

Published in final edited form as:

Mol Microbiol. 2011 September ; 81(5): 1166–1177. doi:10.1111/j.1365-2958.2011.07781.x.

Different routes to the same ending: comparing the N-glycosylation processes of *Haloferax volcanii* and *Haloarcula marismortui*, two halophilic archaea from the Dead Sea

Doron Calo¹, Ziqiang Guan², Shai Naparstek¹, and Jerry Eichler^{1,*}

¹Department of Life Sciences, Ben Gurion University of the Negev, Beersheva 84105, Israel.

²Department of Biochemistry, Duke University Medical Center, Durham, NC 27710, USA

Summary

Recent insight into the N-glycosylation pathway of the haloarchaeon, *Haloferax volcanii*, is helping to bridge the gap between our limited understanding of the archaeal version of this universal post-translational modification and the better-described eukaryal and bacterial processes. To delineate as yet undefined steps of the *Hfx. volcanii* N-glycosylation pathway, a comparative approach was taken with the initial characterization of N-glycosylation in *Haloarcula marismortui*, a second haloarchaeon also originating from the Dead Sea. While both species decorate the reporter glycoprotein, the S-layer glycoprotein, with the same N-linked pentasaccharide and employ dolichol phosphate as lipid glycan carrier, species-specific differences in the two N-glycosylation pathways exist. Specifically, *Har. marismortui* first assembles the complete pentasaccharide on dolichol phosphate and only then transfers the glycan to the target protein, as in the bacterial N-glycosylation pathway. In contrast, *Hfx. volcanii* initially transfers the first four pentasaccharide subunits from a common dolichol phosphate carrier to the target protein and only then delivers the final pentasaccharide subunit from a distinct dolichol phosphate to the N-linked tetrasaccharide, reminiscent of what occurs in eukaryal N-glycosylation. This study further indicates the extraordinary diversity of N-glycosylation pathways in Archaea, as compared with the relatively conserved parallel processes in Eukarya and Bacteria.

Introduction

Proteins from all three domains of life, i.e. Eukarya, Bacteria and Archaea, can experience N-glycosylation, namely the covalent attachment of glycans to select Asn residues (Helenius and Aebi, 2004; Eichler and Adams, 2005; Szymanski and Wren, 2005; Weerapana and Imperiali, 2006; Abu-Qarn *et al.*, 2008a; Calo *et al.*, 2010a). At present, the archaeal version of this universal post-translational modification is not as well understood as the parallel processes in Eukarya or Bacteria. Recent studies on the halophilic archaeon, *Haloferax volcanii*, have, however, begun to close this gap.

Haloferax volcanii cells are surrounded by a proteinaceous S-layer comprising a single component, the S-layer glycoprotein (Kessel *et al.*, 1988; Sumper *et al.*, 1990; Trachtenberg *et al.*, 2000). In addition to other post-translational processing events, such as signal sequence cleavage (Fine *et al.*, 2006) and lipid modification (Konrad and Eichler, 2002), the *Hfx. volcanii* S-layer glycoprotein experiences both N- and O-glycosylation (Sumper *et al.*, 1990). N-glycosylation of the S-layer glycoprotein is catalysed by a series of clustered *agl*

genes involved in the biosynthesis and attachment of a pentasaccharide comprising a hexose, two hexuronic acids, a methyl ester of hexuronic acid and a mannose to at least two Asn residues of the protein. AglJ, AglG, AglI, AglE are glycosyltransferases that sequentially add the first four pentasaccharide residues to a common dolichol phosphate carrier, while the dolichol phosphate mannose synthase, AglD, adds the final pentasaccharide residue to a distinct dolichol phosphate (Abu-Qarn *et al.*, 2007; 2008b; Yurist-Doutsch *et al.*, 2008; 2010; Guan *et al.*, 2010; Kaminski *et al.*, 2010; Magidovich *et al.*, 2010). AglB, the archaeal oligosaccharyl-transferase (Abu-Qarn and Eichler, 2006; Chaban *et al.*, 2006; Igura *et al.*, 2008), subsequently delivers glycans to selected S-layer glycoprotein Asn residues (Abu-Qarn *et al.*, 2007). AglF, AglM and AglP also contribute to the N-glycosylation process, respectively, acting as a glucose-1-phosphate uridylyltransferase, as a UDP-glucose dehydrogenase and as a *S*-adenosyl-L-methionine-dependent methyltransferase (Magidovich *et al.*, 2010; Yurist-Doutsch *et al.*, 2010).

While significant advances have been made in delineating the *Hfx. volcanii* N-glycosylation pathway, much is still unknown. For instance, the route taken by the final pentasaccharide residue, mannose, in reaching the S-layer glycoprotein remains unclear. Two possibilities can be entertained. In the first, the mannose residue is transferred from its dedicated dolichol phosphate carrier to the glycan comprising the first four pentasaccharide subunits already loaded onto a distinct dolichol phosphate. The resulting pentasaccharide would then be delivered to the S-layer glycoprotein *en bloc* via the actions of AglB. Alternatively, the final mannose residue could be added to the tetrasaccharide already N-linked to the S-layer glycoprotein.

Answering this and other outstanding questions related to *Hfx. volcanii* N-glycosylation may benefit from efforts addressing how other Archaea perform this post-translational modification. Also originating from the Dead Sea (Oren *et al.*, 1990), *Haloarcula marismortui* encodes for a predicted S-layer glycoprotein. Like its *Hfx. volcanii* and *Halobacterium salinarum* counterparts (cf. Calo *et al.*, 2010a), the *Har. marismortui* S-layer glycoprotein is also likely to be N-glycosylated, given the presence of sequons, namely the Asn-X-Ser/Thr (X = Pro) motif that serves the site of archaeal N-glycosylation (Abu-Qarn and Eichler, 2007; Igura and Kohda, 2011). Such expected similarities in the *Hfx. volcanii* and *Har. marismortui* S-layer glycoproteins, as well as the availability of complete genome sequences for both organisms (Baliga *et al.*, 2004; Hartman *et al.*, 2010), invite comparison of N-glycosylation in the two geographically and phylogenetically proximal archaeal species in the hope of shedding new light on the archaeal version of this post-translational modification.

In the following, N-glycosylation of the *Har. marismortui* S-layer glycoprotein is addressed in terms of the glycan N-linked to the proteins, the genes encoding putative enzymes involved in N-glycosylation in this species and the dolichol phosphate carriers upon which the N-linked glycan is assembled, relative to the parallel aspects of the *Hfx. volcanii* N-glycosylation pathway. The results reveal that while the *Har. marismortui* S-layer glycoprotein is modified by the identical N-linked glycan that decorates its *Hfx. volcanii* counterpart, differences in the two N-glycosylation pathways exist. Thus, these findings serve not only to further delineate steps of the *Hfx. volcanii* N-glycosylation pathway but also reveal diversity in N-glycosylation pathways across Archaea.

