y. Ding, I. Brockhausen, and K. F. Jarrell, unpublished data). The enzyme acts on both galactose and GalNAc, mediating the conversion to glucose and GlcNAc, respectively. The reverse reaction is much less favored. An *mmp1090* deletion mutant possessed archaellins that migrated faster than wild-type archaellins in Western blots, indicating that MMP1090 plays a role in the biosynthesis of a glycan sugar, although the precise role is still undetermined.

Finally, gene annotations and analysis of protein sequences using the BLAST algorithm to search the NCBI database reveal that at least some of the other genes in the *mmp1089-mmp1094* operon could also be involved in the biosynthesis of sugars comprising the N-linked glycan. The creation of gene deletions and analyses of the resulting glycan structures in these mutants are ongoing.

*(b) Biosynthetic genes in M. maripaludis gene cluster 2.* A second locus on the *M. maripaludis* genome also includes genes involved in the biosynthesis of the sugars that comprise the N-linked glycan-decorating glycoproteins. To date, this locus has been shown to include two adjacent but divergently oriented gene sets. The first set encompasses *mmp0350-mmp0354*, while the second extends from *mmp0355* to *mmp0359*. Some of the genes located in the *mmp0350-mmp0359* region appear to be the *M. maripaludis* equivalent of *Pseudomonas aeruginosa wbp* genes that are involved in the conversion of UDP-GlcNAc to UDP-2,3-diacetamido-2,3-dideoxy-β-D-mannuronic acid [ManNAc(3NAc)A], as needed for the biosynthesis of the B-band O-antigen (131). MMP0353 (a dehydrogenase), MMP0351 (an aminotransferase), MMP0350 (an acetyltransferase), and MMP0357 (an epimerase) share substantial sequence similarity with WpbA, WbpE, WbpD, and WbpI, respectively, while MMP0352 is 48% similar to WbpB albeit with only 37% coverage. A polyhistidine-tagged version of MMP0352 expressed in *E. coli* has been studied *in vitro* (109) and was shown to catalyze the oxidation of UDP-GlcNAc to generate the keto sugar UDP-2-acetamido-3-oxo-2,3-dideoxy-α-D-glucopyranose in a $NAD^+$-dependent fashion. MMP0352 did not oxidize UDP-Glc, UDP-GalNAc, GlcNAc, or glucosamine. It does not appear that UDP-GlcNAcA was tested as a possible substrate for MMP0352 in these studies. These data thus strongly suggest that in *M. maripaludis*, UDP-GlcNAc is successively acted upon by MMP0353, MMP0352, MMP0351, and MMP0350 to yield the UDP-GlcNAc(3NAc)A (UDP-2,3-diacetamido-2,3,-dideoxy-*D*-glucuronic acid) donor for the second sugar of the N-linked glycan, and with the further action of MMP0357, the activated donor for the third sugar is generated. These assignments are consistent with the estimated size of the glycan attached to archaellins, as assessed by Western blotting. In other words, modification of the protein by a glycan comprising a single sugar in the *mmp0350*, *mmp0351*, *mmp0352*, and *mmp0353* deletion strains, all of which are nonarchaellated, is seen. The assignment of MMP0357 as being critical for the biosynthesis of the third sugar is also consistent with Western blot and mass spectrometry data showing that the deletion of *mmp0357* results in a two-sugar-containing glycan being attached to archaellins and cells that are archaellated (Siu et al., unpublished). The deletion of *mmp0354*, considered another potential flippase candidate based on its annotation, did not have any apparent effect on archaellin molecular mass (119).

The adjacent gene cluster, *mmp0355-mmp0359*, is transcribed from the opposite strand from *mmp0350-mmp0354*. The current annotation of the *mmp0355-mmp0359* cluster reveals an interesting mix of genes with strong sequence similarity to genes encoding proteins involved in acetamido sugar biosynthesis and GTs, in addition to those showing similarity only to genes of unknown function. Thus far, only *mmp0357* has been deleted, as described above. In terms of sequence, MMP0356 is highly similar to GT 1 superfamily members, while MMP0358 shares high sequence identity to hypothetical proteins found widely in *Archaea* and has a conserved domain found in members of the GT-GTB-type superfamily. MMP0359 presents high sequence identity to annotated GTs in a wide variety of *Archaea* and possesses a conserved domain found in the GT-GTA-type superfamily, in particular DPM-DPG-synthase-like enzymes belonging to the GT 2 superfamily. In eukaryotes, Dpm1 is the catalytic subunit of DolP-mannose synthase. Finally, the remaining member of this gene cluster is an interesting one. MMP0355 contains no detectable conserved domains, according to either BLAST or InterProScan analysis. A BLAST search identified only a few hypothetical proteins with significant sequence alignment, found in a limited number of diverse methanogens (*Methanoculleus bourgensis*, *Methanococcoides burtonii*, *Methanolobus psychrophilus*, and *Methanocorpusculum labreanum*) and a single *Thermococcus* species (*T. onnurineus*). Interestingly, the search did not identify an *mmp0355* homologue in any other *M. maripaludis* isolate or in *M. voltae*. If this gene is indeed involved in glycan biosynthesis, as its genomic location suggests, it appears to be performing an unusual function that is not at all widespread, even among very closely related organisms.

**(iii) *In vitro* studies of purified *M. maripaludis* enzymes involved in glycan sugar biosynthesis.** In addition to the genetic and mass spectrometry analysis of the *M. maripaludis* archaellin N-glycan, significant insight into the biosynthesis of the component sugars of the glycan have been provided by *in vitro* studies by Namboori and Graham showing that *Methanococcales* use a bacterial pathway to generate GlcNAc (108). The difficulties that were overcome in this study include the problem that key enzymes involved in GlcNAc biosynthesis cannot be identified simply by comparative sequence analysis since many belong to large protein families with diverse functions. In addition, several key enzymes in the pathway have orthologues. Candidate proteins from both *M. maripaludis* and *Methanocaldococcus jannaschii* were expressed and purified from *E. coli* and used in *in vitro* assays designed to test predicted functions. From these studies, the authors identified enzymes involved in the formation of UDP-GlcNAc starting with fructose-6-phosphate (Fru-6-P) and the subsequent conversion of UDP-GlcNAc to UDP-*N*-acetylmannosamine (ManNAc) and UDP-ManNAcA. These UDP-acetamido sugars are the likely pre-

cursors of the N-glycans of the archaellins, pilins, and S-layer proteins in *Methanococcus* as well as for the biosynthesis of glycolipids and coenzyme M.

Based on the results of these *in vitro* assays, the following sequence of events was proposed (108). Starting with Fru-6-P, an isomerizing glutamine-Fru-6-P transaminase, MMP1680 (GlmS), produces glucosamine-6-phosphate (GlcN-6-P). The phosphoglucosamine mutase activity of MMP1077 (GlmM) results in the transfer of the phosphate group from GlcN-6-P to the C-1 position, forming GlcN-1-P. MMP1076 (GlmU) then catalyzes the next two steps, namely, the acetylation of the 2-amino group of GlcN-1-P using acetyl coenzyme A (CoA) to form GlcNAc-1-P and its transfer to UTP to form UDP-GlcNAc. From UDP-GlcNAc, UDP-ManNAc is formed by the UDP-GlcNAc 2-epimerase activity of MMP0705 (WecB). Finally, MMP0706, a UDP-ManNAc 6-dehydrogenase, catalyzes the oxidation of UDP-ManNAc to UDP-ManNAcA in a $NAD^+$-dependent step.

Interestingly, orthologues of *mmp0705* and *mmp0706* are located near each other in the *mmp0355-mmp0359* cluster. Moreover, deletion analysis has shown that the products of both *mmp0353* (the orthologue of *mmp0706*) and *mmp0357* (the orthologue of *mmp0705*) are critical for wild-type glycan formation, indicating that the missing proteins are not compensated for by MMP0705/MMP0706. Initial attempts to delete *mmp0705* or *mmp0705*, however, have been unsuccessful, suggesting that they, unlike *mmp0353* and *mmp0357*, might be essential, possibly due to a role in the biosynthesis of coenzyme M, reported to contain ManNAcA (124). Interestingly, purified, heterologously expressed, polyhistidine-tagged MMP0357 did not show detectable UDP-GlcNAc 2-epimerase activity, even though it is 53% identical and 68% similar to MMP0705 (108). MMP0706 is homologous not only to UDP-ManNAc dehydrogenase (WecC) but also to a number of other enzymes (UDP-Glc dehydrogenase, UDP-acetylgalactosamine dehydrogenase, and UDP-glucose/GDP-mannose dehydrogenase), whereas MMP0353 is 35% identical and 54% similar to MMP0706. Since *mmp0353* deletion leads to the generation of a single-sugar subunit-containing glycan (Siu et al., unpublished), this suggests that the role of MMP0353 is not as an MMP0706/WecC homologue involved in the biosynthesis of the third sugar but rather is in the biosynthesis of the second sugar, the diacetylated glucuronic acid (see above).

**(iv) Lipid carrier of *M. maripaludis* N-glycosylation.** Until recently, there was no direct information on the nature of the lipid carrier(s) for N-glycosylation in any methanogen. Early studies involving the addition of bacitracin to several methanococcal species resulted in the appearance of lower-molecular-weight archaellins in both *Methanococcus deltae* ΔRC (later reclassified as *M. maripaludis* ΔRC [132]) and *M. maripaludis* but not in *M. voltae*, suggesting that this treatment had interfered with N-linked glycosylation in the first two organisms (133). Since bacitracin forms a complex with DolPP carriers, preventing the regeneration of the monophosphate form of the lipid (134), these data and those obtained from similar bacitracin studies with *Hbt. salinarum* (83, 135) led to the conclusion that a DolPP carrier was likely to be involved in N-glycosylation in *M. maripaludis* but not in *M. voltae*. However, a recent mass spectrometry analysis of the lipids extracted from *M. maripaludis* was reported, and a short $C_{55}$ DolP was identified as bearing the first three sugars of the archaellin N-linked tetrasaccharide glycan (110). Like *Hfx. volcanii* DolP, two sites of saturation were found, likely corresponding to the α- and ω-position isoprene units. No analysis of the *M. voltae* lipid carrier was reported, although *in vitro* enzymatic assays of *M. voltae* AglK (revealing DolP-GlcNAc synthase activity) demonstrated that it favored short DolP ($C_{55}$ and $C_{60}$) over longer DolP species ($C_{85}$ to $C_{105}$) as the substrate. In addition, purified AglC, the GT needed for the addition of the second glycan subunit, uses DolP-GlcNAc but not DolPP-GlcNAc as the substrate (110). The use of DolP as the substrate is consistent with the bacitracin data but at odds with the ability of *M. voltae aglH* to functionally complement an *alg7* mutant in yeast, which suggested the use of DolPP as a lipid carrier (112). The *M. voltae* AglK-generated DolP-GlcNAc bears an anomeric α-linkage and not a β-linkage as reported for DolP-Man and DolP-Glc in eukaryotes. However, since the final glycosylated archaellins in *M. voltae* are known to attach the N-linked glycan via a β-GlcNAc subunit (111), inversion of stereochemistry would be consistent with the action of known OSTs (110).

**(v) Current model of N-glycosylation in *M. maripaludis*.** The findings described above can be combined to offer a detailed delineation of the pathway used by *M. maripaludis* for N-glycosylation of the archaellins. The process likely begins with the addition of UDP-GalNAc to DolP (110) by an unidentified GT. The addition of the remaining three nucleotide-activated sugars occurs through the actions of the AglO, AglA, and AglL GTs, respectively (119). The likely pathway for the formation of the UDP-activated second sugar starting from Fru-6-P involves the combined activities of MMP1680, MMP1077, and MMP1076 to form UDP-GlcNAc (108). This precursor is converted to the second sugar of the glycan via the actions of MMP0353, MMP0352, MMP0351, and MMP0350 (Siu et al., unpublished). The epimerase activity of MMP0357 then results in the generation of the third sugar precursor, which is further acted upon by AglXYZ to add the acetamidino functional group (125). AglU adds the threonine to sugar 3, likely only after the fourth sugar has been added by the AglL GT (127). AglV adds a methyl group to sugar 4 (127). The DolP-bound tetrasaccharide is then flipped across the membrane by an unidentified flippase and transferred to target Asn residues by AglB (119). The current model of *M. maripaludis* N-glycosylation is presented in Fig. 4.

## Pathway of N-Linked Glycosylation in the Thermoacidophile *Sulfolobus acidocaldarius*

In contrast to studies elucidating the process of N-glycosylation in the *Euryarchaeota*, specifically, in *Hfx. volcanii*, *M. voltae*, and *M. maripaludis*, details on the parallel process in *Crenarchaeota* were unavailable until only recently. Because *Crenarchaeota* and *Euryarchaeota*, the two best-studied phyla of the *Archaea*, differ significantly in several basic cellular processes (136), such as replication and cell division, investigation of N-glycosylation in a crenarchaeote is essential for exploring the diversity of the process in the archaeal domain as a whole. To analyze the crenarchaeal N-glycosylation process, initial efforts were undertaken to elucidate the structure and composition of the N-glycan covering the S-layer glycoprotein of *S. acidocaldarius*. Mass spectrometry analysis revealed that the S-layer glycoprotein is modified with a heterogeneous family of glycans, with the largest comprising $GlcNAc_2Glc_1Man_2$ and a sulfated sugar called sulfoquinovose (137), consistent with the previously reported tribranched hexasaccharide linked to *S. acidocaldarius* cytochrome $b_{558/566}$ (138). This analysis furthermore revealed a remarkably high glycosylation density of the S-layer glycoprotein in the C-terminal domain,

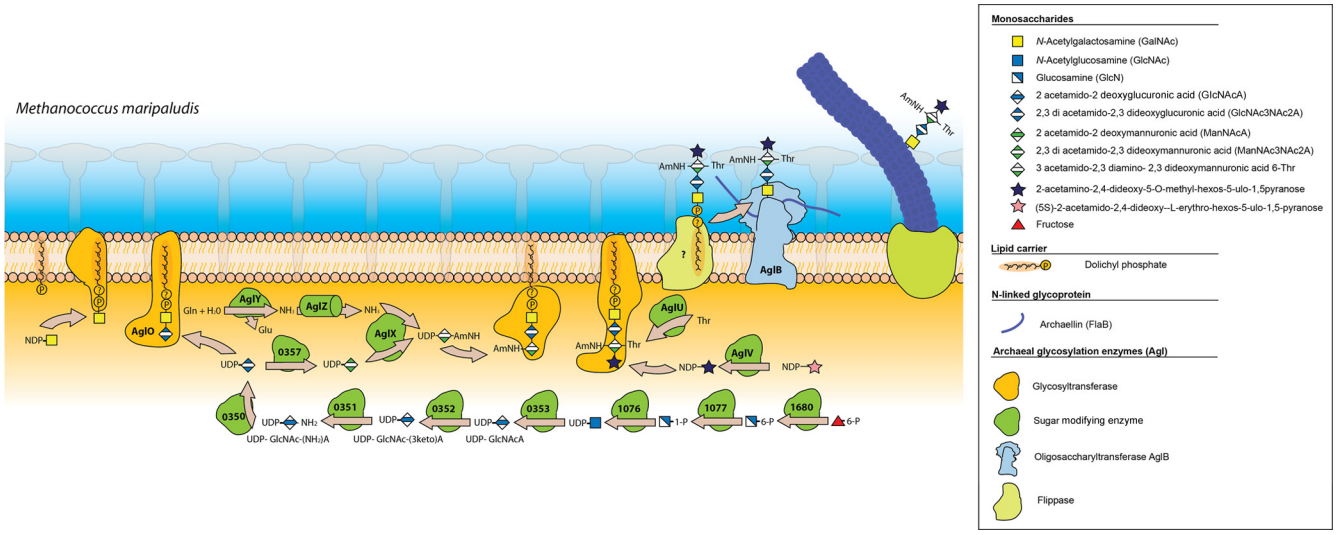



**FIG 4** Working model of the N-glycosylation pathway in *Methanococcus maripaludis*. Shown are the steps at which various known *agl* gene products act in the biosynthesis and assembly of the N-glycan and the steps for which Agl protein involvement remains to be identified. The actual arrangement, size, and shape of the various GTs and AglB and the mechanism by which AglB acts are not known.

averaging one site every 30 to 40 amino acid residues. A high glycosylation density has also been reported for the S-layer glycoprotein of *Methanothermus fervidus*, although of the 20 potential N-glycosylation sites, only about 10 have been calculated to be occupied with glycan (139). The prediction of the N-glycosylation frequency in archaeal S-layer glycoproteins (Table 2) furthermore showed that the majority of thermo(acido)philic archaea, such as *Acidianus ambivalens*, *Thermoplasma acidophilum*, *Mt. fervidus*, and *Methanothermus sociabilis*, also possess S-layer proteins with a remarkably large number of predicted N-glycosylation sites. This high N-glycosylation frequency might underline the necessity of this PTM for a stable and rigid cell wall in the face of elevated temperatures and acidity.

As was the case with halophiles and methanogens, the development of key genetic tools allowing the creation of markerless in-frame deletion mutants (140) propelled analysis of the N-glycosylation process in the crenarchaeote *S. acidocaldarius*. As in other archaeal model systems, here too, the S-layer glycoprotein has proven to be an excellent reporter for studying N-glycosylation. To identify enzymes important for N-glycan biosynthesis in *S.*

**TABLE 2** Overview of the frequency of potential N-glycosylation sites in selected archaeal S-layer glycoproteins and the *S. cerevisiae* cell wall protein

| S-layer gene/protein (GenBank accession no.) | Protein size (aa) | Calculated mass (kDa) | Apparent mass (kDa) | No. of sequons (NXS/T) | Frequency of sequons (aa) |
|---|---|---|---|---|---|
| *Acidianus ambivalens slaA* (CAM84437.1) | 1,016 | 112.3 | 110 | 33 | ~31 |
| *Haloarcula japonica csg* (BAB39352.1) | 862 | 90.6 | 170 | 5 | ~170 |
| *Har. marismortui csg*[a] (AAV44550.1) | 887 | 92.3 | 180 | 5 | ~177 |
| *Hbt. salinarum csg*[a] (AAA72185.1) | 852 | 89.6 | 200 | 13 | ~65 |
| *Hfx. volcanii csg*[a] (AAA72996.1) | 827 | 85.2 | 180 | 7 | ~118 |
| *Mcc. jannaschii slmj1* (CAC84089.1) | 558 | 60.5 | 90 | 7 | ~80 |
| *M. maripaludis slp* (CAF29939.1) | 575 | 58.9 | 88 | 1 | ~575 |
| *Methanococcus vannielii slmv1* (CAC83951.2) | 566 | 59.0 | 60 | 1 | ~566 |
| *M. voltae*[a] *sla* (Q50833.2) | 576 | 60.6 | 76 | 2 | ~288 |
| *Methanococcus thermolithotrophicus slmt1* (CAC83952.2) | 559 | 59.2 | 82.5 | 4 | ~140 |
| *Methanosarcina acetivorans MA0829* (NP_615788.1) | 671 | 74.3 | 119–134 | 4 | ~167 |
| *Mt. fervidus slgA*[a] (CAA41230) | 593 | 65.5 | 92 | 19 | ~30 |
| *Mt. sociabilis slgA* (CAA41229.1) | 593 | 65.5 | 89 | 19 | ~30 |
| *Methanotorris igneus slmi1* (CAD92366.1) | 519 | 55.6 | | 8 | ~64 |
| *P. furiosus* S-layer protein (NP_579128.1) | 605 | 65.7 | | 7 | ~86 |
| *S. acidocaldarius slp1*[a] (CAJ30479.1) | 1,424 | 151.0 | 130 | 32 | ~43 |
| *S. islandicus M1627_1831* (ACP55706.1) | 1,232 | 132.8 | 250 (likely dimer) | 38 | ~32 |
| *S. solfataricus slp1*[a] (CAJ31324.1) | 1,231 | 131.8 | 250 (likely dimer) | 39 | ~32 |
| *S. tokodaii slaA* (NP_378193.1) | 1,442 | 156.5 | 130 | 41 | ~35 |
| *T. acidophilum Ta0280*[a] (cell surface glycoprotein) (NP_393760.1) | 925 | 101.5 | 150 | 43 | ~21 |
| *S. cerevisiae* cell wall protein ECM33[a] (CAA85022.1) | 468 | 48.3 | | 13 | ~36 |

[a] Attached N-linked glycan structures have been determined (see the text).

*acidocaldarius*, initial studies focused on the identification of homologues of known enzymes involved in the biosynthesis of N-glycans in *Bacteria*, *Eukarya*, and other *Archaea*. This approach, however, did not identify any reliable candidates for enzymes involved in the N-glycosylation process, except for the OST AglB. Based on the conserved motif (WWDXG) found in the amino acid sequence of the eukaryal, bacterial, and archaeal OSTs, the gene encoding this central enzyme of N-glycosylation (*Saci1274*) was identified. In contrast to the *aglB*-based clustering of N-glycosylation genes observed, for example, in *Hfx. volcanii* (78), no additional genes encoding predicted GTs are detected near *aglB* in *S. acidocaldarius*. Indeed, the clustering of genes putatively participating in the N-glycosylation process does not seem to be common within the thermophilic archaea (57, 96). In *Crenarchaeota*, only species of *Metallosphaera* seem to present an *aglB*-based cluster of GT-encoding genes (57). The scattering of genes encoding GTs throughout the genome and the lack of clear homologues to known GTs thus made the identification of genes involved in the *S. acidocaldarius* N-glycosylation pathway more difficult than was the case with the euryarchaeal systems studied to date.

Based on previous structural and compositional characterizations of the N-glycan decorating *S. acidocaldarius* glycoproteins (137, 138), it was assumed that the N-glycosylation process began with the transfer of a GlcNAc residue onto the lipid carrier, similar to what occurs in the parallel eukaryal process, where the GlcNAc-1-P transferase Alg7 catalyzes the conversion of UDP-GlcNAc and DolP into UMP and DolPP-GlcNAc (141, 142). A BLAST search aimed at identifying an *S. acidocaldarius* homologue of eukaryal Alg7 led to the identification of Saci0093, which is annotated as a putative UDP-GlcNAc:dolichol phosphate GlcNAc-1-phosphotransferase. The encoding gene, renamed *aglH*, as it is predicted to catalyze the same reaction as that catalyzed by *aglH* in *M. voltae*, is found in a region of the genome where downstream genes are predicted to code for enzymes involved in isoprenoid lipid biosynthesis. *saci0091* (*idi*) codes for a predicted isopentenyl-diphosphate delta-isomerase. The Idi homologue in *Sulfolobus shibatae* has been shown to catalyze the interconversion of isopentenyl diphosphate and dimethylallyl diphosphate, components employed in polyprenoid biogenesis (143, 144). The adjacent gene, *saci0092* (*gds*), codes for a geranylgeranyl pyrophosphate synthase. In *S. shibatae*, this bifunctional enzyme synthesizes both farnesyl pyrophosphate and geranylgeranyl pyrophosphate (145, 146), the latter reportedly serving as a precursor in DolP biosynthesis (147). Such clustering of isoprenoid biosynthesis genes directly upstream of *aglH* could reflect a need for coordinated regulation and expression of genes involved in the biosynthesis of dolichol or DolP.

