Characterization of protein glycosylation in an Asgard archaeon

Satoshi Nakagawa [a,b,c,*], Hiroyuki Imachi [b], Shigeru Shimamura [b], Saeko Yanaka [c,d,e], Hirokazu Yagi [c,d], Maho Yagi-Utsumi [c,d,e], Hiroyuki Sakai [f], Shingo Kato [f,g], Moriya Ohkuma [f], Koichi Kato [c,d,e], Ken Takai [b,c]

[a] *Laboratory of Marine Environmental Microbiology, Division of Applied Biosciences, Graduate School of Agriculture, Kyoto University, Oiwake-cho, Kitashirakawa, Sakyo-ku, Kyoto 606-8502, Japan*
[b] *Institute for Extra-cutting-edge Science and Technology Avant-garde Research (X-star), Japan Agency for Marine-Earth Science and Technology (JAMSTEC), 2-15 Natsushima-cho, Yokosuka 273-0061, Japan*
[c] *Exploratory Research Center on Life and Living Systems (ExCELLS), National Institute of Natural Sciences, 5-1 Higashiyama, Myodaiji, Okazaki, Aichi 444-8787, Japan*
[d] *Graduate School of Pharmaceutical Sciences, Nagoya City University, 3-1 Tanabe-dori, Mizuhoku, Nagoya 467-8603, Japan*
[e] *Institute for Molecular Science (IMS), National Institutes of Natural Sciences, 5-1 Higashiyama, Myodaiji, Okazaki, Aichi 444-8787, Japan*
[f] *Japan Collection of Microorganisms (JCM), RIKEN BioResource Research Center, Tsukuba, Ibaraki, Japan*
[g] *Submarine Resources Research Center, JAMSTEC, Yokosuka 273-0061, Japan*

ARTICLE INFO

ABSTRACT

Archaeal cells are typically enveloped by glycosylated S-layer proteins. Archaeal protein glycosylation provides valuable insights not only into their adaptation to their niches but also into their evolutionary trajectory. Notably, thermophilic *Thermoproteota* modify proteins with *N*-glycans that include two GlcNAc units at the reducing end, resembling the "core structure" preserved across eukaryotes. Recently, Asgard archaea, now classified as members of the phylum *Promethearchaeota,* have offered unprecedented opportunities for under-standing the role of archaea in eukaryogenesis. Despite the presence of genes indicative of protein *N*-glycosyl-ation in this archaeal group, these have not been experimentally investigated. Here we performed a glycoproteome analysis of the firstly isolated Asgard archaeon *Promethearchaeum syntrophicum*. Over 700 different proteins were identified through high-resolution LC-MS/MS analysis, however, there was no evidence of either the presence or glycosylation of putative S-layer proteins. Instead, *N*-glycosylation in this archaeon was primarily observed in an extracellular solute-binding protein, possibly related to chemoreception or trans-membrane transport of oligopeptides. The glycan modification occurred on an asparagine residue located within the conserved N-X-S/T sequon, consistent with the pattern found in other archaea, bacteria, and eukaryotes. Unexpectedly, three structurally different *N*-glycans lacking the conventional core structure were identified in this archaeon, presenting unique compositions that included atypical sugars. Notably, one of these sugars was likely HexNAc modified with a threonine residue, similar to modifications previously observed in mesophilic methanogens within the *Methanobacteriati*. Our findings advance our understanding of Asgard archaea physi-ology and evolutionary dynamics.

Introduction

The emergence of eukaryotes, i.e., eukaryogenesis, has been a prominent topic in biology. A key breakthrough in this field was the identification of the Asgard archaea, now classified as the phylum *Promethearchaeota* [1]. This group is phylogenetically the closest known relative to eukaryotes and possesses a large repertoire of eukaryotic signature proteins (ESPs) [2,3]. This finding sheds light on the potential origin of eukaryotes within the archaea domain [2-4]. The successful cultivation of representative Asgard archaea such as *Promethearchaeum syntrophicum* and "*Candidatus* (*Ca.*) Lokiarchaeum ossiferum*" has pro-vided novel insights into the process of eukaryogenesis, particularly through their complex morphological features, including protrusions [5, 6]. The protrusions are supported by a lokiactin cytoskeleton [6], however, the cell surface properties enabling the structural flexibility are not yet clear.

The outermost structure of most archaeal cells is the protein layer, so-called surface (S)-layer [7]. Comprising S-layer proteins and glycoproteins (SLPs), this crystalline lattice envelops the entire cell surface [8] and may account for 10–30% of the total cellular protein content [9, 10]. Although the functions of S-layer are still not fully understood, it is evident that these protective barriers serve critical roles, conferring cell rigidity, regulating cell shape and division, assisting adhesion, and mediating interactions with the environment and neighboring cells or viruses [11–14]. The known functions of S-layer glycosylation are also diverse, and its specific function is still not fully elucidated. In some prokaryotes, glycosylation is associated with S-layer stabilization, cell shape maintenance, adaptation to changing environments, defense against viral infections, shielding against antimicrobial peptides, biofilm formation, creating lubricating hydration layers, and modulation of host immune responses [11,15–17]. Archaeal protein glycosylation can be classified into two primary categories: *N*-glycosylation and *O*-glycosylation. The latter involving two or more sugars is specific to members of the phylum *Halobacteriota* within the *Methanobacteriati* (formerly *Euryarchaeota*) [18–20]. *N*-glycosylation, which involves the attachment of glycans to specific asparagine residues within a protein motif known as the N-X-S/T sequon, is widespread across many archaea. This sequon is conserved in most *N*-glycosylation instances in archaea, bacteria, and eukaryotes examined to date. However, an extended motif was reported in *Campylobacter jejuni*, and within eukaryotes, glycosylation can also occur at the N-X-C motif [21,22]. *N*-glycosylation in archaea is facilitated by a group of genes known as the *agl* genes. Among these, the oligosaccharyl transferase gene (*aglB*) is particularly well-conserved, and its phylogeny suggested that the last common ancestor of eukaryotes and the Asgard archaeal sister lineage might already glycosylate proteins with similar *N*-glycans [23]. In addition to the evolutionary importance, such cell surface glycans often govern symbiotic [24–29] or pathogenic [30–33] relationships between different organisms or even within a single species. Considering the currently known symbiotic associations of Asgard archaea with methanogens [5], and ancestral eukaryotes and later organelle-like cellular components [34], *N*-glycosylation would provide important aspects for a deeper understanding of Asgard archaea evolution, physiology, and eukaryogenesis. However, our understanding of archaeal glycosylation remains nascent, with only a handful of archaea studied to date. Besides members of the kingdom *Methanobacteriati*, only a limited member of terrestrial *Thermoproteota*, i.e. several species of *Sulfolobus* and *Pyrobaculum*, have been characterized glycobiologically [35]. One notable finding in *Thermoproteota* glycosylation is the presence of a unique sugar pattern, two GlcNAc units at the reducing end, which closely resembles that found across eukaryotes [36,37]. This similarity raises expectations that more eukaryote-like glycan structures might be found in Asgard archaea.

Investigating glycosylation in poorly growing microorganisms like Asgard archaea presents notable challenges. Conventional glycobiological tools, like nuclear magnetic resonance (NMR), are often not applicable. However, recent technological advancements in mass spectrometry combined with open modification search methods have enabled the exploration of both glycan and protein profiles in these microorganisms, even from a limited number of cells with novel or atypical glycans [38–40]. While these techniques do not fully decipher complex glycan structures, such as stereoisomers, they facilitate simultaneous amino acid sequence and glycan structural analyses, making them suitable for analyzing both pure cultures and complex microbial communities.

In this study, we focused on the protein glycosylation of Asgard archaea for two primary reasons: i) glycosylation, in particular glycosylation of cell surface proteins is a key factor in diverse cellular interactions, including symbiosis, and ii) investigating glycosylation could reveal evolutionary similarities between eukaryotes and archaea. Our findings will help to shed light on the physiological significance and evolutionary trajectory of protein glycosylation in archaea.

Experimental procedures

Cultivation

P. syntrophicum strain MK-D1^T was cultured in an anaerobic medium as described previously [5,41]. Approximately 1.3×10^9 cells were obtained from a 4.0-liter culture. Cells were harvested at the late-exponential growth phase by centrifugation, washed three times with the anaerobic basal medium, and stored at $-80\,°C$ until use. *Sulfolobus acidocaldarius* strain 98-3^T (=JCM 8929=DSM639), *Saccharolobus solfataricus* strain P1^T (=JCM 8930=DSM1616), and *Pyrodictium abyssi* strain AV2^T (=JCM 9394=DSM6158) were cultivated according to the respective Japan collection of microorganisms (JCM) guidelines and served as controls. Cells of *S. acidocaldarius* and *S. solfataricus* were washed using the same method as for *P. syntrophicum*, while *P. abyssi* cells were washed with sterile 3% (w/v) NaCl. A single biological replicate was performed for each species.

