Article

Uncovering the prevalence, key biogenesis enzymes, and biological significance of archaeal lipoproteins

Yirui Hong [1], Kira S. Makarova [2], Andy A. Garcia [3], Rachel Xu[1], Friedhelm Pfeiffer [4], Paula V. Welander [3] & Mechthild Pohlschroder [1] ✉

Lipid-anchored proteins are integral components of cell surfaces. In bacteria, lipidation of proteins with a conserved lipobox motif ($[L/V/I]^{-3}$ $[A/S/T/V/I]^{-2}$ $[G/A/S]^{-1}$ $[C]^{+1}$) is catalyzed by prolipoprotein diacylglyceryl transferase (Lgt). Although lipobox-containing proteins, or lipoproteins, are predicted to be abundant in several archaeal species, no archaeal homologs of Lgt have been identified, suggesting distinct lipidation enzymes evolved in archaea to accommodate their unique membrane lipids. Here, we predicted lipoprotein presence for all major archaeal lineages and revealed a high prevalence of lipoproteins across the domain Archaea. Using comparative genomics, we identified a comprehensive set of candidates for archaeal lipoprotein biogenesis components (Ali). Genetic and biochemical characterization in the archaeon *Haloferax volcanii* confirmed that two paralogous genes, *aliA* and *aliB*, are important for lipoprotein lipidation. Moreover, deletion of both genes led to a complete absence of diphytanylglyceryl thioether from lipoprotein extracts, revealing the chemical nature of lipid anchors in *Hfx. volcanii* lipoproteins. Disruption of AliA- and AliB-mediated lipoprotein lipidation caused severe growth defects, decreased motility, and cell-shape alterations, underscoring the importance of lipoproteins in archaeal cell physiology. Notably, AliA and AliB exhibit distinct, non-redundant enzymatic activities with potential substrate selectivity, uncovering a new layer of regulation in prokaryotic lipoprotein lipidation.

Archaea, once thought to be exclusively associated with extreme environments, are ubiquitous[1] and important for ecological processes, biotechnology, and human health[2–6]. Despite their morphological similarities with bacteria, archaea and bacteria differ significantly, including in the composition and stereochemistry of their membrane lipids. Most bacteria have fatty acyl chains ester-linked to glycerol-3-phosphate backbones as their membrane lipids, whereas archaeal membrane lipids predominantly have isoprenoid-based alkyl chains that are ether-linked to glycerol-1-phosphate backbones[7–10]. However, both archaea and bacteria use membrane lipids to anchor surface proteins, such as the C-terminally anchored archaeosortase substrates in archaea and the N-terminally anchored lipoproteins in bacteria[11,12]. Lipoproteins, here referred to those lipidated proteins featuring a conserved motif known as lipobox, are involved in a variety of cellular processes in bacteria, ranging from virulence and nutrient transport to cell architecture maintenance, and have therefore been studied extensively[13,14]. Located at the C-terminus of an N-terminal signal peptide, the lipobox ($[L/V/I]^{-3}$ $[A/S/T/V/I]^{-2}$ $[G/A/S]^{-1}$ $[C]^{+1}$)[15], especially

[1]Department of Biology, University of Pennsylvania, Philadelphia, PA, USA. [2]Division of Intramural Research, National Library of Medicine, National Institutes of Health, Bethesda, MD, USA. [3]Department of Earth System Science, Stanford University, Stanford, CA, USA. [4]Computational Systems Biochemistry, Max Planck Institute of Biochemistry, Martinsried, Germany. ✉e-mail: pohlschr@sas.upenn.edu

the conserved cysteine at the +1 position, has been shown to be essential for lipoprotein biogenesis. Depending on the sequence of the signal peptide, lipoprotein precursors are translocated either in an unfolded state via the general secretory (Sec) pathway or in a folded state via the twin-arginine translocation (Tat) pathway[16]. The bacterial prolipoprotein diacylglyceryl transferase (Lgt) then catalyzes the formation of a thioether bond between the conserved lipobox cysteine and a membrane lipid moiety, thus anchoring the lipoprotein to the membrane[17]. Lipoprotein signal peptidase (Lsp) subsequently cleaves the signal peptide N-terminal to the conserved cysteine, generating the mature form of some lipoproteins[18]. In other cases, the α-amino group of the cysteine can be further acylated by lipoprotein N-acyltransferase (Lnt)[19], lipoprotein N-acyltransferase system A and B (LnsAB)[20], or lipoprotein intramolecular transacylase (Lit)[21].

Lipoproteins are also predicted to be major components of archaeal cell surfaces. In the model archaeon *Haloferax volcanii*, 121 of 316 secreted proteins were predicted to be lipoproteins based on the presence of a lipobox, implying the significant roles of lipoproteins in archaeal cell biology[22]. Substitution of the conserved lipobox cysteine with a serine, a common technique used to determine whether a protein is a lipoprotein[23], prevented the maturation of several predicted *Hfx. volcanii* lipoproteins and led to protein mislocalization in some cases[24,25]. Furthermore, while the chemical nature of the lipid attached to archaeal lipoproteins remains unknown, mass spectrometry analyses of a lipobox-containing halocyanin from *Natronomonas pharaonis* suggest three potential modifications: attachment of a lipid − diphytanylglycerol diether (archaeol) − to the lipobox cysteine, cleavage of the N-terminal signal peptide, and further acetylation of the cysteine[26], similar to the modifications of bacterial lipoproteins. Despite the accumulating evidence for a lipobox-dependent protein anchoring mechanism in archaea, components underlying such a mechanism have remained enigmatic. Notably, no archaeal homologs of bacterial Lgt or Lsp have been identified, suggesting distinct enzymes have evolved as an adaptation to the unique membrane lipids of archaea.

In this study, we aimed to investigate the significance of lipoproteins in archaeal cell biology and to identify key components involved in archaeal lipoprotein biogenesis. Through comparative genomic analyses, we predicted a high prevalence and abundance of lipoproteins across the domain Archaea and identified a comprehensive set of candidate archaeal lipoprotein biogenesis components (Ali). Among these, two paralogous proteins, coined AliA and AliB, were particularly promising due to their conserved residues and visual similarities with Lgt. These proteins were further characterized in *Hfx.*

volcanii because of its high lipoprotein abundance and genetic tractability[27]. Using this model archaeon, we confirmed that AliA and AliB are important for archaeal lipoprotein lipidation. Furthermore, single deletions of *aliA* or *aliB*, as well as the double deletion of both genes, significantly impacted archaeal growth, motility, and cell shape, with Δ*aliA* displaying more severe phenotypes than Δ*aliB*. Overall, our findings establish the pivotal roles of lipoproteins in archaeal cell biology and resolve the long-standing search for lipoprotein biogenesis enzymes in archaea, thereby opening new avenues for the study of prokaryotic lipoproteins.