Results

Initial characterization of the Har. marismortui S-layer glycoprotein

The S-layer glycoprotein is the sole component of the surface layer surrounding *Hfx. volcanii* cells (Kessel *et al.*, 1988; Sumper *et al.*, 1990; Trachtenberg *et al.*, 2000). *Har.*

marismortui encodes for a protein annotated as a S-layer glycoprotein (pNG5138) that is 45% identical to its *Hfx. volcanii* homologue (Fig. 1). The amino acid sequence of the *Har. marismortui* S-layer glycoprotein includes five putative N-glycosylation sites, i.e. N265, N467, N575, N727 and N786 (amino acid numbering before signal peptide cleavage). One of these residues, N467, coincides with a putative N-glycosylation site present but not modified in the *Hfx. volcanii* S-layer glycoprotein, i.e. N370 (Abu-Qarn *et al.*, 2007). Otherwise, there is little resemblance in the distribution of putative N-glycosylation sites in the two proteins. Like the *Hfx. volcanii* S-layer glycoprotein, the *Har. marismortui* S-layer glycoprotein also presents a substantial number of Thr residues near the C-terminal end of the protein. In the *Hfx. volcanii* S-layer glycoprotein, many of these Thr residues are thought to undergo O-glycosylation (Sumper *et al.*, 1990). Finally, both S-layer glycoproteins contain identical putative *trans*-membrane domains predicted to terminate four residues from the C-terminus, assigned to the cell interior.

Biochemical characterization of the *Har. marismortui* S-layer glycoprotein began with separation of the protein content of *Har. marismortui* cells by SDS-PAGE. Subsequent Coomassie staining revealed a major protein species that migrates to a similar position as does the *Hfx. volcanii* S-layer glycoprotein (Fig. 2, left panel, arrow). Glycosylation of the *Har. marismortui* S-layer glycoprotein was demonstrated by PAS glycostaining (Fig. 2, middle panel) and upon incubation of *Har. marismortui* cells with [2-³H]mannose. It was previously been shown that *Hfx. volcanii* cells incubated with [2-³H]mannose incorporate the radiolabelled saccharide into the S-layer glycoprotein, reflecting that mannose is the final residue of the pentasaccharide N-linked to the *Hfx. volcanii* protein (Calo *et al.*, 2011). As with the *Hfx. volcanii* S-layer glycoprotein, radiolabel was also incorporated into the S-layer glycoprotein of *Har. marismortui* cells incubated with [2-³H]mannose (Fig. 2, right panel).

***Har. marismortui* contains pentasaccharide-linked dolichol phosphate**

Guan *et al.* (2010) previously reported that the first four residues of the pentasaccharide N-linked to the *Hfx. volcanii* S-layer glycoprotein are derived from a tetrasaccharide-charged dolichol phosphate carrier, while the final pentasaccharide residue, mannose, is derived from a distinct dolichol phosphate carrier. To assess whether dolichol phosphate also serves as the lipid glycan carrier in *Har. marismortui* N-glycosylation, a total *Har. marismortui* lipid extract was examined by normal-phase liquid chromatography-electrospray ionization/mass spectrometry (LC-ESI/MS) performed in the negative ion mode.

As observed upon examination of a total *Hfx. volcanii* lipid extract (Fig. 3A, upper panel), mass spectrometric analysis of a *Har. marismortui* lipid extract also revealed [M-2H]²⁻ ion peaks at *m/z* 776.44 (this and all subsequently reported *m/z* values given are for the monoisotopic peak) and 810.48, respectively, corresponding to C₅₅ and C₆₀ dolichol phosphate species modified by a hexose, two hexuronic acids and a methyl ester of hexuronic acid (Fig. 3A, lower panel), namely the first four residues of the pentasaccharide N-linked to the *Hfx. volcanii* S-layer glycoprotein. However, in contrast to what was seen in the *Hfx. volcanii* glycan-charged dolichol phosphate pool, the *Har. marismortui* sample additionally included [M-2H]²⁻ ion peaks at *m/z* 857.47 and 891.5 (Fig. 3A, lower panel), representing C₅₅ and C₆₀ dolichol phosphate modified by the same tetrasaccharide plus an additional 162 Da hexose residue.

Analysis of the tandem mass spectrometry (MS/MS) fragmentation pattern obtained from the *m/z* 891.5 species confirmed its assignment as pentasaccharide-modified C₆₀ dolichol phosphate. Such analysis revealed product ions of *m/z* 1079.760, corresponding to C₆₀ dolichol phosphate modified by a hexose (calculated monoisotopic mass of 1079.8 Da, as shown in the fragmentation scheme of Fig. 3B, inset), *m/z* 1255.78, corresponding to C₆₀

dolichol phosphate modified by a hexose-hexuronic acid disaccharide (calculated monoisotopic mass of 1255.8 Da), m/z 527.11, corresponding to a hexose-methyl ester of hexuronic acid-hexuronic acid trisaccharide (calculated monoisotopic mass of 527.1 Da), and m/z 369.10, corresponding to a hexosemethyl ester of hexuronic acid disaccharide (calculated monoisotopic mass of 369.1 Da) (Fig. 3B). Hence, the monoisotopic $[M-2H]^{2-}$ ion peak detected at m/z 891.50 (monoisotopic mass of 1785.02 Da; Fig. 3A, lower panel) is in good agreement with calculated mass of C₆₀ dolichol phosphate modified by a pentasaccharide comprising a hexose, two hexuronic acids, the methyl ester of hexuronic acid and a final hexose (1784.98 Da).

The *Har. marismortui* and *Hfx. volcanii* S-layer glycoproteins are N-glycosylated by the same pentasaccharide

To determine whether the *Har. marismortui* S-layer glycoprotein is N-glycosylated by the same pentasaccharide bound to dolichol phosphate, chymotrypsin-generated fragments of the S-layer glycoprotein were examined by nano-flow, reverse-phase LC-ESI/MS. Such analysis revealed a peak of m/z 1183.86, corresponding to the $[M+2H]^{2+}$ ion of a S-layer glycoprotein-derived peptide containing the N-glycosylation site, Asn-467 [⁴⁶⁵DVNGTANQGVDQVAL⁴⁷⁸ (amino acid numbering before signal peptide cleavage)], and modified by a pentasaccharide comprising two hexoses, two hexuronic acids and the methyl ester of hexuronic acid (Fig. 4). This value is in good agreement with calculated mass of the pentasaccharide-modified peptide (1183.8608 Da). As such, it can be concluded that not only does dolichol phosphate serve as the glycan lipid carrier in *Har. marismortui* N-glycosylation, as is the case in *Hfx. volcanii* (Guan *et al.*, 2010; Kaminski *et al.*, 2010), but that the S-layer glycoproteins of both species are modified by the same N-linked pentasaccharide. However, in contrast to what has been observed in *Hfx. volcanii*, where S-layer glycoprotein modified by pentasaccharide precursors are detected, in *Har. marismortui*, only S-layer glycoprotein modified by the intact pentasaccharide was detected.