Protein extraction and trypsin digestion

Proteins were extracted as described previously [42]. Briefly, cells were resuspended in 100 mM triethylammonium bicarbonate (TEAB) (pH 8.6) containing 2 mM phenylmethylsulfonyl fluoride (PMSF) and disrupted by sonication at 2 °C. A total of 20 μg of protein pellet was dissolved with MPEX PTS reagent (GL Science, Tokyo, Japan), then reduced with dithiothreitol (final concentration, 23.8 mM) at 95 °C for 5 min followed by 25 min at room temperature. Subsequently, alkylation was performed with iodoacetamide (final concentration, 22.7 mM) in dark at room temperature for 30 min. Protein digestion was conducted in solution A of the MPEX PTS reagent, using trypsin (Thermo Fisher Scientific, Waltham, MA) and lysyl endopeptidase (Lys-C) (Fujifilm Wako, Osaka, Japan) at a trypsin/Lys-C to protein ratio of 1:20. After detergent removal using a spin column (Thermo Fisher Scientific), the peptide samples were applied to the LC-MS/MS. The generally recommended glycopeptide enrichment step was omitted due to the limited cell yield of *P. syntrophicum* and potential artifacts and biases this step could introduce [43,44].

LC-MS/MS analysis

Peptide samples were separated using an Ultimate 3000 RSLCnano system (Thermo Fisher Scientific) with a two-column chromatography setup consisting of an L-column ODS trap (5 μm, 0.3 × 5 mm; CERI, Tokyo, Japan) and a Zaplous alpha Pep-C18 analytical column (3 μm, 120 Å, 0.1 × 150 mm; AMR, Tokyo, Japan), as previously described [45]. Samples were loaded onto the trap column at a flow rate of 6 μl/min using a sample buffer (0.1% TFA, 2% acetonitrile) and then eluted into an Orbitrap Fusion™ Tribrid™ Mass Spectrometer (Thermo Fisher Scientific) at a flow rate of 500 nl/min (Buffer A, 0.1% formic acid) through the analytical column. Analytical runs of 230 min were conducted with a gradual change in buffer composition. Initially, the composition was shifted from 5% to 45% Buffer B (100% acetonitrile) over 225 min, then shifted from 45% to 95% Buffer B in the final 5 min.

The mass spectrometer was operated in positive ion mode. Ionization was facilitated with an electrospray voltage of 1.7 kV, and the ion transfer tube temperature was set to 250 °C. MS spectra were acquired in the Orbitrap mass analyzer (*m/z* range: 350–1800, resolution: 120,000 FWHM with a maximum injection time of 50 ms, AGC 5×10^4) with EASY-IC internal mass calibration. Orbitrap MS/MS higher-energy collision dissociation (HCD) scans of precursors were conducted with the following parameters: NCE 28%, maximal injection time of auto, AGC 5×10^4 with a resolution of 50,000, and an isolation window of *m/z* 2. To avoid repetitive analysis of identical components, a dynamic exclusion of 60 s was employed. Specific oxonium ions, such as HexNAc at *m/z* 204.087 and its fragment at *m/z* 138.0545, prompted two additional scans for potential glycopeptides: an Orbitrap electron-transfer/

high-energy collision dissociation (EThcD) scan (NCE 25%, maximal injection time of 86 ms, AGC 5 × 10^4 with a resolution of 50,000, isolation window of m/z 3) and a stepped collision energy HCD scan (NCE at 10%, 25%, and 40%, maximal injection time of 86 ms, AGC 5 × 10^4 with a resolution of 50,000, isolation window of m/z 2).

Data analysis

The LC-MS/MS raw files were analyzed using Byonic software v3.11.3 (Protein Metrics Inc., Cupertino, CA) [46]. Genome sequences were retrieved from accession numbers CP042905.2 for *P. syntrophicum* strain MK-D1, NZ_CP020364.1 for *S. acidocaldarius* JCM8929, NZ_LT549890.1 for *S. solfataricus* JCM11322, and AP028907.1 for *P. abyssi* JCM9394. The cleavage specificity was set as semi-specific N-ragged, and up to three missed cleavage events were allowed. Carbamidomethyl was set as a fixed modification of cysteine, while methionine oxidation and acetylation of the protein N-terminal amino groups were included as variable modifications. A maximum mass precursor tolerance of 10 ppm was allowed whereas a mass tolerance of up to 10 ppm was set for HCD fragments and 20 ppm for EThcD fragments. For *N*-glycosylation open searches, the wildcard parameter was enabled, allowing a delta m/z range of 300–2000 on asparagine residues. For focused searches, all parameters remained consistent except for the disabling of wildcard searching, and specific glycoforms identified from open searches were incorporated as variable modifications. Manual corrections were applied to glycosylation sites misassigned by Byonic, especially in long peptides with multiple asparagine residues, using MS/MS spectral patterns and the N-X-S/T sequon for validation. Protein identifications were filtered with |log prob| ≥ 2. Peptide identifications required a Byonic score ≥ 200, |log prob| ≥ 2, and a two-dimensional false discovery rate (2D FDR) ≤ 0.01. Relative protein abundances were determined using the normalized spectral abundance factors (NSAFs) method, which is based on the number of peptide spectrum matches (PSMs) per protein [47]. NSAF values were subsequently multiplied by 100 to derive the relative protein abundances as percentages. Subcellular localization was predicted using pSORTb version 3.0 [48]. The structures of proteins were predicted using the AlphaFold (version 2.2.0) [49]. Transmembrane helices in proteins were predicted using TMHMM-2.0 [50].

Results and discussion

N-glycosylation genes in P. syntrophicum

To explore the potential and evolutionary aspects of *N*-glycosylation in *P. syntrophicum*, the gene repertoire associated with this process was examined within its genome. An orthologue of the oligosaccharyl transferase (OST) gene (*aglB*, locus tag= DSAG12_01167) containing WWDYG and DEGKWPWM motifs was identified [5]. These motifs are conserved across "Lokiarchaeia" and "Thorarchaeia" members within the Asgard archaea [23]. In contrast to bacteria, *Methanobacteriati*, and *Metallosphaera* species [23], the glycosylation-associated genes were not clustered near *aglB*, suggesting a dispersed organization of *agl* genes within the Asgard archaeal genome. A BLASTp search, using *agl* genes of *Haloferax volcanii* and *Methanococcus maripaludis* as a query, identified homologues for *aglD* (DSAG12_01846), *aglF* (DSAG12_01619, DSAG12_01835, DSAG12_02719), *aglJ* (DSAG12_01866), *aglL* (DSAG12_01899), *aglM* (DSAG12_01705), *agl1* (DSAG12_01836), and *agl12* (DSAG12_01837). Although these genes suggested the occurrence of *N*-glycosylation in *P. syntrophicum*, it is not feasible to deduce the specific glycoforms or glycosites from this dataset. Besides *aglB*, which is homologous to the catalytic Stt3 subunit of eukaryotic OST, a gene encoding the non-catalytic components Ost3/Ost6 was identified (DSAG12_01196), consistent with observations in other "Lokiarchaeia" [23,51]. Furthermore, a homolog of the translocon-associated protein (TRAP) complex, especially TRAP beta (DSAG12_01731), was identified

within its genome. However, Ost1 (ribophorin I), typically present in "Lokiarchaeia" genomes [23,51], was absent in *P. syntrophicum*, suggesting a possible divergence in glycosylation processes.

Protein diversity in P. syntrophicum

Our glycoproteome analysis using *N*-glycosylation open search identified 732 different proteins from *P. syntrophicum*, corresponding to approximately 19.1% of its total coding sequence (Table 1). Based on their Byonic score (|log P value| [LPV]), the detected proteins contained 510 high-confidence hits (≥+2 above the top decoy LPV). In comparison, for *S. acidocaldarius, S. solfataricus*, and *P. abyssi,* 1300 (56.5% of CDSs), 1348 (42.0% of CDSs), and 1404 (64.9% of CDSs) proteins were detected, respectively. This suggests that *P. syntrophicum* might utilize a relatively limited variety of proteins. Additionally, the proportion of detected proteins relative to the total coding sequences was significantly lower in *P. syntrophicum*. This reduction may also be attributed to the comparatively larger genome sizes of Asgard archaea, which result from increased gene duplications and reduced gene losses compared to other archaeal lineages [51]. Further analysis is needed to determine whether other proteins are more actively expressed in *P. syntrophicum* under different cultivation conditions.

For each archaeon investigated, the top ten proteins, ranked by LPVs, are listed in Tables 1 and S1. In *P. syntrophicum*, an ABC transporter substrate-binding protein (SBP) (DSAG12_03090) was detected with significantly high LPV and NSAF values, supporting the earlier transcriptomic data [5]. A similar SBP (B6A19_RS05040) was also among its top ten proteins in *S. acidocaldarius* (Table S1). These SBPs bind solutes with high affinity at the outer cell surface and then transfer the substrate to the membrane integral permease domain [52]. From InterProScan search, DSAG12_03090 was classified into SBP family 5 (IPR039424) and Mpp-A type, responsible for the transport of oligopeptides and nickel [53]. Similar to specific binding proteins in bacteria, archaeal SBPs have also been suggested to act as receptors in chemotaxis [54], potentially serving a dual function in both nutrient uptake and sensory mechanisms. BLASTp search identified homologous genes of DSAG12_03090 in "*Ca.* L. ossiferum" cultivated from estuarine sediment [6], as well as Lokiarchaeia metagenome-assembled genomes (MAGs) from estuarine sediments [55] and freshwater sediments [56], suggesting the common significance of this SBP in the nutritional strategies of Asgard archaea. Based on NSAF values, the top 10-ranking proteins in *P. syntrophicum* are provided in Table S2. Six of these proteins were histone domain-containing proteins (NFYB/HAP3 family transcription factor), which are prevalent in archaea and play roles in genome compaction and organization [57]. Some archaeal histone paralogs appear to have evolved distinct and conserved functional roles, reminiscent of eukaryotic histone variants [58]. Despite detecting 40 out of the 80 known eukaryotic signature proteins (ESPs) of *P. syntrophicum*

Table 1
Proteins identified in *P. syntrophicum* listed according to their Byonic scores (|log P value| [LPV]), showing the top 10 highest-ranked proteins.