Results

Lipoproteins are widespread across archaea

Previous in silico studies have predicted a high abundance of lipoproteins in several species within Euryarchaeota[22,24] (now referred to as Methanobacteriota under the International Code of Nomenclature of Prokaryotes[28]), but the lipoprotein prevalence across archaea remains unknown. To address this gap, we analyzed 524 archaeal genomes from the Archaeal Clusters of Orthologous Genes (arCOGs) database[29,30], covering all major archaeal lineages. Putative Sec and Tat lipoproteins as well as other secreted proteins were identified using SignalP 6.0[31], and the percentage of secreted proteins that were predicted as lipoproteins was calculated for each genome. The results show that lipoproteins are prevalent not only in Euryarchaeota but also in the DPANN superphylum (Fig. 1a, Supplementary Fig. 1 and Supplementary Data 1). Conversely, relatively few are found in TACK and Asgard−the two phyla more closely related to eukaryotes, in which lipobox-containing proteins have not been identified. In some archaeal species, especially the Euryarchaeal halobacteria species, lipoproteins are predicted to constitute more than 50% of secreted proteins, implying their significant roles in archaeal physiology. Since both Sec and Tat pathways can be used for archaeal lipoprotein transport[24,25], we investigated whether different archaeal species show distinct preferences for these two pathways. Consistent with previous findings[24], the prediction results indicate that halobacterial species predominantly use the Tat pathway for lipoprotein translocation (Fig. 1b), potentially as a strategy to avoid protein misfolding in their high-salt environment. All other archaeal species primarily use the Sec pathway for lipoprotein translocation.

Computational identification of candidate enzymes for archaeal lipoprotein lipidation

Considering that no archaeal homologs of bacterial Lgt or Lsp were previously identified using PSI-BLAST[32] or Dali[33], we reasoned that
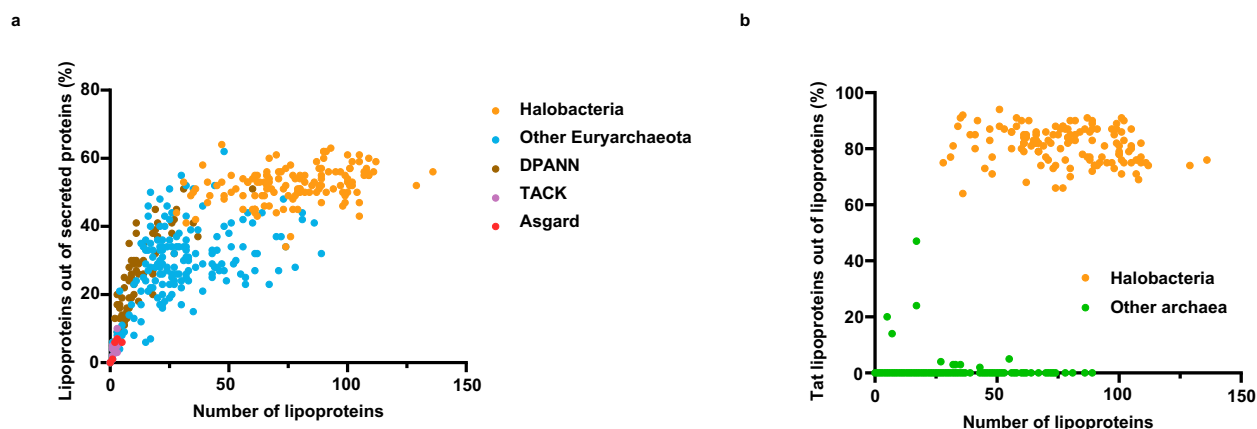

Fig. 1 | Lipoproteins are widespread across archaea and can be translocated via both Sec and Tat pathways. a The percentage of lipoproteins out of total secreted proteins for all archaeal genomes analyzed (n = 524). Each point represents the data for one genome. Euryarchaeota is an archaeal phylum that does not fall within a superphylum, while DPANN, TACK, and Asgard are the three archaeal superphyla. Halobacteria is a class in the Euryarchaeota phylum. **b** The percentage of Tat lipoproteins out of all lipoproteins for all archaeal genomes analyzed (n = 524).

a

b

E.coli Lgt, PDB: 5AZB   HVO_2859, AliA, AF2 model   HVO_2611, AliB, AF2 model

c

d

uncharacterized archaeal membrane proteins, which frequently co-occur with predicted lipoproteins across archaeal genomes, could be promising candidates for further investigation. Therefore, we calculated a lipoprotein co-occurrence score for each arCOG by comparing its presence and absence across genomes with that of lipoproteins (Supplementary Data 2). Among the 25 arCOGs with the highest scores, only two membrane proteins have unassigned functions: arCOG02142,

a predicted membrane-associated metalloprotease of the Tiki superfamily[34], and arCOG02177, an uncharacterized membrane protein. arCOG02177 proteins are broadly conserved in Euryarchaeota with several gaps in Thermoplasmata and Methanobacteria, and are also present in the DPANN superphylum with some gaps (Fig. 2a, Supplementary Data 1 and 3). Examination of multiple alignments of arCOG02177 representatives revealed several conserved residues,

Fig. 2 | Bioinformatic analyses identified arCOG02177 and arCOG02178 proteins as potential lipoprotein biogenesis components. a Phyletic pattern of predicted lipoproteins compared with phyletic patterns of arCOG02177 and arCOG02178. For each genome, gene presence is shown as a vertical bar color-coded according to the major archaeal lineages indicated above. **b** Comparison of *E. coli* Lgt structure (PDB: 5AZB) and AlphaFold2 models of HVO_2859 and HVO_2611. Conserved residues in two different pockets are highlighted in red for one and magenta for the other. Dashed circles indicate similar subdomains between the Lgt structure and Alpha-Fold2 models. **c** HMM-HMM alignment of arCOG02177 (query) and arCOG02178 (target) showing the conserved residues. Complete alignments are shown in Datasets S4 and S5. The lower case corresponds to less conserved positions. The symbols between the query and the target reflect amino acid similarity between the two multiple alignments as follows: | indicates the mostly identical amino acids in respective alignment columns; +, very similar amino acids;. indicates positively scored amino acids; · indicates negatively scored amino acids; =, very different amino acids. **d** Conserved gene neighborhoods for arCOG02177 and arCOG02178. For each gene neighborhood, the species name, genome partition, and coordinates of the locus are indicated. Genes are depicted by block arrows, with the length roughly proportional to the gene size. arCOG02177 (*aliA*) and arCOG02178 (*aliB*) are shown in blue with a red outline. Lipoproteins are indicated by an asterisk. The genes in the neighborhoods are designated by respective arCOG numbers shown below the arrows, and gene or protein family names are indicated within the arrows. GldA glycerol dehydrogenase, IMP inositol monophosphatase family, MB uncharacterized metal binding protein, DeoC deoxyribose-phosphate aldolase, SpoIIM uncharacterized membrane protein, a component of a putative membrane remodeling system, YcaO ribosomal protein S12 methylthiotransferase accessory factor, RbsK sugar kinase, HTH helix-turn-helix protein.