Besides showing sequence homology to its eukaryal counterparts, *S. acidocaldarius* AglH apparently also presents a topology similar to those of its homologues (148, 149). *S. acidocaldarius* AglH possesses 10 predicted *trans*-membrane domains (TMD) separated by five cytoplasmic and four external hydrophilic loops. Still, attempts to confirm the proposed function of AglH by generating an *aglH* deletion mutant have proven fruitless. This was also the case with *M. voltae*, where attempts to delete *aglH* were unsuccessful (71) despite the fact that the N-glycosylation pathway is not essential in the organism, as reflected by the ability to delete *aglB* (71). Hence, to verify the proposed function of *S. acidocaldarius* AglH as a UDP-GlcNAc:dolichol phosphate GlcNAc-1-phosphotransferase, a cross-domain complementation study was performed, as previously accomplished in studies of *M. voltae* AglH (112). Although *S. acidocaldarius* grows at temperatures near 75°C, *aglH* from this organism was still able to complement a conditional lethal *S. cerevisiae alg7* mutant and restore N-glycosylation at 28°C (B. H. Meyer and S.-V. Albers, unpublished data). Based on this interdomain complementation as well as sequence and topology similarity, it is most likely that N-glycosylation in *S. acidocaldarius* begins with the action of AglH, reminiscent of what occurs in eukaryotes, where DolPP-GlcNAc is generated and serves as the precursor for further N-glycan assembly (66). Still, despite the complementation data showing that two different AglH proteins can rescue a conditional lethal Alg7 yeast mutant, *in vitro* data do not support the ability of at least *M. voltae* AglH to catalyze this reaction (110).

Very little is known about the second and third steps in N-glycan assembly, but in the fourth step, the unusual sulfated sugar sulfoquinovose is transferred onto the presumably lipid-linked trisaccharide. Sulfoquinovose is found mostly as the nonphosphorous head group of sulfoquinovosyldiacylglycerol (SQDG), which is localized exclusively in the photosynthetic membranes of all higher plants, mosses, ferns, and algae as well as in most photosynthetic bacteria (150, 151). Although SQDG seems to have a tight correlation with phototrophic membranes, the presence of SQDG does not correlate with photosynthetic ability and is not required for photosynthesis (152). Chemical mutagenesis revealed four genes (*sqdA*, *sqdB*, *sqdC*, and *sqdD*) essential for SQDG biosynthesis in *Rhodobacter sphaeroides* (153, 154). Sequence comparisons to SqdB, which is the only enzyme that has apparent orthologues in every SQDG-producing organism, led to the identification of Saci0423, which shares 40% sequence identity with SqdB (155). *saci0423*, now renamed *agl3*, codes for the predicted UDP-sulfoquinovose synthase and is found in a gene cluster in which the neighboring genes code for enzymes involved in the proposed biosynthesis pathway of UDP-sulfoquinovose in *S. acidocaldarius* (155). *agl4* (*saci0424*) is annotated as a glucokinase, which is predicted to phosphorylate a glucose precursor in the first step of the pathway. Phosphorylated glucose could be further activated by Agl2 (Saci0422), a nucleotidyltransferase, to create NDP-glucose. UDP-glucose and sodium sulfite have been shown to be substrates of the heterologously expressed UDP-sulfoquinovose synthase Agl3 (155). The successful markerless in-frame deletion of *agl3* demonstrated the importance of Agl3 for N-glycosylation. The S-layer glycoprotein SlaA and the archaellin FlaB each had a decreased molecular mass in the *agl3* deletion mutant, indicative of a change in N-glycan composition. Indeed, nano-LC tandem mass spectrometry (MS/MS) of the S-layer glycoprotein-derived peptides from the *agl3* deletion mutant confirmed the reduced N-glycan size, with the largest N-glycan in the *agl3* deletion mutant being composed of the $Man_1GlcNAc_2$ trisaccharide. The composition of this trisaccharide not only confirmed the participation of Agl3 in the N-glycosylation process but also demonstrated the highly ordered process of *S. acidocaldarius* N-glycan assembly. In addition to the lack of sulfoquinovose and its linked glucose, the second mannose residue was also absent in the N-glycan in this mutant. This demonstrates that sulfoquinovose and, presumably, its attached glucose have to be added before the second mannose can be transferred onto the glycan.

In the fifth step in the assembly of the *S. acidocaldarius* N-glycan, one of two terminating hexose residues is transferred onto the presumably lipid-linked glycan precursor. Based on the *agl16* de-

letion mutant (see below), transfer of the terminal mannose is most likely. Here too, however, the GTase responsible for this reaction remains unknown.

In the last step in the biosynthesis of the tribranched hexasaccharide, the other terminal hexose is added in a reaction catalyzed by the GT Agl16 (156). Indeed, Agl16 was the first GT involved in crenarchaeal N-glycan biosynthesis to be identified. It is a soluble 40.6-kDa protein and contains a GT 1 domain (PF00534). The successful deletion of *agl16* resulted in small changes in the molecular weights of SlaA and FlaB, although such alterations were less than those noted for the *agl3* deletion strain, reflecting the loss of only a terminal hexose residue in cells lacking Agl16. Indeed, nano-LC-ESI MS/MS analysis of SlaA-derived glycopeptides from the *agl16* deletion strain confirmed the largest N-glycan to be a pentasaccharide lacking a terminal hexose. High-performance liquid chromatography (HPLC) analysis of the N-glycan isolated from the S-layer protein of the *agl3* deletion strain likewise revealed a reduced glucose concentration compared to that of the N-glycan of the background strain (155). Based on these results, it was proposed that Agl16 is involved in the transfer of the last terminal glucose subunit onto the sulfoquinovose residue in the lipid-linked glycan ultimately transferred to select Asn residues of target proteins in this species.

In contrast to other *Archaea* studied, N-glycosylation of extracellular proteins in *S. acidocaldarius* is essential. Thus far, every attempt to create a deletion mutant in *aglB* in *S. acidocaldarius* has failed. These attempts included the use of the markerless in-frame deletion procedure under standard (76°C) and low-temperature (60°C) conditions as well as direct homologous recombination with the selection marker *pyrEF*, designed to disrupt the *aglB* gene. Only the integration of a second copy of *aglB* in the genome in exchange for *saci1162*, encoding an α-amylase, allowed deletion of the original *aglB* sequence (Meyer and Albers, unpublished). These results demonstrate that the unsuccessful attempts to delete *aglB* were not due to polar effects or other experimental problems but rather to the importance of N-glycosylation for the survival of *S. acidocaldarius*. Furthermore, the loss of even one terminal hexose of the N-glycan has an effect on the growth and motility of *S. acidocaldarius*. These effects were increased with further decreases in the levels of N-glycan sugars (155, 156). The essentiality of the N-glycosylation process by *S. acidocaldarius* might reflect a stronger need for the thermostabilization of proteins, as realized by this PTM. Indeed, the enhanced number of glycosylation sites found in thermophilic archaeal S-layer glycoproteins relative to their mesophilic counterparts underlines this idea (118).

**Lipid carrier of *S. acidocaldarius* N-glycosylation.** The *aglH* complementation data suggested that *S. acidocaldarius* uses DolPP as a lipid carrier, yet mass spectrometry-based analysis of a total *S. acidocaldarius* lipid extract identified both dolichol and DolP but, surprisingly, not DolPP (157). *S. acidocaldarius* dolichol and DolP are considerably shorter than any known archaeal or eukaryal counterpart, mostly containing nine isoprene subunits (i.e., $C_{45}$), with smaller amounts of $C_{40}$ and $C_{50}$ also being detected. More strikingly, these *S. acidocaldarius* lipids not only incorporate saturated isoprene subunits at the α- and ω-positions, as in other archaeal DolPs (see above), but also present an additional two to four saturated isoprenes at more internal positions. This unusually high degree of saturation may reflect an adaptation to the high temperature and low pH with which this organism must cope. Still, when glycan-charged versions of DolP were sought, only species bearing different hexoses were detected. No DolP carriers bearing more complex glycans were discerned. Thus, possibly due to factors related to the lipid extraction protocol employed in these studies, it is too early to draw functional conclusions on the N-glycosylation pathway in *S. acidocaldarius* based on the failure to detect DolPP or complex glycan-charged DolP.

**Current model of N-glycosylation in *S. acidocaldarius*.** The biosynthesis of the N-glycan, apparently first assembled as a lipid-linked tribranched hexasaccharide (137), begins on the cytoplasmic side of the cell membrane. Here, one GlcNAc derived from a nucleotide-activated precursor is transferred onto the unusually short and highly saturated DolP lipid carrier (157) by the proposed action of AglH, an annotated UDP-GlcNAc: dolichol-phosphate GlcNAc-1-P transferase. Information on second and third steps is lacking, but in the fourth assembly step, Agl3 converts UDP-glucose and sodium sulfite into UDP-sulfoquinovose, an N-glycan precursor, which is then transferred onto the lipid-linked trisaccharide by an unknown GT (155). In the final steps, the terminal mannose and glucose moieties are transferred onto the lipid-linked glycan. The transfer of the final glucose residue is mediated by Agl16, a soluble GT (156). The fully assembled lipid-linked heptasaccharide is then flipped across the cytoplasmic membrane by an unidentified flippase and transferred by the OST AglB (Saci1274) onto a target protein. The current model of *S. acidocaldarius* N-glycosylation is presented in Fig. 5.

A depiction of the *agl* gene regions in *Hfx. volcanii*, *M. maripaludis*, and *S. solfataricus* is presented in Fig. 6. A summary of the known features of the N-glycosylation systems in archaeal model organisms is shown in Table 3.

## ARCHAEAL GLYCAN STRUCTURES

In the archaeal N-linked glycan structures that have been determined, wide diversity is seen in terms of size; degree of branching; identity of the linking sugar; modification of sugar components by amino acids, sulfate, and methyl groups; and the presence of unique sugars (6, 56). Such diversity far outstrips that observed thus far in *Bacteria* or *Eukarya* (26, 52, 56).

There are numerous examples in the literature suggesting that many proteins in a large variety of *Archaea*, often S-layer proteins and archaellins, are glycoproteins (2). The presumptive evidence for glycosylation includes glycoprotein-specific staining procedures, lectin binding, deglycosylation experiments, and studies with glycosylation inhibitors (2, 158–167). In most of these cases, the nature of the glycan linkages to protein and the chemical structure of the glycan were not determined. Consequently, these examples are not further discussed. Instead, we describe below those archaeal N-linked glycans for which the chemical structures have been determined.

### Halophiles

***Halobacterium salinarum*.** The S-layer glycoprotein of *Hbt. salinarum* was the first prokaryotic protein shown to be glycosylated (44). Further studies conducted during the 1980s by Sumper, Wieland, and Lechner revealed a remarkable and complex tale of the glycans attached (135, 168, 169). The S-layer glycoprotein presents three types of attached glycan, two that are N-linked and a third that is O-linked. In the latter case, some 15 Glc-Gal disac-

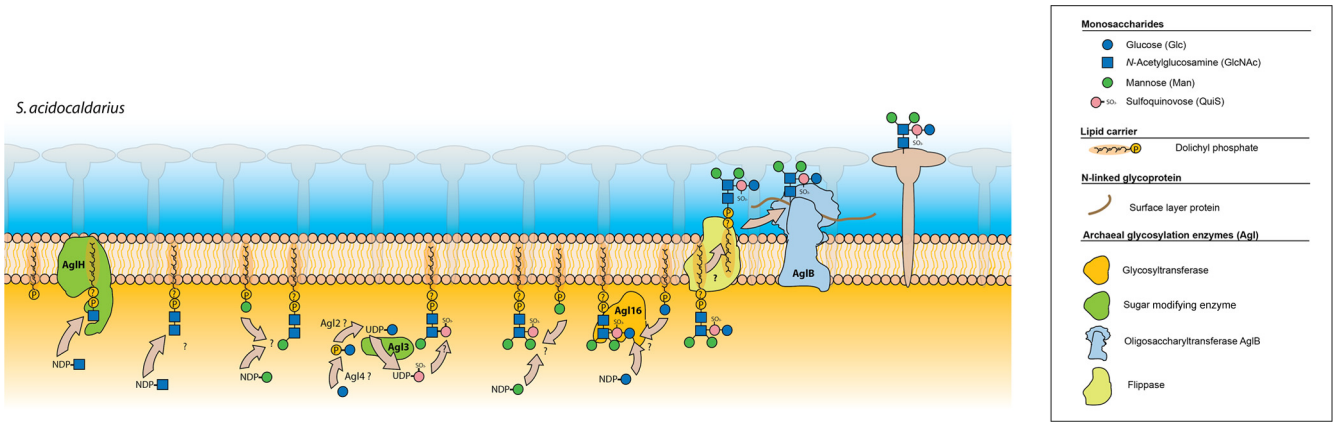



**FIG 5** Working model of the N-glycosylation pathway in *Sulfolobus acidocaldarius*. Shown are the steps at which various known *agl* gene products act in the biosynthesis and assembly of the N-glycan and the steps for which Agl protein involvement remains to be identified. The actual arrangement, size, and shape of the various GTs and AglB and the mechanism by which AglB acts are not known.

charides are O-linked to a cluster of threonine residues found near the C terminus of the protein. Of the two different N-linked glycans, one is a repeating-unit pentasaccharide found only at one position, an asparagine at position 2 from the N terminus, while the second is a sulfated oligosaccharide found at 10 positions scattered throughout the protein. The two N-linked glycans are formed on different lipid carriers and rely on different linking sugars. Moreover, the repeating-unit glycan is found exclusively

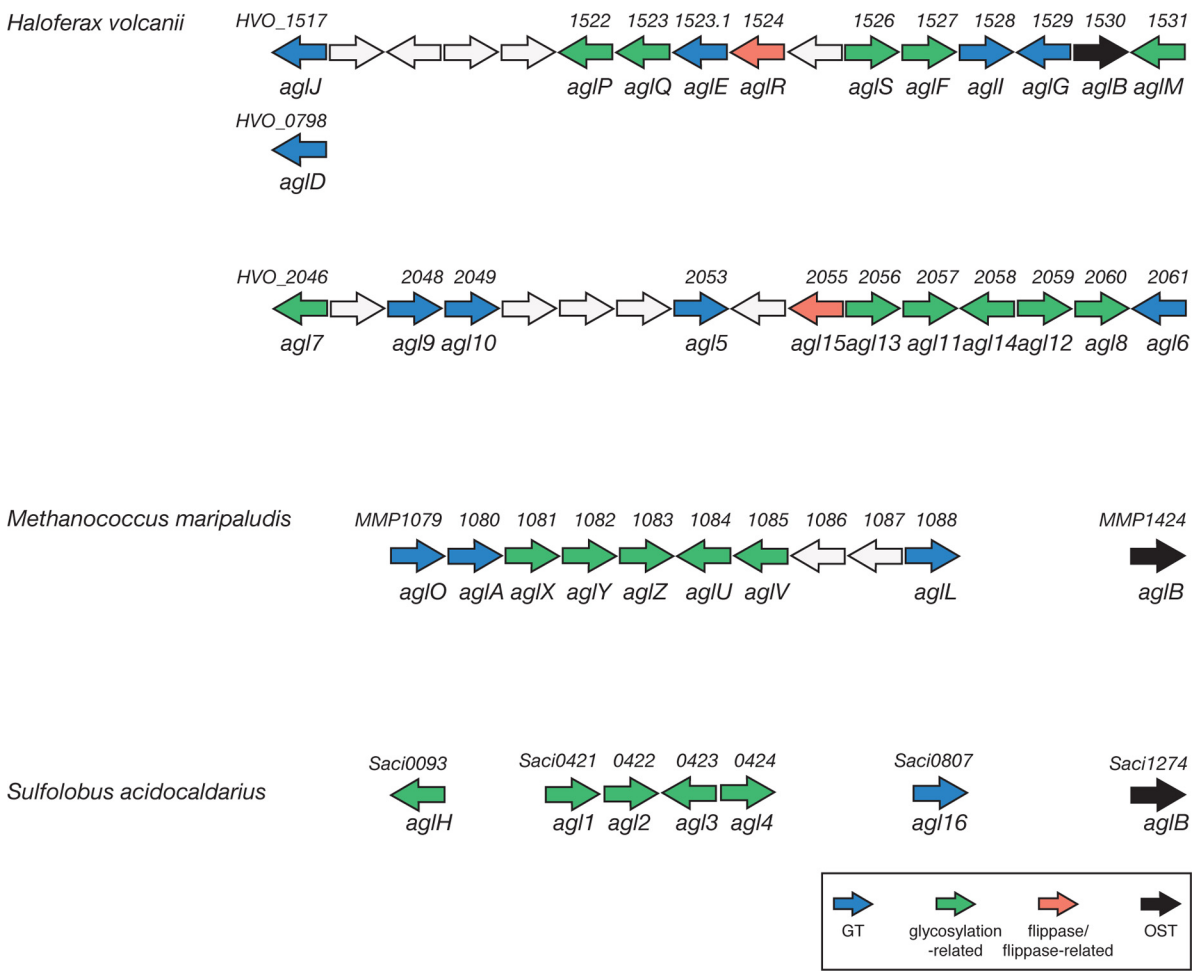



**FIG 6** Schematic depiction of known *agl* genes in model *Archaea*. Genes are colored to reflect demonstrated functions, as shown in the inset.

**TABLE 3** Summary of the known features of N-glycosylation across *Archaea*[a]

| Organism | Lipid carrier | Degree of lipid carrier saturation | Glycan structure | Linking sugar(s) | OST | Flippase or flippase related | GTases | Other N-glycosylation-related enzymes | Gene organization | Essentiality |
|---|---|---|---|---|---|---|---|---|---|---|
| *Hfx. volcanii*[b] | $C_{55}/C_{60}$ dolichol phosphate | Saturated at α- and ω-position isoprenes | Linear pentasaccharide or linear tetrasaccharide depending on NaCl | Hex, sulfated Hex | AglB | AglR | AglD, AglE, AglG, AglI, AglJ, AglS | AglF, AglM, AglP | Clustered in genome with *aglB* | Nonessential |
| *M. voltae* | $C_{55}/C_{60}$ dolichol phosphate | Not determined | Linear tri- or tetrasaccharide depending on strain | GlcNAc | AglB | Unidentified | AglA, AglC, AglH, AglK | | Not clustered | Nonessential |
| *M. maripaludis* | $C_{55}$ dolichol phosphate | Saturated at α- and ω-position isoprenes | Linear tetrasaccharide (archaellins) or branched pentasaccharide (pilins) | GalNAc | AglB | Unidentified | AgA, AglL, AglO | AglU, AglV, AglX, AglY, AglZ, MMP0350, MMP0351, MMP0352, MMP0353, MMP0357, MMP1076, MMP1077, MMP1680 | Clustered in genome but not with *aglB* | Nonessential |
| *S. acidocaldarius* | $C_{40}/C_{45}/C_{50}$ dolichol phosphate | Multisaturated at α-, ω-, and internal-position isoprenes | Tribranched hexasaccharide | GlcNAc | AglB | Unidentified | AglH, Agl16 | Agl1, Agl2, Agl3, Agl4 | Not clustered | Essential |

[a] Abbreviations: Agl, archaeal glycosylation; Hex, hexose; GlcNAc, *N*-acetylglucosamine; GalNAc, *N*-acetylgalactosamine.
[b] Agl5 to Agl15 are not included in this summary.

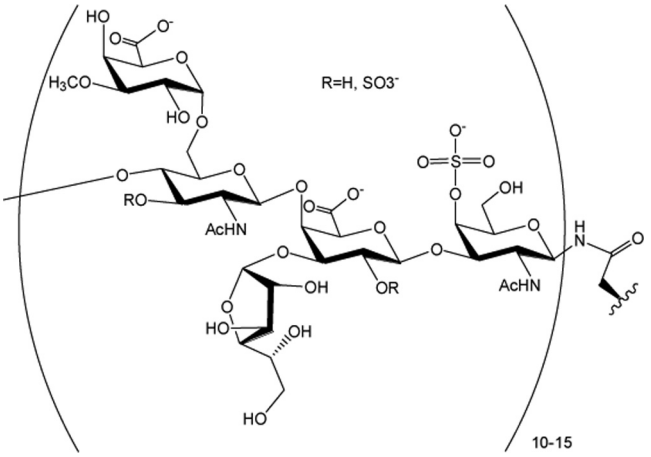



**FIG 7** Structure of the repeating-unit N-linked glycan of *Halobacterium salinarum* (171, 172).

on the S-layer glycoprotein, while the sulfated oligosaccharide is also found attached to archaellins (135, 168, 169).

The structure of the repeating-unit glycan was determined to consist of a pentasaccharide polymerized 10 times on average, with a total molecular mass of approximately 10 to 12 kDa (171, 172). This is the only repeating-unit N-linked glycan reported to date for *Archaea*. The repeating unit was reported to be a linear chain of GlcNAc-(1-4)-GalA-(1-3)-4-sulfo-GalNAc, with a galactofuranosyl branch 1,3-linked to the GalA (galacturonic acid) and a 3-*O*-methylgalacturonic acid branch 1,6-linked to the GlcNAc at the nonreducing end. The repeating units are polymerized via a 1,4-linkage. Two sulfates are found in the structure, but the precise location of only one was determined, namely, at the 4 position of the first GlcNAc subunit. The methylated sugar is a stable component of the glycan and is not the same as the transiently methylated component observed during the biosynthesis of the sulfated oligosaccharide, as described below (173). An asparaginyl-GalNAc bond links the glycan to the protein (Fig. 7). This was the first example of this type of N-glycosyl linkage to be reported (172). The same linkage was subsequently also found to attach the N-linked glycan of *M. maripaludis* archaellins (120).

Transfer of the repeating unit to the protein is inhibited by bacitracin, suggesting that the lipid carrier involved is DolPP (83, 135). However, it remains unknown whether the full glycan is assembled inside the cell and transported through the cytoplasmic membrane or whether the chain of repeating units is polymerized on the outer face of the cytoplasmic membrane, as can occur with lipopolysaccharide O-chains (174).

The second N-linked glycan decorating the *Hbt. salinarum* S-layer protein, a sulfated oligosaccharide, is found at 10 sites located within the C-terminal two-thirds of the protein (169, 175). The oligosaccharide (Fig. 8) was shown to comprise a linear chain of 1,4-linked monosaccharides, with glucose being found at the reducing end [HexA(1-4)HexA(1-4)HexA(1-4)Glc and Glc(1-4)HexA(1-4)HexA(1-4)Glc or Glc(1-4)HexA(1-4)HexA(1-4)HexA(1-4)Glc, where one of the HexA molecules in each glycan is iduronic acid (84)]. An ester-linked sulfate was found at the 3 position of each hexuronic acid (82, 169). The iduronic acid residue was found at the lipid-linked intermediate stage. This is unusual, since in other systems, iduronic acid is formed by the epi-

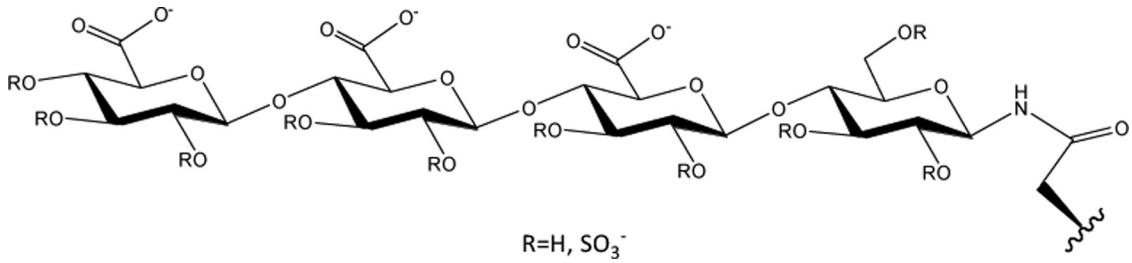



**FIG 8** Structure of the N-linked sulfated oligosaccharide glycan of *Halobacterium salinarum* (84).

merization of GlcA after the sugar has already been attached to protein (84).