Rank	Locus tag	Description	|Log Prob|	NSAF
1	DSAG12_03090	ABC transporter substrate-binding protein	881.65	6.00
2	DSAG12_02329	GTPase	286.58	0.81
3	DSAG12_02465	NADP-dependent isocitrate dehydrogenase	249.28	0.36
4	DSAG12_02468	aconitate hydratase	230.55	0.40
5	DSAG12_02001	Glu/Leu/Phe/Val dehydrogenase	222.39	0.84
6	DSAG12_03802	pyruvate, phosphate dikinase	208.8	0.24
7	DSAG12_02466	citrate/2-methylcitrate synthase	204.56	0.35
8	DSAG12_01664	hypothetical protein	203.83	0.30
9	DSAG12_02681	acetate–CoA ligase	195.07	0.31
10	DSAG12_03376	TrmB family transcriptional regulator	192.71	0.15

[5], their NSAF values were generally low (Table S3). Previous research identified potential S-layer genes encoding PKD domain proteins in *P. syntrophicum* (DSAG12_02910, DSAG12_00685, DSAG12_01918, and DSAG12_03009), however, none of these proteins were detected in this study. This result supports earlier suggestions that Asgard archaea cells may not be covered with S-layer proteins [6]. The absence of a rigid S-layer could hypothetically allow for more cellular flexibility and movement in *P. syntrophicum*, an intriguing possibility that warrants further exploration given its unique morphological features.

In comparison, S-layer proteins were identified in all other archaea investigated with high LPVs (Table S1). The top-ranking protein in *S. acidocaldarius* (B6A19_RS11570), originally annotated as a hypothetical protein in the genome analysis, was previously identified as the S-layer protein SlaA [59]. Similarly, SSOP1_RS01945 of *S. solfataricus* was identified as an S-layer protein SlaA [60]. For PABY_19310 of *P. abyssi*, the structural prediction using AlphaFold2 and the abundance of hydrophobic amino acids suggested its association with S-layer outer component SlaA (Fig. S1). This is consistent with prior findings that archaea generally produce a substantial quantity of S-layer proteins [9, 10]. In general, the detected protein list highlighted a distinct proteomic profile of *P. syntrophicum* compared to the other three archaea. Although the selection of (hyper)thermophilic archaea may explain common and high LPVs and NSAF for thermosome and other chaperonin family proteins, thermosome appeared, ranked 22nd based on LPV, in *P. syntrophicum*. Similarly, elongation factors were prevalently detected in all examined archaea, ranking 28th (based on LPV) in *P. syntrophicum*.

From detection to validation of N-glycosylation

While open searching allows the detection of heterogeneous peptide modifications by *N*-linked glycans without predefining their structures [39], this approach often results in identifying pseudo-positive scans with random modification masses, despite filtering out low-confidence peptide MS/MS scans. Therefore, a histogram of peptide modification values was created, binning the observed values in 0.01 Da increments (Figs. 1 and S2). Examination of the modifications in *P. syntrophicum* revealed three predominant modification values: 801.30, 909.30, and 1212.40 Da (Fig. 1). Although these masses were the foci of this study, the possibility that other less prevalent delta masses correspond to minor
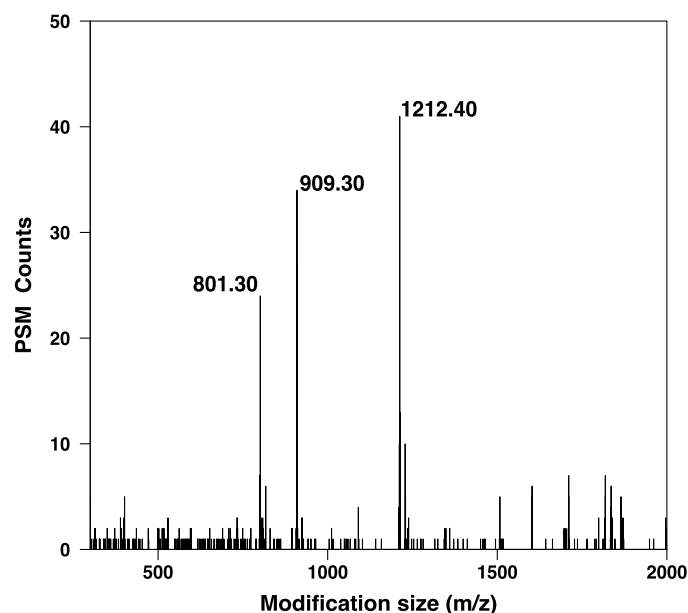

Fig. 1. Delta mass distribution of glycopeptides in 0.1 Da increments, showing detected peptide-spectrum matches (PSMs) with modifications ranging from *m/z* 300 to 2000 Da, derived from open searching data of *P. syntrophicum*.

glycan structures cannot be excluded. In addition, the EThcD scan, triggered by specific oxonium ions, may also limit the detection of more unusual and less frequent glycans. Unlike *P. syntrophicum, S. acidocaldarius* and *S. solfataricus* showed a uniform and frequent modification at 1118.33 Da (Fig. S2), suggesting that the presence of multiple modification values may be characteristic of *P. syntrophicum*. Within eukaryotes, glycosylation systems are known to generate highly heterogeneous glycan structures, potentially pointing a similarity between the glycosylation patterns of *P. syntrophicum* and eukaryotes. Previously, a specific *N*-glycan was identified in various *S. acidocaldarius* proteins including cytochrome $b_{558/566}$, archaellin FlaB, and S-layer proteins SlaA and SlaB [37,59,61]. This glycan, a tribranched hexasaccharide comprising $GlcNAc_2Glc_1Man_2$ and one sulfated sugar, 6-sulfoquinovose (QuiS), has a theoretical modification mass of 1118.33194 Da [37,59, 61], which corresponded to the predominant delta mass detected in this study (Fig. S1). Typically, glycoproteome studies have been conducted using enriched glycopeptides. However, due to the limited biomass from *P. syntrophicum*, this study utilized all peptides without glycopeptide enrichment. The abundant detection of the expected glycoform in *S. acidocaldarius* confirms that our protocol is satisfactory under these constraints. In addition, for another strain of *S. solfataricus*, i.e., strain P2, a heptasaccharide comprising $Hex_4HexNAc_2QuiS_1$ (theoretical modification value, 1280.38476 Da) was reported [62], which includes one more Hex than that observed in strain P1 in this study. Within the *Sulfolobaceae* family, variability of the *N*-glycans exists in the number of Hex units at the non-reducing end [63]. The delta mass profile for *P. abyssi* exhibited three clusters in the histogram, distinct from those seen in other studied archaea (Fig. S2). However, due to low PSM counts for these modifications and a noisier histogram, further verification of oxonium ions and glycan fragments is required in MS/MS spectra (described below).

N-glycosylation analysis using focused searching

Given the efficiency of focused searching for detecting glycopeptides and glycoproteins [39], we performed this analysis targeting the predominant delta masses identified in each archaeon. This approach, despite potentially overlooking minor glycopeptides with unique glycans, generally improved the detection of glycopeptide-spectrum matches (glyco-PSMs) in *P. syntrophicum* and other archaea (Fig. S3). For the predominant SBP in *P. syntrophicum* (DSAG12_03090), glycopeptides with one of the three aforementioned modifications were detected in 82 PSMs during open searching, which increased to 301 PSMs during focused searching. Nevertheless, overall glycosylation in *P. syntrophicum* was notably infrequent. Glyco-PSM analysis revealed that 88.3% (113/128) of the 801.30 Da modification, 84.4% (141/167) of the 909.30 Da modification, and 67.0% (120/179) of the 1212.40 Da modification were primarily associated with the SBP (DSAG12_03090). Additional glycoproteins, including various hypothetical proteins and different SBPs, were identified (Table S4), although their glyco-PSM counts were low. The functions of these hypothetical proteins remain speculative; however, a BLASTp search indicated their widespread presence among various "Lokiarchaeia", suggesting a conserved function. No consistent pattern was observed in the putative subcellular localization of these glycoproteins. Although glycosylation frequency could be confirmed through gel electrophoresis and glycoprotein-staining experiments, the rarity of glycosylation may be attributed to the complex and resource-demanding glycan biosynthetic processes [40], particularly given the slow growth strategy of *P. syntrophicum* in nutrient-deficient environments [5].