suggesting potential enzymatic activities of these proteins (Fig. 2b and Supplementary Data 4). A HHpred search[35] showed only weak similarity (probability 29%) between arCOG02177 and Lgt (PF01790). A Dali search[33] against PDB database, using the AlphaFold2 model of *Hfx. volcanii* arCOG02177 representative HVO_2859 as a query, did not detect any structural similarity with Lgt. However, a visual comparison between the HVO_2859 model and the *Escherichia coli* Lgt structure revealed several similarities in the architecture of the two proteins (Fig. 2b). Lgt has two periplasmic subdomains, arm-1 and head[36], and similar structures are visible in HVO_2859. Moreover, the potential catalytic pocket of HVO_2859 contains two conserved arginines, which are also found in Lgt and are known to be catalytically important (Fig. 2b)[36,37]. This evidence suggests that arCOG02177 proteins might derive from the Lgt family. Based on comparison of HVO_2859 and Lgt topologies, we proposed two equally parsimonious evolutionary scenarios for the origin of arCOG02177 from Lgt. Both scenarios maintain the connectivity of arms and head subdomains while involving a few structural rearrangements and the emergence of alternative residues in the catalytic pocket (Supplementary Fig. 2). A distant paralog of arCOG02177 is encoded in all halobacterial genomes but is classified into a separate arCOG, arCOG02178, due to divergence (Fig. 2a, c and Supplementary Figs. 2, 3). Interestingly, both conserved arginine residues in arCOG02177 are replaced by histidines in arCOG02178, suggesting a functional difference between the two paralogs (Fig. 2b, c and Supplementary Data 4 and 5).

Phylogenetic analyses of arCOG02177 and arCOG02178 representatives showed that major clades are compatible with archaeal taxonomy with limited horizontal transfer (Supplementary Fig. 3). Such evolutionary behavior is characteristic of proteins involved in important cellular functions[38]. Additionally, gene neighborhood analyses show that arCOG02177 genes are often found nearby, and have the potential to be co-expressed with inositol monophosphatase family enzymes, predicted metal binding proteins, and glycerol dehydrogenase GldA, a key enzyme of archaeal lipid biosynthesis (Fig. 2d and Supplementary Data 6). arCOG02177 is also associated with arCOG01994, an uncharacterized membrane protein SpoIIM, which was predicted to be involved in membrane remodeling or vesicle formation[39]. arCOG02178 genes are linked to deoxyribose-phosphate aldolase DeoC and are often encoded in the vicinity of predicted lipoproteins (Fig. 2d and Supplementary Data 6). Taken together, these findings justified the continued investigation of arCOG02177 and arCOG02178 proteins as archaeal lipoprotein biogenesis components.

AliA and AliB are important for the biogenesis of both Sec and Tat lipoproteins in *Hfx. volcanii*

To verify the roles of arCOG02177 and arCOG02178 in lipoprotein biogenesis, we examined their functions in vivo using the model archaeon *Hfx. volcanii*. Three predicted lipoproteins were used as substrates, including one predicted to be translocated via the Sec pathway (HVO_1176) and two translocated via the Tat pathway (HVO_B0139 and HVO_1705) (Fig. 3a and Supplementary Fig. 4)[24,25]. Additionally, we included a Tat substrate (HVO_0844) that is anchored by a C-terminal transmembrane segment as a control (Fig. 3a)[25]. To confirm that HVO_1176, HVO_B0139, and HVO_1705 are lipoproteins, we mutated their lipobox cysteine to a serine, the amino acid with the most similar structural properties to cysteine (Supplementary Fig. 4). We then tagged both the mutant and wild-type proteins with C-terminal myc tags and overexpressed them in *Hfx. volcanii*. Western blots of cell lysates showed a significantly decreased amount of the three proteins with the Cys-to-Ser mutation (Fig. 3b). These mutant proteins were also not detected in the supernatant (Supplementary Fig. 5), suggesting that the Cys-to-Ser mutation may compromise protein stability. Additionally, mutant HVO_1176 and HVO_B0139 migrated more slowly than the mature wild-type proteins (Fig. 3b), suggesting the Cys-to-Ser mutation hindered their maturation. In summary, substituting the lipobox cysteine with a serine disrupted the maturation and reduced the stability of all three lipoproteins, consistent with the essential role of this conserved cysteine in the established bacterial lipoprotein biogenesis pathway.