The sulfated oligosaccharide is attached to protein via an unusual asparaginyl-glucose linkage, a linkage different from that used by the repeating-unit glycan (82) (see above). Furthermore, the addition of this sulfated oligosaccharide is not prevented by bacitracin, indicating the involvement of a DolP carrier (83). Indeed, $C_{60}$ DolP carrying the oligosaccharide was isolated, although the lipid-linked glycan contained a methylglucose, which is not found on the final protein-bound product (173). This transient methylation was shown to be important for N-glycosylation of the S-layer protein. Finally, the oligosaccharide is already sulfated on the lipid carrier, unlike the case in higher eukaryotes, where sulfation occurs at the protein level (135).

***Haloferax volcanii.*** For *Hfx. volcanii*, a complicated picture has emerged. At least three N-glycan structures have been reported to be attached to the S-layer glycoprotein when the organism is cultured at different salt concentrations and in different laboratories (64, 85, 94, 176). When cultured in medium containing 3.5 M NaCl, Mengele and Sumper (176) described S-layer protein N-glycans composed of different sugars. Asn-13 and Asn-498 were reported to be modified with a linear polymer consisting of 10 β1,4-linked glucose residues, while the glycan attached to Asn-274 and Asn-279 contained additional galactose and idose subunits. Other, more recent findings contradict this initial description. When cells were grown under high-salinity conditions (3.4 M NaCl), a pentasaccharide consisting of a hexose, two hexuronic acids, a methyl ester of a hexuronic acid, and a terminal mannose (Man-MetHexA-HexA-HexA-Hex) decorated at least two sites on the S-layer glycoprotein, namely, Asn-13 and Asn-83 (72, 79). When the cells were, however, grown in medium containing only 1.75 M NaCl, a very different pattern of N-glycosylation was seen. Now, a rhamnose (Rha)-Hex-Hex-sulfated Hex tetrasaccharide was found attached to Asn-498 of the S-layer glycoprotein, a position not modified at higher salinities (94). In addition, the pentasaccharide Man-MetHexA-HexA-HexA-Hex was still bound to Asn-13 and Asn-83 in cells grown at lower salinity albeit to a smaller proportion of these residues than when growth took place with higher salt concentrations. Like the *Hbt. salinarum* S-layer glycoprotein, this corresponds to a rare example of a single protein being modified with two different N-linked glycans (86, 94).

Recently, the identification of N-linked glycoproteins in *Hfx. volcanii* was extended to include the archaellin FlgA1. In this protein, three Asn residues within typical N-glycosylation sequons were confirmed to be modified by the same pentasaccharide as that found linked to the S-layer protein (99) (see above).

***Haloarcula marismortui.*** In *Har. marismortui*, the S-layer glycoprotein is modified by the same or a similar pentasaccharide as the one that decorates the *Hfx. volcanii* S-layer glycoprotein, although the route of assembly differs slightly. Unlike in *Hfx. volcanii*, in *Har. marismortui*, the entire N-linked pentasaccharide is assembled on a single DolP carrier, from where it is transferred to the S-layer glycoprotein (101) (see above).

## Methanogens

***Methanococcus.*** N-linked glycans have been determined for two different *Methanococcus* species, *M. voltae* (111) and *M. maripaludis* (120). While the structures determined share similarities, they also show important differences. The glycan N-linked to the archaellins and S-layer protein of *M. voltae* PS is a trisaccharide (Fig. 9), with a threonine linked to the terminal sugar residue [β-ManpNAcA6Thr-(1-4)-β-GlcpNAc3NAcA-(1-3)-β-GlcpNAc] (111). Subsequently, a tetrasaccharide was identified in other versions of *M. voltae* PS with the same first three sugar residues but

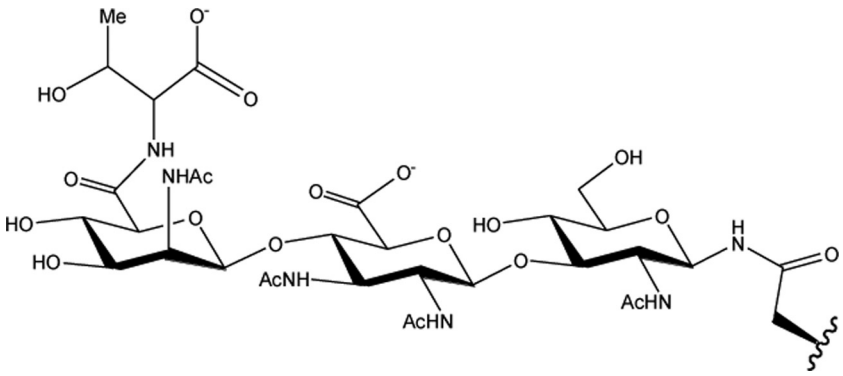


**FIG 9** Structure of the N-linked glycan of *Methanococcus voltae* (111).

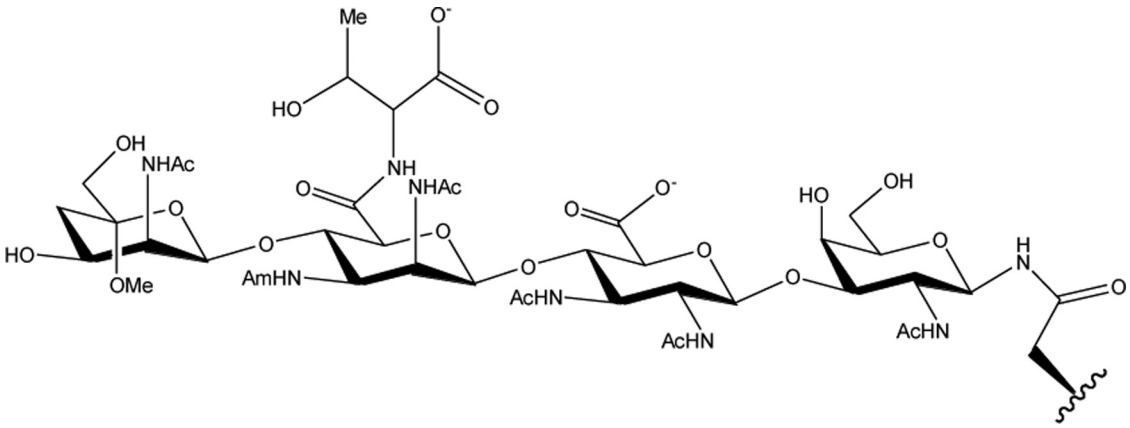



**FIG 10** Structure of the archaellin N-linked glycan of *Methanococcus maripaludis* (120).

with an extra 220- or 260-Da sugar at the nonreducing end that was not defined (113). In *M. maripaludis* S2, archaellins are also decorated with a tetrasaccharide N-glycan. Here, the glycan structure is Sug-(1-4)-β-ManNAc3NAmA6Thr-(1-4)-β-GlcNAc3NAcA-(1-3)-β-GalNAc, where Sug is (5S)-2-acetamido-2,4-dideoxy-5-*O*-methyl-α-L-erythro-hexos-5-ulo-1,5-pyranose (Fig. 10). This sugar has thus far been conclusively identified only in this species (120), although an unidentified sugar of the same mass (217 Da) was found in the N-linked glycan of the bacterium *Helicobacter pullorum* (177). In addition to the presence of this unique terminal sugar, the *M. maripaludis* glycan is linked to protein via GalNAc and not GlcNAc, as is the *M. voltae* glycan. Finally, the third sugar of the *M. maripaludis* glycan, while similar to that of the *M. voltae* glycan, is a 3-acetamidino derivative. A further unusual feature of the N-glycosylation system in *M. maripaludis* was revealed when the major structural protein of the pili was analyzed (121). The glycosylated pilin subunit (MMP1685; EpdE [178]) is modified with a pentasaccharide corresponding to the archaellin-bound tetrasaccharide containing an extra hexose branch extending from the linking GalNAc residue (121). However, since nonglycosylated versions of both archaellins and pilin are found in an *aglB* mutant, it seems that both glycans are assembled and attached by the same N-glycosylation system (119; D. B. Nair and K. F. Jarrell, unpublished data).

***Methanothermus fervidus.*** A hexasaccharide is N-linked to the S-layer protein of the hyperthermophilic methanogen *Methanothermus fervidus* (139). The linking sugar was reported to be GalNAc, with two α1,6-linked *O*-methyl sugars being found at the nonreducing end. Sandwiched between these subunits are three α1,2-linked mannose residues, thus yielding a final structure of α-3-*O*-MetMan-(1-6)-α-3-*O*-MetMan-[(1-2)-α-Man]3-(1-4)-GalNAc (139). The S-layer protein in this species contains 20 potential N-linked glycosylation sites, although it was estimated that only about half are actually glycan modified. In this hyperthermophile, both dolichol and undecaprenol were reported, with dolichol likely being the carrier used in glycoprotein synthesis and undecaprenol likely being used in the biosynthesis of the pseudomurein wall polymer (179). Unfortunately, genetic tools that would allow analysis to follow up these interesting findings are not available at this time.

***Methanosaeta concilii.*** The structurally unusual cell sheath of the methanogen *Methanosaeta concilii* (*Methanothrix soehngenii*) was shown to include a glycoprotein component (180). The glycan was, moreover, shown to be attached to the protein via an asparaginyl-rhamnose linkage. The glycan-modified peptide sequence reported is unusual, being Asn-Glu-Gly-Ser-Lys, rather than the classic Asn-X-Ser/Thr sequon. The glycan was reported to comprise 15 to 30 residues, with rhamnose being the major sugar and with ribose, fucose, and other hexoses being detected in smaller amounts. Branching from the rhamnose and ribose subunits was also reported.

## Thermoacidophiles

***Sulfolobus.*** Structural work on N-linked glycans from the two best-studied *Sulfolobus* species, *S. acidocaldarius* and *S. solfataricus*, has also been reported. Initially, it was reported that *S. acidocaldarius* cytochrome $b_{558/566}$ is a highly glycosylated protein (181). Subsequently, the N-linked glycan attached to this protein was shown to be a tribranched hexasaccharide (Fig. 11) with the structure β-Glc-(1-4)-β-6-sulfo-Qui-(1-3)-β-GlcNAc[-α-Man-(1-6)][-α-Man-(1-4)]-(1-4)-β-GlcNAc (138), where sulfo-Qui is the rare sulfated sugar sulfoquinovose, previously found in photosynthetic membranes in chloroplasts and bacteria (155). More recently, this same glycan was found to be N-linked to a variety of other surface proteins in this species, including the S-layer glycoprotein and archaellin FlaB (137, 155, 156). The N-glycan of *Sulfolobus* is unique within the *Archaea* in that it shows similarities to the eukaryal N-linked glycan. Specifically, glycans in these organisms as well as in another thermoacidophilic archaeon, *Thermoplasma acidophilum* (see below), offer archaeal examples where glycans are linked to proteins via a chitobiose core (137, 138, 182, 184), a common feature of all eukaryal N-linked glycans (66). Additionally, the tribranched topology of the N-glycan in *Sulfolobus* is reminiscent of eukaryotic N-glycans, which are generally multiantennary. Interestingly, in *Archaea*, branching of the N-glycan appears to be a common feature of a thermophilic life-style, with all characterized thermophilic archaeal N-glycans, namely, from *S. acidocaldarius* (137), *S. solfataricus* (182), *Pyrococcus furiosus* (30), and *T. acidophilum* (183), being branched.

Recently, the N-glycan decorating the surface protein SSO1273 from another *Sulfolobus* species, *S. solfataricus*, was partially solved and was shown to be one monosaccharide larger than that found in *S. acidocaldarius* (182). The branched heptasaccharide thus has the structure $[Hex]_4(GlcNAc)_2$ plus sulfoquinovose,

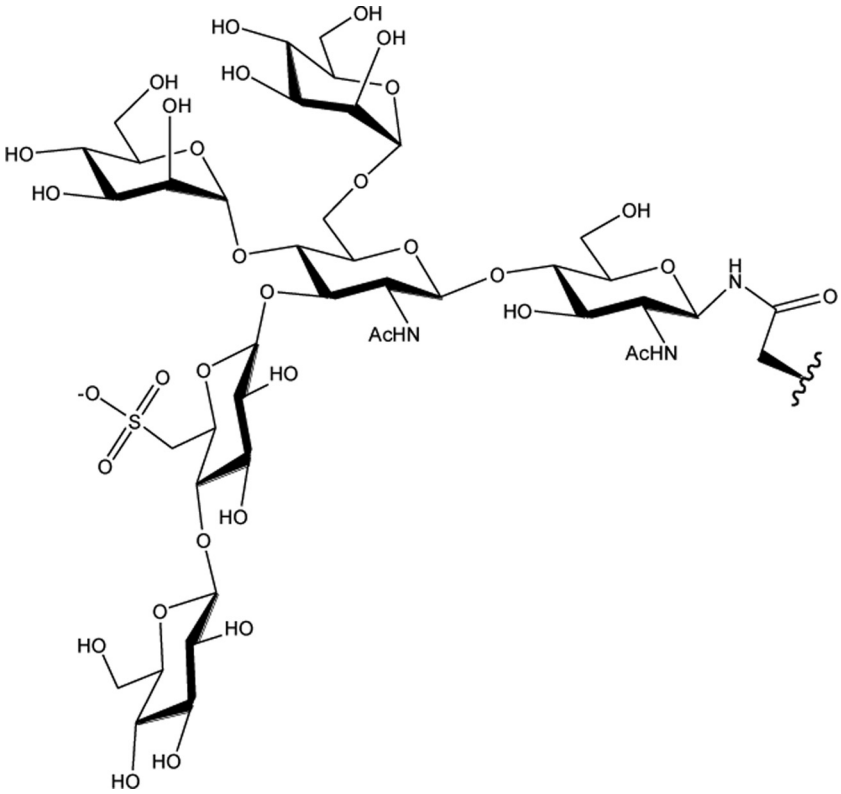



**FIG 11** Structure of the N-linked glycan of *Sulfolobus acidocaldarius* (137, 138).

where the four hexoses are mannose and glucose. Similar to the previous structure reported for *S. acidocaldarius*, this glycan contains a chitobiose core and a sulfoquinovose, and like the S-layer protein of *S. acidocaldarius* (see above), SSO1273 is also heavily N-glycosylated. Twenty of the 39 predicted glycosylation sites were analyzed, and all sites were confirmed to be occupied. Numerous other surface proteins in *S. solfataricus*, including binding proteins of ABC transporters and proteases, are modified by the same N-linked glycan. An *in silico* analysis of the modified proteins suggested that the different glycoproteins were secreted from the cells by a variety of mechanisms, since these authors included examples possessing Sec system, Tat system, and type IV pilin-like signal peptides (182).

***Thermoplasma acidophilus.*** N-linked glycoproteins from *Thermoplasma acidophilum* were among the first prokaryotic glycoproteins reported. The initial structure of the N-linked glycan attached to such proteins reported by Yang and Haug (184) showed a large high-mannose-type glycan, highly branched at the nonreducing end and linked to asparagine via a chitobiose core to the major membrane protein (Fig. 12). More recently, a very different branched glycan structure was determined for the same strain (183). This glycan is much smaller than that previously reported and comprises the structure R-(1-6)-α-Man-(1-3)-α-Man[(1-6)-α-Man]-(1-3)-β-Fuc6S[(1-2)-α-Hep]-(1-4)-β-Glc-(1-3)-β-Gal, where R refers to either α-Man or a hydrogen atom, Hep is D-glycero-D-galacto-heptose, and Fuc6S is 6-C-sulfo-D-fucose or 6-deoxy-6-C-sulfo-D-galactose. This heptasaccharide is found in roughly equal amounts as an octasaccharide corresponding to the same structure containing an extra terminal mannose residue. Whether the structures determined in these two studies indeed represent distinct glycans or instead result from the significantly different methodological approaches employed is unknown. The more recent structure is linked to a 150-kDa surface protein via a glucose-galactose disaccharide and contains an unusual sulfohexose, 6-C-sulfo-fucose (183). The presence of sulfohexoses in *Thermoplasma* and *Sulfolobus* glycans has led to speculation that they might contribute to cell stability in the challenging high-temperature and acidic growth conditions that these organisms encounter (183). Interestingly, homologues of the UDP-sulfoquinovose synthase (Agl3) from *S. acidocaldarius* can be detected in *T. acidophilum* (Ta1073) as well as in the closely related species *Ferroplasma acidamanus* (Faci_050500002306) and *Picrophilus torridus* (POT1078).

## Hyperthermophiles

***Pyrococcus furiosus.*** N-glycan structures from *Pyrococcus furiosus* and *Archaeoglobus fulgidus* were partially solved by using an N-glycosylated peptide isolated from *in vitro* glycosylation assays as the starting material (30, 185). In the *P. furiosus* assay, membrane extracts were used as the source of both AglB and lipid-linked oligosaccharide sugar donors, while a short synthetic peptide containing a sequon served as the acceptor molecule (186). The N-glycan attached to the peptide following the reaction was determined to be a heptasaccharide consisting of two HexNAcs, two Hexs, one HexA, and two pentoses (30). The mass spectrometry fragmentation pattern suggested that the two pentoses are attached as branches, one to the HexA and the other to an adjacent Hex. The linkage to protein was not confirmed but was deduced not to be the eukaryotic type (GlcNAc–β-asparagine). Instead, it

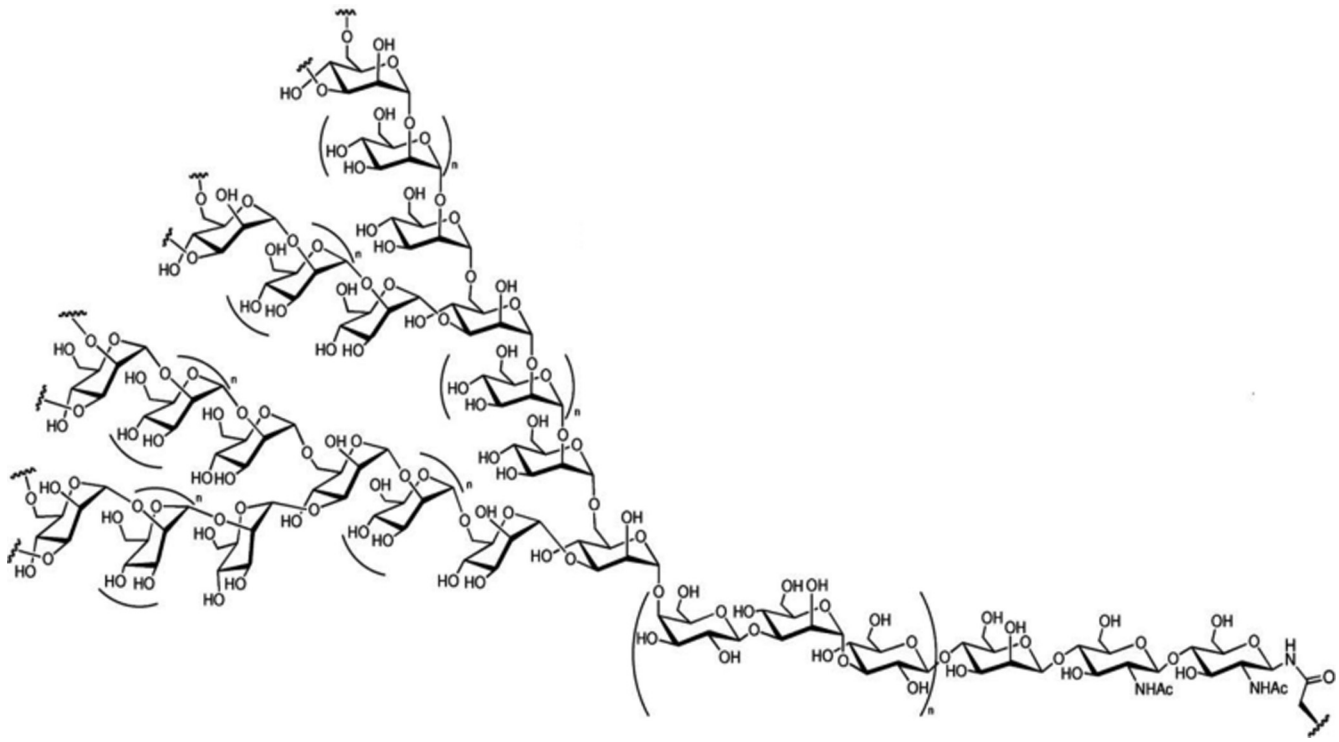


**FIG 12** Structure of the large high-mannose-type N-linked glycan of *Thermoplasma acidophilum* (184).

is likely GalNAc-asparagine, as found in *Hbt. salinarum* and *M. maripaludis* glycoproteins (30, 120, 172).

***Archaeoglobus fulgidus.*** AglB-S1, one of the three paralogues of *A. fulgidus* AglB, was expressed in *E. coli* and used in *in vitro* assays together with a synthetic peptide acceptor carrying a sequon and lipid-linked oligosaccharides prepared from *A. fulgidus* cells as the donor (185). The N-linked glycan transferred to the peptide in this reaction was found to be a linear heptasaccharide consisting of four hexoses and three deoxyhexoses (dHexs), with the structure Hex-Hex-dHex-dHex-dHex-Hex-Hex. The linking sugar is a hexose, as first reported for the *Hbt. salinarum* sulfated oligosaccharide glycan (82).

## Archaeal Virus Glycoproteins

Recently, a novel N-linked glycan was isolated from a haloarchaeal viral protein. In haloarchaeal pleomorphic virus 1 (HRPV-1), the VP4 spike protein, one of two major structural proteins, was shown to be N-glycosylated (187). Two asparagine residues within an NTT sequon, Asn-182 and Asn-427, were linked to the pentasaccharide α-5FmLeg-(2-4)-βGlcA(2$SO_3$)-(1-2)-α-Man-(1-4)-β-(1-4)-β-Glc, where 5FmLeg is 5-*N*-formyl-legionaminic acid (Fig. 13). This is the first report of a legionaminic acid modification in an archaeon-derived glycoprotein. Legionaminic acid and pseudaminic acid are 9-carbon sialic acid-like sugars found in lipopolysaccharides and in O-linked glycans of certain bacteria (188). Interestingly, when heterologously expressed in *Hfx. volcanii*, VP4 was still produced as a glycoprotein but was instead modified with the pentasaccharide that decorates the *Hfx. volcanii* S-layer glycoprotein and archaellin (72, 79, 99). Indeed, genome scanning studies suggest that the ability of *Archaea* to synthesize non-ulosonic acid sugars like sialic, legionaminic, and pseudaminic acids is limited (189).