The glycosites within the *P. syntrophicum* SBP (DSAG12_03090) were predominantly located at N157, N398, N474, and N483 (Fig. 2a). Of the twelve predicted *N*-glycosites, eight were glycosylated, all preserving the conserved N-X-S/T motif, where X represents any amino acid except but proline (Fig. 2b). Although methodological factors such as limited trypsin digestion sites may potentially influence the detection of

(a)

(b)

(c)
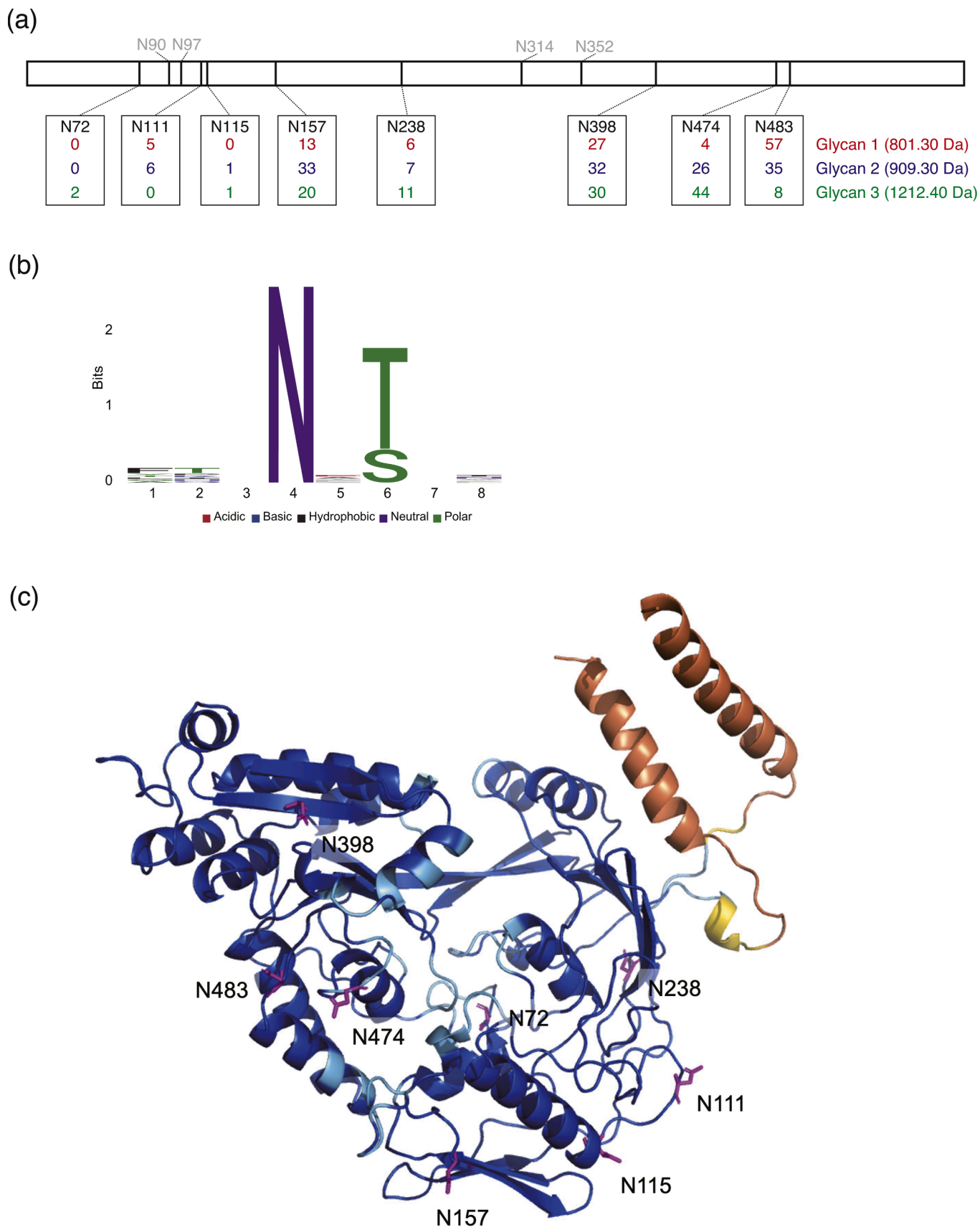
Fig. 2. (a) Distribution of glycosylation sites and identified glycans on the predominant substrate-binding protein (SBP) (DSAG12_03090) from *P. syntrophicum*. Since N474 and N483 were located in the same peptide, it was not possible to determine exact glycosite for some PSMs. (b) Sequence logo representing consensus sequences for *N*-glycosylation sites on the DSAG12_03090 protein. (c) Predicted structure of DSAG12_03090 protein using AlphaFold2. Amino acid residues are colored by Confidence score, pLDDT (blue <100, light blue <70, yellow <50, orange <30). Potential glycosylation sites are shown in Magenta.

glycosites, this denotes a *N*-glycosylation frequency of about one site per 74 residues. In comparison, a homologous protein from "*Ca.* L. ossiferum" (NEF87_000722) possessed seven sequons, suggesting a maximum expected glycosylation frequency of one site per 85 residues. Notably, the most frequently glycosylated site, N483 of DSAG12_030905, was conserved in the sequence alignment with the "*Ca.* L. ossiferum" protein. According to AlphaFold2, this glycosite was located at the apex of the protein, opposite the transmembrane and membrane-anchoring region identified in the orange N-terminal and C-terminal helices, as illustrated on the upper right side of Fig. 2c. An unusual feature of *N*-glycosylation

systems in archaea is the ability of certain species to attach multiple different glycans to the same protein [35]. In the S-layer glycoprotein of *Halobacterium salinarum*, an organism where only a single AglB has been identified [64], two distinct *N*-glycans are attached [65]. In *Methanococcus maripaludis*, a tetrasaccharide is attached to archaellins [66], whereas a pentasaccharide is attached to pilins [67]. Glycosylated SBPs have been reported in various archaea [19,37,68-70], and the SBP from *S. acidocaldarius* detected in this study (B6A19_RS05040) also exhibited significant glycosylation (glyco-PSM count: 402) (Table S5). However, the precise role of SBP glycosylation remains unclear. Glycosylation is
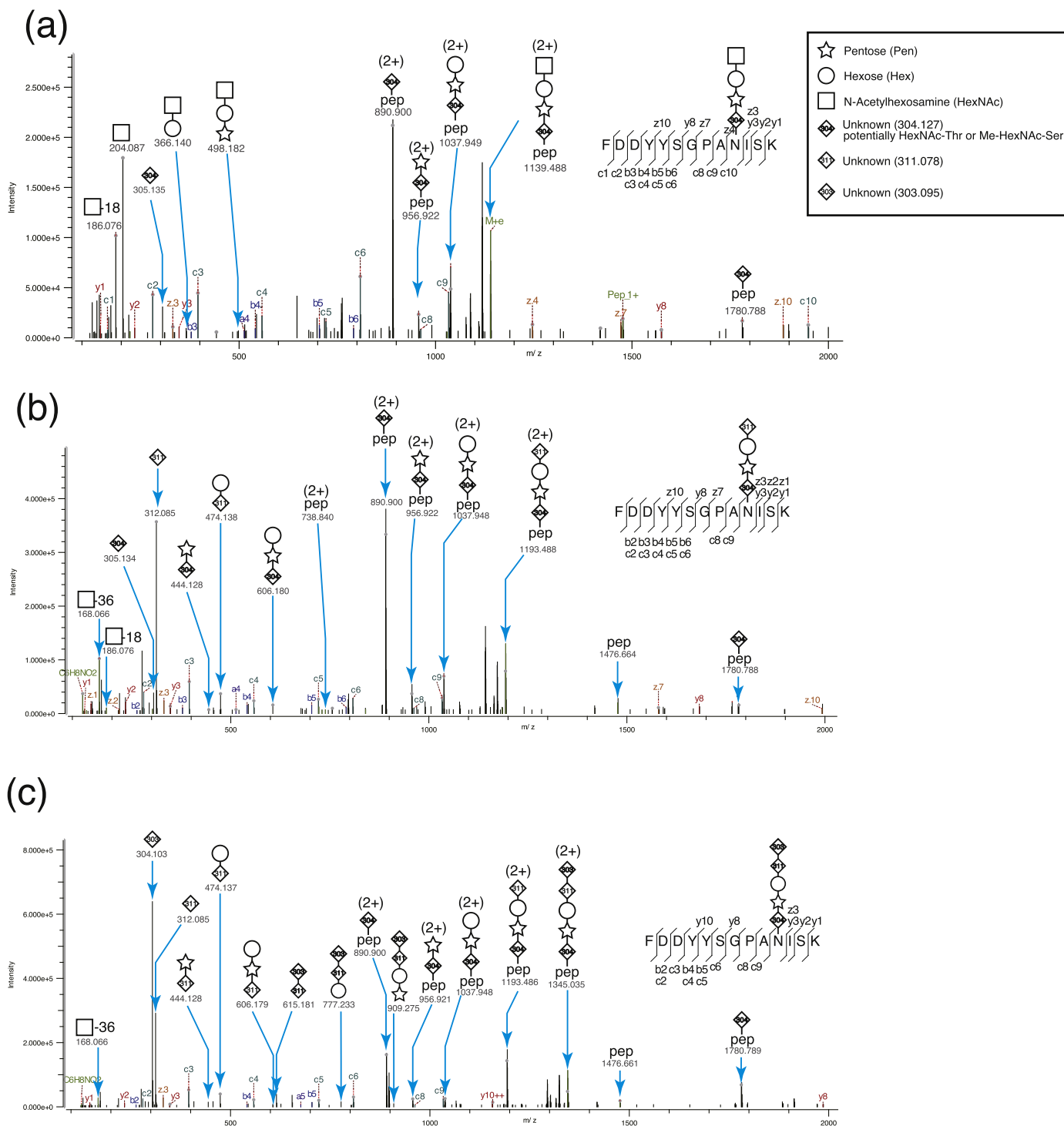
Fig. 3. Electron-transfer/higher-energy collision dissociation (EThcD) MS/MS analyses supporting the assignment of linear glycans of 801 Da (a), 909 Da (b), and 1212 Da (c) attached to the DSAG12_03090 peptide.

not essential for substrate recognition, as demonstrated by the successful heterologous expression of these proteins in *E. coli* [71,72] and their retained substrate-binding activity following deglycosylation [73,74]. The glycosylation may contribute to the stability of SBPs under extreme environmental conditions. Additionally, glycan-solute interactions may influence the solute-binding process.