Next, we generated single and double deletion strains of the *Hfx. volcanii* arCOG02177 gene (*hvo_2859*) and arCOG02178 gene (*hvo_2611*) to investigate the effect of these deletions on lipoprotein biogenesis. Western blot analyses of myc-tagged HVO_1176 showed that in the deletion strains, HVO_1176 migrated more slowly than in wild-type *Hfx. volcanii* (Fig. 3c). If archaeal lipoprotein biogenesis follows a pathway similar to that in bacteria, this slower migration could be attributed to defects in lipoprotein lipidation or signal peptide cleavage, which we collectively refer to as a lipoprotein biogenesis defect. Notably, HVO_1176 in the deletion strains migrated to a position comparable to that of the Cys-to-Ser HVO_1176 mutant (Fig. 3b,c), suggesting deletion of *hvo_2859* or *hvo_2611* completely disrupted the cysteine-associated modifications mentioned above. Similarly, deletion of *hvo_2859* or *hvo_2611* prevented the maturation of HVO_B0139, as indicated by the absence of mature proteins in the deletion strains (Fig. 3c). Interestingly, although the maturation of HVO_1705 seemed to be abolished in Δ*hvo_2859*, it was only partially affected in Δ*hvo_2611* (Fig. 3c), implying functional diversification of the two paralogous proteins. The biogenesis defects of these proteins were successfully complemented by expressing *hvo_2859* or *hvo_2611* in trans (Supplementary Fig. 6), confirming their roles in lipoprotein biogenesis. It is noteworthy that overexpressing *hvo_2859* in Δ*hvo_2611*, and vice versa, did not complement the lipoprotein biogenesis defect, confirming the non-redundant roles of *hvo_2611* and *hvo_2859* (Supplementary Fig. 6). As a control, the transmembrane protein HVO_0844 was also analyzed for its mobility rate. Although reduced levels of mature HVO_0844 were observed upon deletion of *hvo_2859*, maturation of the protein was not abolished in any of the mutants (Fig. 3c), underscoring that the observed biogenesis defects in the mutants are specific to lipoproteins. In summary, deletions of the *Hfx. volcanii* arCOG02177 gene
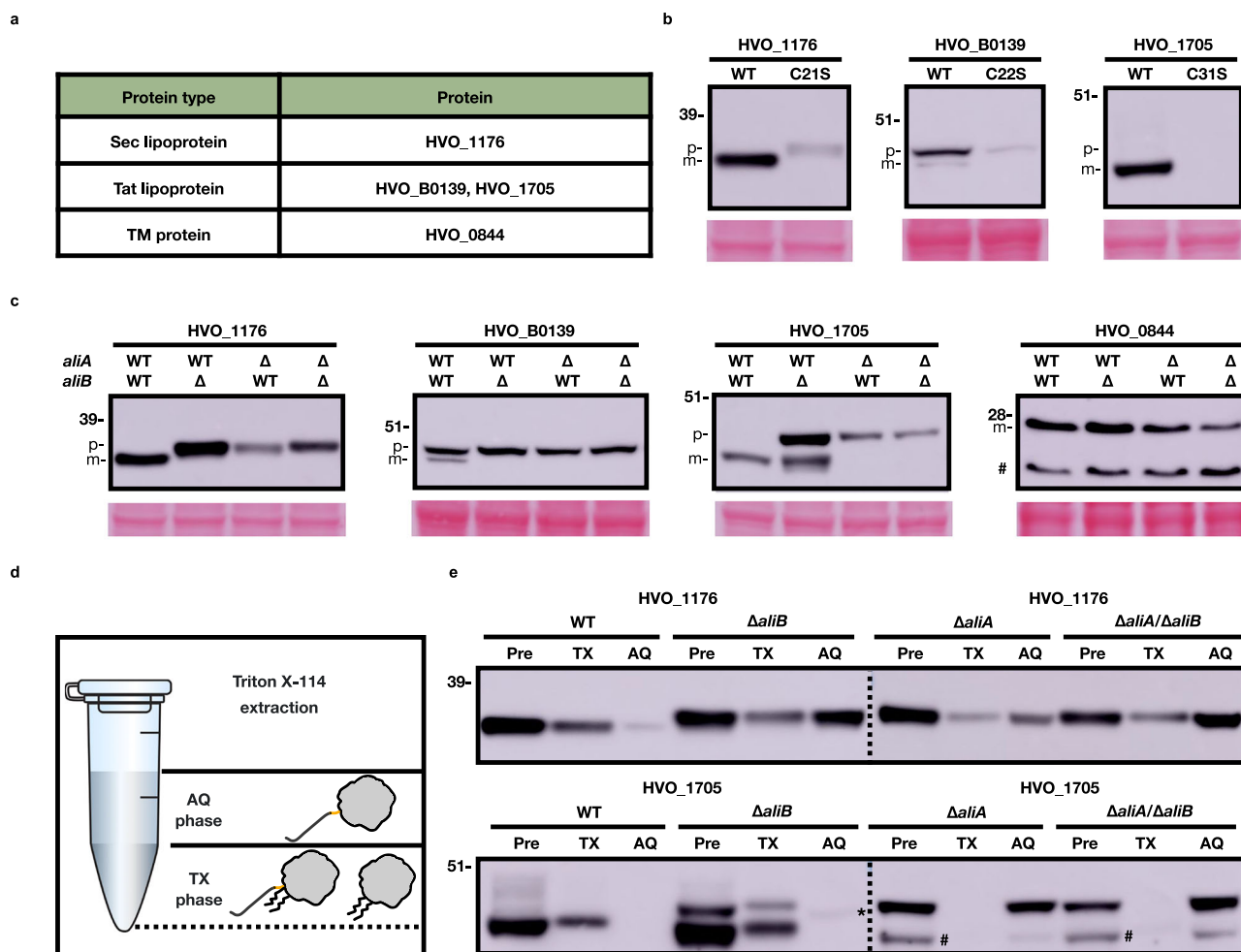
Fig. 3 | AliA and AliB are involved in Sec and Tat lipoprotein lipidation.
a Proteins studied in the mobility shift assay. TM, transmembrane. **b** Western blots of overexpressed myc-tagged wild-type lipoproteins (WT) and lipoproteins with the cysteine mutated to a serine (C21S, C22S, or C31S) in *Hfx. volcanii* whole cell lysates. For each lipoprotein, the same amount of total protein was loaded onto the gel across strains. Ponceau S staining of the PVDF membrane is shown at the bottom as a loading control. Results are representative of at least two independent experiments. Precursors (p) and mature proteins (m) are labeled accordingly. The positions of molecular weight markers are indicated by numbers on the left side of the blots. **c** Western blots of overexpressed myc-tagged lipoproteins in the whole cell lysate of different *Hfx. volcanii* strains. Results are representative of at least two independent experiments. *AliA, hvo_2859; aliB, hvo_2611*. WT indicates the presence of wild-type *ali*, and Δ indicates the deletion of the wild-type *ali* gene. #, suspected

protein degradation product. **d** The anticipated distribution of precursors and mature lipoproteins in the Triton X-114 extraction assay. Non-lipidated and non-cleaved precursors are expected to be in the aqueous (AQ) phase, while lipidated precursors and mature lipoproteins are expected to localize in the detergent (TX) phase. The lipobox was highlighted in yellow. The Eppendorf tube was adapted from a stock image created in BioRender. Hong, Y. (2025) https://BioRender.com/6bjbjo5. **e** Western blots of samples before and after the Triton X-114 extraction. For each lipoprotein, the same volume of TX and AQ samples were loaded onto the gel. Results are representative of at least two independent experiments. Pre, protein samples before extraction. Dotted lines indicate the border between two different gels. * indicates the protein position in the AQ phase of Δ*aliB*. Corresponding representative blots are included in the Source Data file.

(hvo_2859) and the arCOG02178 gene *(hvo_2611)* affect the maturation of all three lipoproteins investigated, suggesting both proteins are functional in *Hfx. volcanii* and are important for Sec and Tat lipoprotein biogenesis. Accordingly, we named arCOG02177 and arCOG02178 archaeal lipoprotein biogenesis components A and B (AliA and AliB), respectively.

AliA and AliB are involved in lipoprotein lipidation
Based on the bioinformatic and biochemical results above, we hypothesized that AliA and AliB are involved at the initial stage of the lipoprotein biogenesis pathway, specifically in lipoprotein lipidation. Thus, deletion of *aliA* or *aliB* would prevent both the lipidation and the subsequent signal peptide cleavage of lipoproteins. To test this hypothesis, we analyzed the lipidation status of the highly expressed proteins HVO_1176 and HVO_1705 in different strains using a Triton X-114 extraction assay, which is routinely used to assess protein

lipidation[16,40–43]. In this assay, the mixture of cell lysates and Triton X-114 solution separates into two phases: lipidated hydrophobic proteins are mainly associated with the detergent phase (TX phase), while non-lipidated hydrophilic proteins remain in the aqueous phase (AQ phase; Fig. 3d)[40,41]. The results indicated that in wild-type *Hfx. volcanii*, the lipidated mature HVO_1176 was predominantly found in the TX phase, with a minor fraction in the AQ phase (Fig. 3e), possibly due to carryover from the TX phase. In contrast, HVO_1176 precursors in Δ*ali* strains were primarily detected in the AQ phase (Fig. 3e), consistent with our prediction that these proteins lack lipid modifications and confirming that both AliA and AliB are required for HVO_1176 lipidation. A smaller, yet still detectable, portion of the non-lipidated precursors was present in the TX phase, possibly due to the strong hydrophobicity of their non-cleaved signal peptide (Supplementary Fig. 4), which promotes partitioning into the TX phase. Analysis of HVO_1705 revealed a more complex pattern. In wild type, mature

HVO_1705 was found in the TX phase (Fig. 3e), whereas in *aliA* deletion strains, HVO_1705 precursors were detected in the AQ phase, suggesting the precursors were both non-lipidated and non-cleaved. Interestingly, two distinct forms of HVO_1705 precursors were observed in Δ*aliB*. One form, which was lipidated but non-cleaved, remained in the TX phase. The other, which migrated faster, likely represented a non-lipidated and non-cleaved form present in the AQ phase (Fig. 3e). Collectively, these results revealed that deletion of *aliA* and *aliB* similarly affects the lipidation of HVO_1176 but has differing impacts on HVO_1705, confirming the important yet distinct roles of AliA and AliB in lipoprotein lipidation.