A schematic diagram summarizing the structures and distribu-

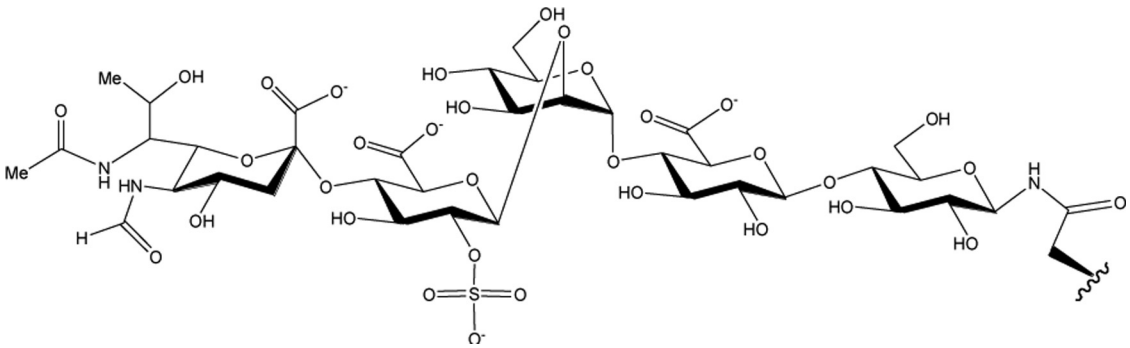


**FIG 13** Structure of the N-linked glycan of haloarchaeal pleomorphic virus 1 (HRPV-1) VP4 (187).

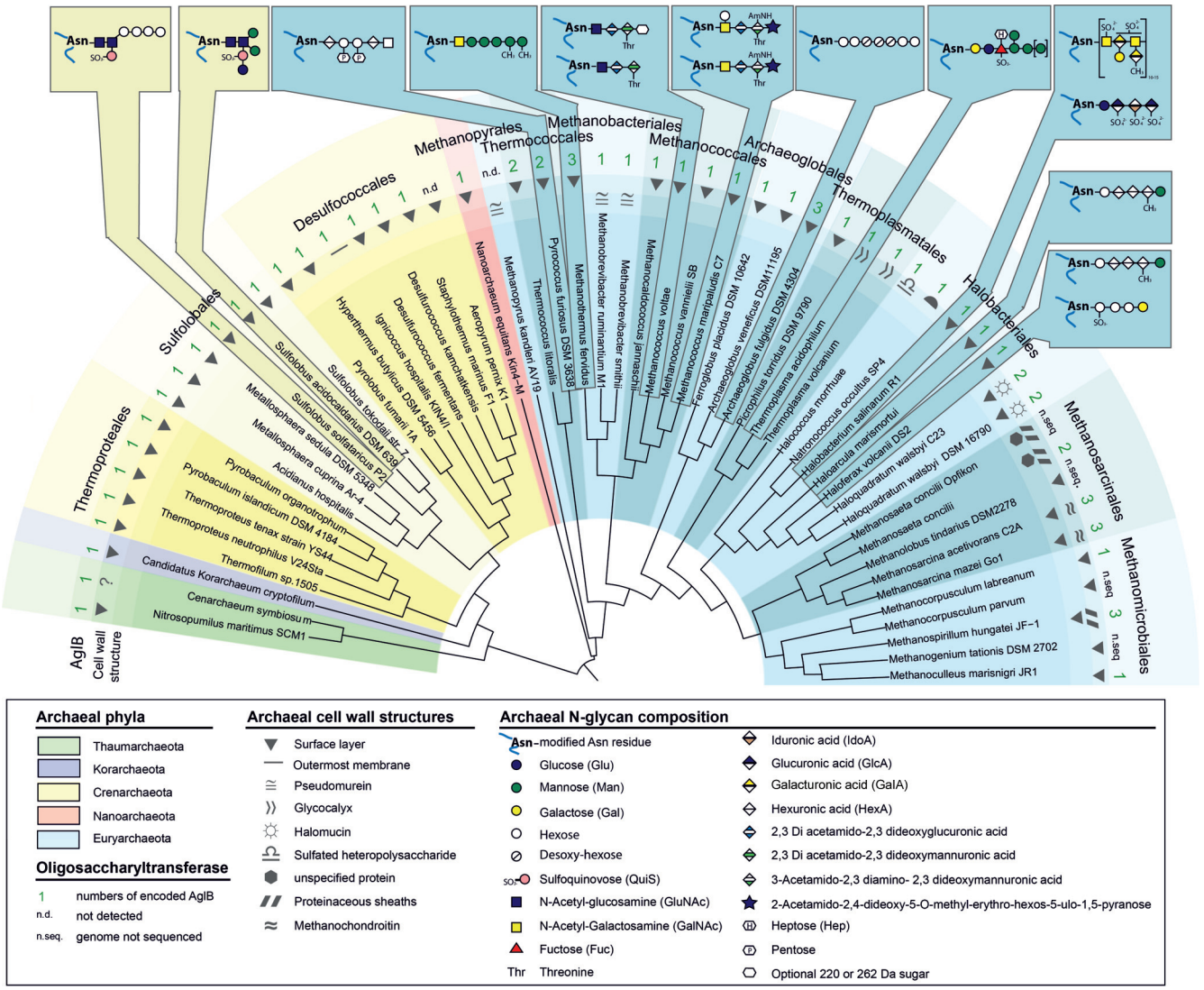



**FIG 14** Diversity of N-linked glycans in the domain *Archaea*. The phylogenetic tree is based on the alignment of the full-length 16S rRNA sequence. Cell wall structures are indicated. The number of copies of the *aglB* gene is shown. n.d., not determined. (Adapted from reference 224.)

tion of known glycan structures in the domain *Archaea* is presented in Fig. 14.

## *N*-OLIGOSACCHARYLTRANSFERASES

The most highly conserved component of the N-glycosylation pathway across the three domains of life is the OST. In *Archaea*, *Bacteria*, and some lower eukaryotes such as the protozoans *Leishmania major* and *Trypanosoma brucei*, OSTs are single-subunit enzymes, in contrast to the large heteromeric OST complexes seen in higher eukaryotes, containing up to eight nonidentical subunits (29, 56, 190). Indeed, no homologues of any of these other eukaryotic OST complex proteins, thought to influence the glycosylation rate and efficiency and to fine-tune the process, are found in *Bacteria* or *Archaea* (56). The OST catalytic subunit in eukaryotes is called Stt3 (*st*aurosporine and *t*emperature sensitive 3), while the single-subunit enzymes in *Bacteria* and *Archaea* are termed PglB (*p*rotein *gl*ycosylation B) and AglB (*a*rchaeal *gl*ycosylation B), respectively (48, 71, 191). Interestingly, though, the single-subunit Stt3 OST from *L. major* has been found to substitute for the entire OST complex in yeast (192). OSTs in the three domains of life share little primary sequence identity, with the exception of the region immediately surrounding the universally conserved WWDXG motif (191, 193–195). All of the OSTs, however, share an architecture (191, 196, 197). Stt3, PglB, and AglB can be functionally divided into an N-terminal region of 400 to 600 residues containing 13 TMD (198) and a C-terminal globular domain of 150 to 500 amino acids. The 13 TM helices are connected by short cytoplasmic or external loops (ELs) except for the extended external loop 1 (EL1) and EL5. The WWDXG motif lies at the beginning of the C-terminal domain, which is located in the periplasm (PglB [196] and AglB [198]) or ER lumen (Stt3 [197]). Evidence for the extracellular location of the catalytic site of AglB was presented in early studies in which an artificial acceptor peptide unable to penetrate the membrane of *Hbt. salinarum* was nonetheless modified with a sulfated glycan (173). While full-length sequence alignments failed to identify homology in the membrane-span-

ning N-terminal regions of OSTs, Jaffee and Imperiali (199) recently developed a program that compared the sequences of the various OSTs in terms of a specific topographical feature, such as the first TMD or first loop region. With this approach, those authors were able to identify several conserved motifs in OSTs across all three domains.

### Requirement for AglB Varies

Recently, 168 completely sequenced genomes of archaea were searched for the presence of *aglB* (57). An *aglB* gene was located in all but two cases, with some species seemingly carrying multiple copies of the gene. In many cases, *aglB* was found in a cluster of other genes putatively involved in the N-glycosylation process. The two species for which an *aglB* gene was not found are *Aeropyrum pernix* and *Methanopyrus kandleri*. Recently, Palmieri et al. (182) analyzed the proteomes of several archaea, including *A. pernix*, for potential N-linked glycoproteins with NetNGlyc prediction software. While the number of potential N-linked proteins was the lowest for *A. pernix* of the four species analyzed, the frequency of the predicted N-glycosylation in the proteome of this organism was nevertheless still 30%. Analysis of the two outliers in genomic searches for *aglB*, *A. pernix* and *M. kandleri*, is of obvious interest since if N-glycosylation does in fact occur in these species, the mechanism or OST involved would likely be unusual. The use of the *in vitro* OST assay developed by the Kohda group in Japan may help resolve this question (186).

Still, despite the seemingly widespread nature of N-glycosylation in the archaeal domain, it is not essential for cell viability in all species. Successful deletion or insertional inactivation of *aglB* was reported for *M. voltae* (71), *M. maripaludis* (119), and *Hfx. volcanii* (72). Nonetheless, strains carrying *aglB* deletions had significant defects in archaellation, S-layer stability, and growth in high salt (see below). In contrast to the *Euryarchaeota*, the N-glycosylation process in the crenarchaeote *S. acidocaldarius* does seem to be essential. Thus far, every attempt to create a deletion mutant in *aglB* in *S. acidocaldarius* has failed, and strains deleted for *agl3* or *agl16* show a drastic reduction in fitness (see below).

In the eukaryotic parasite *T. brucei*, OST activity is catalyzed by single-subunit enzymes encoded by three paralogous genes. It is of note that only two of these can complement a yeast *stt3* deletion strain. However, each of the OSTs transfers different variants of the same N-linked glycan of this organism to different sites on acceptor proteins (200). This finding has implications for *Archaea*, since several archaeal species possess multiple *aglB* genes. While it is not yet known whether these multiple genes all code for functional enzymes, they all possess an identifiable WWDXG motif, although the multiple versions do present differences in this important motif (57). For example, in *Methanocella arvoryzae*, two AglB proteins possess the WWDYG motif, while a third has the modified motif WWDDG. If functional, the different AglBs in a given species may have different substrate or target preferences. However, for those organisms where different N-glycans are attached to proteins, only one *aglB* gene has been reported (see below).

AglBs from *P. furiosus* (30, 201), *A. fulgidus* (185), and *M. voltae* (110) have been expressed and purified from *E. coli* and were shown *in vitro* to catalyze glycan transfer to target acceptor peptides. Since the *E. coli*-produced versions are functional, no archaeon-specific PTM is necessary for activity, nor is the presence of any other archaeon-specific subunit (201).

### AglB Structure

An understanding of the critical step in N-linked glycosylation catalyzed by OST should be easier in systems where a single enzyme rather than a component of a large complex is involved. Significant advances in defining OST structures using PglB and AglB have recently been reported, including the use of a protocol designed to optimize the expression and purification of active *C. jejuni* PglB in milligram amounts from *E. coli* (202); crystallography studies from the Kohda group on AglB from *P. furiosus* and *A. fulgidus* (30, 185, 203, 204), including the first full-length archaeal OST (205); and solution of the X-ray structure of full-length *Campylobacter lari* PglB in complex with an acceptor peptide (195). The latter study showed that when bound with peptide, PglB forms two cavities on opposite sides of the protein. One cavity contained the bound peptide, indicating that it provided access for the acceptor protein to the enzyme, while the other cavity was hypothesized to bind the incoming lipid-linked oligosaccharide (195).

Recent efforts have provided solved crystal structures for the catalytic domains of several AglB proteins. *Pyrococcus horikoshii* and *P. furiosus* each contain two versions of AglB, termed AglB-L (long) and AglB-S (short), while *A. fulgidus* encodes three versions of the protein, designated AglB-L, AglB-S1, and AglB-S2 (203). Crystal structures were obtained for the soluble globular catalytic C-terminal domain of AglB-L of *P. furiosus* (30), AglB-L of *P. horikoshii* (203), and all three paralogues in *A. fulgidus* (185, 203, 206). While these structures were catalytically inactive, they nonetheless led to the identification of additional motifs involved in the catalytic activity of AglB, as well as in OSTs of the other domains, and provide an excellent example of how research focused on archaeal models has significantly contributed to an understanding of the corresponding systems in other domains. Recently, a second crystal structure for a full-length OST was reported to complement that of *C. lari* PglB, namely, *A. fulgidus* AglB-L (198). Remarkable structural similarity was found between full-length PglB and AglB-L, despite the low sequence similarity of these distantly related proteins. This furthers insights into the N-glycosylation catalysis reaction gained from the crystal structures of the catalytically inactive C-terminal domains.

Like all OST catalytic subunits, AglB possesses the universally conserved WWDXG motif. This motif was thought initially to be directly involved in catalysis since mutations in this motif abolished or greatly reduced activity in both yeast Stt3 (194) and *C. jejuni* PglB (26, 207). However, the X-ray structure of full-length *C. lari* PglB in complex with an acceptor peptide (195) revealed that the β-hydroxyl group of the threonine at position +2 of the sequon in the bound acceptor peptide formed specific hydrogen bonds with the WWD motif of PglB. It would thus seem that the function of the WWD motif is to recognize the Thr/Ser at position +2 and help orientate the asparagine residue in the sequon into the catalytic pocket, rather than activation of the amide nitrogen of the acceptor protein, as originally hypothesized (185, 195).

In addition to the WWDXG motif, analysis of the archaeal AglB structures revealed novel features of this enzyme. For instance, a novel DK motif (DXXKXXX[M/I], where X is any amino acid) is found in a kinked helix adjacent to the WWDYG motif in the tertiary structure of AglB from *P. furiosus* (30). The functional importance of this novel motif was demonstrated by mutating the parallel region in yeast Stt3 (30) and, later, in AglBs from *A. fulgi-*

*dus* (185) and *P. furiosus* (201). While the DK motif is found in eukaryotic Stt3 and most AglB proteins, other AglBs and bacterial PglBs contain a so-called MI motif (MXXIXXX[I/V/W]) at the corresponding location. The MI motif was shown by mutational studies of PglB to contribute to catalytic function (193). A variation of the DK motif was observed in some *Archaea*, including members of the classes *Halobacteria*, *Methanomicrobia*, and *Archaeoglobi*, where an insertion of 4 to 14 residues interrupts the DK motif (185, 206). All three *A. fulgidus* AglB paralogues contain the variant DK motif with an insertion in support of the proposed rule that all OSTs of any single organism contain the same type of DK/MI motif (185). The DK and MI motifs were determined from the crystal structures to contribute, along with the WWDXG motif, to the recognition pocket surrounding the Ser/Thr part of the sequon, with the properties of the pocket depending on the lysine in the DK motif or the isoleucine in the MI motif (185, 195). The recently reported full-length structure of AglB-L shows the lysine-type pocket (198). Site-directed mutation of the lysine in the DK motif to alanine (D574A) was also found to affect the amino acid preference at the X position of the sequon (201). This mutational change led to an enhanced preference for arginine and lysine at the X position. Another motif, EXD or DXD, located in the loop region between the first and second TMD, was also shown to be important for catalytic activity by mutational studies in yeast (193). Yet another motif, GXXDX(D/E), was discovered when it was found that a D167A mutation impaired *P. furiosus* AglB activity (201). Alanine mutations in the corresponding region in yeast Stt3 had already been reported to be lethal (208), supporting the importance of this motif. The acidic residues in the two DxD motifs were shown in the crystal structure of *C. lari* PglB to be involved in the coordination of the divalent cation required for the activity of all OSTs (195). Using analysis of the sequence corresponding to topographical features of OSTs, a further motif was discovered in EL5, namely, [I/V]XXX[S/T][I/V]XE (199). The glutamic acid of this motif is thought, along with the conserved aspartic acid residues in ELs 1 and 3, to help form the pocket that accommodates the divalent cation and the nucleophilic asparagine. Mutation of the corresponding glutamic acid (E319) in PglB resulted in significantly lower or no activity, indicating the importance of this residue in OST activity (199). Analysis of the crystal structure of full-length AglB-L showed that D47 of the first DXD motif of EL1 (equivalent to *C. lari* PglB D56) and D161 and H163 of the second DXD motif of EL2 (*C. lari* D154 and D156) were all involved in the coordination of the divalent metal ion. For AglB-L, it is hypothesized that sequon binding in the Ser/Thr-binding pocket results in a conformational change of EL5, which brings the side chain of E360 (*C. lari* E319) close to the metal ion. E360 has been proposed to be a key residue in the activation mechanism of the acceptor amide nitrogen (209) and is a highly conserved residue in the TIXE motif found in *Archaea* and *Bacteria* (SVSE motif in *Eukarya*).

Alanine scanning mutagenesis was performed on each of the three conserved motifs, namely, WWDYG, DXD, and DK, in AglB from *P. furiosus* (201). Changing each aspartate residue in the DXD or WWDYG motif resulted in an almost complete loss of glycan transfer activity to a synthetic peptide, as measured *in vitro*. Likewise, mutation of all the other residues in the WWDYG motif also led to an inhibition of enzyme activity, as did mutations of aspartic acid, lysine, and isoleucine residues in the DK motif, confirming roles for all three motifs in OST function.

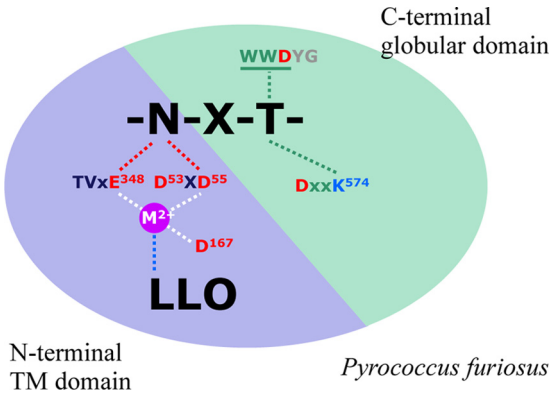



**FIG 15** Schematic model of the catalytic site of *Pyrococcus furiosus* AglB and the presumed functions of the conserved motifs. Four interactions are assumed to occur, at least transiently, during the oligosaccharyltransferase reaction. The Ser/Thr residue in the +2 position of the acceptor sequon interacts with the Ser/Thr-binding pocket in the C-terminal globular domain (green). The pocket consists of the WWD part of the highly conserved WWDYG motif and the lysine residue of the DK or DKi (DK with insertion) motif (the isoleucine residue of the MI motif in other AglBs). A divalent metal ion ($M^{2+}$[$Mg^{2+}$ or $Mn^{2+}$]) is coordinated by three acidic residues: D55 contained in the DXD motif located in EL1, D167 located in EL2, and the glutamic acid (E348) of the TIXE motif located in EL5 (white). The asparagine residue in the sequon is recognized by the two acidic residues in the DXD and TIXE motifs (red). The pyrophosphate or monophosphate group of the lipid-linked oligosaccharide interacts with the divalent metal ion via electrostatic interactions (blue). (Courtesy of D. Kohda; reprinted with permission.)

In the case of *A. fulgidus*, site-directed mutagenesis of the metal-coordinating residues of AglB-L was recently reported (198). D47A and D47N mutations resulted in a complete loss of enzymatic activity, while a D47E mutant retaining the acidic charge still showed significant activity (20%). Interestingly, mutation of the corresponding residue in *C. lari* PglB (D54) to alanine, asparagine, or glutamic acid led to an inactive enzyme (199). E360A and E360Q mutations resulted in an inactive enzyme, although E360D showed 30% activity. Thus, as seen for *C. lari* PglB, *A. fulgidus* D47 and E360 were proposed to activate the amide group of the acceptor Asn (198). Both the D161A and H163A mutations abolished *A. fulgidus* enzyme activity.

Based on extensive mutagenesis studies and crystallographic analyses, presumed roles for the conserved motifs of AglB were proposed, as presented in the model for AglB of *P. furiosus* depicted in Fig. 15 (201). These motifs are thought to interact with the two substrates of AglB, namely, the lipid-linked oligosaccharide (LLO) and the N-glycosylation sequon (N-X-S/T) of the target protein. The WWDYG motif is thought to form the Ser/Thr-binding pocket in conjunction with the lysine of the DK motif (or the isoleucine in those AglBs that possess the MI motif) for the interaction with the serine/threonine residue in position +2 of the sequon. The acidic residues in the DXD motif as well as Asp-167 and the glutamic acid of the TIXE motif are proposed to interact with the LLO pyrophosphate group by binding to a divalent cation. The asparagine residue in the sequon is recognized by the acidic residues in the first DXD motif and the glutamic acid of the TIXE motif.

### Amino Acid Preferences in and around Archaeal N-Glycosylation Sequons

Abu-Qarn and Eichler (87) were the first to report amino acid preferences in the immediate neighborhood of N-glycosylation

sites that had been experimentally proven to be occupied in *Archaea*. As described elsewhere (22), the asparagine residue of a consensus sequon is recognized by AglB, with the X position in the sequon being any residue but proline. Moreover, in contrast to *C. jejuni*, where an extended sequon (D/E-Z-N-X-S/T, where neither Z nor X is proline [210]) is seen, no requirement for an acidic residue at the −2 position was observed in archaeal glycosylation sites. More recently, *in vitro* studies examined the amino acid preferences at positions flanking N-glycosylation sites, using AglB from *P. furiosus* and peptide libraries (191). These studies confirmed that AglB does not require an acidic amino acid at the −2 position and that proline could not be in the X position of the sequon. Moreover, proline was highly unfavored at the −1 position and following the sequon Ser/Thr position. Proline was, however, the most favored residue at the −2 position. Sequons containing threonine were shown to have better acceptor activity than those containing serine. Examination of the nature of the amino acid in position X showed wide variation, with valine being the most preferred. In that same study, a mutational change of the lysine within the DK motif to alanine led to an unexpected effect on AglB activity (201). Here, the preference of the enzyme for the amino acid at the X position was altered, leading to higher activity in *in vitro* assays toward peptides having an Asn-Arg-Thr sequon rather than the otherwise optimal Asn-Val-Thr motif. The explanation for this observation lies in the removal of the positive charge of the lysine, thus preventing unfavorable electrostatic interactions toward an arginine (and lysine) at position X. Moreover, this suggested a role for the AglB DK motif in sequon recognition. Finally, many of these observations are supported by further analyses of occupied glycosylation sites in cell surface proteins of *T. acidophilum* (183).

Examination of occupied glycosylation sites in archaeal glycoproteins demonstrated that glycan attachment can occur near both the N and C termini of modified proteins (e.g., positions 2 and 6 in the mature S-layer glycoproteins of *Hbt. salinarum* [135] and *Mt. fervidus* [211], respectively, as well as an asparagine residue 9 amino acids from the C terminus of FlaB2 in *M. voltae* [111]). Glycosylation at sites near the N terminus suggests that this modification occurs following signal peptide cleavage in these proteins, while the latter data suggest that glycosylation can occur posttranslationally, since it is difficult to imagine how the C-terminal end of a glycoprotein could simultaneously be associated with the ribosome on the cytoplasmic face of the membrane and the N-glycosylation machinery on the external membrane surface. However, it has been observed that the archaellins of *M. voltae* and *M. maripaludis* are still glycosylated even when signal peptide cleavage is prevented, as in *flaK* mutants that lack the archaellin-specific prepilin-like peptidase, indicating that at least for type IV pilin-like proteins, such as archaellins, signal peptide removal is not an essential prerequisite for glycan attachment (212). It is also clear that occupied glycosylation sites can be located extremely close together, as observed for Asn-60 and Asn-65 of FlaB2 of *M. voltae* (111), Asn-119 and Asn-124 of FlaB2 of *M. maripaludis* (120), and Asn-274 and Asn-279 of the *Hfx. volcanii* S-layer glycoprotein (176).