***agl* genes are clustered in the *Hfx. volcanii* and the *Har. marismortui* genomes**

In *Hfx. volcanii*, all but one of *agl* genes involved in the assembly and attachment of the pentasaccharide N-linked to the S-layer glycoprotein are sequestered in a region of the genome between *HVO_1517* (*aglJ*) and *HVO_1531* (*aglM*); *aglD* (*HVO_0798*) lies elsewhere in the genome (Yurist-Doutsch and Eichler, 2009; Yurist-Doutsch *et al.*, 2010). Given the similarity in N-linked glycosylation, the existence of similar clustering of putative N-glycosylation genes in *Har. marismortui* was considered. Accordingly, the region surrounding *rrnAC0431*, encoding the homologue of the oligosaccharyltransferase, AglB, was analysed. When *Har. marismortui* *rrnAC0419*, *rrnAC0427* and *rrnAC0430* were used to query the *Hfx. volcanii* genome in a BLAST search, these were revealed to respectively encode homologues of *Hfx. volcanii* AglJ, AglR and AglG. Likewise, when the reverse was performed and *Hfx. volcanii* *agl* sequences were used to query the *Har. marismortui* genome in a BLAST search, it was revealed that *Har. marismortui* *rrnAC0443*, annotated as a hypothetical protein, is a homologue of *Hfx. volcanii* AglP (*e*-value of $8e^{-21}$). The *Har. marismortui* gene cluster encoding AglB, AglG, AglJ and AglR homologues also includes *rrnAC0421* and *rrnAC0429*, encoding proteins annotated as glycosyltransferases. Using the same bioinformatics strategy, it was further revealed that *Har. marismortui* also encodes for homologues of *Hfx. volcanii* AglE (pNG7005, *e*-value of $1e^{-15}$), AglII (pNG7005, *e*-value of $1e^{-11}$), AglF (*rrnAC3238*, *e*-value of $3e^{-119}$) and AglM (*rrnAC3239*, *e*-value of $9e^{-159}$) (best hits, according to *e*-values, are listed). Finally, no *Har. marismortui* homologue of *Hfx. volcanii* *aglQ* was detected. Figure 5A shows a schematic comparison of the distribution and orientation of *agl* genes and other putative N-glycosylation sequences near *aglB* and elsewhere in the two genomes.

As noted above, *aglD* (*HVO_0798*) does not cluster with genes encoding the other known *Hfx. volcanii* Agl proteins (Yurist-Doutsch and Eichler, 2009). The *Har. marismortui* homologue of AglD, *rrnAC1873* (Calo *et al.*, 2010b), is also found in a region of the genome distinct from the *aglB*-based cluster or other *agl* homologues. Moreover, comparing the sequences near *Hfx. volcanii aglD* and *Har. marismortui rrnAC1873* reveals that homologous genes are found clustered in the immediate vicinities of the two, albeit in opposing orientations. Specifically, *Hfx. volcanii* *HVO_0787*, *HVO_0788*, *HVO_0789*, *HVO_0790*, *HVO_0792*, *HVO_0794* and *HVO_0795* correspond to *Har. marismortui* *rrnAC1886*, *rrnAC1885*, *rrnAC1884*, *rrnAC1881*, *rrnAC1879*, *rrnAC1876* and *rrnAC1875* (Fig. 5B).

Differences exist in the *Har. marismortui* and *Hfx. volcanii* N-glycosylation processes

Although pentasaccharide-modified dolichol phosphate was readily detected in *Har. marismortui*, no such species was detected in *Hfx. volcanii* (Fig. 3 and Guan *et al.*, 2010), even when a fivefold larger aliquot of the total *Hfx. volcanii* lipid extract was considered (not shown). To test whether this was due to the rapid delivery of lipid-bound pentasaccharide to the S-layer glycoprotein in *Hfx. volcanii*, the glycan-charged dolichol phosphate pool was considered in *Hfx. volcanii* cells deleted of *aglB*. In the absence of the oligosaccharyltransferase, AglB, N-glycosylation of the *Hfx. volcanii* S-layer glycoprotein does not transpire (Abu-Qarn *et al.*, 2007), presumably because the complete pentasaccharide or its tetrasaccharide precursor is not delivered to the protein from a dolichol phosphate carrier. Hence, any pentasaccharide-charged dolichol phosphate formed in *Hfx. volcanii* would be expected to accumulate in cells deleted of *aglB*. However, when total lipid extracts were prepared from *Hfx. volcanii* $\Delta aglB$ cells and analysed by normal-phase LC-ESI/MS, it was noted that as with the parent strain, only tetrasaccharide-modified C₅₅ and C₆₀ dolichol phosphate species were detected (Fig. 6A).

To eliminate the possibility that no pentasaccharide-modified *Hfx. volcanii* dolichol phosphate was detected as this entity simply could not elute from the silica resin used for the liquid chromatography step, LC-ESI/MS was repeated, this time using less hydrophilic aminopropyl resin for the initial separation step. While such analysis of the total *Hfx. volcanii* lipid extract yielded singly charged [M-H]⁻ ion peaks at *m/z* 1553.91 and 1621.97, respectively, corresponding to tetrasaccharide-charged C₅₅ and C₆₀ dolichol phosphate, no pentasaccharide-modified lipid carriers were detected (Fig. 6B, top panel). Likewise, no pentasaccharide-modified dolichol phosphate was detected when the total lipid extract prepared from $\Delta aglB$ cells was similarly examined (Fig. 6B, middle panel). In contrast, a total *Har. marismortui* lipid extract analysed by aminopropyl resin-based LC-ESI/MS not only contained the tetrasaccharide-charged species but also singly charged [M-H]⁻ ion peaks at *m/z* 1715.95 and 1781.91, corresponding to pentasaccharide-charged C₅₅ and C₆₀ dolichol phosphate respectively (Fig. 6B, bottom panel).

Finally, to determine whether the final sugar residue of the pentasaccharide-modified dolichol phosphate detected in *Har. marismortui*, i.e. mannose, is derived from a soluble nucleoside-charged species or from a dolichol phosphate-bound precursor, the monosaccharide-charged C₆₀ dolichol phosphate ion peak was detected in the *Har. marismortui* and *Hfx. volcanii* lipid extracts during a single LC/MS run, over a 40 min span. As reflected in Fig. 7, the single *m/z* 1079.8 ion peak from each strain could be resolved into three distinct species, distinguished by their slightly different LC retention times. Earlier efforts (Guan *et al.*, 2010; Kaminski *et al.*, 2010) had shown the third of these *Hfx. volcanii* peaks as corresponding to mannose-charged dolichol phosphate. As such, it is likely that the 16.80 min peak in the *Har. marismortui* sample also corresponds to mannose-charged dolichol phosphate.

Discussion

In the present study, it was revealed that like *Hfx. volcanii* and *Hbt. salinarum* (cf. Calo *et al.*, 2010a), *Har. marismortui* is also capable of performing N-glycosylation. Specifically, examination of the *Hfx. volcanii* and *Har. marismortui* S-layer glycoproteins showed that both reporter glycoproteins are modified by the same N-linked pentasaccharide comprising two hexoses, two hexuronic acids and the methyl ester of hexuronic acid. Of course, more detailed future analysis of each glycan may reveal the use of distinct hexoses or hexuronic acids in each case, although both N-linked pentasaccharides were shown to include at least one mannose.