Lipoprotein lipidation by thioether-linked archaeol is abolished in Δ*aliA*/Δ*aliB*

Although previous studies have suggested the existence of a diphytanylglyceryl (archaeol) thioether modification on cysteine residues of archaeal proteins[44], including a lipobox-containing halocyanin from *Natronomonas pharaonis*[26], the substrate selectivity of this lipidation and the enzymes responsible have not been previously defined. Moreover, the chemical nature of the lipid anchor in archaeal lipoproteins has remained unknown. Based on the data above, we hypothesized that lipoproteins in *Hfx. volcanii* are modified by archaeol through a thioether linkage, and that AliA and AliB mediate this lipidation reaction. To test this hypothesis, we first extracted and analyzed core lipids from wild-type and Δ*aliA*/Δ*aliB* strains of *Hfx. volcanii*. Liquid chromatography-mass spectrometry (LC-MS) analysis revealed no significant differences in lipid types or their relative abundances between the strains (Fig. 4a,b and Supplementary Figs 7,8), suggesting that archaeol synthesis is not impaired in the mutant. We then extracted lipoproteins from both strains using Triton X-114 and treated the extracts with methyl iodide to cleave thioether bonds[45,46] and release the attached lipids while preserving their covalent bond to the cysteine-derived sulfur atom for downstream LC-MS analysis. Comparison of the resulting spectra with that of synthetic methylthioarchaeol (Supplementary Methods), the expected cleavage product[44], confirmed the presence of thioether-linked archaeol lipidation in lipoprotein extracts from wild-type *Hfx. volcanii*, and its complete absence in Δ*aliA*/Δ*aliB* (Fig. 4c,d and Supplementary Fig. 9). These findings identify AliA and/or AliB as key components required for thioether-linked archaeol lipidation of archaeal lipoproteins, revealing the previously elusive molecular machinery for archaeal lipoprotein lipidation and opening new directions for understanding protein lipidation and membrane anchoring in archaea.

AliA- and AliB-mediated lipoprotein biogenesis is critical for archaeal cell physiology

Sporadic studies have investigated the function of individual archaeal lipoproteins[47,48], but systematic analyses of lipoprotein functions in archaea are still lacking. Here, we performed an in silico functional categorization of the 93 predicted lipoproteins in *Hfx. volcanii*, revealing their involvement in essential cellular processes including nutrient transport and metabolism, energy production and conversion, and signal transduction (Fig. 5a and Supplementary Data 7). Given these broad roles, we hypothesized that disruption of lipoprotein biogenesis in Δ*ali* strains would impair *Hfx. volcanii* fitness. Consistent with the high number of lipoproteins involved in nutrient uptake (Fig. 5a), Δ*ali* strains formed smaller, lighter colonies on semi-defined Hv-Cab agar plates (Fig. 5b) and showed slower growth in the exponential phase in Hv-Cab liquid medium compared to the wild type (Fig. 5c). Their growth defects were more pronounced in minimal medium with glucose and alanine as the sole carbon and nitrogen sources, respectively (Fig. 5d). Δ*aliB* showed a strong reduction in growth compared to the wild type, while Δ*aliA* and Δ*aliA*/Δ*aliB* exhibited minimal growth, highlighting again the functional distinction between AliA and AliB.

Previous studies have shown that wild-type *Hfx. volcanii* transitions from rod-shaped cells in early log to disk-shaped cells in late log when grown in the Hv-Cab medium[49]. While rod-shaped cells are associated with better swimming ability[50], disk-shaped cells are hypothesized to have enhanced nutrient uptake due to a higher surface-area-to-volume ratio[51]. Considering the abundance of lipoproteins predicted to function in nutrient uptake and the severe growth defects of mutant strains in minimal medium, these mutants may be impaired in nutrient acquisition. Therefore, we examined whether they exhibit altered cell shape compared to the wild type, potentially as a compensatory adaptation to improve nutrient uptake. Microscopic analyses showed that early-log wild-type *Hfx. volcanii* culture contained mostly rods with some disks, whereas Δ*ali* cells were primarily disk-shaped in early log (Fig. 5e), suggesting they transitioned to the disk shape earlier than the wild type. Given the link between cell shape and motility, we further analyzed the motility of Δ*ali* mutants. All three mutants formed smaller motility halos than wild type (Fig. 5f), consistent with their early transition to non-motile disks. Moreover, deletion of *aliA* resulted in more severe phenotypes than deletion of *aliB*, including smaller motility halos and a higher proportion of disk-shaped cells (Fig. 5e,f). In summary, these results represent the first systematic investigation into lipoprotein functions in archaea, revealing the crucial role of AliA- and AliB-mediated lipoprotein biogenesis in archaeal cell physiology. The distinct phenotypes of Δ*aliA* and Δ*aliB* also suggest differences in their substrate specificities, warranting further investigation.

Discussion

Despite the high prevalence and importance of archaeal lipoproteins, they have been largely understudied, with no components of their biogenesis pathway identified until now. In this study, using comparative genomics, reverse genetics, and biochemical approaches, we identified the first two components of an archaeal lipoprotein biogenesis pathway, AliA and AliB, which were shown to be crucial for archaeal lipoprotein lipidation.

Our bioinformatic analyses suggest that a family of membrane proteins, arCOG02177, referred to as AliA, exhibits a phyletic pattern closely matching that of predicted lipoproteins across archaea. It is present in Euryarcheota and DPANN and is hypothesized to have been present in the last archaeal common ancestor[39,52]. AliA has a marginal sequence similarity with bacterial Lgt. Visual comparison of the *E. coli* Lgt crystal structure and AliA AlphaFold2 model revealed several common structural features and a remarkable conservation of two arginines, which are known to be important for Lgt activity[36,37]. However, a structural similarity between the AliA model and Lgt could not be detected in a Dali search, suggesting that AliA either originated from Lgt by a radical protein rearrangement or evolved independently. A diverged AliA paralog, arCOG02178 or AliB, possesses a distinct set of putative catalytic residues and is only present in halobacterial species. Interestingly, several archaeal genomes have a substantial number of predicted lipoproteins but lack both AliA and AliB, suggesting that additional lipoprotein lipidation pathways are yet to be discovered.