In comparison, *S. acidocaldarius* and *S. solfataricus* exhibited significantly higher glycosylation frequencies (Table S5). Their S-layer proteins exhibited high glyco-PSMs, which is consistent with previous findings that many thermoacidophilic archaea possess S-layer proteins with an exceptionally large number of *N*-glycosites, typically averaging one glycosite per 30–40 residues [59]. In *Sulfolobus* and *Saccharolobus*, studied as model *Thermoproteota* for glycobiological research, the focus has been predominantly on glycoproteins such as S-layer proteins, archaellins, and cytochromes. Our study uncovered more diverse glycoproteins, including hypothetical proteins and ABC transporter solute-binding proteins, all modified by the homologous glycan (Table S5). In general, these glycoproteins tended to be localized to the cytoplasmic membrane. Among the three predominant delta masses identified in *P. abyssi*, only the 1176.43 Da modification was reliable due to the presence of oxonium ions of HexNAc and a consensus sequon. Other delta masses were not focused because the MS/MS spectra of peptides with the 1117.58 Da or 1266.66 Da modifications lacked potential sugar oxonium ions, which are commonly present in multiple putative glycopeptides, as well as the conserved N-X-S/T sequon. These are consistent with previous findings that the structural heterogeneity within one proteome is generally low in prokaryotes, although prokaryotic protein glycosylation shows a large species- and strain-level variability [40]. For *P. abyssi*, glycosylation was observed in the putative S-layer proteins, ABC transporter permease, and hypothetical proteins (Table S5). The glyco-PSM counts were notably lower compared to those observed in thermoacidophilic *Thermoproteota*.

N-glycosylation in P. syntrophicum

The MS/MS examination of the three predominant delta masses from *P. syntrophicum* identified the consistent presence of the oxonium ion of HexNAc at 204.08 Da. The MS/MS examination of PSMs linked to the 801.30 Da delta mass uncovered a presumed linear glycan consisting of 304.13-Pen-Hex-HexNAc, where the 304.13 Da is a moiety of unknown composition (Fig. 3a). Likewise, MS/MS spectra corresponding to the 909.30 Da delta mass showed a linear glycan of 304.13-Pen-Hex-311.08, with the 304.13 Da and 311.08 Da components being of unknown composition (Fig. 3b). Additionally, MS/MS spectra for the 1212.40 Da delta mass also indicated a linear glycan structure of 304.13-Pen-Hex-311.08-303.10, with the 304.13 Da, 311.08 Da, and 303.10 Da segments also remaining unknown (Fig. 3c). The occurrence of linear *N*-glycans may be reasonable, since branching in the *N*-glycans appears to be uncommon in mesophilic archaea [35], though further studies are required to confirm their association with growth temperatures.

Archaeal protein *N*-glycosylation is distinguished by a diverse array of linking sugars [35]. An unusual linking sugar with an *m/z* of 304.127 Da, identified in all three glycans from *P. syntrophicum*, was not previously reported in other archaea. However, since we could precisely identify the mass up to two to three decimal places, we identified a candidate sugar with a theoretical mass of 304.127051 Da, which could potentially correspond to HexNAc modified with threonine or Me-HexNAc modified with serine. Although the same mass could theoretically result from HexA modified with lysine, this is unlikely given that previously identified linking sugars predominantly consist of HexNAc or Hex, along with their derivatives. Additionally, fragments of HexNAc were observed in glycopeptides lacking HexNAc residues (Figs. 3b and 3c), supporting the presence of a HexNAc derivative. No matching structure could be identified for the modification mass of 311.08 Da and 303.10 Da. For archaea, threonine-modified sugar-containing glycan was previously reported in mesophilic methanogens

belonging to the genus *Methanococcus*, where the *N*-glycan contains ManNAcA or GlcNAcNAc linked with threonine [66,75]. Asgard archaea include hydrogen/formate-producing fermentative mesophiles that coexist closely with other microorganisms, such as methanogens that scavenge hydrogen gas and formate [5]. Such intimate symbiotic interactions might foster frequent lateral gene transfers between Asgard archaea and proto-eukaryotic ancestors [5,76]. The second sugar unit from the reducing end in *P. syntrophicum* was identified as a pentose. Such pentose-containing *N*-glycan was reported in the hyperthermophilic *Pyrococcus furiosus* (*Methanobacteriati*) [77]. To our knowledge, glycans from non-*Methanobacteriati* archaea are characterized by the structure of reducing end, having two GlcNAc residues or its modified forms, as those found in eukaryotes [36]. Thus, the discovery of unique glycan structures in *P. syntrophicum* was unexpected, particularly considering its evolutionary proximity to eukaryotes.

In comparison, expected glycan structures were confirmed for *S. acidocaldarius* and *S. solfataricus* (Fig. S4), validating our analytical methodology. For *P. abyssi*, MS/MS analysis of PSMs corresponding to the 1176.43 Da delta mass revealed a presumed linear glycan structure composed of 290.11-HexNAc-Hex-304.10-217.10 (Fig. S4). The candidate structure for the modification mass of 304.10 Da remained unidentified. Nonetheless, the 217.10 Da might correspond to Me-HexNAc, as previously reported in "*Ca.* Kuenenia stuttgartiensis" [40]. In addition, the candidate linking sugar corresponding to a mass of 290.11 Da was identified as HexNAc modified with serine (theoretical mass, 290.1114 Da), although the possibility of it being an atypical sugar cannot be excluded. While theoretically, this mass could represent Hex modified by glutamine, this is unlikely considering the typical linking sugars in *Thermoproteota N*-glycans are HexNAc or its variants [36,59, 62,63]. The identification of HexNAc as the second sugar from the reducing end suggests that the *N*-glycan possesses a chitobiose core, consistent with those found in other *Thermoproteota*.

Conclusion

This study unveiled the glycoproteomic characteristics of an evolutionarily important Asgard archaeon for the first time. Although detailed structural analyses, such as NMR and GC–MS, for *N*-glycans are needed for future studies, our study highlights the effectiveness of high-resolution LC-MS/MS in the glycobiological analysis of fastidious prokaryotes. While it is hypothesized that the last common ancestor of eukaryotes and the Asgard archaeal lineage might have shared protein *N*-glycosylation patterns [23], our findings showed the glycan properties of Asgard archaea diverge from those of eukaryotes, *Thermoproteota, Methanobacteriota, Halobacteriota,* and *Thermoplasmatota* (Fig. 4). The origin and timing of eukaryotic *N*-glycosylation, whether arising from symbiotic events or independent evolution before or after alphaproteobacterial endosymbiosis, remain unclear. According to the syntrophy hypothesis by Lopez-Garcia and Moreira [34], the deltaproteobacterial cell surface might have transitioned into eukaryotic one, suggesting Asgard archaeal cell surfaces might not directly resemble those of eukaryotes, which could explain the glycobiological discrepancies. Additionally, considering the 2-billion-year timeline of eukaryogenesis, current Asgard archaea have likely changed their glycosylation profiles to adapt to their environments. The cell surface glycan, as the interface with the environment, is probably subject to evolutionary pressures, as reflected by the diverse glycan structures currently observed in archaea. While the initial Asgard ancestors might have shared a chitobiose core typical of eukaryotes, subsequent environmental adaptations could have led to unique glycan configurations. The specific function of SBP glycosylation in *P. syntrophicum* remains to be elucidated. Nevertheless, the prevalent association of *N*-glycan with a single SBP suggests its significant role, potentially in facilitating nutrient uptake and sensory processes. Future glycoproteomic investigations on members of the "*Heimdallarchaeia*" (particularly "Hodarchaeales"), which have a phylogenetic relationship closer to eukaryotes than that of *Promethearchaeum*,
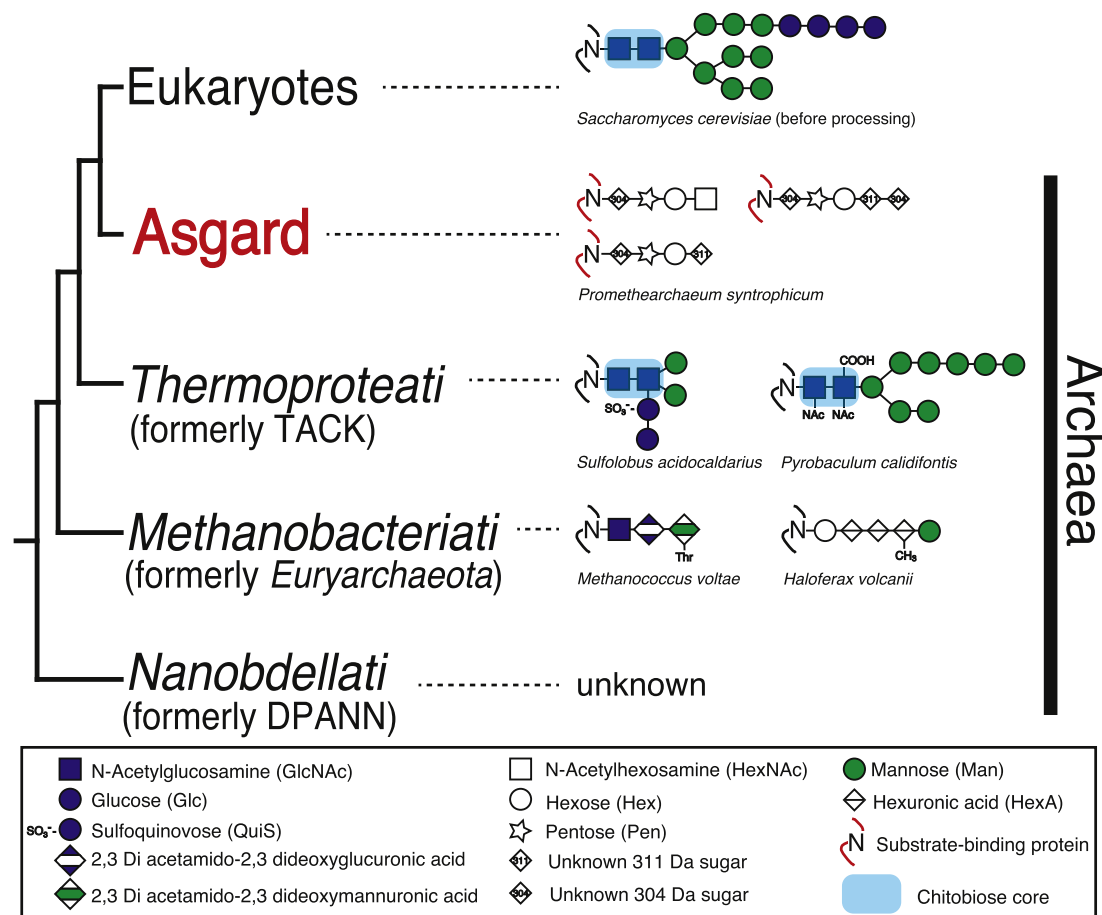
Fig. 4. The structural diversity of *N*-linked glycans in archaeal kingdoms.