In *Hfx. volcanii*, a series of clustered *agl* gene products are involved in the assembly and attachment of the N-linked pentasaccharide (cf. Calo *et al.*, 2010a). While the *Har. marismortui* genes involved in N-glycosylation remain to be defined, it appears that *Har. marismortui* has clustered genes encoding homologues of several Agl proteins in the immediate region of the oligosaccharyltransferase-encoding *aglB* sequence. Such clustering of *agl* genes allows for more co-ordinated control of N-glycosylation gene expression in response to changes in growth environment, as suggested by earlier quantitative PCR analysis assessing the transcription of a subset of *Hfx. volcanii* *agl* genes in response to changing growth conditions (Yurist-Doutsch *et al.*, 2008; 2010). Previous efforts have also shown that *Hfx. volcanii* strains lacking either *aglD* or *aglJ* can be functionally complemented by introduction of their respective *Har. marismortui* homologues, i.e. rrnAC1873 and rrnAC0149 (Calo *et al.*, 2010b; 2011; D. Calo, unpublished), pointing to the common activities of these proteins in each organism. Moreover, the fact that those genes encoding *Har. marismortui* homologues of AglF and AglM are found adjacent to each other is of note, given *in vitro* studies showing that *Hfx. volcanii* AglF and AglM work in concert (Yurist-Doutsch *et al.*, 2010). Still, the number of putative *Har. marismortui* N-glycosylation genes in the vicinity of *aglB* does not match the number of clustered *agl* sequences in *Hfx. volcanii*. In fact, the *Har. marismortui* genome contains sequences encoding homologues of other Agl proteins scattered throughout the genome.

Despite the various N-glycosylation-related similarities of *Hfx. volcanii* and *Har. marismortui*, significant species-specific differences nonetheless exist. The most striking difference is seen at the level of the dolichol phosphates that serve as lipid glycan carriers in archaeal N-glycosylation (Mescher *et al.*, 1976; Lechner *et al.*, 1985; Hartmann and König, 1989; Guan *et al.*, 2010). The present study finds that the pentasaccharide N-linked to the *Har. marismortui* S-layer glycoprotein is first fully assembled on a single dolichol phosphate and only then transferred to the protein target. This strategy differs from what is seen in *Hfx. volcanii*, where the same S-layer glycoprotein N-linked pentasaccharide is derived from a tetrasaccharide sequentially assembled on a single dolichol phosphate and a final mannose residue derived from a distinct dolichol phosphate carrier (Guan *et al.*, 2010). As such, *Har. marismortui* N-glycosylation is reminiscent of the parallel bacterial process, where a heptasaccharide is assembled through the sequential addition of seven soluble nucleotide-activated sugars onto a common undecaprenol phosphate carrier (Szymanski and Wren, 2005; Weerapana and Imperiali, 2006; Abu-Qarn *et al.*, 2008a). In contrast, the *Hfx. volcanii* N-glycosylation process, involving multiple glycan-charged dolichol phosphates, is reminiscent of the parallel eukaryal process. In higher Eukarya, the first seven sugar subunits of the 14-member oligosaccharide eventually formed in the endoplasmic reticulum are sequentially added to a common dolichol phosphate carrier, while the second set of seven sugar subunits are derived from individual hexose-charged dolichol phosphates (Burda and Aebi, 1999; Helenius and Aebi, 2004). However, despite these pathway differences, the final mannose subunit of the pentasaccharide N-linked to the *Har.*

marismortui S-layer glycoprotein is derived from a distinct dolichol phosphate carrier, as in *Hfx. volcanii*.

The *Hfx. volcanii* N-glycosylation process further relies on a degree of sophistication until now only seen in Eukarya. The fact that tetrasaccharide-modified but no pentasaccharide-modified dolichol phosphate could be detected in *Hfx. volcanii* cells, even in cells lacking the oligosaccharyltransferase, AgIB, points to the final pentasaccharide subunit, mannose, as being added to the tetrasaccharide already N-linked to the S-layer glycoprotein. This same general strategy is employed in Eukarya, where in the Golgi, additional sugar subunits are attached to oligosaccharides already N-linked to the target polypeptide. However, since no S-layer glycoprotein modified by pentasaccharide precursors was detected in *Har. marismortui*, and since significantly more tetrasaccharide- than pentasaccharide-modified S-layer glycoprotein is seen in *Hfx. volcanii* (Abu-Qarn *et al.*, 2007), it would appear that Archaea pay a price in N-glycosylation efficiency for adopting the eukaryal strategy of adding sugar subunits to protein-linked glycans. This could explain why Eukarya have spatially separated the two N-glycosylation strategies, assigning transfer of the dolichol phosphate-bound oligosaccharide to a protein target to the endoplasmic reticulum, and then having the subsequent introduction of more protein-specific sugar subunits occur in the Golgi (Helenius and Aebi, 2004). Of course, the inability to detect *Har. marismortui* S-layer glycoprotein-derived glycopeptides modified by pentasaccharide precursors could be due to other considerations, such as AgIB specificity or differing efficiencies in modifying the different N-glycosylation sites in the protein.

In a recent analysis of 56 archaeal genome sequences, *agIB*, encoding the archaeal oligosaccharyltransferase, was detected in all but two species (Magidovich and Eichler, 2009). Although experimental verification for N-glycosylation is still lacking for the majority of these species, it would appear that this post-translational modification is common in Archaea. Continued analysis of the archaeal version of N-glycosylation will thus not only provide additional insight into this universal proteinprocessing event but will also elucidate strategies adopted by extremophiles to cope with the physically challenging environments in which they live.

Experimental procedures

Strains and growth conditions

Haloflex volcanii WR536 (H53) cells were grown in medium containing 3.4 M NaCl, 0.15 M MgSO₄·7H₂O, 1 mM MnCl₂, 4 mM KCl, 3 mM CaCl₂, 0.3% (w/v) yeast extract, 0.5% (w/v) tryptone, 50 mM Tris-HCl, pH 7.2, at 42°C (Mevarech and Werczberger, 1985). *Har. marismortui* cells were grown in 3.5 M NaCl, 0.19 M MgSO₄·7H₂O, 3 mM CaCl₂, 1 mM MnCl₂ and 1% (w/v) yeast extract (Mevarech *et al.*, 1976).

[2-³H]-mannose radiolabelling

Haloflex volcanii and *Har. marismortui* cells grown to midexponential phase were incubated with 6 μl of [2-³H]-mannose (23.8 mCi mmol⁻¹; PerkinElmer, Boston, MA) in a volume of 100 μl. Sixty minutes later, the samples were precipitated with 15% (w/v) trichloroacetic acid and separated by SDS-PAGE. The S-layer glycoprotein was identified by Coomassie staining and fluorography and exposure to film.

LC/MS and MS/MS

LC-ESI/MS/MS analysis of total *Hfx. volcanii* and *Har. marismortui* lipid extracts was performed as described previously (Kaminski *et al.*, 2010), with one modification. In some

cases, liquid chromatographic separation was achieved using a Unison UK-Amino column (250 × 2 mm; Imtakt, Philadelphia, PA).