To confirm the involvement of AliA and AliB in lipoprotein biogenesis, we carried out in vivo experiments in *Hfx. volcanii* using three predicted lipoproteins as substrates. Cys-to-Ser substitutions in the conserved lipobox motif rendered all three lipoproteins unstable, as shown by their significantly decreased abundance in cell lysates and absence from the supernatant. When overexpressed in Δ*ali* strains, lipoproteins HVO_1176 and HVO_B0139 only existed as precursors, indicating that both AliA and AliB are important for their biogenesis. In contrast, HVO_1705 existed as both mature proteins and precursors in Δ*aliB* but only as precursors in Δ*aliA*, implying a functional difference between AliA and AliB, as well as a substrate selectivity of AliB. While the biogenesis defect of HVO_1705 could be complemented by the in

a

b

c

d
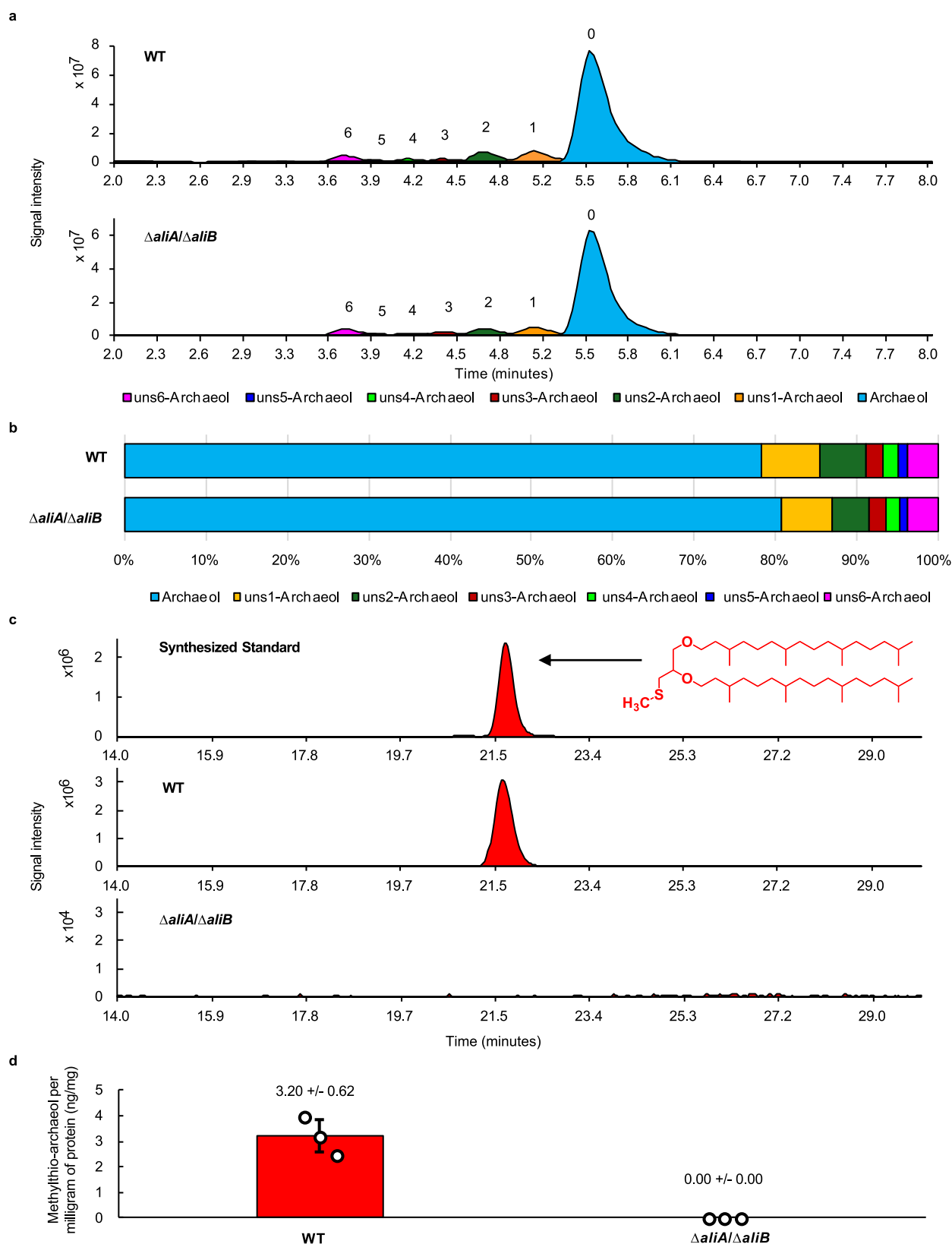
trans expression of the deleted genes, overexpressing *aliA* in Δ*aliB* or vice versa could not restore the phenotype to wild type, confirming the non-redundant roles of AliA and AliB in lipoprotein biogenesis. Using a Triton X-114 extraction assay, we demonstrated that AliA and AliB play key roles in lipoprotein lipidation. Both are equally critical for HVO_1176 lipidation, as deletion of either led to the accumulation of non-lipidated HVO_1176. AliA is also vital for HVO_1705 lipidation, whereas deletion of *aliB* had only minimal effects. Notably, although the three lipoproteins showed a significantly decreased abundance with the cysteine mutation, this decrease was not seen in Δ*ali* strains. This result suggests that in addition to AliA- and AliB-mediated lipidation, the lipobox cysteine is also involved in other processes that

Fig. 4 | Deletion of both *aliA* and *aliB* led to complete absence of thioether-linked archaeol from lipoprotein extracts. a Representative merged extracted ion chromatograms (EICs) (m/z values in methods) showing the presence of saturated and unsaturated diether core lipids (archaeols) in base hydrolyzed biomass from the wild-type and Δ*aliA*/Δ*aliB* strains. **b** Average (n = 3) relative abundance of the saturated and unsaturated diether core lipid species found in wild-type and Δ*aliA*/Δ*aliB* strains. **c** Representative merged EICs (m/z = 683.7, 700.8, and 705.7, corresponding to the protonated, ammoniated, and sodiated adducts of methylthio-archaeol) showing that the synthesized methylthio-archaeol standard is readily detectable with our methods and that the same methylthio-archaeol compound is detected in extracts from methyl iodide treated lipoproteins from the wild type, while methylthio-archaeol is notably absent in that from the Δ*aliA*/Δ*aliB* strain. Note that the y-axis of the EIC for the Δ*aliA*/Δ*aliB* strain is shown two orders of magnitude zoomed in compared to the wild type. **d** Bar chart showing the average mass of methylthio-archaeol recovered from methyl iodide treated lipoproteins per milligram of starting protein, based on three biological replicates. An average of 3.20 ng +/− 0.62 ng of methylthio-archaeol per milligram of protein was recovered from the wild type, while no methylthio-archaeol was recovered from the Δ*aliA*/Δ*aliB* strain. Error bars represent standard deviations.

maintain lipoprotein stability. One possibility is that the Cys-to-Ser mutant lipoproteins were recognized and processed by signal peptidase I (Supplementary Table 1), and subsequently degraded by extracellular proteases. Other cysteine modifications, such as acetylation of its α-amino group[26], might also contribute to lipoprotein stability. This strategy is not uncommon; for instance, methylation of a lipidated cysteine has been shown to significantly enhance the half-life of certain proteins[53,54]. Previous studies raised the possibility that archaeal lipoproteins are modified by thioether-linked archaeol[26,44]. To test this hypothesis and assess the role of AliA and AliB in this process, we treated lipoprotein extracts with methyl iodide to release covalently attached lipids and analyzed them using LC-MS. Consistent with the hypothesis, the results revealed the presence of a thioether-linked archaeol lipidation in wild-type *Hfx. volcanii* lipoprotein extracts and its complete absence in the Δ*aliA*/Δ*aliB* strain. These data strongly support archaeol as the lipid anchor of *Hfx. volcanii* lipoproteins and confirm the critical roles of AliA and/or AliB in forming the thioether bond between archaeol and target proteins.