could elucidate the origins and evolution of eukaryotic *N*-glycosylation. Considering that eukaryotic *N*-glycosylation starts in the endoplasmic reticulum and progresses to the Golgi apparatus, a detailed investigation of the *N*-glycosylation machineries and their cellular localizations in Asgard archaeal cells is essential. Given the multiple hypotheses on eukaryogenesis, it may be oversimplified to explain the emergence of eukaryotic *N*-glycosylation through a mere symbiotic event between an archaeon and a bacterium. In conclusion, from a glycobiological standpoint, the Asgard archaeon does not exhibit clearly ancestral traits similar to eukaryotes; rather, its unique characteristics may result in part from horizontal gene transfers with cohabiting mesophilic species like methanogens, as well as adaptations to their oligotrophic environments. The physiological functions and evolutionary backgrounds of their distinctive glycans, which include atypical sugars potentially modified with threonine, remain to be elucidated. Our findings, along with novel glycan identifications in *P. abyssi*, shed light on the largely unexplored frontier of archaeal glycobiology and establish a foundation for upcoming inquiries into glycosylation variations within prokaryotes.

Data availability

All LC-MS/MS raw files, Byonic result files (.byrslt), and complete protein identification lists have been deposited in the ProteomeXchange Consortium via jPOSTrepo [78], under the data set identifiers PXD051006 (JPST003012) for *P. syntrophicum* strain MK-D1[T], PXD051007 (JPST003014) for *S. acidocaldarius* strain 98-3[T], PXD051008 (JPST003015) for *S. solfataricus* strain P1[T], and PXD051009 (JPST003016) for *S. abyssi* strain AV2[T].

CRediT authorship contribution statement

Satoshi Nakagawa: Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Hiroyuki Imachi:** Writing – review & editing, Resources, Funding acquisition. **Shigeru Shimamura:** Writing – review & editing, Methodology, Data curation. **Saeko Yanaka:** Writing – review & editing, Visualization, Software. **Hirokazu Yagi:** Writing – review & editing, Validation. **Maho Yagi-Utsumi:** Writing – review & editing. **Hiroyuki Sakai:** Writing – review & editing, Investigation. **Shingo Kato:** Writing – review & editing, Investigation. **Moriya Ohkuma:** Writing – review & editing. **Koichi Kato:** Writing – review & editing. **Ken Takai:** Writing – review & editing, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Living Systems (ExCELLS) (ExCELLS program No. 22EXC601).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.bbadva.2024.100118.

References

[1] H. Imachi, M.K. Nobu, S. Kato, Y. Takaki, M. Miyazaki, M. Miyata, M. Ogawara, Y. Saito, S. Sakai, Y.O. Tahara, Y. Takano, E. Tasumi, K. Uematsu, T. Yoshimura, T. Itoh, M. Ohkuma, K. Takai, *Promethearchaeum syntrophicum* gen. nov., sp. nov., an anaerobic, obligately syntrophic archaeon, the first isolate of the lineage 'Asgard' archaea, and proposal of the new archaeal phylum *Promethearchaeota* phyl. nov. and kingdom *Promethearchaeati* regn. nov, Int. J. Syst. Evol. Microbiol. (2024) in press.

[2] A. Spang, J.H. Saw, S.L. Jorgensen, K. Zaremba-Niedzwiedzka, J. Martijn, A. E. Lind, R. van Eijk, C. Schleper, L. Guy, T.J. Ettema, Complex archaea that bridge the gap between prokaryotes and eukaryotes, Nature 521 (2015) 173–179.

[3] K. Zaremba-Niedzwiedzka, E.F. Caceres, J.H. Saw, D. Backstrom, L. Juzokaite, E. Vancaester, K.W. Seitz, K. Anantharaman, P. Starnawski, K.U. Kjeldsen, M. B. Stott, T. Nunoura, J.F. Banfield, A. Schramm, B.J. Baker, A. Spang, T.J. Ettema, Asgard archaea illuminate the origin of eukaryotic cellular complexity, Nature 541 (2017) 353–358.

[4] K.W. Seitz, C.S. Lazar, K.U. Hinrichs, A.P. Teske, B.J. Baker, Genomic reconstruction of a novel, deeply branched sediment archaeal phylum with pathways for acetogenesis and sulfur reduction, ISME J. 10 (2016) 1696–1705.

[5] H. Imachi, M.K. Nobu, N. Nakahara, Y. Morono, M. Ogawara, Y. Takaki, Y. Takano, K. Uematsu, T. Ikuta, M. Ito, Y. Matsui, M. Miyazaki, K. Murata, Y. Saito, S. Sakai, C. Song, E. Tasumi, Y. Yamanaka, T. Yamaguchi, Y. Kamagata, H. Tamaki, K. Takai, Isolation of an archaeon at the prokaryote-eukaryote interface, Nature 577 (2020) 519–525.

[6] T. Rodrigues-Oliveira, F. Wollweber, R.I. Ponce-Toledo, J. Xu, S.K.R. Rittmann, A. Klingl, M. Pilhofer, C. Schleper, Actin cytoskeleton and complex cell architecture in an Asgard archaeon, Nature 613 (2023) 332–339.

[7] U.B. Sleytr, B. Schuster, E.M. Egelseer, D. Pum, S-layers: principles and applications, FEMS Microbiol. Rev. 38 (2014) 823–864.

[8] U.B. Sleytr, A. Breitwieser, D. Pum, Crystalline cell surface layers (S-layers, in: T. M. Schmidt (Ed.), Encyclopedia of Microbiology, Academic Press, 2019, pp. 783–792.

[9] J.H.Y. Lau, J.F. Nomellini, J. Smit, Analysis of high-level S-layer protein secretion in *Caulobacter crescentus*, Can. J. Microbiol. 56 (2010) 501–514.

[10] M. Sara, U.B. Sleytr, S-layer proteins, J. Bacteriol. 182 (2000) 859–868.

[11] T.A.M. Bharat, A. von Kugelgen, V. Alva, Molecular logic of prokaryotic surface layer structures, Trends. Microbiol. 29 (2021) 405–415.

[12] R.P. Fagan, N.F. Fairweather, Biogenesis and functions of bacterial S-layers, Nat. Rev. Microbiol. 12 (2014) 211–222.

[13] I.A. Zink, K. Pfeifer, E. Wimmer, U.B. Sleytr, B. Schuster, C. Schleper, CRISPR-mediated gene silencing reveals involvement of the archaeal S-layer in cell division and virus infection, Nat. Commun. 10 (2019) 4797.

[14] H. Engelhardt, Are S-layers exoskeletons? The basic function of protein surface layers revisited, J. Struct. Biol. 160 (2007) 115–124.

[15] L.J. Gebhard, Z. Vershinin, T. Alarcon-Schumacher, J. Eichler, S. Erdmann, Influence of N-glycosylation on virus-host interactions in *Halorubrum lacusprofundi*, Viruses 15 (2023) 1469.