As in *Bacteria* and *Eukarya* (213, 214), noncanonical N-glycosylation sequons can also undergo glycosylation in certain *Archaea* as well. When the Asn-Ala-Ser sequon at positions 2 to 4 of the S-layer glycoprotein of *Hbt. salinarum* was modified so that the serine residue was replaced with valine, leucine, or asparagine, the Asn-2 residue was still decorated by the repeating-unit glycan (215). Interestingly, attempts to construct a mutant carrying a change in the S-layer glycoprotein Asn-2 residue that would have prevented the attachment of the repeating-unit glycan failed, suggesting that cells without this modification to their S-layer glycoprotein are not viable. Changing of Ser-481 in the 479Asn-Ser-Ser481 sequon to valine prevented N-glycosylation with the sulfated oligosaccharide normally attached to Asn-479.

### A Variety of Linking Sugars Is Attached by AglB

One distinguishing feature of the archaeal N-glycosylation system is the variety of different linking sugars used. While both bacterial and eukaryotic OSTs require a sugar with a C-2 acetamido group as the linking sugar, this is not always the case for *Archaea*. Tai and Imperiali (216) showed that the acetamido group at C-2 played a significant role in yeast OST activity, with even minor changes in this group resulting in drastically reduced or no activity. Investigation of the substrate specificity of *C. jejuni* PglB showed that it has a relaxed specificity and can transfer not only its native heptasaccharide glycan but also a wide variety of nonnative O-antigen polysaccharides assembled on UndPP. However, all of the transferred glycans had linking sugars that contained a C-2 acetamido group (i.e., bacillosamine, GlcNAc, GalNAc, or fucosamine [FucNAc] [209, 217]), whereas attempts to transfer O-antigen polysaccharide that contained galactose at the linking position were unsuccessful (209). In *Haemophilus influenzae*, a unique N-linked glycosylation system was reported, where glycosylation occurs in the cytoplasm and is independent of PglB. Instead, UDP-Glc or UDP-Gal is linked directly to Asn sites of a classic sequon in the high-molecular-weight adhesion HMW1 protein by the HMW1C glycosyltransferase without a lipid-linked intermediate (32). In this PglB-independent case, the N-linked glycan does not contain an N-acetylated linking sugar (33). More recently, a similar system was reported for *Actinobacillus pleuropneumoniae*, where an *N*-glycosyltransferase homologue of HMW1C transferred a glucose or galactose to asparagine (218). In *Archaea*, many N-linked glycan structures have been determined, and in several cases, the linking sugar is a hexose, indicating that AglB does not require a C-2 acetamido sugar at the linking position, as do Stt3 and PglB. Such glycans are found in *A. fulgidus* (185), *T. acidophilus* (183), *Hfx. volcanii* (6), and *Hbt. salinarum* (135). The presence of the C-2 acetamido group of the linking sugar has been implicated in the catalytic activity of OST, with suggestions that it may interact with the catalytic site through hydrogen bonds (216), although it was also suggested that the acetamido group could act as a neighboring group that stabilizes the oxonium intermediate prior to the nucleophilic attack of the amide group of the asparagine target (209). The lack of this requirement in *Archaea* suggests that AglB may employ a different mechanism of catalysis.

### Multiple N-Glycans in a Single Species/Glycoprotein

An unusual feature of N-glycosylation systems in *Archaea* is the ability of certain species to attach two different glycans to the same protein. Remarkably, in the S-layer glycoprotein of *Hbt. salinarum*, an organism where only a single AglB has been identified (57), two distinct N-glycans are attached (135). A repeating-unit saccharide is found only once at Asn-2, linked via GalNAc, whereas an estimated 10 asparagines bear a lower-molecular-weight saccharide linked via glucose (219). How AglB determines the unique location of the one repeating-unit glycan is unknown.

Also unique is that the repeating-unit glycan is assembled on and transferred from a DolPP carrier, while the lower-molecular-weight glycan is transferred from a DolP carrier (82, 135, 168).

More recently, it was shown that *Hfx. volcanii* can vary the glycan structure and site of glycosylation in the S-layer glycoprotein depending on the salt concentration of the growth medium (94). If the cells are grown in medium containing 3.4 M NaCl, Asn-13 and Asn-83 are both modified with a pentasaccharide glycan, while Asn-498 is unmodified. However, in medium containing only 1.75 M NaCl, Asn-498 now becomes glycosylated but with a different glycan, a novel tetrasaccharide, while the pentasaccharide is still found attached to Asn-13 and Asn-83 albeit at reduced levels. Here too, only a single AglB-encoding gene can be detected (57). Genes involved in this second N-glycosylation pathway were recently defined (95), as described above.

In *M. maripaludis*, a tetrasaccharide is attached to archaellins (120), whereas a pentasaccharide is attached to pilins (121). How this occurs is not yet known. It may be that the same tetrasaccharide is transferred by AglB to both types of structural proteins and the branching hexose is then added to the existing tetrasaccharide on the pilins independently of AglB, as occurs for the terminal mannose in the glycan of the S-layer glycoprotein of *Hfx. volcanii* (90, 92). Alternatively, both a tetrasaccharide-bearing dolichol carrier and a pentasaccharide-bearing dolichol carrier may be synthesized, with AglB being able to transfer both glycans. However, this raises the question of how AglB distinguishes archaellins and pilins so as to deliver the proper glycan to the correct protein. On the other hand, *in vitro* studies demonstrated that AglB from *A. fulgidus* could not use a lipid-linked oligosaccharide from *P. furiosus* as the substrate and that AglB from *P. furiosus* likewise could not use *A. fulgidus* lipid-linked oligosaccharides, showing the specificity of the OST for the glycan moiety (185). Interestingly, the glycan from *A. fulgidus* uses a hexose as a linking sugar, while that of *P. furiosus* uses a C-2 acetamido sugar.

## BIOLOGICAL EFFECTS OF N-LINKED GLYCOSYLATION PERTURBATIONS IN *ARCHAEA*

Although *Archaea* can be found in numerous "normal" biological niches, such as seawater, soil, and even our own intestinal flora, these microorganisms remain best known for their abilities to thrive in extreme environments (205). Cell surface proteins found in direct contact with the harsh conditions that *Archaea* endure may thus rely on their N-linked glycans to cope with the associated physical challenges. This could explain the importance of N-glycosylation for normal cell morphology and glycoprotein function and stability, as discussed below. Indeed, the need for species-specific means of dealing with so many different extreme surroundings offers a possible explanation for the enormous diversity seen in archaeal N-linked glycan content (6, 56).

Evidence for the functional importance of N-glycosylation in *Archaea* initially came from studies conducted well before the genomic age offered insight into the mechanisms of this PTM. For example, enhanced surface charge in the face of increased salinity was offered as the reason for the high content of sulfated sugars in the N-linked glycans decorating the S-layer glycoprotein of the extreme halophile *Hbt. salinarum* (176). Furthermore, and again prior to the discovery of *agl* genes, there were indications in the literature suggesting that interference with the N-glycosylation pathway had important implications for archaellum assembly. Examination of an *Hbt. salinarum* mutant able to overproduce archaella that were recovered in the growth medium showed that the component archaellins were shifted to a lower molecular weight than the archaellins from the parent strain. This was thought to reflect a scenario whereby these proteins were underglycosylated in the mutant and that the incorporation of such archaellins into archaella yielded structures that were not properly anchored in the cell envelope (220). Similarly, treatment of *M. maripaludis* ΔRC with bacitracin, known to interfere with glycosylation, led to lower-molecular-weight, likely nonglycosylated archaellins and to cells that were no longer archaellated (133). Also, in *S. acidocaldarius*, bacitracin treatment resulted in the arrest of cell proliferation (221).

A direct relationship between N-glycosylation and the ability of a given archaeal species to survive extreme surroundings was later more clearly demonstrated with the identification of Agl proteins and the generation and examination of strains harboring deletions of the encoding genes. The most striking example of this approach serving as notice of the importance of N-glycosylation in *Archaea* is provided by *S. acidocaldarius*, where the process is essential (118). At the same time, however, N-glycosylation is not needed for the viability in *M. voltae* (71), *M. maripaludis* (119), or *Hfx. volcanii* (72). These conclusions are based on the viability of *aglB* mutants lacking the OST. However, even in those *Euryarchaeota* in which *aglB* has been successfully deleted, significant effects on the cell phenotype have been reported, related mostly to archaellation and motility, S-layer stability, and growth defects in high-salinity medium, indicating that even where N-glycosylation is not essential, it is still advantageous.

In the cases of *Hfx. volcanii* and *S. acidocaldarius*, the S-layer glycoprotein has been used extensively as the model reporter of N-glycosylation (6, 107, 118, 222). This protein is particularly important in *Archaea*, as opposed to *Bacteria*, since most *Archaea*, including these two organisms (63, 223), have a cell envelope consisting solely of an S-layer overlying a cytoplasmic membrane without the benefit of an intermediate peptidoglycan equivalent, a situation never found in S-layer-bearing *Bacteria* (224, 225). Hence, in *Archaea*, the S-layer, and, by extension, the S-layer glycoprotein, plays a significant role in terms of its interactions with the environment, shape maintenance, and, importantly, protection of the cell.

For *Hfx. volcanii*, detailed studies of the impact of absent or perturbed N-glycosylation on S-layer stability, architecture (72), as well as resistance to added protease (75, 76, 80) have been conducted. When mutants deleted for *aglJ*, responsible for adding the first sugar of the N-linked pentasaccharide decorating the S-layer glycoprotein, were challenged with proteinase K, S-layer stability, reflected as the level of intact S-layer glycoprotein, was >4-fold lower than that of parent strain cells treated similarly (75). A comparable susceptibility of the S-layer to added protease was also observed for cells lacking AglF, AglG, AglI, or AglM (76, 80). Previous experiments likewise showed increased susceptibility of the S-layer to trypsin in an *aglD* deletion mutant but, unexpectedly, not in *aglB* mutants, where no N-linked pentasaccharide decorates the S-layer glycoprotein (72). However, significantly more S-layer glycoprotein was released into the growth medium by AglB-lacking cells. The release of more S-layer glycoprotein into the growth medium of Δ*aglB* mutants could reflect a requirement for N-glycosylation in the stability of the protein or in its incorporation into the S-layer. Alternatively, the absence of N-glycosylation could increase susceptibility to attack on the cell surface.

Taken together, these results indicate that N-glycosylation is important for the proper assembly and maintenance of the S-layer on the *Hfx. volcanii* cell surface.

In *Hfx. volcanii*, N-glycosylation not only seems to be important for the architecture and stability of the S-layer but apparently also confers to the cells the ability to respond to changes in the salinity of their growth medium. Specifically, recent studies have shown that *Hfx. volcanii* modifies aspects of its N-glycosylation process in response to changing growth conditions, as revealed when the N-glycosylation profile of the S-layer glycoprotein in cells grown in medium containing 3.4 M NaCl was compared to that for medium containing 1.75 M NaCl (94). At the higher salinity, S-layer glycoprotein Asn-13 and Asn-83 were modified by the pentasaccharide described above, while DolP was shown to be modified by the tetrasaccharide comprising the first four pentasaccharide residues, again as discussed above. However, in cells grown at the lower salinity, Asn-13 and Asn-83 were still modified by the pentasaccharide albeit to a lesser extent. More strikingly, DolP charged with a distinct tetrasaccharide comprising a sulfated hexose, two hexoses, and a rhamnose was observed. This is likely the same DolP-bound tetrasaccharide reported previously when *Hfx. volcanii* cells were grown in 1.25 M NaCl-containing medium (85). The same tetrasaccharide was shown to decorate S-layer glycoprotein Asn-498 in cells grown in low-salt medium, a position that was not modified in cells grown in high-salt medium. Under both salt conditions, a potential glycosylation site at Asn-370 was unmodified. Hence, in response to changes in environmental salinity, *Hfx. volcanii* modulates not only the N-linked glycans decorating the S-layer glycoprotein but also the sites of this PTM. Presently, it is not clear what advantage this differential N-glycosylation offers *Hfx. volcanii* in the face of changes in environmental salinity, nor is it clear which properties of a glycan make it favored under specific growth conditions. While no studies examining possible glycosylation changes as a function of growth medium salinity have been conducted with other archaea, it seems unlikely that such variations in response to environmental conditions are limited to this species alone. Accordingly, glycosylation of *Methanospirillum hungatei* archaellins was reported to occur only with low-phosphate-containing medium (158).

In *S. acidocaldarius*, disruptions in N-glycosylation achieved by the deletion of *agl* genes also affected growth at higher salinities. While an *agl16* deletion strain, lacking only a single sugar from the N-glycan, grew at the same rate as the parent strain in standard medium containing no NaCl, an increase of the salinity to 300 mM NaCl led to a 50% decrease in the growth rate of the deletion strain (156). At the same salinity, an *agl3* mutant, where three subunits of the N-linked glycan are missing, had a doubling time that was 5 times longer than that of the parent strain (155). In medium containing 400 mM NaCl, only parent strain cells were able to grow although only after a long lag period.

One striking aspect of the N-glycosylation process in *Sulfolobus* is the high glycosylation density per glycoprotein. Specifically, an examination of the number of putative N-glycosylation sites in the large S-layer glycoprotein from different *Sulfolobus* species made by using the NetNGlyc 1.0 server (http://www.cbs.dtu.dk/services/NetNGlyc) predicted an extremely high glycosylation density of 32 to 41 N-glycosylated sequons per protein (118). Nano-LC-ESI MS/MS analysis of the C-terminal portion of the *S. acidocaldarius* S-layer glycoprotein confirmed this high glycosylation density, where, on average, an N-linked glycan is detected per 30 to 40 residues (137). Such high-density N-glycosylation has not been proven for the S-layer glycoproteins of any other archaeal species. However, it seems that thermophilic archaea tend to have a higher N-glycosylation frequency than mesophilic archaea, which typically possess an average of 2 to 10 predicted N-glycosylated sequons per protein (Table 2). For instance, the S-layer glycoproteins from *Hfx. volcanii* and *M. voltae* present 7 and 2 sequons, respectively (63, 111). The C-terminal domain of the *S. acidocaldarius* S-layer glycoprotein, spanning residues L1004 to Q1395, has 11 predicted N-linked glycosylation sites, 9 of which were experimentally confirmed to be occupied (137). Therefore, at least for the *Sulfolobus* species that have a high glycosylation density of their extracellular proteins, the experimentally determined glycosylation sites match very well to the predicted glycosylation sites. The high glycosylation density of the *S. acidocaldarius* S-layer glycoprotein could represent an adaptation of this organism to the high temperatures and acidic environment that it naturally encounters. As glycosylation plays an important role in protein stability, the N-glycosylation of proteins in thermophilic archaea might have a greater impact on the maintenance of cell integrity at the high growth temperature that they require than N-glycosylation of proteins in mesophilic archaea. This could also explain why N-glycosylation is essential for *S. acidocaldarius*, in contrast to the other *Archaea* studied to date. Currently, there have been two studies comparing the glycosylation of S-layer glycoproteins from hyperthermophilic species and their mesophilic relatives. In these studies, it was noted that the S-layer glycoproteins of the hyperthermophiles *Mcc. jannaschii* and *Methanotorris igneus* contain a higher number of potential N-glycosylation sites than the same protein in mesophilic species (167, 226). While this observation underscores the potential contribution that N-glycosylation makes to protein thermostability, thermal stabilization of S-layer glycoproteins may not rely solely on glycosylation density and/or the composition of the N-glycan but may also be attributed to phosphorylation, salt bridging, or covalent cross-linking (227). Indeed, examination of the amino acid sequences of the S-layer proteins of mesophilic, thermophilic, and hyperthermophilic methanococci revealed the highest ratio of cysteines for the hyperthermophilic *Methanococcales* (228).

The creation of various *agl* gene deletion strains that resulted in truncated N-glycans of known composition led to a more defined correlation between N-linked glycosylation and archaellation in the methanogens. This correlation is very well defined for both *M. voltae* and *M. maripaludis* (71, 119, 229). Any deletion in a GT, OST, or biosynthetic gene that resulted in nonglycosylated archaellins or archaellins decorated with a glycan consisting of only the first sugar led to cells that were nonarchaellated. Any mutant strain carrying an *agl* gene mutation that resulted in a truncated glycan larger than one sugar was still archaellated, but the motility of the mutant cells was impaired compared to the motility of wild-type cells, with impairment in swimming being related directly to the degree of truncation in the glycan. It is not known whether a minimum-length glycan is needed for recognition by the secretion/assembly system or whether those structures assembled by archaellins bearing severely truncated glycans are simply unstable.

The defects in assembly of archaella in various *agl* mutants have not been observed for pili. While the major pilin that comprises the type IV-like pili on the surface of *M. maripaludis* cells has been shown to also be decorated with N-glycans (121), the assembly of wild-type-looking pili is usually unaffected by dele-

tions in *agl* genes that inhibit archaellum formation. There is no more drastic example of this than that of an *aglB* mutant, where the pilins are completely nonglycosylated and yet the cells remain piliated (119). However, whether these pili are fully functional (230) has not been addressed. There is one exception involving *agl* genes and piliation, and this is for *mmp0350* mutants. MMP0350 is likely an acetyltransferase for the second sugar, as the deletion of *mmp0350* results in a one-sugar glycan (229). These cells are nonarchaellated, as expected, but also have difficulty in maintaining pilus attachment to the cell surface. Pili are assembled but are instead released into the culture medium.

Deletion of *agl* genes had similarly dramatic effects on *Hfx. volcanii* and *S. acidocaldarius* archaellation as was found initially for *Methanococcus*. In *Hfx. volcanii*, three predicted N-glycosylation sequons of the major structural archaellin FlgA1 were shown to be decorated with the same pentasaccharide as the one first reported to modify the S-layer glycoprotein (99). An *aglB* deletion strain was nonmotile and nonarchaellated. In addition, a variety of mutants carrying deletions in other *agl* genes involved in the biosynthesis or assembly of the glycan also showed motility defects in semisolid media. In *Hfx volcanii*, it seems that an N-linked glycan containing a minimum of three sugar subunits is necessary for archaellum assembly, since *aglJ*, *aglG*, *aglM*, *aglI*, or *aglF* deletion mutants, where gene products necessary for the biosynthesis and/or assembly of the first three glycan subunits are absent (75, 76, 80), were nonmotile. An *algE* deletion mutant, which was unable to code for a protein needed for the attachment of the fourth subunit of the N-linked pentasaccharide (74), was very poorly motile compared to parent strain cells. Unexpectedly, a mutant carrying a deletion of *aglD*, coding for the GT involved in the addition of the final N-linked pentasaccharide subunit, mannose (72), was nonmotile. When the asparagine of any of three examined sequons of the major archaellin FlgA was replaced by glutamine, the mutant forms, unlike wild-type *flgA*, could not rescue the swimming motility phenotype of a *flgA* deletion (99), suggesting that glycosylation at each of these sites is necessary for archaellation. For *S. acidocaldarius*, a connection between N-glycosylation and archaellation was also recently found. A strain deleted for *agl16*, the GT needed for the addition of the terminal hexose of the N-linked glycan decorating archaellins in this species (156), was severely impaired in its motility. No motility was detected in a strain carrying a deletion of *agl3*, coding for the UDP-sulfoquinovose synthase, which leads to a truncated glycan comprising a trisaccharide alone (155). Finally, electron microscopy showed that cells lacking either of these two *agl* genes were nonarchaellated, indicating that the loss of even a single sugar from the N-linked hexasaccharide attached to archaellins led to the generation of unstable archaella.

## FUTURE OUTLOOK

Remarkable progress has been made in the elucidation of the N-glycosylation pathway in model *Archaea* in less than a decade. However, it is clear that much more is still unknown. While protein N-glycosylation is thought to be almost universal in *Archaea*, glycan structures are known for only a few organisms, and genetic studies on the pathways involved have been performed with even fewer organisms, due mainly to a lack of proper tools. Indeed, many key steps remain a mystery for any archaeon, the foremost being the identity and mechanisms of the proposed flippases. Biosynthetic pathways for N-glycan constituents also remain patchy. This is especially true for the unique and unusual sugars found in archaeal N-glycans. Even basic information about the lipid carriers involved is lacking, such as their biosynthetic pathway and whether both DolP and DolPP carriers are involved.

Another whole set of challenges arises because it is often the case that important findings presented for one archaeon remain unstudied for others. It thus remains unknown whether such findings are applicable to the entire domain or are unique to a single or limited number of species. Recent findings showing the ability of *Hfx. volcanii* to vary the structure of its N-linked glycan depending on growth conditions are a case in point. Variation in the N-glycosylation of proteins as an adaptation mechanism is likely to be widespread but is currently unproven. The features of a certain glycan that would make it favored under certain growth conditions are also unknown. Likewise, it is not known how many different glycans one organism can make as various aspects of the growth conditions are modified. It is possible that certain *Archaea* may be able to synthesize many different glycans depending on the nature of a particular environmental stress. Where two different glycans are attached to the same protein, it is not known how the cell determines which glycan goes to which sequon and how this could be accomplished with a single OST.

In fact, the link between extremophilic habitats and protein N-glycosylation is currently more conjecture than fact. It is not known if certain archaeal proteins are more stable in their extreme niches because they are glycosylated. Certainly, some *Archaea* can readily survive without this PTM, as shown by viable *aglB* deletion strains of both *Methanococcus* and *Haloferax* species. However, for *Sulfolobus*, which lives in a more extreme environment in terms of high temperature and low pH, N-glycosylation is necessary, possibly to stabilize proteins under such harsh conditions. An understanding of this connection, of course, has broad implications in biotechnology, as it may offer perspective on tailoring specific N-glycans to synthesize glycoproteins that are both more active and more stable under specific physically challenging conditions.

It is not yet known if archaeal N-glycosylation pathways or parts thereof can be reconstituted in bacterial or eukaryal systems to create glycoproteins in other domains that incorporate archaeon-specific sugars as components. There is limited evidence that archaeal enzymes involved in N-glycosylation can complement eukaryotic deletions. It seems possible that archaeal enzymes, in particular AglB, might be useful for glycoengineering of vaccines. A limitation of the use of PglB in protein glycan coupling technology is the requirement of this OST for a linking sugar presenting an acetamido group in the C-2 position (231, 232). The employment of AglB in these efforts may help overcome this requirement, since *Archaea* use a broader variety of linking sugars in their N-glycans, including those lacking this C-2 modification. The unusual and sometimes unique sugars found in archaeal N-linked glycans open up the possibility of identifying potentially useful archaeal enzymes for specific roles in glycoengineering, in much the same way that a eukaryotic glycosylation pathway was recently engineered in *E. coli* (233). Continued study of N-glycosylation in *Archaea* should add to our knowledge of the evolutionary history of the pathway, which is thought to have archaeal origins (56). This may be especially fruitful for *Sulfolobus*, where the core structure of the glycan is the chitobiose observed in eukaryotic glycans, such that homologous pathways may be found.

One key to progress in archaeal N-glycosylation studies is to break the bottleneck caused by the current access to too few ar-

chaeal organisms with appropriate genetic systems (234). With apparently so many other *Archaea* with N-glycosylation systems to study, the chances of discovery of unique variations of the conserved system are high. As with many properties of *Archaea*, the N-glycosylation pathway appears to share common features with parallel systems in other domains but with added archaeon-specific variations. Even with the limited studies conducted to date, it is clear that the archaeal domain holds the potential for vast diversity in N-glycan structures and biosynthetic enzymes that could be used for biotechnology. Continued study is certain to not only expand our knowledge of N-glycosylation in the *Archaea* but also shed new light on the process throughout all domains of life.