For LC-ESI/MS analysis of *Hfx. volcanii* and *Har. marismortui* S-layer glycoproteins, the protein contents of *Hfx. volcanii* and *Har. marismortui* cells were separated on 7.5% polyacrylamide gels and stained with Coomassie R-250 (Fluka, St. Louis, MO). For in-gel digestion of the *Hfx. volcanii* S-layer glycoprotein, the relevant bands (identified via the unique SDS-PAGE migration and staining pattern of the protein) were excised, destained in 400 μ l of 50% (v/v) acetonitrile (Sigma, St Louis, MO) in 40 mM NH_4HCO_3 , pH 8.4, dehydrated with 100% acetonitrile and dried using a Speed-Vac drying apparatus. The S-layer glycoprotein was reduced with 10 mM dithiothreitol (Sigma) in 40 mM NH_4HCO_3 at 56°C for 60 min and then alkylated for 45 min at room temperature with 55 mM iodoacetamide in 40 mM NH_4HCO_3 . The gel pieces were washed with 40 mM NH_4HCO_3 for 15 min, dehydrated with 100% acetonitrile, and SpeedVac dried. The gel slices were rehydrated with 12.5 ng μl^{-1} mass spectrometry (MS)-grade Trypsin Gold (Promega, Madison, WI) in 40 mM NH_4HCO_3 . In the case of the *Har. marismortui* S-layer glycoprotein, the same protocol was followed except that the protein was digested with 12.5 ng μl^{-1} sequencing-grade chymotrypsin (Promega). In both cases, protease-generated peptides were extracted with 0.1% (v/v) formic acid in 20 mM NH_4HCO_3 , followed by sonication for 20 min at room temperature, dehydration with 50% (v/v) acetonitrile and additional sonication. After three rounds of extraction, the gel pieces were dehydrated with 100% acetonitrile, dried completely with a SpeedVac, resuspended in 5% (v/v) acetonitrile containing 1% formic acid (v/v) and infused into the mass spectrometer using static nanospray Econotips (New Objective, Woburn, MA). The protein digests were separated on-line by nano-flow reverse-phase liquid chromatography (LC) by loading onto a 150 mm by 75 μm (internal diameter) by 365 μm (external diameter) Jupifer preppacked fused silica 5 μm C_{18} 300 Å reverse-phase column (Thermo Fisher Scientific, Bremen, Germany). The sample was eluted into the LTQ Orbitrap XL mass spectrometer (Thermo Fisher Scientific) using a 60 min linear gradient of 0.1% formic acid (v/v) in acetonitrile/0.1% formic acid (1:19, by volume) to 0.1% formic acid in acetonitrile/0.1% formic acid (4:1, by volume) at a flow rate of 300 nl min^{-1} .

Other methods

Glycostaining of the S-layer glycoprotein by periodic acidSchiff reagent was performed as previously described (Dubray and Bezar, 1982).

Acknowledgments

J.E. is supported by the Israel Science Foundation (Grant 30/07). The mass spectrometry facility in the Department of Biochemistry of the Duke University Medical Center and Z.G. are supported by the LIPID MAPS Large Scale Collaborative Grant No. GM-069338 from NIH. S.N. is the recipient of a Negev-Zin Associates Scholarship.

References

- Abu-Qarn M, Eichler J. Protein N-glycosylation in Archaea: defining *Haloferox volcanii* genes involved in S-layer glycoprotein glycosylation. *Mol Microbiol.* 2006; 61:511–525. [PubMed: 16762024]
- Abu-Qarn M, Eichler J. An analysis of amino acid sequences surrounding archaeal glycoprotein sequons. *Archaea.* 2007; 2:73–81. [PubMed: 17350928]
- Abu-Qarn M, Yurist-Doutch S, Giordano A, Trauner A, Morris HR, Hitchen P, et al. *Haloferox volcanii* AglB and AglD are involved in N-glycosylation of the S-layer glycoprotein and proper assembly of the surface layer. *J Mol Biol.* 2007; 14:1224–1236. [PubMed: 17996897]

- Abu-Qarn M, Eichler J, Sharon N. Not just for Eukarya anymore: N-glycosylation in Bacteria and Archaea. *Curr Opin Struct Biol.* 2008a; 18:544–550. [PubMed: 18694827]
- Abu-Qarn M, Giordano A, Battaglia F, Trauner A, Morris HR, Hitchen P, et al. Identification of AglE, a second glycosyltransferase involved in N-glycosylation of the *Haloflex volcanii* S-layer glycoprotein. *J Bacteriol.* 2008b; 190:3140–3146. [PubMed: 18310347]
- Baliga NS, Bonneau R, Facciotti MT, Pan M, Glusman G, Deutsch EW, et al. Genome sequence of *Haloarcula marismortui*: a halophilic archaeon from the Dead Sea. *Genome Res.* 2004; 14:2221–2234. [PubMed: 15520287]
- Burda P, Aebi M. The dolichol pathway of N-linked glycosylation. *Biochim Biophys Acta.* 1999; 1426:239–257. [PubMed: 9878760]
- Calo D, Kaminski L, Eichler J. Protein glycosylation in Archaea: sweet and extreme. *Glycobiology.* 2010a; 20:1065–1079. [PubMed: 20371512]
- Calo D, Eilam Y, Lichtenstein RG, Eichler J. Towards glyco-engineering in Archaea: replacing *Haloflex volcanii* AglD with homologous glycosyltransferases from other halophilic archaea. *Appl Environ Microbiol.* 2010b; 76:5684–5692. [PubMed: 20601508]
- Calo D, Guan Z, Eichler J. Glyco-engineering in Archaea: differential N-glycosylation of the S-layer glycoprotein in a transformed *Haloflex volcanii* strain. *Microbiol Biotechnol.* 2011; 4:461–470.
- Chaban B, Voisin S, Kelly J, Logan SM, Jarrell KF. 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.* 2006; 61:259–268. [PubMed: 16824110]
- Dubray G, Bezar G. A highly sensitive periodic acid-silver stain for 1,2-diol groups of glycoproteins and polysaccharides in polyacrylamide gels. *Anal Biochem.* 1982; 119:325–329. [PubMed: 6176144]
- Eichler J, Adams MWW. Post-translational protein modification in Archaea. *Microbiol Mol Biol Rev.* 2005; 69:393–425. [PubMed: 16148304]
- Fine A, Irihimovitch V, Dahan I, Konrad Z, Eichler J. Cloning, expression and purification of functional Sec11a and Sec11b, type I signal peptidases of the archaeon *Haloflex volcanii*. *J Bacteriol.* 2006; 188:1911–1919. [PubMed: 16484202]
- Guan Z, Naparstek S, Kaminski L, Konrad Z, Eichler J. Distinct glycan-charged phosphodolichol carriers are required for the assembly of the pentasaccharide N-linked to the *Haloflex volcanii* S-layer glycoprotein. *Mol Microbiol.* 2010; 78:1294–1303. [PubMed: 21091511]
- Hartman AL, Norais C, Badger JH, Delmas S, Haldenby S, Madupu R, et al. The complete genome sequence of *Haloflex volcanii* DS2, a model archaeon. *PLoS ONE.* 2010; 5:e9605. [PubMed: 20333302]
- Hartmann E, Konig H. Uridine and dolichyl diphosphate activated oligosaccharides are intermediates in the biosynthesis of the S-layer glycoprotein of *Methanothermobacter ferredoxinus*. *Arch Microbiol.* 1989; 151:274–281.
- Helenius A, Aebi M. Roles of N-linked glycans in the endoplasmic reticulum. *Annu Rev Biochem.* 2004; 73:1019–1049. [PubMed: 15189166]
- Igura M, Kohda D. Quantitative assessment of the preferences for the amino acid residues flanking archaeal N-linked glycosylation sites. *Glycobiology.* 2011; 21:575–583. [PubMed: 21115605]
- Igura M, Maita N, Kamishikiryo J, Yamada M, Obita T, Maenaka K, Kohda D. Structure-guided identification of a new catalytic motif of oligosaccharyltransferase. *EMBO J.* 2008; 27:234–243. [PubMed: 18046457]
- Kaminski L, Abu-Qarn M, Guan Z, Naparstek S, Ventura VV, Raetz CRH, et al. AglJ adds the first sugar of the N-linked pentasaccharide decorating the *Haloflex volcanii* S-layer glycoprotein. *J Bacteriol.* 2010; 192:5572–5579. [PubMed: 20802039]
- Kessel M, Wildhaber I, Cohen S, Baumeister W. Three-dimensional structure of the regular surface glycoprotein layer of *Halobacterium volcanii* from the Dead Sea. *EMBO J.* 1988; 7:1549–1554. [PubMed: 16453843]
- Konrad Z, Eichler J. Lipid modification of proteins in Archaea: attachment of a mevalonic acid-based lipid moiety to the surface-layer glycoprotein of *Haloflex volcanii* follows protein translocation. *Biochem J.* 2002; 366:959–964. [PubMed: 12069685]