Additionally, disruption of lipoprotein biogenesis by deleting *aliA* or *aliB* affected multiple aspects of *Hfx. volcanii* physiology. The mutant strains formed smaller and lighter colonies on Hv-Cab agar plates and exhibited slower exponential growth in Hv-Cab liquid medium compared to the wild type. However, their growth defect was not severe. Previous studies indicate that some non-lipidated lipoproteins in *Hfx. volcanii* can still be anchored to cell membranes via their hydrophobic signal peptides[24] and might therefore remain functional. Similarly, non-lipidated lipoproteins in Δ*ali* strains might still support cell growth in semi-defined medium by contributing to the transport and metabolism of certain carbon and nitrogen sources. However, in minimal medium using glucose and alanine as the sole carbon and nitrogen sources, respectively, Δ*ali* strains showed severe growth defects, likely due to the impaired biogenesis of lipoproteins responsible for glucose and alanine transport or metabolism. *Ali* deletions also affect cell shape and motility. Δ*ali* strains transitioned earlier from rods to disks compared to the wild type, potentially as a compensation mechanism for their defective nutrient uptake capability. Moreover, since no known cell-shape regulatory components are predicted as lipoproteins (Supplementary Data 7), uncharacterized regulatory mechanisms involving lipoproteins might exist and warrant further investigation. Consistent with the association between disk shape and reduced motility, Δ*ali* strains also showed smaller motility halos compared to the wild type.

Considering the phenotypic differences between Δ*aliA* and Δ*aliB* in lipoprotein biogenesis and cell physiology, AliA and AliB likely possess different enzymatic activities and substrate specificities. This observation is further supported by the replacement of two conserved arginine residues in AliA with histidines in AliB. In *E. coli* Lgt, the two conserved arginines (R143 and R239) are essential for the lipidation reaction and are proposed to facilitate specific binding of Lgt to the negatively charged membrane lipids[36]. Therefore, AliA and AliB may differ in lipid-binding capabilities, resulting in varying efficiencies of lipoprotein lipidation or even different functions. Such divergence has been observed in the model bacterium *Mycobacterium smegmatis*[55], where only one *M. smegmatis* Lgt homolog complemented a *Corynebacterium glutamicum* Δ*lgt* strain[56]. The other homolog carries mutations in several conserved residues and is proposed to either be inactive or have decreased enzymatic activity[55,56]. Despite this finding, closer examinations of the functional differentiation of the two *M. smegmatis* Lgt paralogs are still lacking. Therefore, future structure-function analyses of AliA and AliB using site-directed mutagenesis will provide valuable insights into the lipoprotein biogenesis in both archaea and bacteria. Similarly, in vitro lipidation assays, which are currently limited by low expression levels of the two proteins, will help determine whether AliA and AliB have stand-alone lipidation activities and different lipid-binding capabilities.

Overall, our study underscores the crucial roles of lipoproteins in archaeal cell physiology and reported the identification and characterization of the first two components of an archaeal lipoprotein biogenesis pathway. The successful identification of AliA and AliB using comparative genomics and sequence analyses also paves the way for identifying and characterizing other components in this pathway, including the downstream lipoprotein signal peptidase. Furthermore, we highlighted the distinct roles of AliA and AliB in lipoprotein lipidation, raising exciting new questions about the delicate regulation of lipoprotein lipidation in both archaea and bacteria.

Methods

Plasmids, strains, primers, and reagents
All plasmids, strains, and primers used in this study are listed in Supplementary Tables 2-4. DNA Phusion Taq polymerase, restriction enzymes, and DNA ligase were purchased from New England BioLabs. The RQ1 RNase-Free DNase was purchased from Promega.

Sequence comparison, phylogenetic analysis, and gene neighborhood analysis
Secreted proteins encoded in 524 archaeal genomes from the arCOG database were identified using SignalP 6.0[31]. SignalP 6.0 identifies five types of signal peptides, including distinct predictions for lipoproteins secreted via Tat (TATLIPO) and Sec (LIPO) pathways. We considered only confident predictions (probability ≥90%) for all types of signal peptides. The arCOG database[29,30] that includes annotated clusters of orthologous genes for 524 archaeal genomes covering all major archaeal lineages is available at https://ftp.ncbi.nih.gov/pub/wolf/COGs/arCOG/tmp.ar18/.