## ACKNOWLEDGMENTS

K.F.J. is supported by a Discovery Grant from the Natural Sciences and Engineering Research Council of Canada (NSERC). Y.D. is supported by the China Scholarship Council and an Ontario Graduate Scholarship. S.-V.A. was supported by intramural funds of the Max Planck Society, and B.H.M. was supported by an ERC starting grant (number 311523, Archaellum). J.E. is supported by the Israel Science Foundation (grant 8/11) and the U.S. Army Research Office (W911NF-11-1-520). L.K. is the recipient of a Negev-Zin Associates Scholarship.

## REFERENCES

1. **Prabakaran S, Lippens G, Steen H, Gunawardena J.** 2012. Post-translational modification: nature's escape from genetic imprisonment and the basis for dynamic information encoding. Wiley Interdiscip. Rev. Syst. Biol. Med. **4:**565–583. http://dx.doi.org/10.1002/wsbm.1185.
2. **Eichler J, Adams MW.** 2005. Posttranslational protein modification in Archaea. Microbiol. Mol. Biol. Rev. **69:**393–425. http://dx.doi.org/10.1128/MMBR.69.3.393-425.2005.
3. **Apweiler R, Hermjakob H, Sharon N.** 1999. On the frequency of protein glycosylation, as deduced from analysis of the SWISS-PROT database. Biochim. Biophys. Acta **1473:**4–8. http://dx.doi.org/10.1016/S0304-4165(99)00165-8.
4. **Aebi M.** 2013. N-linked protein glycosylation in the ER. Biochim. Biophys. Acta **1833:**2430–2437. http://dx.doi.org/10.1016/j.bbamcr.2013.04.001.
5. **Breitling J, Aebi M.** 2013. N-linked protein glycosylation in the endoplasmic reticulum. Cold Spring Harb. Perspect. Biol. **5:**a013359. http://dx.doi.org/10.1101/cshperspect.a013359.
6. **Eichler J.** 2013. Extreme sweetness: protein glycosylation in Archaea. Nat. Rev. Microbiol. **11:**151–156. http://dx.doi.org/10.1038/nrmicro2957.
7. **Larkin A, Imperiali B.** 2011. The expanding horizons of asparagine-linked glycosylation. Biochemistry **50:**4411–4426. http://dx.doi.org/10.1021/bi200346n.
8. **Zarschler K, Janesch B, Pabst M, Altmann F, Messner P, Schäffer C.** 2010. Protein tyrosine O-glycosylation—a rather unexplored prokaryotic glycosylation system. Glycobiology **20:**787–798. http://dx.doi.org/10.1093/glycob/cwq035.
9. **Iwashkiw JA, Vozza NF, Kinsella RL, Feldman MF.** 2013. Pour some sugar on it: the expanding world of bacterial protein O-linked glycosylation. Mol. Microbiol. **89:**14–28. http://dx.doi.org/10.1111/mmi.12265.
10. **Kinoshita T, Fujita M, Maeda Y.** 2008. Biosynthesis, remodelling and functions of mammalian GPI-anchored proteins: recent progress. J. Biochem. **144:**287–294. http://dx.doi.org/10.1093/jb/mvn090.
11. **Eisenhaber B, Bork P, Eisenhaber F.** 2001. Post-translational GPI lipid anchor modification of proteins in kingdoms of life: analysis of protein sequence data from complete genomes. Protein Eng. **14:**17–25. http://dx.doi.org/10.1093/protein/14.1.17.
12. **Kobayashi T, Nishizaki R, Ikezawa H.** 1997. The presence of GPI-linked protein(s) in an archaeobacterium, Sulfolobus acidocaldarius, closely related to eukaryotes. Biochim. Biophys. Acta **1334:**1–4. http://dx.doi.org/10.1016/S0304-4165(96)00099-2.
13. **Ilg T.** 2000. Proteophosphoglycans of *Leishmania*. Parasitol. Today **16:**489–497. http://dx.doi.org/10.1016/S0169-4758(00)01791-9.
14. **Haynes PA.** 1998. Phosphoglycosylation: a new structural class of glycosylation? Glycobiology **8:**1–5. http://dx.doi.org/10.1093/glycob/8.1.1.
15. **Lippert DN, Dwyer DW, Li F, Olafson RW.** 1999. Phosphoglycosylation of a secreted acid phosphatase from *Leishmania donovani*. Glycobiology **9:**627–636. http://dx.doi.org/10.1093/glycob/9.6.627.
16. **Freeze HH, Ichikawa M.** 1995. Identification of N-acetylglucosamine-alpha-1-phosphate transferase activity in *Dictyostelium discoideum*: an enzyme that initiates phosphoglycosylation. Biochem. Biophys. Res. Commun. **208:**384–389.
17. **Krieg J, Hartmann S, Vicentini A, Gläsner W, Hess D, Hofsteenge J.** 1998. Recognition signal for C-mannosylation of Trp-7 in RNase 2 consists of sequence Trp-x-x-Trp. Mol. Biol. Cell **9:**301–309. http://dx.doi.org/10.1091/mbc.9.2.301.
18. **Doucey MA, Hess D, Cacan R, Hofsteenge J.** 1998. Protein C-mannosylation is enzyme-catalysed and uses dolichyl-phosphate-mannose as a precursor. Mol. Biol. Cell **9:**291–300. http://dx.doi.org/10.1091/mbc.9.2.291.
19. **Julenius K.** 2007. NetCGlyc 1.0: prediction of mammalian C-mannosylation sites. Glycobiology **17:**868–876. http://dx.doi.org/10.1093/glycob/cwm050.
20. **Furmanek A, Hofsteenge J.** 2000. Protein C-mannosylation: facts and questions. Acta Biochim. Pol. **47:**781–789. http://www.actabp.pl/pdf/3_2000/781.pdf.
21. **Spiro RG.** 2002. Protein glycosylation: nature, distribution, enzymatic formation, and disease implications of glycopeptide bonds. Glycobiology **12:**43R–56R. http://dx.doi.org/10.1093/glycob/12.4.43R.
22. **Marshall RD.** 1972. Glycoproteins. Annu. Rev. Biochem. **41:**673–702.
23. **Gavel Y, 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. http://dx.doi.org/10.1093/protein/3.5.433.
24. **Bause E, Hettkamp H.** 1979. Primary structural requirements for N-glycosylation of peptides in rat liver. FEBS Lett. **108:**341–344. http://dx.doi.org/10.1016/0014-5793(79)80559-1.
25. **Nothaft H, Szymanski CM.** 2010. Protein glycosylation in bacteria: sweeter than ever. Nat. Rev. Microbiol. **8:**765–778. http://dx.doi.org/10.1038/nrmicro2383.
26. **Wacker M, Linton D, Hitchen PG, Nita-Lazar M, Haslam SM, North SJ, Panico M, Morris HR, Dell A, Wren BW, Aebi M.** 2002. N-linked glycosylation in *Campylobacter jejuni* and its functional transfer into *E. coli*. Science **298:**1790–1793. http://dx.doi.org/10.1126/science.298.5599.1790.
27. **Karaoglu D, Kelleher DJ, Gilmore R.** 1997. The highly conserved Stt3 protein is a subunit of the yeast oligosaccharyltransferase and forms a subcomplex with Ost3p and Ost4p. J. Biol. Chem. **272:**32513–32520. http://dx.doi.org/10.1074/jbc.272.51.32513.
28. **Spirig U, Glavas M, Bodmer D, Reiss G, Burda P, Lippuner V, te Heesen S, Aebi M.** 1997. The STT3 protein is a component of the yeast oligosaccharyltransferase complex. Mol. Gen. Genet. **256:**628–637. http://dx.doi.org/10.1007/s004380050611.
29. **Kelleher DJ, Gilmore R.** 2006. An evolving view of the eukaryotic oligosaccharyltransferase. Glycobiology **16:**47R–62R. http://dx.doi.org/10.1093/glycob/cwj066.
30. **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. http://dx.doi.org/10.1038/sj.emboj.7601940.
31. **Yan Q, Lennarz WJ.** 1999. Oligosaccharyltransferase: a complex multisubunit enzyme of the endoplasmic reticulum. Biochem. Biophys. Res. Commun. **266:**684–689. http://dx.doi.org/10.1006/bbrc.1999.1886.
32. **Grass S, Lichti CF, Townsend RR, Gross J, St Geme JW, III.** 2010. The *Haemophilus influenzae* HMW1C protein is a glycosyltransferase that transfers hexose residues to asparagine sites in the HMW1 adhesin. PLoS One **6:**e1000919. http://dx.doi.org/10.1371/journal.ppat.1000919.
33. **Gross J, Grass S, Davis AE, Gilmore-Erdmann P, Townsend RR, St Geme JW, III.** 2008. The *Haemophilus influenzae* HMW1 adhesin is a glycoprotein with an unusual N-linked carbohydrate modification. J. Biol. Chem. **283:**26010–26015. http://dx.doi.org/10.1074/jbc.M801819200.
34. **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. http://dx.doi.org/10.1038/415447a.

35. **Frank CG, Sanyal S, Rush JS, Waechter CJ, Menon AK.** 2008. Does Rft1 flip an N-glycan lipid precursor? Nature **454:**E3–E4; discussion E4–E5. http://dx.doi.org/10.1038/nature07165.
36. **Rush JS, Gao N, Lehrman MA, Matveev S, Waechter CJ.** 2009. Suppression of Rft1 expression does not impair the transbilayer movement of Man5GlcNAc2-P-P-dolichol in sealed microsomes from yeast. J. Biol. Chem. **284:**19835–19842. http://dx.doi.org/10.1074/jbc.M109.000893.
37. **Jelk J, Gao N, Serricchio M, Signorell A, Schmidt RS, Bangs JD, Acosta-Serrano A, Lehrman MA, Bütikofer P, Menon AK.** 2013. Glycoprotein biosynthesis in a eukaryote lacking the membrane protein Rft1. J. Biol. Chem. **288:**20616–20623. http://dx.doi.org/10.1074/jbc.M113.479642.
38. **Sanyal S, Menon AK.** 2010. Stereoselective transbilayer translocation of mannosyl phosphoryl dolichol by an endoplasmic reticulum flippase. Proc. Natl. Acad. Sci. U. S. A. **107:**11289–11294. http://dx.doi.org/10.1073/pnas.1002408107.
39. **Gemmill TR, Trimble RB.** 1999. Overview of N- and O-linked oligosaccharide structures found in various yeast species. Biochim. Biophys. Acta **1426:**227–237. http://dx.doi.org/10.1016/S0304-4165(98)00126-3.
40. **Neuberger A.** 1938. Carbohydrates in protein: the carbohydrate component of crystalline egg albumin. Biochem. J. **32:**1435–1451.
41. **Ventosa A, Oren A.** 1996. *Halobacterium salinarum* nom. corrig., a name to replace *Halobacterium salinarium* (Elazari-Volcani) and to include *Halobacterium halobium* and *Halobacterium cutirubrum*. Int. J. Syst. Evol. Microbiol. **46:**347.
42. **Koncewicz MA.** 1972. Glycoproteins in the cell envelope of *Halobacterium halobium*. Biochem. J. **128:**124P.
43. **Mescher MF, Strominger JL, Watson SW.** 1974. Protein and carbohydrate composition of the cell envelope of *Halobacterium salinarium*. J. Bacteriol. **120:**945–954.
44. **Mescher MF, Strominger JL.** 1976. Purification and characterization of a prokaryotic glucoprotein from the cell envelope of *Halobacterium salinarium*. J. Biol. Chem. **251:**2005–2014.
45. **Sleytr UB, Thorne KJ.** 1976. Chemical characterization of the regularly arranged surface layers of *Clostridium thermosaccharolyticum* and *Clostridium thermohydrosulfuricum*. J. Bacteriol. **126:**377–383.
46. **Schäffer C, Messner P.** 2004. Surface-layer glycoproteins: an example for the diversity of bacterial glycosylation with promising impacts on nanobiotechnology. Glycobiology **14:**31R–42R. http://dx.doi.org/10.1093/glycob/cwh064.
47. **Messner P.** 2009. Prokaryotic protein glycosylation is rapidly expanding from "curiosity" to "ubiquity." Chembiochem **10:**2151–2154. http://dx.doi.org/10.1002/cbic.200900388.
48. **Szymanski CM, Yao R, Ewing CP, Trust TJ, Guerry P.** 1999. Evidence for a system of general protein glycosylation in *Campylobacter jejuni*. Mol. Microbiol. **32:**1022–1030. http://dx.doi.org/10.1046/j.1365-2958.1999.01415.x.
49. **Young NM, Brisson JR, Kelly J, Watson DC, Tessier L, Lanthier PH, Jarrell HC, Cadotte N, St Michael F, Aberg E, Szymanski CM.** 2002. Structure of the N-linked glycan present on multiple glycoproteins in the Gram-negative bacterium, *Campylobacter jejuni*. J. Biol. Chem. **277:**42530–42539. http://dx.doi.org/10.1074/jbc.M206114200.
50. **Nothaft H, Liu X, McNally DJ, Szymanski CM.** 2010. N-linked protein glycosylation in a bacterial system. Methods Mol. Biol. **600:**227–243. http://dx.doi.org/10.1007/978-1-60761-454-8_16.
51. **Nothaft H, Szymanski CM.** 2013. Bacterial protein N-glycosylation: new perspectives and applications. J. Biol. Chem. **288:**6912–6920. http://dx.doi.org/10.1074/jbc.R112.417857.
52. **Szymanski CM, Wren BW.** 2005. Protein glycosylation in bacterial mucosal pathogens. Nat. Rev. Microbiol. **3:**225–237. http://dx.doi.org/10.1038/nrmicro1100.
53. **Alaimo C, Catrein I, Morf L, Marolda CL, Callewaert N, Valvano MA, Feldman MF, Aebi M.** 2006. Two distinct but interchangeable mechanisms for flipping of lipid-linked oligosaccharides. EMBO J. **25:**967–976. http://dx.doi.org/10.1038/sj.emboj.7601024.
54. **Woese CR, Fox GE.** 1977. Phylogenetic structure of the prokaryotic domain: the primary kingdoms. Proc. Natl. Acad. Sci. U. S. A. **74:**5088–5090. http://dx.doi.org/10.1073/pnas.74.11.5088.
55. **Fox GE, Magrum LJ, Balch WE, Wolfe RS, Woese CR.** 1977. Classification of methanogenic bacteria by 16S ribosomal RNA characterization. Proc. Natl. Acad. Sci. U. S. A. **74:**4537–4541. http://dx.doi.org/10.1073/pnas.74.10.4537.
56. **Schwarz F, Aebi M.** 2011. Mechanisms and principles of N-linked protein glycosylation. Curr. Opin. Struct. Biol. **21:**576–582. http://dx.doi.org/10.1016/j.sbi.2011.08.005.
57. **Kaminski L, Lurie-Weinberger MN, Allers T, Gophna U, Eichler J.** 2013. Phylogenetic- and genome-derived insight into the evolution of N-glycosylation in Archaea. Mol. Phylogenet. Evol. **68:**327–339. http://dx.doi.org/10.1016/j.ympev.2013.03.024.
58. **Dell A, Galadari A, Sastre F, Hitchen P.** 2010. Similarities and differences in the glycosylation mechanisms in prokaryotes and eukaryotes. Int. J. Microbiol. **2010:**148178. http://dx.doi.org/10.1155/2010/148178.
59. **Calo D, Kaminski L, Eichler J.** 2010. Protein glycosylation in Archaea: sweet and extreme. Glycobiology **20:**1065–1076. http://dx.doi.org/10.1093/glycob/cwq055.
60. **Weerapana E, Imperiali B.** 2006. Asparagine-linked protein glycosylation: from eukaryotic to prokaryotic systems. Glycobiology **16:**91R–101R. http://dx.doi.org/10.1093/glycob/cwj099.
61. **Eichler J, Jarrell K, Albers S.** 2013. A proposal for the naming of N-glycosylation pathway components in Archaea. Glycobiology **23:**620–621. http://dx.doi.org/10.1093/glycob/cwt034.
62. **Bhat AH, Mondal H, Chauhan JS, Raghava GP, Methi A, Rao A.** 2012. ProGlycProt: a repository of experimentally characterized prokaryotic glycoproteins. Nucleic Acids Res. **40:**D388–D393. http://dx.doi.org/10.1093/nar/gkr911.
63. **Sumper M, Berg E, Mengele R, Strobel I.** 1990. Primary structure and glycosylation of the S-layer protein of *Haloferax volcanii*. J. Bacteriol. **172:**7111–7118.
64. **Eichler J, Maupin-Furlow J.** 2013. Post-translation modification in Archaea: lessons from *Haloferax volcanii* and other haloarchaea. FEMS Microbiol. Rev. **37:**583–606. http://dx.doi.org/10.1111/1574-6976.12012.
65. **Abu-Qarn M, Eichler J.** 2006. Protein N-glycosylation in Archaea: defining *Haloferax volcanii* genes involved in S-layer glycoprotein glycosylation. Mol. Microbiol. **61:**511–525. http://dx.doi.org/10.1111/j.1365-2958.2006.05252.x.
66. **Burda P, Aebi M.** 1999. The dolichol pathway of N-linked glycosylation. Biochim. Biophys. Acta **1426:**239–257. http://dx.doi.org/10.1016/S0304-4165(98)00127-5.
67. **Kornfeld R, Kornfeld S.** 1985. Assembly of asparagine-linked oligosaccharides. Annu. Rev. Biochem. **54:**631–664. http://dx.doi.org/10.1146/annurev.bi.54.070185.003215.
68. **Yan A, Lennarz WJ.** 2005. Unraveling the mechanism of protein N-glycosylation. J. Biol. Chem. **280:**3121–3124. http://dx.doi.org/10.1074/jbc.R400036200.
69. **Linton D, Dorrell N, Hitchen PG, Amber S, Karlyshev AV, Morris HR, Dell A, Valvano MA, Aebi M, Wren BW.** 2005. Functional analysis of the *Campylobacter jejuni* N-linked protein glycosylation pathway. Mol. Microbiol. **55:**1695–1703. http://dx.doi.org/10.1111/j.1365-2958.2005.04519.x.
70. **Allers T, Ngo HP, Mevarech M, Lloyd RG.** 2004. Development of additional selectable markers for the halophilic archaeon *Haloferax volcanii* based on the *leuB* and *trpA* genes. Appl. Environ. Microbiol. **70:**943–953. http://dx.doi.org/10.1128/AEM.70.2.943-953.2004.
71. **Chaban B, Voisin S, Kelly J, Logan SM, Jarrell KF.** 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. http://dx.doi.org/10.1111/j.1365-2958.2006.05226.x.
72. **Abu-Qarn M, Yurist-Doutsch S, Giordano A, Trauner A, Morris HR, Hitchen P, Medalia O, Dell A, Eichler J.** 2007. *Haloferax volcanii* AglB and AglD are involved in N-glycosylation of the S-layer glycoprotein and proper assembly of the surface layer. J. Mol. Biol. **374:**1224–1236. http://dx.doi.org/10.1016/j.jmb.2007.10.042.
73. **Kaminski L, Eichler J.** 2010. Identification of residues important for the activity of *Haloferax volcanii* AglD, a component of the archaeal N-glycosylation pathway. Archaea **2010:**315108. http://dx.doi.org/10.1155/2010/315108.
74. **Abu-Qarn M, Giordano A, Battaglia F, Trauner A, Hitchen PG, Morris HR, Dell A, Eichler J.** 2008. Identification of AglE, a second glycosyltransferase involved in N-glycosylation of the *Haloferax volcanii* S-layer glycoprotein. J. Bacteriol. **190:**3140–3146. http://dx.doi.org/10.1128/JB.00056-08.
75. **Kaminski L, Abu-Qarn M, Guan Z, Naparstek S, Ventura VV, Raetz CR, Hitchen PG, Dell A, Eichler J.** 2010. AglJ adds the first sugar of the N-linked pentasaccharide decorating the *Haloferax volcanii* S-layer gly-