- Lechner J, Wieland F, Sumper M. 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.* 1985; 260:860–866. [PubMed: 2857171]
- Magidovich H, Eichler J. Glycosyltransferases and oligosaccharyltransferases in Archaea: putative components of the N-glycosylation pathway in the third domain of life. *FEMS Microbiol Lett.* 2009; 300:122–130. [PubMed: 19765088]
- Magidovich H, Yurist-Doutsch S, Konrad Z, Ventura VV, Hitchen PG, Dell A, Eichler J. AgIP is a S-adenosyl-L-methionine-dependent methyltransferase that participates in the N-glycosylation pathway of *Haloferax volcanii*. *Mol Microbiol.* 2010; 76:190–199. [PubMed: 20149102]
- Mescher MF, Hansen U, Strominger JL. Formation of lipid-linked sugar compounds in *Halobacterium salinarum*. Presumed intermediates in glycoprotein synthesis. *J Biol Chem.* 1976; 251:7289–7294. [PubMed: 1002689]
- Mevarech M, Werczberger R. Genetic transfer in *Halobacterium volcanii*. *J Bacteriol.* 1985; 162:461–462. [PubMed: 3980444]
- Mevarech M, Leicht W, Werber MM. Hydrophobic chromatography and fractionation of enzymes from extremely halophilic bacteria using decreasing concentration gradients of ammonium sulfate. *Biochemistry.* 1976; 15:2383–2387. [PubMed: 1276145]
- Oren A, Ginzburg M, Ginzburg BZ, Hochstein LI, Volcani BE. *Haloarcula marismortui* (*Volcani*) sp. nov., nom. rev., an extremely halophilic bacterium from the Dead Sea. *Int J Syst Bacteriol.* 1990; 40:209–210. [PubMed: 11536469]
- Sumper M, Berg E, Mengele R, Strobel I. Primary structure and glycosylation of the S-layer protein of *Haloferax volcanii*. *J Bacteriol.* 1990; 172:7111–7118. [PubMed: 2123862]
- Szymanski CM, Wren BW. Protein glycosylation in bacterial mucosal pathogens. *Nat Rev Microbiol.* 2005; 3:225–237. [PubMed: 15738950]
- Trachtenberg S, Pinnick B, Kessel M. The cell surface glycoprotein layer of the extreme halophile *Halobacterium salinarum* and its relation to *Haloferax volcanii*: cryo-electron tomography of freeze-substituted cells and projection studies of negatively stained envelopes. *J Struct Biol.* 2000; 130:10–26. [PubMed: 10806087]
- Weerapana E, Imperiali B. Asparagine-linked protein glycosylation: from eukaryotic to prokaryotic systems. *Glycobiology.* 2006; 16:91R–101R.
- Yurist-Doutsch S, Eichler J. 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.* 2009; 191:3068–3075. [PubMed: 19251857]
- Yurist-Doutsch S, Abu-Qarn M, Battaglia F, Morris HR, Hitchen PG, Dell A, Eichler J. *aglF*, *aglG* and *aglH*, novel members of a gene cluster involved in the N-glycosylation of the *Haloferax volcanii* S-layer glycoprotein. *Mol Microbiol.* 2008; 69:1234–1245. [PubMed: 18631242]
- Yurist-Doutsch S, Magidovich H, Ventura VV, Hitchen PG, Dell A, Eichler J. N-glycosylation in Archaea: on the coordinated actions of *Haloferax volcanii* AgIF and AgIM. *Mol Microbiol.* 2010; 75:1047–1058. [PubMed: 20487296]

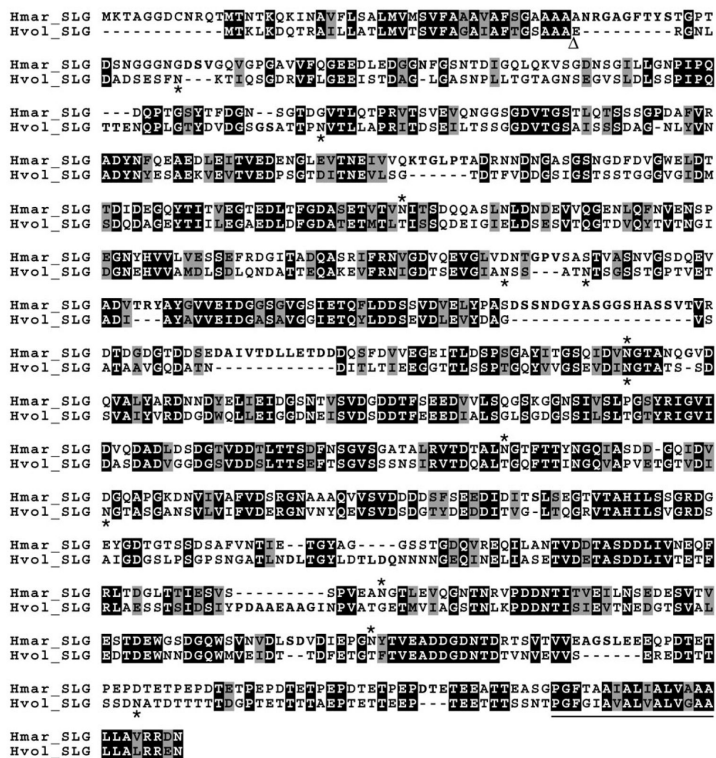


Fig. 1. Alignment of the *Har. marismortui* and *Hfx. volcanii* S-layer glycoprotein amino acid sequences. The sequences of the *Har. marismortui* and *Hfx. volcanii* S-layer glycoproteins [AAV44550 (Hmar_SLG) and ADE05144 (Hvol_SLG) respectively] were aligned using CLUSTALW and drawn using Boxshade, programs both found at the Biology Workbench (<http://workbench.sdsc.edu/>). In the figure, identical positions are drawn on a black background, while similar residues are drawn on a grey background. Verified or predicted N-glycosylation sites are identified by asterisks either above (*Har. marismortui*) or below (*Hfx. volcanii*) the sequences. The predicted *trans*-membrane domain of each sequence is underlined, while the known site of signal sequence cleavage in the *Hfx. volcanii* S-layer glycoprotein is depicted by an open triangle.