[16] C. de la Fuente-Nunez, J. Mertens, J. Smit, R.E. Hancock, The bacterial surface layer provides protection against antimicrobial peptides, Appl. Environ. Microbiol. 78 (2012) 5452–5456.

[17] J. Eichler, N-glycosylation in Archaea-new roles for an ancient posttranslational modification, Mol. Microbiol. 114 (2020) 735–741.

[18] M.F. Mescher, J.L. Strominger, Purification and characterization of a prokaryotic glycoprotein from the cell envelope of *Halobacterium salinarium*, J. Biol. Chem. 251 (1976) 2005–2014.

[19] M. Sumper, E. Berg, R. Mengele, I. Strobel, Primary structure and glycosylation of the S-layer protein of *Haloferax volcanii*, J. Bacteriol. 172 (1990) 7111–7118.

[20] H. Lu, Y. Lu, J. Ren, Z. Wang, Q. Wang, Y. Luo, J. Han, H. Xiang, Y. Du, C. Jin, Identification of the S-layer glycoproteins and their covalently linked glycans in the halophilic archaeon *Haloarcula hispanica*, Glycobiology 25 (2015) 1150–1162.

[21] M. Kowarik, N. Young, S. Numao, B. Schulz, I. Hug, N. Callewaert, D. Mills, D. Watson, M. Hernandez, J. Kelly, M. Wacker, M. Aebi, Definition of the bacterial N-glycosylation site consensus sequence, EMBO J. 25 (2006) 1957–1966.

[22] M.S. Lowenthal, K.S. Davis, T. Formolo, L.E. Kilpatrick, K.W. Phinney, Identification of novel N-glycosylation sites at noncanonical protein consensus motifs, J. Proteome Res. 15 (2016) 2087–2101.

[23] S. Nikolayev, C. Cohen-Rosenzweig, J. Eichler, Evolutionary considerations of the oligosaccharyltransferase AglB and other aspects of N-glycosylation across Archaea, Mol. Phylogenet. Evol. 153 (2020) 106951.

[24] J.P. Hymes, B.R. Johnson, R. Barrangou, T.R. Klaenhammer, Functional analysis of an S-Layer-associated fibronectin-binding protein in *Lactobacillus acidophilus* NCFM, Appl. Environ. Microbiol. 82 (2016) 2676–2685.

[25] O. Goto, S. Kunisawa, L. Ivanov, K. Takeyama, M. Sakamoto, I. Setoyama, A. Uematsu, K. Domino, R. Becher, U. Sasakawa, K. Benno, Innate lymphoid cells regulate intestinal epithelial cell glycosylation, Science 345 (2014) 1254009.

[26] M.I. Kanipes, S.R. Kalb, R.J. Cotter, D.F. Hozbor, A. Lagares, C.R. Raetz, Relaxed sugar donor selectivity of a *Sinorhizobium meliloti* ortholog of the *Rhizobium leguminosarum* mannosyl transferase LpcC. Role of the lipopolysaccharide core in symbiosis of *Rhizobiaceae* with plants, J. Biol. Chem. 278 (2003) 16365–16371.

[27] I. Lerouge, J. Vanderleyden, O-antigen structural variation: mechanisms and possible roles in animal/plant-microbe interactions, FEMS Microbiol. Rev. 26 (2002) 17–47.

[28] M. Hu, Y. Bai, X. Zheng, Y. Zheng, Coral-algal endosymbiosis characterized using RNAi and single-cell RNA-seq, Nat. Microbiol. 8 (2023) 1240–1251.

[29] E.M. Wood-Charlson, L.L. Hollingsworth, D.A. Krupp, V.M. Weis, Lectin/glycan interactions play a role in recognition in a coral/dinoflagellate symbiosis, Cell Microbiol. 8 (2006) 1985–1993.

[30] A. Fioravanti, F. Van Hauwermeiren, S.E. Van der Verren, W. Jonckheere, A. Goncalves, E. Pardon, J. Steyaert, H. De Greve, M. Lamkanfi, H. Remaut, Structure of S-layer protein Sap reveals a mechanism for therapeutic intervention in anthrax, Nat. Microbiol. 4 (2019) 1805–1814.

[31] M.J. Blaser, Role of the S-layer proteins of *Campylobacter fetus* in serum-resistance and antigenic variation: a model of bacterial pathogenesis, Am. J. Med. Sci. 306 (1993) 325–329.

[32] Z. Pei, M.J. Blaser, Pathogenesis of *Campylobacter fetus* infections. Role of surface array proteins in virulence in a mouse model, J. Clin. Invest. 85 (1990) 1036–1043.

[33] F. Nasher, B.W. Wren, Flagellin *O*-linked glycans are required for the interactions between *Campylobacter jejuni* and *Acanthamoeba castellanii*, Microbiology 169 (2023) 001386.

[34] P. Lopez-Garcia, D. Moreira, The Syntrophy hypothesis for the origin of eukaryotes revisited, Nat. Microbiol. 5 (2020) 655–667.

[35] K.F. Jarrell, Y. Ding, B.H. Meyer, S.V. Albers, L. Kaminski, J. Eichler, N-linked glycosylation in Archaea: a structural, functional, and genetic analysis, Microbiol. Mol. Biol. Rev. 78 (2014) 304–341.

[36] D. Fujinami, Y. Taguchi, D. Kohda, Asn-linked oligosaccharide chain of a crenarchaeon, *Pyrobaculum calidifontis*, is reminiscent of the eukaryotic high-mannose-type glycan, Glycobiology 27 (2017) 701–712.

[37] U. Zahringer, H. Moll, T. Hettmann, V.A. Knirel, G. Schafer, Cytochrome b(558/566) from the archaeon *Sulfolobus acidocaldarius* has a unique Asn-linked highly branched hexasaccharide chain containing 6-sulfoquinovose, Eur. J. Biochem. 267 (2000) 4144–4149.

[38] M. Thaysen-Andersen, N.H. Packer, B.L. Schulz, Maturing glycoproteomics technologies provide unique structural insights into the *N*-glycoproteome and its regulation in health and disease, Mol. Cell Proteomics 15 (2016) 1773–1790.

[39] A.R. Ahmad Izaham, N.E. Scott, Open database searching enables the identification and comparison of bacterial glycoproteomes without defining glycan compositions prior to searching, Mol. Cell Proteomics 19 (2020) 1561–1574.

[40] M. Pabst, D.S. Grouzdev, C.E. Lawson, H.B.C. Kleikamp, C. de Ram, R. Louwen, Y. M. Lin, S. Lucker, M.C.M. van Loosdrecht, M. Laureni, A general approach to explore prokaryotic protein glycosylation reveals the unique surface layer modulation of an anammox bacterium, ISME J. 16 (2022) 346–357.

[41] H. Imachi, M.K. Nobu, M. Miyazaki, E. Tasumi, Y. Saito, S. Sakai, M. Ogawara, A. Ohashi, K. Takai, Cultivation of previously uncultured microorganisms with a continuous-flow down-flow hanging sponge (DHS) bioreactor, using a syntrophic archaeon culture obtained from deep marine sediment as a case study, Nat. Protoc. 17 (2022) 2784–2814.

[42] Y. Hashimoto, S. Shimamura, A. Tame, S. Sawayama, J. Miyazaki, K. Takai, S. Nakagawa, Physiological and comparative proteomic characterization of *Desulfolithobacter dissulfuricans* gen. nov., sp. nov., a novel mesophilic, sulfur-disproportionating chemolithoautotroph from a deep-sea hydrothermal vent, Front. Microbiol. 13 (2022) 1042116.

[43] N.E. Scott, B.L. Parker, A.M. Connolly, J. Paulech, A.V. Edwards, B. Crossett, L. Falconer, D. Kolarich, S.P. Djordjevic, P. Hojrup, N.H. Packer, M.R. Larsen, S. J. Cordwell, Simultaneous glycan-peptide characterization using hydrophilic interaction chromatography and parallel fragmentation by CID, higher energy collisional dissociation, and electron transfer dissociation MS applied to the N-linked glycoproteome of *Campylobacter jejuni*, Mol. Cell Proteomics 10 (2011) M000031–MCP000201.

[44] S. Mysling, G. Palmisano, P. Hojrup, M. Thaysen-Andersen, Utilizing ion-pairing hydrophilic interaction chromatography solid phase extraction for efficient glycopeptide enrichment in glycoproteomics, Anal. Chem. 82 (2010) 5598–5609.

[45] S. Kawai, S. Shimamura, Y. Shimane, Y. Tsukatani, Proteomic time-course analysis of the filamentous anoxygenic phototrophic bacterium, *Chloroflexus aurantiacus*, during the transition from respiration to phototrophy, Microorganisms 10 (2022) 1288.

[46] M. Bern, Y.J. Kil, C. Becker, Byonic: advanced peptide and protein identification software, Curr. Protoc. Bioinformatics Chapter 13 (2012). Unit13 20.

[47] B. Zybailov, A.L. Mosley, M.E. Sardiu, M.K. Coleman, L. Florens, M.P. Washburn, Statistical analysis of membrane proteome expression changes in *Saccharomyces cerevisiae*, J. Proteome Res. 5 (2006) 2339–2347.

[48] N.Y. Yu, J.R. Wagner, M.R. Laird, G. Melli, S. Rey, R. Lo, P. Dao, S.C. Sahinalp, M. Ester, L.J. Foster, F.S. Brinkman, PSORTb 3.0: improved protein subcellular localization prediction with refined localization subcategories and predictive capabilities for all prokaryotes, Bioinformatics 26 (2010) 1608–1615.