coprotein. J. Bacteriol. **192:**5572–5579. http://dx.doi.org/10.1128/JB.00705-10.
76. **Yurist-Doutsch S, Abu-Qarn M, Battaglia F, Morris HR, Hitchen PG, Dell A, Eichler J.** 2008. aglF, aglG and aglI, novel members of a gene island involved in the N-glycosylation of the *Haloferax volcanii* S-layer glycoprotein. Mol. Microbiol. **69:**1234–1245. http://dx.doi.org/10.1111/j.1365-2958.2008.06352.x.
77. **Glucksmann MA, Reuber TL, Walker GC.** 1993. Genes needed for the modification, polymerization, export, and processing of succinoglycan by *Rhizobium meliloti*: a model for succinoglycan biosynthesis. J. Bacteriol. **175:**7045–7055.
78. **Yurist-Doutsch S, Eichler J.** 2009. Manual annotation, transcriptional analysis, and protein expression studies reveal novel genes in the agl cluster responsible for N glycosylation in the halophilic archaeon *Haloferax volcanii*. J. Bacteriol. **191:**3068–3075. http://dx.doi.org/10.1128/JB.01838-08.
79. **Magidovich H, Yurist-Doutsch S, Konrad Z, Ventura VV, Dell A, Hitchen PG, Eichler J.** 2010. AglP is an S-adenosyl-L-methionine-dependent methyltransferase that participates in the N-glycosylation pathway in *Haloferax volcanii*. Mol. Microbiol. **76:**190–199. http://dx.doi.org/10.1111/j.1365-2958.2010.07090.x.
80. **Yurist-Doutsch S, Magidovich H, Ventura VV, Hitchen PG, Dell A, Eichler J.** 2010. N-glycosylation in Archaea: on the coordinated actions of *Haloferax volcanii* AglF and AglM. Mol. Microbiol. **75:**1047–1058. http://dx.doi.org/10.1111/j.1365-2958.2009.07045.x.
81. **Plavner N, Eichler J.** 2008. Defining the topology of the N-glycosylation pathway in the halophilic archaeon *Haloferax volcanii*. J. Bacteriol. **190:**8045–8052. http://dx.doi.org/10.1128/JB.01200-08.
82. **Lechner J, Wieland F, 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.
83. **Wieland F, Dompert W, Bernhardt G, Sumper M.** 1980. Halobacterial glycoprotein saccharides contain covalently linked sulphate. FEBS Lett. **120:**110–114. http://dx.doi.org/10.1016/0014-5793(80)81058-1.
84. **Wieland F, Lechner J, Sumper M.** 1986. Iduronic acid: constituent of sulphated dolichyl phosphate oligosaccharides in halobacteria. FEBS Lett. **195:**77–81. http://dx.doi.org/10.1016/0014-5793(86)80134-X.
85. **Kuntz C, Sonnenbichler J, Sonnenbichler I, Sumper M, Zeitler R.** 1997. Isolation and characterization of dolichol-linked oligosaccharides from *Haloferax volcanii*. Glycobiology **7:**897–904. http://dx.doi.org/10.1093/glycob/7.7.897.
86. **Guan Z, Naparstek S, Kaminski L, Konrad Z, Eichler J.** 2010. Distinct glycan-charged phosphodolichol carriers are required for the assembly of the pentasaccharide N-linked to the *Haloferax volcanii* S-layer glycoprotein. Mol. Microbiol. **78:**1294–1303. http://dx.doi.org/10.1111/j.1365-2958.2010.07405.x.
87. **Abu-Qarn M, Eichler J.** 2007. An analysis of amino acid sequences surrounding archaeal glycoprotein sequons. Archaea **2:**73–81. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2686383/.
88. **Swiezewska E, Danikiewicz W.** 2005. Polyisoprenoids: structure, biosynthesis and function. Prog. Lipid Res. **44:**235–258. http://dx.doi.org/10.1016/j.plipres.2005.05.002.
89. **Naparstek S, Guan Z, Eichler J.** 2012. A predicted geranylgeranyl reductase reduces the ω-position isoprene of dolichol phosphate in the halophilic archaeon, *Haloferax volcanii*. Biochim. Biophys. Acta **1821:**923–933. http://dx.doi.org/10.1016/j.bbalip.2012.03.002.
90. **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. http://dx.doi.org/10.1016/j.bbagen.2012.06.014.
91. **Liu D, Cole RA, Reeves PR.** 1996. An O-antigen processing function for Wzx (RfbX): a promising candidate for O-unit flippase. J. Bacteriol. **178:**2102–2107.
92. **Cohen-Rosenzweig C, Yurist-Doutsch S, Eichler J.** 2012. AglS, a novel component of the *Haloferax volcanii* N-glycosylation pathway, is a dolichol phosphate-mannose mannosyltransferase. J. Bacteriol. **194:**6909–6916. http://dx.doi.org/10.1128/JB.01716-12.
93. **Mullakhanbhai MF, Larsen H.** 1975. *Halobacterium volcanii* spec. nov., a Dead Sea halobacterium with a moderate salt requirement. Arch. Microbiol. **104:**207–214. http://dx.doi.org/10.1007/BF00447326.
94. **Guan Z, Naparstek S, Calo D, Eichler J.** 2012. Protein glycosylation as an adaptive response in Archaea: growth at different salt concentrations leads to alterations in *Haloferax volcanii* S-layer glycoprotein N-glycosylation. Environ. Microbiol. **14:**743–753. http://dx.doi.org/10.1111/j.1462-2920.2011.02625.x.
95. **Kaminski L, Guan Z, Yurist-Doutsch S, Eichler J.** 2013. Two distinct N-glycosylation pathways process the *Haloferax volcanii* S-layer glycoprotein upon changes in environmental salinity. mBio **4**(6):e00716–13. http://dx.doi.org/10.1128/mBio.00716-13.
96. **Magidovich H, 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:**122–130. http://dx.doi.org/10.1111/j.1574-6968.2009.01775.x.
97. **Tripepi M, Imam S, Pohlschroder M.** 2010. *Haloferax volcanii* flagella are required for motility but are not involved in PibD-dependent surface adhesion. J. Bacteriol. **192:**3093–3102. http://dx.doi.org/10.1128/JB.00133-10.
98. **Jarrell KF, Albers SV.** 2012. The archaellum: an old motility structure with a new name. Trends Microbiol. **20:**307–312. http://dx.doi.org/10.1016/j.tim.2012.04.007.
99. **Tripepi M, You J, Temel S, Önder Ö, Brisson D, Pohlschröder M.** 2012. N-glycosylation of *Haloferax volcanii* flagellins requires known Agl proteins and is essential for biosynthesis of stable flagella. J. Bacteriol. **194:**4876–4887. http://dx.doi.org/10.1128/JB.00731-12.
100. **Oren A, Ginzburg M, Ginzburg BZ, Hochstein LI, Volcani BE.** 1990. *Haloarcula marismortui* (Volcani) sp. nov., nom. rev., an extremely halophilic bacterium from the Dead Sea. Int. J. Syst. Bacteriol. **40:**209–210. http://dx.doi.org/10.1099/00207713-40-2-209.
101. **Calo D, Guan Z, Naparstek S, Eichler J.** 2011. Different routes to the same ending: comparing the N-glycosylation processes of *Haloferax volcanii* and *Haloarcula marismortui*, two halophilic archaea from the Dead Sea. Mol. Microbiol. **81:**1166–1177. http://dx.doi.org/10.1111/j.1365-2958.2011.07781.x.
102. **Calo D, Eilam Y, Lichtenstein RG, Eichler J.** 2010. Towards glycoengineering in archaea: replacement of *Haloferax volcanii* AglD with homologous glycosyltransferases from other halophilic archaea. Appl. Environ. Microbiol. **76:**5684–5692. http://dx.doi.org/10.1128/AEM.00681-10.
103. **Calo D, Guan Z, Eichler J.** 2011. Glyco-engineering in Archaea: differential N-glycosylation of the S-layer glycoprotein in a transformed *Haloferax volcanii* strain. Microb. Biotechnol. **4:**461–470. http://dx.doi.org/10.1111/j.1751-7915.2011.00250.x.
104. **Jarrell KF, Ding Y, Nair DB, Siu S.** 2013. Surface appendages of Archaea: structure, function, genetics and assembly. Life **3:**86–117. http://dx.doi.org/10.3390/life3010086.
105. **Kelly JF, Jarrell KF.** 2012. Protein glycosylation in the third domain of life: the Archaea, p 108–126. *In* Reid CW, Twine SM, Reid AN (ed), Bacterial glycomics. Caister Academic Press, Cumberland, RI.
106. **Jarrell KF, Jones GM, Nair DB.** 2010. Role of N-linked glycosylation in cell surface structures of Archaea with a focus on flagella and S layers. Int. J. Microbiol. **2010:**470138. http://dx.doi.org/10.1155/2010/470138.
107. **Jarrell KF, Jones GM, Kandiba L, Nair DB, Eichler J.** 2010. S-layer glycoproteins and flagellins: reporters of archaeal posttranslational modifications. Archaea **2010:**612948. http://dx.doi.org/10.1155/2010/612948.
108. **Namboori SC, Graham DE.** 2008. Acetamido sugar biosynthesis in the Euryarchaea. J. Bacteriol. **190:**2987–2996. http://dx.doi.org/10.1128/JB.01970-07.
109. **Namboori SC, Graham DE.** 2008. Enzymatic analysis of uridine diphosphate N-acetyl-D-glucosamine. Anal. Biochem. **301:**94–100. http://dx.doi.org/10.1016/j.ab.2008.06.034.
110. **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. http://dx.doi.org/10.1038/nchembio.1249.
111. **Voisin S, Houliston RS, Kelly J, Brisson JR, Watson D, Bardy SL, Jarrell KF, Logan SM.** 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. http://dx.doi.org/10.1074/jbc.M500329200.
112. **Shams-Eldin H, Chaban B, Niehus S, Schwarz RT, Jarrell KF.** 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. http://dx.doi.org/10.1128/JB.01778-07.

113. **Chaban B, Logan SM, Kelly JF, Jarrell KF.** 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. http://dx.doi.org/10.1128/JB.00885-08.
114. **Kalmokoff ML, Jarrell KF.** 1991. Cloning and sequencing of a multigene family encoding the flagellins of *Methanococcus voltae*. J. Bacteriol. **173:**7113–7125.
115. **Bardy SL, Mori T, Komoriya K, Aizawa S, Jarrell KF.** 2002. Identification and localization of flagellins FlaA and FlaB3 within flagella of *Methanococcus voltae*. J. Bacteriol. **184:**5223–5233. http://dx.doi.org/10.1128/JB.184.19.5223-5233.2002.
116. **Thomas NA, Jarrell KF.** 2001. Characterization of flagellum gene families of methanogenic archaea and localization of novel flagellum accessory proteins. J. Bacteriol. **183:**7154–7164. http://dx.doi.org/10.1128/JB.183.24.7154-7164.2001.
117. **Price NP, Momany FA.** 2005. Modeling bacterial UDP-HexNAc: polyprenol-P HexNAc-1-P transferases. Glycobiology **15:**29R–42R. http://dx.doi.org/10.1093/glycob/cwi065.
118. **Meyer BH, Albers SV.** 2013. Hot and sweet: protein glycosylation in Crenarchaeota. Biochem. Soc. Trans. **41:**384–392. http://dx.doi.org/10.1042/BST20120296.
119. **Vandyke DJ, Wu J, Logan SM, Kelly JF, Mizuno S, Aizawa SI, Jarrell KF.** 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. http://dx.doi.org/10.1111/j.1365-2958.2009.06671.x.
120. **Kelly J, Logan SM, Jarrell KF, Vandyke DJ, Vinogradov E.** 2009. A novel N-linked flagellar glycan from *Methanococcus maripaludis*. Carbohydr. Res. **344:**648–653. http://dx.doi.org/10.1016/j.carres.2009.01.006.
121. **Ng SYM, Wu J, Nair DB, Logan SM, Robotham A, Tessier L, Kelly JF, Uchida K, Aizawa S, Jarrell KF.** 2011. Genetic and mass spectrometry analysis of the unusual type IV-like pili of the archaeon *Methanococcus maripaludis*. J. Bacteriol. **193:**804–814. http://dx.doi.org/10.1128/JB.00822-10.
122. **Pohlschroder M, Gimenez MI, Jarrell KF.** 2005. Protein transport in Archaea: Sec and twin arginine translocation pathways. Curr. Opin. Microbiol. **8:**713–719. http://dx.doi.org/10.1016/j.mib.2005.10.006.
123. **DiMarco AA, Bobik TA, Wolfe RS.** 1990. Unusual coenzymes of methanogenesis. Annu. Rev. Biochem. **59:**355–394. http://dx.doi.org/10.1146/annurev.bi.59.070190.002035.
124. **Sauer FD, Blackwell BA, Kramer JK, Marsden BJ.** 1990. Structure of a novel cofactor containing N-(7-mercaptoheptanoyl)-O-3-phosphothreonine. Biochemistry **29:**7593–7600. http://dx.doi.org/10.1021/bi00485a008.
125. **Jones GM, Wu J, Ding Y, Uchida K, Aizawa S, Robotham A, Logan SM, Kelly J, Jarrell KF.** 2012. Identification of genes involved in the acetamidino group modification of the flagellin N-linked glycan of *Methanococcus maripaludis*. J. Bacteriol. **194:**2693–2702. http://dx.doi.org/10.1128/JB.06686-11.
126. **Feng L, Senchenkova SN, Tao J, Shashkov AS, Liu B, Shevelev SD, Reeves PR, Xu J, Knirel YA, Wang L.** 2005. Structural and genetic characterization of enterohemorrhagic *Escherichia coli* O145 O antigen and development of an O145 serogroup-specific PCR assay. J. Bacteriol. **187:**758–764. http://dx.doi.org/10.1128/JB.187.2.758-764.2005.
127. **Ding Y, Jones GM, Uchida K, Aizawa SI, Robotham A, Logan SM, Kelly J, Jarrell KF.** 2013. Identification of genes involved in the biosynthesis of the third and fourth sugars of the *Methanococcus maripaludis* archaellin N-linked tetrasaccharide. J. Bacteriol. **195:**4094–4104. http://dx.doi.org/10.1128/JB.00668-13.
128. **Amor PA, Yethon JA, Monteiro MA, Whitfield C.** 1999. Assembly of the K40 antigen in *Escherichia coli*: identification of a novel enzyme responsible for addition of L-serine residues to the glycan backbone and its requirement for K40 polymerization. J. Bacteriol. **181:**772–780.
129. **Menon I, Huber T, Sanyal S, Banerjee S, Barré P, Canis S, Warren JD, Hwa J, Sakmar TP, Menon AK.** 2011. Opsin is a phospholipid flippase. Curr. Biol. **21:**149–153. http://dx.doi.org/10.1016/j.cub.2010.12.031.
130. **Jeffery CJ.** 2009. Moonlighting proteins—an update. Mol. Biosyst. **5:**345–350. http://dx.doi.org/10.1039/b900658n.
131. **Larkin A, Olivier NB, Imperiali B.** 2010. Structural analysis of WbpE from *Pseudomonas aeruginosa* PAO1: a nucleotide sugar aminotransferase involved in O-antigen assembly. Biochemistry **49:**7227–7237. http://dx.doi.org/10.1021/bi100805b.
132. **Keswani J, Orkand S, Premachandran U, Mandelco L, Franklin MJ, Whitman WB.** 1996. Phylogeny and taxonomy of mesophilic *Methanococcus* spp. and comparison of rRNA, DNA hybridization, and phenotypic methods. Int. J. Syst. Bacteriol. **46:**727–735. http://dx.doi.org/10.1099/00207713-46-3-727.
133. **Bayley DP, Kalmokoff ML, Jarrell KF.** 1993. Effect of bacitracin on flagellar assembly and presumed glycosylation of the flagellins of *Methanococcus deltae*. Arch. Microbiol. **160:**179–185.
134. **Siewert G, Strominger JL.** 1967. Bacitracin: an inhibitor of the dephosphorylation of lipid pyrophosphate, an intermediate in the biosynthesis of the peptidoglycan of bacterial cell walls. Proc. Natl. Acad. Sci. U. S. A. **57:**767–773. http://dx.doi.org/10.1073/pnas.57.3.767.
135. **Sumper M.** 1987. Halobacterial glycoprotein biosynthesis. Biochim. Biophys. Acta **906:**69–79. http://dx.doi.org/10.1016/0304-4157(87)90005-0.
136. **Forterre P, Brochier C, Philippe H.** 2002. Evolution of the Archaea. Theor. Popul. Biol. **61:**409–422. http://dx.doi.org/10.1006/tpbi.2002.1592.
137. **Peyfoon E, Meyer B, Hitchen PG, Panico M, Morris HR, Haslam SM, Albers SV, Dell A.** 2010. The S-layer glycoprotein of the crenarchaeote *Sulfolobus acidocaldarius* is glycosylated at multiple sites with the chitobiose-linked N-glycans. Archaea **2010:**754101. http://dx.doi.org/10.1155/2010/754101.
138. **Zahringer U, Moll H, Hettmann T, Knirel YA, Schafer G.** 2000. Cytochrome b558/566 from the archaeon *Sulfolobus acidocaldarius* has a unique Asn-linked highly branched hexasaccharide chain containing 6-sulfoquinovose. Eur. J. Biochem. **267:**4144–4149. http://dx.doi.org/10.1046/j.1432-1327.2000.01446.x.
139. **Karcher U, Schroder H, Haslinger E, Allmaier G, Schreiner R, Wieland F, Haselbeck A, Konig H.** 1993. Primary structure of the heterosaccharide of the surface glycoprotein of *Methanothermus fervidus*. J. Biol. Chem. **268:**26821–26826.
140. **Wagner M, van Wolferen M, Wagner A, Lassak K, Meyer B, Reimann J, Albers SV.** 2012. Versatile genetic toolbox for the crenarchaeote *Sulfolobus acidocaldarius*. Front. Microbiol. **3:**214. http://dx.doi.org/10.3389/fmicb.2012.00214.
141. **McLachlan KR, Krag SS.** 1992. Substrate specificity of N-acetylglucosamine 1-phosphate transferase activity in Chinese hamster ovary cells. Glycobiology **2:**313–319. http://dx.doi.org/10.1093/glycob/2.4.313.
142. **Lehrman MA.** 1991. Biosynthesis of N-acetylglucosamine-P-P-dolichol, the committed step of asparagine-linked oligosaccharide assembly. Glycobiology **1:**553–562. http://dx.doi.org/10.1093/glycob/1.6.553.
143. **Yamashita S, Hemmi H, Ikeda Y, Nakayama T, Nishino T.** 2004. Type 2 isopentenyl diphosphate isomerase from a thermoacidophilic archaeon *Sulfolobus shibatae*. Eur. J. Biochem. **271:**1087–1093. http://dx.doi.org/10.1111/j.1432-1033.2004.04010.x.
144. **Nakatani H, Goda S, Unno H, Nagai T, Yoshimura T, Hemmi H.** 2012. Substrate-induced change in the quaternary structure of type 2 isopentenyl diphosphate isomerase from *Sulfolobus shibatae*. J. Bacteriol. **194:**3216–3224. http://dx.doi.org/10.1128/JB.00068-12.
145. **Ohnuma S, Hemmi H, Koyama T, Ogura K, Nishino T.** 1998. Recognition of allylic substrates in *Sulfolobus acidocaldarius* geranylgeranyl diphosphate synthase: analysis using mutated enzymes and artificial allylic substrates. J. Biochem. **123:**1036–1040. http://dx.doi.org/10.1093/oxfordjournals.jbchem.a022040.
146. **Ohnuma S, Hirooka K, Hemmi H, Ishida C, Ohto C, Nishino T.** 1996. Conversion of product specificity of archaebacterial geranylgeranyl-diphosphate synthase. Identification of essential amino acid residues for chain length determination of prenyltransferase reaction. J. Biol. Chem. **271:**18831–18837.
147. **Guan Z, Eichler J.** 2011. Liquid chromatography/tandem mass spectrometry of dolichols and polyprenols, lipid sugar carriers across evolution. Biochim. Biophys. Acta **1811:**800–806. http://dx.doi.org/10.1016/j.bbalip.2011.04.009.
148. **Anderson MS, Eveland SS, Price NP.** 2000. Conserved cytoplasmic motifs that distinguish sub-groups of the polyprenol phosphate:N-acetylhexosamine-1-phosphate transferase family. FEMS Microbiol. Lett. **191:**169–175. http://dx.doi.org/10.1111/j.1574-6968.2000.tb09335.x.
149. **Lehrer J, Vigeant KA, Tatar LD, Valvano MA.** 2007. Functional characterization and membrane topology of *Escherichia coli* WecA, a sugar-phosphate transferase initiating the biosynthesis of enterobacterial com-

mon antigen and O-antigen lipopolysaccharide. J. Bacteriol. **189:**2618–2628. http://dx.doi.org/10.1128/JB.01905-06.

150. **Heinz E.** 1993. Recent investigations on the biosynthesis of the plant sulfolipid, p 163–178. *In* De Kok LJ (ed), Sulfur nutrition and assimilation in higher plants. SPB Academic Publishers, The Hague, The Netherlands.
151. **Benning C.** 1998. Membrane lipids in anoxygenic bacteria, p 83–101. *In* Siegenthaler PA, Murata N (ed), Lipids in photosynthesis: structure, function and genetics. Kluwer Academic Publishers, Boston, MA.
152. **Benning C, Beatty JT, Prince RC, Somerville CR.** 1993. The sulfolipid sulfoquinovosyldiacylglycerol is not required for photosynthetic electron transport in *Rhodobacter sphaeroides* but enhances growth under phosphate limitation. Proc. Natl. Acad. Sci. U. S. A. **90:**1561–1565. http://dx.doi.org/10.1073/pnas.90.4.1561.
153. **Benning C, Somerville CR.** 1992. Identification of an operon involved in sulfolipid biosynthesis in *Rhodobacter sphaeroides*. J. Bacteriol. **174:**6479–6487.
154. **Benning C, Somerville CR.** 1992. Isolation and genetic complementation of a sulfolipid-deficient mutant of *Rhodobacter sphaeroides*. J. Bacteriol. **147:**2352–2360.
155. **Meyer BH, Zolghadr B, Peyfoon E, Pabst M, Panico M, Morris HR, Haslam SM, Messner P, Schäffer C, Dell A, Albers SV.** 2011. Sulfoquinovose synthase—an important enzyme in the N-glycosylation pathway of *Sulfolobus acidocaldarius*. Mol. Microbiol. **82:**1150–1163. http://dx.doi.org/10.1111/j.1365-2958.2011.07875.x.
156. **Meyer BH, Peyfoon E, Dietrich C, Hitchen P, Dell A, Albers SV.** 2013. Agl16, a thermophilic glycosyltransferase, mediating the last step of the N-glycan biosynthesis in the thermoacidophilic crenarchaeon *Sulfolobus acidocaldarius*. J. Bacteriol. **195:**2177–2186. http://dx.doi.org/10.1128/JB.00035-13.
157. **Guan Z, Meyer BH, Albers SV, Eichler J.** 2011. The thermoacidophilic archaeon *Sulfolobus acidocaldarius* contains an unusually short, highly reduced dolichyl phosphate. Biochim. Biophys. Acta **1811:**607–616. http://dx.doi.org/10.1016/j.bbalip.2011.06.022.
158. **Southam G, Kalmokoff ML, Jarrell KF, Koval SF, Beveridge TJ.** 1990. Isolation, characterization, and cellular insertion of the flagella from two strains of the archaebacterium *Methanospirillum hungatei*. J. Bacteriol. **172:**3221–3228.
159. **Kalmokoff ML, Koval SF, Jarrell KF.** 1992. Relatedness of the flagellins from methanogens. Arch. Microbiol. **157:**481–487.
160. **Francoleon DR, Boontheung P, Yang Y, Kin U, Ytterberg AJ, Denny PA, Denny PC, Loo JA, Gunsalus RP, Loo RR.** 2009. S-layer, surface-accessible, and concanavalin A binding proteins of *Methanosarcina acetivorans* and *Methanosarcina mazei*. J. Proteome Res. **8:**1972–1982. http://dx.doi.org/10.1021/pr800923e.
161. **Rohlin L, Leon DR, Kim U, Loo JA, Ogorzalek-Loo RR, Gunsalus RP.** 2012. Identification of the major expressed S-layer and cell surface-layer-related proteins in the model methanogenic archaea: *Methanosarcina barkeri* Fusaro and *Methanosarcina acetivorans* C2A. Archaea **2012:**873589. http://dx.doi.org/10.1155/2012/873589.
162. **Faguy DM, Bayley DP, Kostyukova AS, Thomas NA, Jarrell KF.** 1996. Isolation and characterization of flagella and flagellin proteins from the thermoacidophilic archaea *Thermoplasma volcanium* and *Sulfolobus shibatae*. J. Bacteriol. **178:**902–905.
163. **Sergenova IS, Polosina YY, Kostyukova AS, Metlina AL, Pyatibratov MG, Fedorov OV.** 1995. Flagella of halophilic archaea: biochemical and genetic analysis. Biochemistry (Mosc.) **60:**953–957.
164. **Peters J, Nitsch M, Kuhlmorgen B, Golbik R, Lupas A, Kellermann J, Engelhardt H, Pfander JP, Muller S, Goldie K.** 1995. Tetrabrachion: a filamentous archaebacterial surface protein assembly of unusual structure and extreme stability. J. Mol. Biol. **245:**385–401. http://dx.doi.org/10.1006/jmbi.1994.0032.
165. **Kamekura M, Dyall-Smith M, Upasani V, Ventosa A, Kates M.** 1997. Diversity of alkaliphilic halobacteria: proposals for transfer of *Natronobacterium vacuolatum*, *Natronobacterium magadii*, and *Natronobacterium pharaonis* to *Halorubrum*, *Natrialba*, and *Natronomonas* gen. nov., respectively, as *Halorubrum vacuolatum* comb. nov., *Natrialba magadii* comb. nov., and *Natronomonas pharaonis* comb. nov., respectively. Int. J. Syst. Bacteriol. **47:**853–857. http://dx.doi.org/10.1099/00207713-47-3-853.
166. **Kim BK, Pihl TD, Reeve JN, Daniels L.** 1995. Purification of the copper response extracellular proteins secreted by the copper-resistant methanogen *Methanobacterium bryantii* BKYH and cloning, sequencing, and transcription of the gene encoding these proteins. J. Bacteriol. **177:**7178–7185.
167. **Claus H, Akça E, Debaerdemaeker T, Evrard C, Declercq JP, König H.** 2002. Primary structure of selected archaeal mesophilic and extremely thermophilic outer surface layer proteins. Syst. Appl. Microbiol. **25:**3–12. http://dx.doi.org/10.1078/0723-2020-00100.
168. **Lechner J, Wieland F.** 1989. Structure and biosynthesis of prokaryotic glycoproteins. Annu. Rev. Biochem. **58:**173–194. http://dx.doi.org/10.1146/annurev.bi.58.070189.001133.
169. **Wieland F.** 1988. Structure and biosynthesis of prokaryotic glycoproteins. Biochimie. **70:**1493–1504. http://dx.doi.org/10.1016/0300-9084(88)90286-6.
170. Reference deleted.
171. **Paul G, Wieland F.** 1987. Sequence of the halobacterial glycosaminoglycan. J. Biol. Chem. **262:**9587–9593.
172. **Paul G, Lottspeich F, Wieland F.** 1986. Asparaginyl-*N*-acetylgalactosamine: linkage unit of halobacterial glycosaminoglycan. J. Biol. Chem. **261:**1020–1024.
173. **Lechner J, Wieland F, Sumper M.** 1985. Transient methylation of dolichyl oligosaccharides is an obligatory step in halobacterial sulfated glycoprotein biosynthesis. J. Biol. Chem. **260:**8984–8989.
174. **Raetz CR, Whitfield C.** 2002. Lipopolysaccharide endotoxins. Annu. Rev. Biochem. **71:**635–700. http://dx.doi.org/10.1146/annurev.biochem.71.110601.135414.
175. **Lechner J, Sumper M.** 1987. The primary structure of a procaryotic glycoprotein. Cloning and sequencing of the cell surface glycoprotein gene of halobacteria. J. Biol. Chem. **262:**9724–9729.
176. **Mengele R, Sumper M.** 1992. Drastic differences in glycosylation of related S-layer glycoproteins from moderate and extreme halophiles. J. Biol. Chem. **267:**8182–8185.
177. **Jervis AJ, Langdon R, Hitchen P, Lawson AJ, Wood A, Fothergill JL, Morris HR, Dell A, Wren B, Linton D.** 2010. Characterization of N-linked protein glycosylation in *Helicobacter pullorum*. J. Bacteriol. **192:**5228–5236. http://dx.doi.org/10.1128/JB.00211-10.
178. **Nair DB, Uchida K, Aizawa SI, Jarrell KF.** 2014. Genetic analysis of a type IV pili-like locus in the archaeon *Methanococcus mariplaudis*. Arch. Microbiol. **196:**179–191. http://dx.doi.org/10.1007/s00203-014-0956-4.
179. **Hartmann E, Konig H.** 1989. Uridine and dolichyl diphosphate activated oligosaccharides are intermediates in the biosynthesis of the S-layer glycoprotein of *Methanothermus fervidus*. Arch. Microbiol. **151:**274–281. http://dx.doi.org/10.1007/BF00413142.
180. **Pellerin P, Fournet B, Debeire P.** 1990. Evidence for the glycoprotein nature of the cell sheath of *Methanosaeta*-like cells in the culture of *Methanothrix soehngenii* strain FE. Can. J. Microbiol. **39:**631–636.
181. **Hettmann T, Schmidt CL, Anemüller S, Zähringer U, Moll H, Petersen A, Schäfer G.** 1998. Cytochrome $b_{558/566}$ from the archaeon *Sulfolobus acidocaldarius*: a novel highly glycosylated, membrane-bound B-type hemoprotein. J. Biol. Chem. **273:**12032–12040. http://dx.doi.org/10.1074/jbc.273.20.12032.
182. **Palmieri G, Balestrieri M, Peter-Katalinić J, Pohlentz G, Rossi M, Fiume I, Pocsfalvi G.** 2013. Surface-exposed glycoproteins of hyperthermophilic *Sulfolobus solfataricus* P2 show a common N-glycosylation profile. J. Proteome Res. **12:**2779–2790. http://dx.doi.org/10.1021/pr400123z.
183. **Vinogradov E, Deschatelets L, Lamoureux M, Patel GB, Tremblay TL, Robotham A, Goneau MF, Cummings-Lorbetskie C, Watson DC, Brisson JR, Kelly JF, Gilbert M.** 2012. Cell surface glycoproteins from *Thermoplasma acidophilum* are modified with an N-linked glycan containing 6-C-sulfofucose. Glycobiology **22:**1256–1267. http://dx.doi.org/10.1093/glycob/cws094.
184. **Yang LL, Haug A.** 1979. Purification and partial characterization of a procaryotic glycoprotein from the plasma membrane of *Thermoplasma acidophilum*. Biochim. Biophys. Acta **556:**265–277. http://dx.doi.org/10.1016/0005-2736(79)90047-6.
185. **Matsumoto S, Igura M, Nyirenda J, Matsumoto M, Yuzawa S, Noda N, Inagaki F, Kohda D.** 2012. Crystal structure of the C-terminal globular domain of oligosaccharyltransferase from *Archaeoglobus fulgidus* at 1.75 Å resolution. Biochemistry **51:**4157–4166. http://dx.doi.org/10.1021/bi300076u.
186. **Kohda D, Yamada M, Igura M, Kamishikiryo J, Maenaka K.** 2007. New oligosaccharyltransferase assay method. Glycobiology **17:**1175–1182. http://dx.doi.org/10.1093/glycob/cwm087.