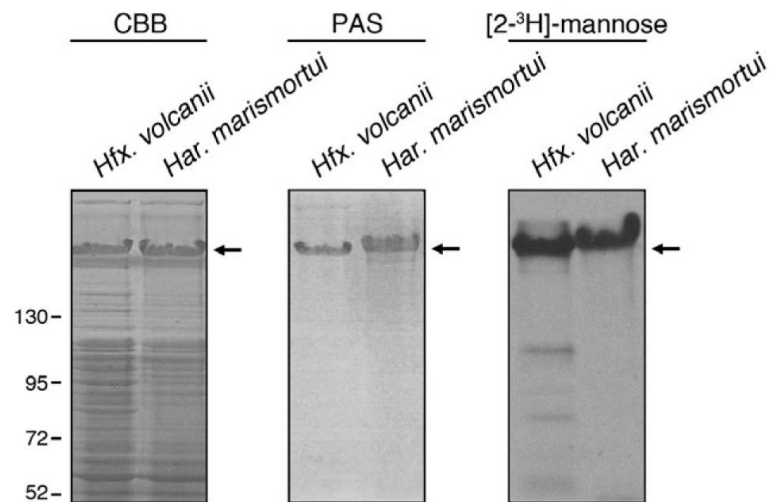


Fig. 2.

Like its *Hfx. volcanii* counterpart, the *Har. marismortui* S-layer protein is glycosylated. The protein contents of *Hfx. volcanii* (left) and *Har. marismortui* (right) cells grown to identical OD550 values separated by 7.5% SDS-PAGE and either Coomassie or PAS-stained [left (CBB) and middle panels respectively]. In addition, the protein contents of *Hfx. volcanii* (left) and *Har. marismortui* (right) cells challenged with [2-³H]-mannose were separated by 7.5% SDS-PAGE and processed for fluorography (right panel). In each panel, the *Hfx. volcanii* and *Har. marismortui* S-layer glycoproteins are indicated by an arrow. Molecular mass markers are shown on the left of the figure.

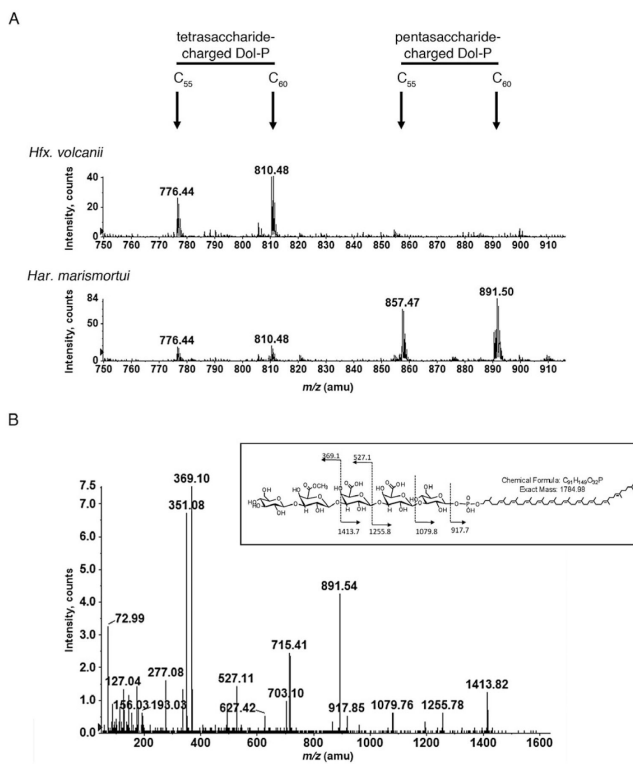


Fig. 3. Pentasaccharide-modified dolichol phosphate is detected in a total lipid extract from *Har. marismortui* but not from *Hfx. volcanii*. A. LC-ESI/MS analysis of the total lipid extract from *Hfx. volcanii* (upper panel) and *Har. marismortui* (lower panel) reveals the presence of tetrasaccharide-modified C55 and C60 dolichol phosphate in both extracts. In contrast, pentasaccharide-modified dolichol phosphate is only seen in the *Har. marismortui* sample. The peaks shown correspond to $[M-2H]^{2-}$ ions. The positions of the different glycan-modified dolichol phosphates are indicated. B. MS/MS analysis of *Har. marismortui* pentasaccharide-modified C60 dolichol phosphate. The fragmentation profile is schematically represented in the inset.

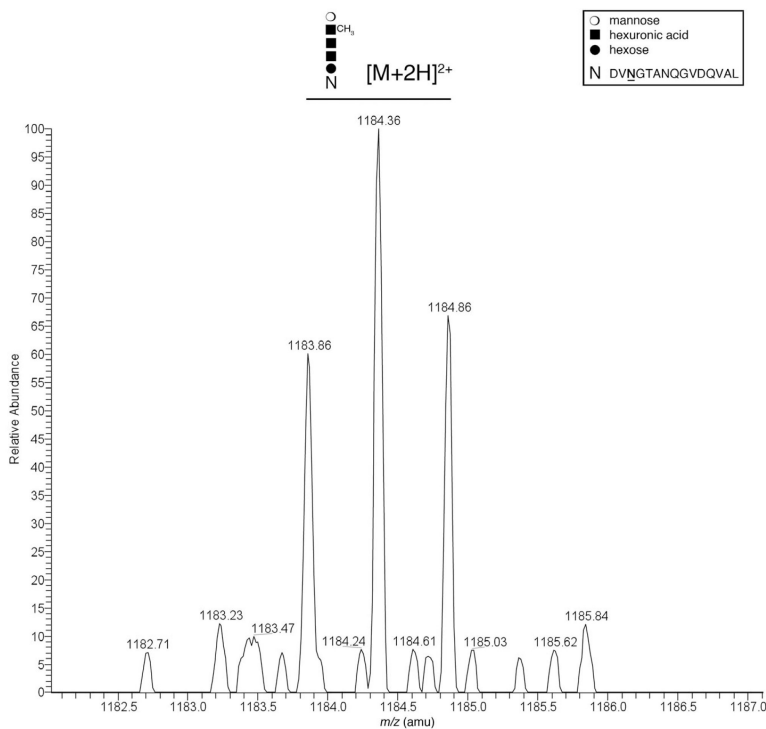


Fig. 4.

The *Har. marismortui* S-layer protein is modified by the same N-linked pentasaccharide as bound to dolichol phosphate and the *Hfx. volcanii* S-layer protein. A monoisotopic $[M+2H]^{2+}$ ion peak corresponding to a *Har. marismortui* S-layer glycoprotein Asn-467 (amino acid numbering before signal peptide cleavage)-containing chymotrypsin-generated fragment, as detected by LC-ESI/MS. The mass-to-charge ratio of the doubly charged peak observed (1183.86 Da) is in good agreement with calculated mass of the peptide modified by a pentasaccharide comprising two hexoses, two hexuronic acids and the methyl ester of hexuronic acid (1183.8608 Da), schematically depicted above the peak. The inset depicts the symbols used for each sugar residue as well as the peptide sequence represented by 'N'. The modified Asn residue in the peptide is indicated in bold and underlined.