[49] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Zidek, A. Potapenko, A. Bridgland, C. Meyer, S.A. A. Kohl, A.J. Ballard, A. Cowie, B. Romera-Paredes, S. Nikolov, R. Jain, J. Adler, T. Back, S. Petersen, D. Reiman, E. Clancy, M. Zielinski, M. Steinegger, M. Pacholska, T. Berghammer, S. Bodenstein, D. Silver, O. Vinyals, A.W. Senior, K. Kavukcuoglu, P. Kohli, D. Hassabis, Highly accurate protein structure prediction with AlphaFold, Nature 596 (2021) 583–589.

[50] A. Krogh, B. Larsson, G. von Heijne, E.L. Sonnhammer, Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes, J. Mol. Biol. 305 (2001) 567–580.

[51] L. Eme, D. Tamarit, E.F. Caceres, C.W. Stairs, V. De Anda, M.E. Schon, K.W. Seitz, N. Dombrowski, W.H. Lewis, F. Homa, J.H. Saw, J. Lombard, T. Nunoura, W.J. Li, Z.S. Hua, L.X. Chen, J.F. Banfield, E.S. John, A.L. Reysenbach, M.B. Stott, A. Schramm, K.U. Kjeldsen, A.P. Teske, B.J. Baker, T.J.G. Ettema, Inference and reconstruction of the heimdallarchaeial ancestry of eukaryotes, Nature 618 (2023) 992–999.

[52] E.L. Borths, K.P. Locher, A.T. Lee, D.C. Rees, The structure of Escherichia coli BtuF and binding to its cognate ATP binding cassette transporter, Proc. Natl. Acad. Sci. U.S.A. 99 (2002) 16642–16647.

[53] J.T. Park, D. Raychaudhuri, H. Li, S. Normark, D. Mengin-Lecreulx, MppA, a periplasmic binding protein essential for import of the bacterial cell wall peptide ʟ-alanyl-gamma-ᴅ-glutamyl-meso-diaminopimelate, J. Bacteriol. 180 (1998) 1215–1223.

[54] M.V. Kokoeva, K.F. Storch, C. Klein, D. Oesterhelt, A novel mode of sensory transduction in archaea: binding protein-mediated chemotaxis towards osmoprotectants and amino acids, EMBO J. 21 (2002) 2312–2322.

[55] W. Wang, J. Tao, K. Yu, C. He, J. Wang, P. Li, H. Chen, B. Xu, Q. Shi, C. Zhang, Vertical stratification of dissolved organic matter linked to distinct microbial communities in subtropic estuarine sediments, Front. Microbiol. 12 (2021) 697860.

[56] C.R. Hahn, I.F. Farag, C.L. Murphy, M. Podar, M.S. Elshahed, N.H. Youssef, Microbial diversity and sulfur cycling in an early earth analogue: from ancient novelty to modern commonality, mBio 13 (2022) e0001622.

[57] B. Henneman, C. van Emmerik, H. van Ingen, R.T. Dame, Structure and function of archaeal histones, PLoS Genet. 14 (2018) e1007582.

[58] K.M. Stevens, J.B. Swadling, A. Hocher, C. Bang, S. Gribaldo, R.A. Schmitz, T. Warnecke, Histone variants in archaea and the evolution of combinatorial chromatin complexity, Proc. Natl. Acad. Sci. U.S.A. 117 (2020) 33384–33395.

[59] E. Peyfoon, B. Meyer, P.G. Hitchen, M. Panico, H.R. Morris, S.M. Haslam, S.V. Albers, A. Dell, The S-layer glycoprotein of the crenarchaeote Sulfolobus acidocaldarius is glycosylated at multiple sites with chitobiose-linked N-glycans, Archaea 2010 (2010) 754101.

[60] H. Claus, E. Akca, T. Debaerdemaeker, C. Evrard, J.P. Declercq, J.R. Harris, B. Schlott, H. Konig, Molecular organization of selected prokaryotic S-layer proteins, Can. J. Microbiol. 51 (2005) 731–743.

[61] Z. Guan, A. Delago, P. Nussbaum, B. Meyer, S.V. Albers, J. Eichler, N-glycosylation in the thermoacidophilic archaeon Sulfolobus acidocaldarius involves a short dolichol pyrophosphate carrier, FEBS Lett. 590 (2016) 3168–3178.

[62] G. Palmieri, M. Balestrieri, J. Peter-Katalinic, G. Pohlentz, M. Rossi, I. Fiume, G. Pocsfalvi, Surface-exposed glycoproteins of hyperthermophilic Sulfolobus solfataricus P2 show a common N-glycosylation profile, J. Proteome Res. 12 (2013) 2779–2790.

[63] M. van Wolferen, A. Shajahan, K. Heinrich, S. Brenzinger, I.M. Black, A. Wagner, A. Briegel, P. Azadi, S.V. Albers, Species-specific recognition of Sulfolobales mediated by UV-inducible pili and S-layer glycosylation patterns, mBio 11 (2020) e03014–e03019.

[64] L. Kaminski, M.N. Lurie-Weinberger, T. Allers, U. Gophna, J. Eichler, Phylogenetic- and genome-derived insight into the evolution of N-glycosylation in Archaea, Mol. Phylogenet. Evol. 68 (2013) 327–339.

[65] M. Sumper, Halobacterial glycoprotein biosynthesis, Biochim. Biophys. Acta 906 (1987) 69–79.

[66] J. Kelly, S.M. Logan, K.F. Jarrell, D.J. VanDyke, E. Vinogradov, A novel N-linked flagellar glycan from Methanococcus maripaludis, Carbohydr. Res. 344 (2009) 648–653.

[67] S.Y. Ng, J. Wu, D.B. Nair, S.M. Logan, A. Robotham, L. Tessier, J.F. Kelly, K. Uchida, S. Aizawa, K.F. Jarrell, Genetic and mass spectrometry analyses of the unusual type IV-like pili of the archaeon Methanococcus maripaludis, J. Bacteriol. 193 (2011) 804–814.

[68] M. Erra-Pujada, P. Debeire, F. Duchiron, M.J. O'Donohue, The type II pullulanase of Thermococcus hydrothermalis: molecular characterization of the gene and expression of the catalytic domain, J. Bacteriol. 181 (1999) 3284–3287.

[69] G. Greller, R. Riek, W. Boos, Purification and characterization of the heterologously expressed trehalose/maltose ABC transporter complex of the hyperthermophilic archaeon Thermococcus litoralis, Eur. J. Biochem. 268 (2001) 4011–4018.

[70] T. Hettmann, C.L. Schmidt, S. Anemuller, U. Zahringer, H. Moll, A. Petersen, G. Schafer, Cytochrome b558/566 from the archaeon Sulfolobus acidocaldarius. A novel highly glycosylated, membrane-bound b-type hemoprotein, J. Biol. Chem. 273 (1998) 12032–12040.

[71] R. Horlacher, K.B. Xavier, H. Santos, J. DiRuggiero, M. Kossmann, W. Boos, Archaeal binding protein-dependent ABC transporter: molecular and biochemical analysis of the trehalose/maltose transport system of the hyperthermophilic archaeon Thermococcus litoralis, J. Bacteriol. 180 (1998) 680–689.

[72] S.M. Koning, M.G. Elferink, W.N. Konings, A.J. Driessen, Cellobiose uptake in the hyperthermophilic archaeon Pyrococcus furiosus is mediated by an inducible, high-affinity ABC transporter, J. Bacteriol. 183 (2001) 4979–4984.

[73] S.V. Albers, M.G. Elferink, R.L. Charlebois, C.W. Sensen, A.J. Driessen, W.N. Konings, Glucose transport in the extremely thermoacidophilic Sulfolobus solfataricus involves a high-affinity membrane-integrated binding protein, J. Bacteriol. 181 (1999) 4285–4291.

[74] M.G. Elferink, S.V. Albers, W.N. Konings, A.J. Driessen, Sugar transport in Sulfolobus solfataricus is mediated by two families of binding protein-dependent ABC transporters, Mol. Microbiol. 39 (2001) 1494–1503.

[75] S. Voisin, R.S. Houliston, J. Kelly, J.R. Brisson, D. Watson, S.L. Bardy, K.F. Jarrell, S.M. Logan, Identification and characterization of the unique N-linked glycan common to the flagellins and S-layer glycoprotein of Methanococcus voltae, J. Biol. Chem. 280 (2005) 16586–16593.

[76] V.Da Cunha, M. Gaïa, P. Forterre, The expanding Asgard archaea and their elusive relationships with Eukarya, mLife 1 (2022) 3–12.

[77] M. Igura, N. Maita, J. Kamishikiryo, M. Yamada, T. Obita, K. Maenaka, D. Kohda, Structure-guided identification of a new catalytic motif of oligosaccharyltransferase, EMBO J. 27 (2008) 234–243.

[78] S. Okuda, Y. Watanabe, Y. Moriya, S. Kawano, T. Yamamoto, M. Matsumoto, T. Takami, D. Kobayashi, N. Araki, A.C. Yoshizawa, T. Tabata, N. Sugiyama, S. Goto, Y. Ishihama, jPOSTrepo: an international standard data repository for proteomes, Nucleic Acids Res. 45 (2017) D1107–D1111.