187. **Kandiba L, Aitio O, Helin J, Guan Z, Permi P, Bamford DH, Eichler J, Roine E.** 2012. Diversity in prokaryotic glycosylation: an archaeal-derived N-linked glycan contains legionaminic acid. Mol. Microbiol. **84:** 578–593. http://dx.doi.org/10.1111/j.1365-2958.2012.08045.x.
188. **McNally DJ, Aubry AJ, Hui JP, Khieu NH, Whitfield D, Ewing CP, Guerry P, Brisson JR, Logan SM, Soo EC.** 2007. Targeted metabolomics analysis of *Campylobacter coli* VC167 reveals legionaminic acid derivatives as novel flagellar glycans. J. Biol. Chem. **282:**14463–14475. http://dx.doi.org/10.1074/jbc.M611027200.
189. **Kandiba L, Eichler J.** 2013. Analysis of putative nonulosonic acid biosynthesis pathways in Archaea reveals a complex evolutionary history. FEMS Microbiol. Lett. **345:**110–120. http://dx.doi.org/10.1111/1574-6968.12193.
190. **Mohorko E, Glockshuber R, Aebi M.** 2011. Oligosaccharyltransferase: the central enzyme of N-linked protein glycosylation. J. Inherit. Metab. Dis. **34:**869–878. http://dx.doi.org/10.1007/s10545-011-9337-1.
191. **Igura M, Kohda D.** 2011. Quantitative assessment of the preferences for the amino acid residues flanking archaeal N-linked glycosylation sites. Glycobiology **21:**575–583. http://dx.doi.org/10.1093/glycob/cwq196.
192. **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. http://dx.doi.org/10.1093/glycob/cwn118.
193. **Maita N, Nyirenda J, Igura M, Kamishikiryo J, Kohda D.** 2010. Comparative structural biology of eubacterial and archaeal oligosaccharyltransferases. J. Biol. Chem. **285:**4941–4950. http://dx.doi.org/10.1074/jbc.M109.081752.
194. **Yan Q, Lennarz WJ.** 2002. Studies on the function of oligosaccharyl transferase subunits. Stt3p is directly involved in the glycosylation process. J. Biol. Chem. **277:**47692–47700. http://dx.doi.org/10.1074/jbc.M208136200.
195. **Lizak C, Gerber S, Numao S, Aebi M, Locher KP.** 2011. X-ray structure of a bacterial oligosaccharyltransferase. Nature **474:**350–355. http://dx.doi.org/10.1038/nature10151.
196. **Li L, Woodward R, Ding Y, Liu XW, Yi W, Bhatt VS, Chen M, Zhang LW, Wang PG.** 2010. Overexpression and topology of bacterial oligosaccharyltransferase PglB. Biochem. Biophys. Res. Commun. **394:**1069–1074. http://dx.doi.org/10.1016/j.bbrc.2010.03.126.
197. **Kim H, von Heijne G, Nilsson I.** 2005. Membrane topology of the STT3 subunit of the oligosaccharyl transferase complex. J. Biol. Chem. **280:** 20261–20267. http://dx.doi.org/10.1074/jbc.M412213200.
198. **Matsumoto S, Shimada A, Nyirenda J, Igura M, Kawano Y, Kohda D.** 2013. Crystal structures of an archaeal oligosaccharyltransferase provides insights into the catalytic cycle of N-linked protein glycosylation. Proc. Natl. Acad. Sci. U. S. A. **110:**17868–17873. http://dx.doi.org/10.1073/pnas.1309777110.
199. **Jaffee MB, Imperiali B.** 2011. Exploiting topological constraints to reveal buried sequence motifs in the membrane-bound N-linked oligosaccharyl transferases. Biochemistry **50:**7557–7567. http://dx.doi.org/10.1021/bi201018d.
200. **Izquierdo L, Schulz BL, Rodrigues JA, Güther ML, Procter JB, Barton GJ, Aebi M, Ferguson MA.** 2009. Distinct donor and acceptor specificities of *Trypanosoma brucei* oligosaccharyltransferases. EMBO J. **28:** 2650–2661. http://dx.doi.org/10.1038/emboj.2009.203.
201. **Igura M, Kohda D.** 2011. Selective control of oligosaccharide transfer efficiency for the N-glycosylation sequon by a point mutation in oligosaccharyltransferase. J. Biol. Chem. **286:**13255–13260. http://dx.doi.org/10.1074/jbc.M110.213900.
202. **Jaffee MB, Imperiali B.** 2013. Optimized protocol for expression and purification of membrane-bound PglB, a bacterial oligosaccharyl transferase. Protein Expr. Purif. **89:**241–250. http://dx.doi.org/10.1016/j.pep.2013.04.001.
203. **Nyirenda J, Matsumoto S, Saitoh T, Maita N, Noda NN, Inagaki F, Kohda D.** 2013. Crystallographic and NMR evidence for flexibility in oligosaccharyltransferases and its catalytic significance. Structure **21:**32–41. http://dx.doi.org/10.1016/j.str.2012.10.011.
204. **Igura M, Maita N, Obita T, Kamishikiryo J, Maenaka K, Kohda D.** 2007. Purification, crystallization and preliminary X-ray diffraction studies of the soluble domain of the oligosaccharyltransferase STT3 subunit from the thermophilic archaeon *Pyrococcus furiosus*. Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun. **63:**798–801. http://dx.doi.org/10.1107/S1744309107040134.
205. **Chaban B, Ng SY, Jarrell KF.** 2006. Archaeal habitats—from the extreme to the ordinary. Can. J. Microbiol. **52:**73–116. http://dx.doi.org/10.1139/w05-147.
206. **Matsumoto S, Shimada A, Kohda D.** 2013. Crystal structure of the C-terminal globular domain of the third paralog of the *Archaeoglobus fulgidus* oligosaccharyltransferases. BMC Struct. Biol. **13:**11. http://dx.doi.org/10.1186/1472-6807-13-11.
207. **Glover KJ, Weerapana E, Numao S, Imperiali B.** 2005. Chemoenzymatic synthesis of glycopeptides with PglB, a bacterial oligosaccharyl transferase from *Campylobacter jejuni*. Chem. Biol. **12:**1311–1315. http://dx.doi.org/10.1016/j.chembiol.2005.10.004.
208. **Chavan M, Rekowicz M, Lennarz W.** 2003. Insight into functional aspects of Stt3p, a subunit of the oligosaccharyl transferase. Evidence for interaction of the N-terminal domain of Stt3p with the protein kinase C cascade. J. Biol. Chem. **278:**51441–51447. http://dx.doi.org/10.1074/jbc.M310456200.
209. **Wacker M, Feldman MF, Callewaert N, Kowarik M, Clarke BR, Pohl NL, Hernandez M, Vines ED, Valvano MA, Whitfield C, 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:**7088–7093. http://dx.doi.org/10.1073/pnas.0509207103.
210. **Kowarik M, Young NM, Numao S, Schulz BL, Hug I, Callewaert N, Mills DC, Watson DC, Hernandez M, Kelly JF, Wacker M, Aebi M.** 2006. Definition of the bacterial N-glycosylation site consensus sequence. EMBO J. **25:**1957–1966. http://dx.doi.org/10.1038/sj.emboj.7601087.
211. **Brockl G, Behr M, Fabry S, Hensel R, Kaudewitz H, Biendl E, Konig H.** 1991. Analysis and nucleotide sequence of the genes encoding the surface-layer glycoproteins of the hyperthermophilic methanogens *Methanothermus fervidus* and *Methanothermus sociabilis*. Eur. J. Biochem. **199:**147–152. http://dx.doi.org/10.1111/j.1432-1033.1991.tb16102.x.
212. **Bardy SL, Jarrell KF.** 2003. Cleavage of preflagellins by an aspartic acid signal peptidase is essential for flagellation in the archaeon *Methanococcus voltae*. Mol. Microbiol. **50:**1339–1347. http://dx.doi.org/10.1046/j.1365-2958.2003.03758.x.
213. **Valliere-Douglass JF, Eakin CM, Wallace A, Ketchem RR, Wang W, Treuheit MJ, Balland A.** 2010. Glutamine-linked and non-consensus asparagine-linked oligosaccharides present in human recombinant antibodies define novel protein glycosylation motifs. J. Biol. Chem. **285:** 16012–16022. http://dx.doi.org/10.1074/jbc.M109.096412.
214. **Schwarz F, Lizak C, Fan YY, Fleurkens S, Kowarik M, Aebi M.** 2011. Relaxed acceptor site specificity of bacterial oligosaccharyltransferase in vivo. Glycobiology **21:**45–54. http://dx.doi.org/10.1093/glycob/cwq130.
215. **Zeitler R, Hochmuth E, Deutzmann R, Sumper M.** 1998. Exchange of Ser-4 for Val, Leu or Asn in the sequon Asn-Ala-Ser does not prevent N-glycosylation of the cell surface glycoprotein from *Halobacterium halobium*. Glycobiology **8:**1157–1164.
216. **Tai VW, Imperiali B.** 2001. Substrate specificity of the glycosyl donor for oligosaccharyl transferase. J. Org. Chem. **66:**6217–6228. http://dx.doi.org/10.1021/jo0100345.
217. **Feldman MF, Wacker M, Hernandez M, Hitchen PG, Marolda CL, Kowarik M, Morris HR, Dell A, Valvano MA, 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. http://dx.doi.org/10.1073/pnas.0500044102.
218. **Schwarz F, Fan YY, Schubert M, Aebi M.** 2011. Cytoplasmic N-glycosyltransferase of *Actinobacillus pleuropneumoniae* is an inverting enzyme and recognizes the NX(S/T) consensus sequence. J. Biol. Chem. **286:**35267–35274. http://dx.doi.org/10.1074/jbc.M111.277160.
219. **Wieland F, Heitzer R, Schaefer W.** 1983. Asparaginylglucose: novel type of carbohydrate linkage. Proc. Natl. Acad. Sci. U. S. A. **80:**5470–5474. http://dx.doi.org/10.1073/pnas.80.18.5470.
220. **Wieland F, Paul G, Sumper M.** 1985. Halobacterial flagellins are sulfated glycoproteins. J. Biol. Chem. **260:**15180–15185.
221. **Meyer W, Schäfer G.** 1992. Characterization and purification of a membrane-bound archaebacterial pyrophosphatase from *Sulfolobus acidocaldarius*. Eur. J. Biochem. **207:**741–746. http://dx.doi.org/10.1111/j.1432-1033.1992.tb17104.x.
222. **Kaminski L, Naparstek S, Kandiba L, Cohen-Rosenzweig C, Arbiv A, Konrad Z, Eichler J.** 2013. Add salt, add sugar: N-glycosylation in *Haloferax volcanii*. Biochem. Soc. Trans. **41:**432–435. http://dx.doi.org/10.1042/BST20120142.
223. **Deatherage JF, Taylor KA, Amos LA.** 1983. Three-dimensional

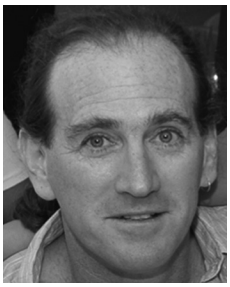
arrangement of the cell wall protein of *Sulfolobus acidocaldarius*. J. Mol. Biol. **167:**823–848. http://dx.doi.org/10.1016/S0022-2836(83)80113-2.
224. **Albers SV, Meyer BH.** 2011. The archaeal cell envelope. Nat. Rev. Microbiol. **9:**414–426. http://dx.doi.org/10.1038/nrmicro2576.
225. **Sleytr UB, Beveridge TJ.** 1999. Bacterial S-layers. Trends Microbiol. **7:**253–260. http://dx.doi.org/10.1016/S0966-842X(99)01513-9.
226. **Claus H, Akca E, Debaerdemaeker T, Evrard C, Declercq JP, Harris JR, Schlott B, König H.** 2005. Molecular organization of selected prokaryotic S-layer proteins. Can. J. Microbiol. **51:**731–743. http://dx.doi.org/10.1139/w05-093.
227. **Engelhardt H, Peters J.** 1998. Structural research on surface layers: a focus on stability, surface layer homology domains, and surface layer-cell wall interactions. J. Struct. Biol. **124:**276–302. http://dx.doi.org/10.1006/jsbi.1998.4070.
228. **Akca E, Claus H, Schultz N, Karbach G, Schlott B, Debaerdemaeker T, Declercq JP, Konig H.** 2002. Genes and derived amino acid sequences of S-layer proteins from mesophilic, thermophilic, and extremely thermophilic methanococci. Extremophiles **6:**351–358. http://dx.doi.org/10.1007/s00792-001-0264-1.
229. **VanDyke DJ, Wu J, Ng SY, Kanbe M, Chaban B, Aizawa SI, Jarrell KF.** 2008. Identification of putative acetyltransferase gene, MMP0350, which affects proper assembly of both flagella and pili in the archaeon *Methanococcus maripaludis*. J. Bacteriol. **190:**5300–5307. http://dx.doi.org/10.1128/JB.00474-08.
230. **Jarrell KF, Stark M, Nair DB, Chong JPJ.** 2011. Flagella and pili are both necessary for efficient attachment of *Methanococcus maripaludis* to surfaces. FEMS Microbiol. Lett. **319:**44–50. http://dx.doi.org/10.1111/j.1574-6968.2011.02264.x.
231. **Terra VS, Mills DC, Yates LE, Abouelhadid S, Cuccui J, Wren BW.** 2012. Recent developments in bacterial protein glycan coupling technology and glycoconjugate vaccine design. J. Med. Microbiol. **61:**919–926. http://dx.doi.org/10.1099/jmm.0.039438-0.
232. **Jervis AJ, Butler JA, Lawson AJ, Langdon R, Wren BW, Linton D.** 2012. Characterization of the structurally diverse N-linked glycans of *Campylobacter* species. J. Bacteriol. **194:**2355–2362. http://dx.doi.org/10.1128/JB.00042-12.
233. **Valderrama-Rincon JD, Fisher AC, Merritt JH, Fan YY, Reading CA, Chhiba K, Heiss C, Azadi P, Aebi M, DeLisa MP.** 2012. An engineered eukaryotic protein glycosylation pathway in *Escherichia coli*. Nat. Chem. Biol. **8:**434–436. http://dx.doi.org/10.1038/nchembio.921.
234. **Leigh JA, Albers SV, Atomi H, Allers T.** 2011. Model organisms for genetics in the domain archaea: methanogens, halophiles, Thermococcales and Sulfolobales. FEMS Microbiol. Rev. **35:**577–608. http://dx.doi.org/10.1111/j.1574-6976.2011.00265.x.

**Ken F. Jarrell** received his Ph.D. from Queen's University, Kingston, Canada, for work on lipopolysaccharide structure in *Pseudomonas aeruginosa* under the supervision of Dr. Andrew Kropinski. Following this, he worked as a Research Associate in the laboratory of Dr. Dennis Sprott at the National Research Council in Ottawa on various aspects of methanogen physiology and structure. He was recruited to the Department of Microbiology and Immunology at Queen's University in 1984 and is now a tenured Full Professor. He has spent two sabbatical leaves as an Alexander von Humboldt Fellow in the laboratory of Professor Albrecht Klein in the Department of Molecular Genetics, Phillips University, Marburg, Germany, and a third sabbatical in the laboratory of Dr. James Chong at the University of York, York, United Kingdom, as a Leverhulme Visiting Professor. He has worked on various aspects of methanogen surface structures, including S-layers, pili, and, mainly, archaella for over 25 years.

**Yan Ding** did her undergraduate degree at Shandong University, China, and graduated with a bachelor of science in 2007. She continued her studies as a graduate student under the supervision of Dr. Peng George Wang and received a master of science in Shandong University, China, in 2010 for work on engineering the length of N-glycans in glycoprotein in an *E. coli* N-glycosylation system. Since 2011, she has been a Ph.D. candidate at Queen's University, Kingston, Ontario, Canada, under the supervision of Dr. Ken F. Jarrell, working on the N-glycosylation pathway involved in archaellation in *Methanococcus maripaludis* while being supported by the China Scholarship Council. In 2013, she won an Ontario Graduate Scholarship as an international student.

**Sonja-Verena Albers** studied biology at the University of Würzburg, Germany. She received her Ph.D. from the University of Groningen, The Netherlands, in 2001, studying sugar transport systems in *Sulfolobus solfataricus*. After that, she received a VENI and VIDI grant from the Dutch Science Organization (NWO) to study protein secretion and appendage assembly in *S. solfataricus* at the University of Groningen, The Netherlands. In 2008, she became a Max Planck Research group leader at the Max Planck Institute for Terrestrial Microbiology, Marburg, Germany. Her group, Molecular Biology of Archaea, works on the assembly and function of archaeal surface structures, their role in biofilm formation, and their regulation by protein kinase networks. As most extracellular archaeal proteins are glycosylated, it was important to elucidate this pathway in *Sulfolobus* and to understand what role these glycans play in these species.

**Benjamin H. Meyer** studied biology at Goethe University, Frankfurt, Germany, where he started his diploma studies elucidating the biosynthesis of the compatible solute $N^{\epsilon}$-acetyl-β-lysin and S-layer proteins in the euryarchaeon *Methanosarcina mazei* Gö1. After receiving his diploma in 2009, he continued his studies on archaeal S-layer proteins at the Max Planck Institute for Terrestrial Microbiology, Marburg, Germany, where he focused on the N-glycosylation of the S-layer glycoprotein in *Sulfolobus*. He received his Ph.D. from the Philipps Universität, Marburg, Germany, in 2012. He is now a Postdoctoral Fellow in the Sonja-Verena Albers research group.



**Jerry Eichler** received his B.Sc. (Physiology, 1982) from McGill University in Montreal, Canada, and his M.Sc. (Neurobiology, 1988) and Ph.D. (Neurobiology, 1994) from the Weizmann Institute of Science in Rehovot, Israel. He was a Postdoctoral Fellow (1994 to 1998) in the Department of Biochemistry at the Dartmouth Medical School in Hanover, NH. He is now a Professor and Chair of the Department of Life Sciences at Ben Gurion University in Beersheva, Israel, where he accepted a position in 1998. Since 2005, his research has addressed N-glycosylation and other posttranslational modifications in *Archaea*.

**Lina Kaminski** received her degrees (B.Sc. in 2006 and M.Sc. in 2009) from the Department of Life Sciences at Ben Gurion University of the Negev in Beersheva, Israel. She is currently a Ph.D. candidate at the same department, working under the supervision of Professor Jerry Eichler. Her research interests include analysis of Agl proteins involved in the N-glycosylation process in *Haloferax volcanii*.