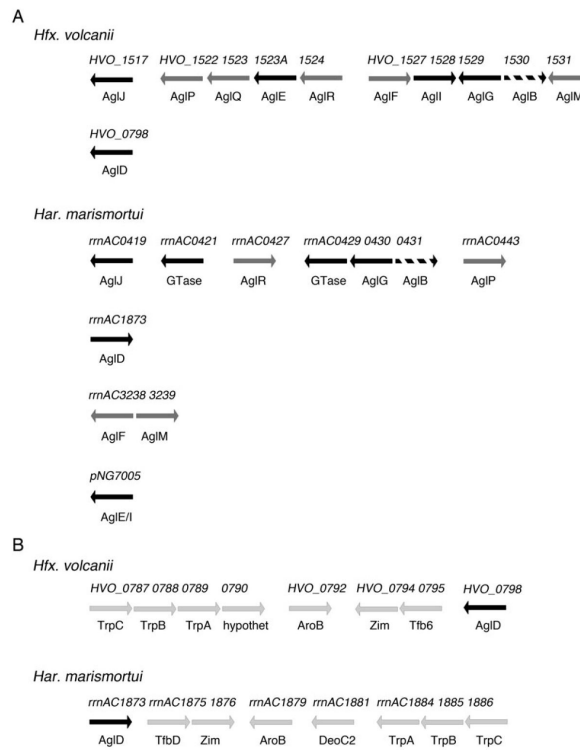
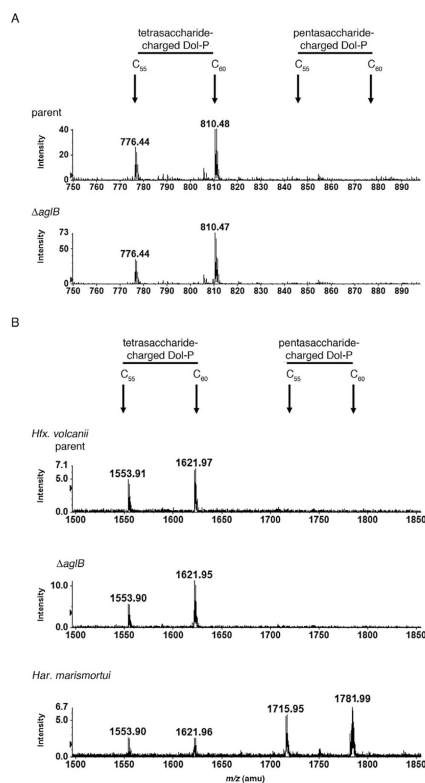


Fig. 5. *agl* genes in *Hfx. volcanii* and their putative *Har. marismortui* homologues are clustered to different degrees.

A. Regions of the *Hfx. volcanii* genome encoding Agl proteins and of the *Har. marismortui* genome encoding predicted Agl protein homologues are schematically presented. In addition, genes encoding predicted glycosyltransferases (GTase) near *Har. marismortui aglB* are shown. In each case, the genome identifier is provided above the gene while the known (*Hfx. volcanii*) or predicted (*Har. marismortui*) product is given below.

B. Homologous genes in the immediate vicinity of *Hfx. volcanii aglD* and its *Har. marismortui* homologue, *rrnAC1873*, are schematically shown. Below each gene, the current annotation is provided. In both (A) and (B), the full identifier is provided for only the first gene of a group of adjacent genes; only numbers are provided for the subsequent genes. Otherwise, the full identifier is provided. The size of each gene is considered to be equal in the scheme, as are the distances between proximal yet non-adjacent genes. Genes encoding the known (*Hfx. volcanii*) or predicted (*Har. marismortui*) oligosaccharyltransferase, AglB, are hatched, genes encoding known or predicted glycosyltransferases are shown in black and genes encoding proteins known or predicted to contribute to N-glycosylation are shown in dark grey. Genes encoding proteins apparently not related to N-glycosylation are shown in light grey.

**Fig. 6.**

No pentasaccharide-modified dolichol phosphate accumulates in *Hfx. volcanii* cells lacking AglB nor is detected using aminopropyl resin for liquid chromatography.

A. Tetrasaccharide-modified C₅₅ and C₆₀ dolichol phosphate are detected upon LC-ESI/MS analysis of the total lipid extract from *Hfx. volcanii* parent strain (upper panel) and $\Delta aglB$ cells (lower panel). In neither case is pentasaccharide-modified dolichol phosphate detected. The peaks shown correspond to doubly charged $[M-2H]^{2-}$ ions, with the positions of the different glycan-modified dolichol phosphates indicated.

B. Tetrasaccharide-modified C₅₅ and C₆₀ dolichol phosphate are detected upon LC-ESI/MS analysis of the total lipid extract from *Hfx. volcanii* parent strain (upper panel) and $\Delta aglB$ cells (middle panel) using aminopropyl resin for the liquid chromatography step. In neither case is pentasaccharide-modified dolichol phosphate detected. In contrast, pentasaccharide-modified C₅₅ and C₆₀ dolichol phosphates are detected upon similar analysis of the total lipid extract from *Har. marismortui* cells (lower panel). The peaks shown correspond to $[M-H]^{-}$ ions, with the positions of the different glycan-modified dolichol phosphates indicated.

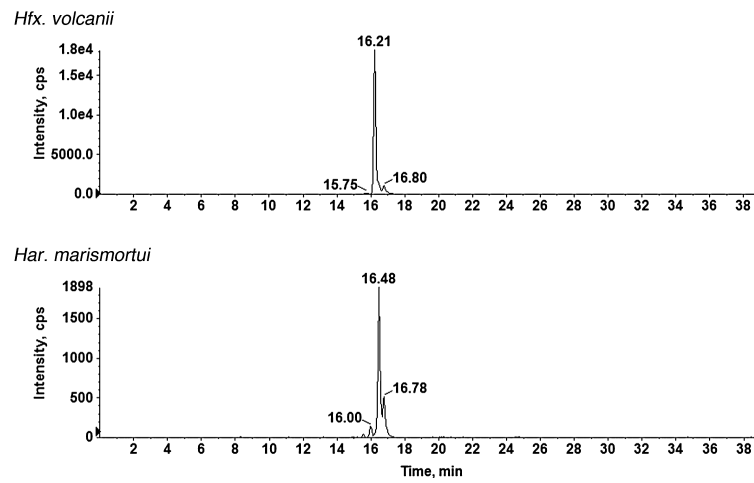


Fig. 7.

Har. marismortui contains mannose-charged dolichol phosphate. Normal-phase LC extracted ion chromatograms (EIC) of the dolichylphosphate-hexose $[M-H]^-$ ion at m/z 1079.8 from *Hfx. volcanii* (upper panel) and from *Har. marismortui* (lower panel) are shown. The *Hfx. volcanii* peak retained at 16.80 min was previously shown to correspond to mannose-charged dolichol phosphate (Guan *et al.*, 2010).