HHpred online tool[35] was used to search for sequence similarity with default parameters against HMM profiles derived from PDB, PFAM, and CDD databases. Muscle5 program[57] with default parameters was used to construct multiple sequence alignments. For both arCOG02177 and arCOG02178 alignments, a consensus sequence was calculated as described previously[58]. Briefly, each amino acid in the consensus sequence corresponds to the best-scoring amino acid against all amino acids in the respective alignment column, calculated using the BLOSUM62 matrix. The HHalign software[59] automatically calculates conserved residues. For phylogenetic analysis, poorly aligned sequences or fragments were discarded. Columns in the multiple alignment were filtered for homogeneity value[58] 0.05 or higher and gap fraction less than 0.667. This filtered alignment was used as an input for the FastTree program[60] to construct an approximate
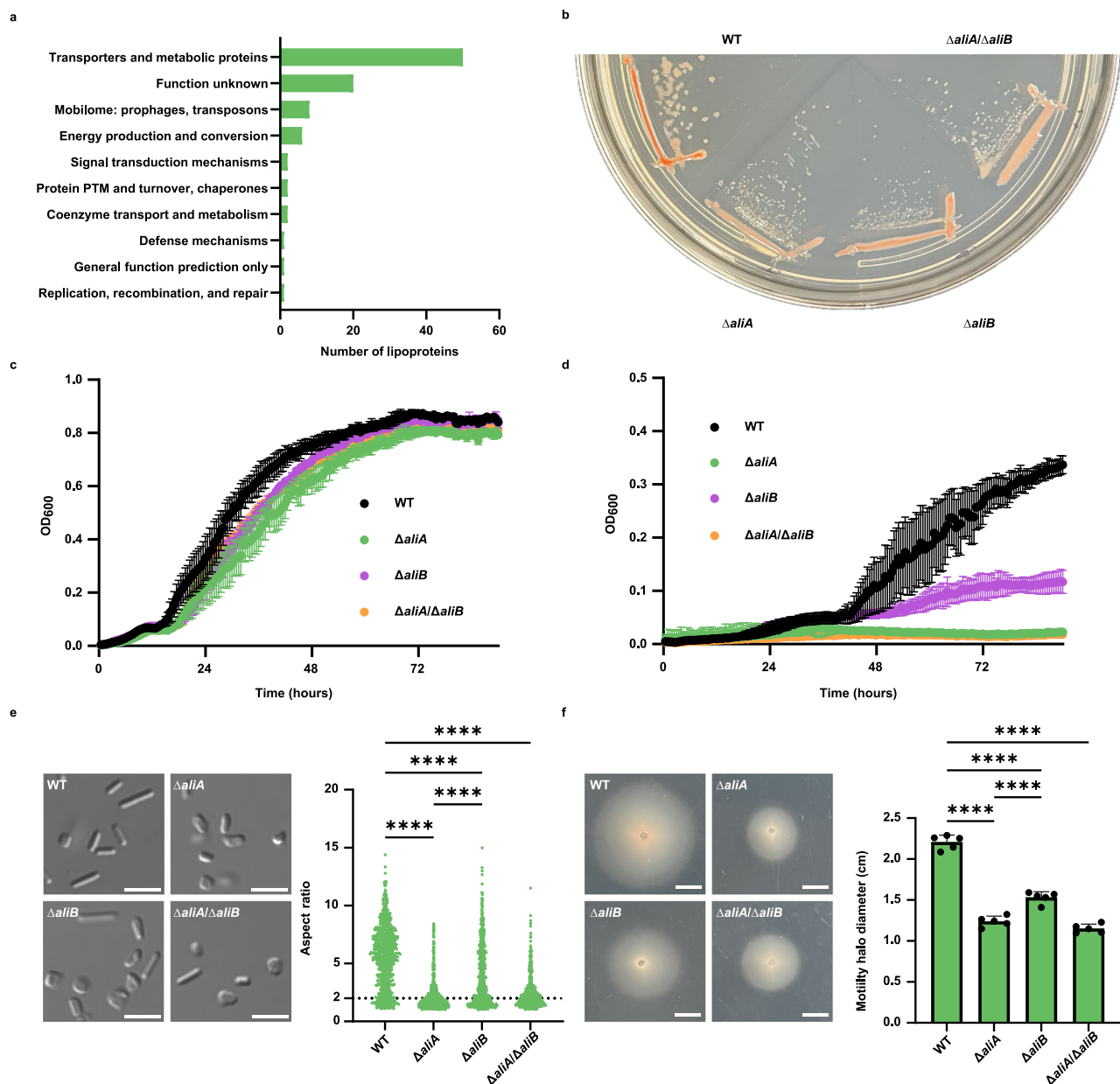
Fig. 5 | AliA and AliB are important for the growth, cell shape, and motility of *Hfx. volcanii*. a Functional classification of 93 predicted *Hfx. volcanii* lipoproteins. PTM, posttranslational modification. **b** Colony morphology of *Hfx. volcanii* wild type and Δ*ali* mutants. Strains were streaked out on Hv-Cab agar plates and incubated at 45 °C for 4 days prior to imaging. **c** Growth curves of *Hfx. volcanii* wild type and Δ*ali* mutants in Hv-Cab medium. Data are presented as mean values of four biological replicates +/− standard deviations. **d** Growth curves of *Hfx. volcanii* wild type and Δ*ali* mutants in minimal medium with glucose and alanine as the sole carbon and nitrogen sources, respectively. Data are presented as mean values of four biological replicates +/− standard deviations. **e** Cell morphology of *Hfx.* *volcanii* wild type and Δ*ali* mutants in early log and the quantification. Cells with aspect ratios (ratio of major to minor axes) <2 are considered disks and short rods, while others are considered regular rods. Three biological replicates were analyzed for each strain, and 257 cells were analyzed for each biological replicate. Scale bars indicate 5 μm. Data were analyzed using a one-way ANOVA test. ****P < 0.0001. **f** Motility halos of *Hfx. volcanii* wild type and Δ*ali* mutants and the quantification. Scale bars indicate 0.5 cm. The quantification shows the motility halo diameters and their mean of five biological replicates for each strain. Data were analyzed using a one-way ANOVA test. Error bars represent standard deviations. ****P < 0.0001.

maximum likelihood phylogenetic tree with the WAG evolutionary model and gamma-distributed site rates. The same program was used to calculate support values. DeepTMHMM[61] was used to search for transmembrane segments. For genome context analysis and search for putative operons, neighborhoods containing five upstream and five downstream genes were constructed for all identified arCOG02177 and arCOG02178 genes. Structural models for HVO_2859 and HVO_2611 proteins were predicted using ColabFold v1.5.5 (AlphaFold2)[62] and visualized using UCSF ChimeraX[63]. DALI server[33] was used to compare predicted models with PDB database.

Growth conditions

Hfx. volcanii strains were grown at 45 °C in liquid (orbital shaker at 250 rpm with an orbital diameter of 2.54 cm) or on solid agar (containing 1.5% [w/v] agar) semi-defined Hv-Cab medium[49], supplemented with tryptophan (Fisher Scientific) and uracil (Sigma) at a final concentration of 50 μg mL⁻¹ unless otherwise noted. Uracil was left out for strains carrying pTA963 or pTA963-based plasmids. 200 μg mL⁻¹ tryptophan was used to induce pTA963 expression for western blots, while 50 μg mL⁻¹ tryptophan was used in all other experiments. For growth curves in minimal medium, CDM minimal medium was made as

previously described[64], using the same 18% salt water, trace elements, thiamine, and biotin solution as for the Hv-Cab medium. Additionally, 20 mM glucose and 25 mM alanine were added as the sole carbon and nitrogen sources, respectively. *E. coli* strains used for cloning were grown at 37 °C in NZCYM (RPI) medium, either as liquid cultures (shaken at 250 rpm with an orbital diameter of 2.54 cm) or on solid agar containing 1.5% [w/v] agar, and supplemented, if required, with ampicillin at a final concentration of 100 μg mL^{-1}.

Construction of *Hfx. volcanii* overexpression strains

Genes of interest were first amplified by PCR using specific primers that contain artificial restriction sites and then digested using BamHI and EcoRI. The PCR products were then ligated to the BamHI and EcoRI digested expression vector pTA963. The ligation products were later transformed into *E. coli* DH5α cells, and the plasmids were extracted from the DH5α cells using the PureLink Quick Plasmid Miniprep Kit (Invitrogen). Site-directed mutagenesis of *hvo_BO139*, *hvo_1176*, and *hvo_1705* was performed on the DH5α overexpression plasmids using the New England BioLabs Q5 site-directed mutagenesis kit with gene-specific primers. The DH5α plasmids were then transformed into *E. coli* Dam$^-$ strain DL739 to get demethylated plasmids. Plasmid sequences were confirmed by Sanger sequencing or whole plasmid sequencing provided by Eurofins Genomics. The verified demethylated plasmids were transformed into *Hfx. volcanii* using the polyethylene glycol (PEG) method[65].

To construct vectors overexpressing both *aliB* and a lipoprotein gene, the pYH10 (pTA963-based vector expressing the *aliB*-His construct) was linearized with HindIII. The linearized vector was then treated with Klenow DNA Polymerase to remove the overhangs created by the HindIII digestion. In parallel, pYH3 (pTA963-based vector expressing the C-terminally myc-tagged HVO_1176) was digested with PvuII to isolate the *hvo_1176*-myc fragment with the tryptophan-inducible promoter (p.*tnaA*-*hvo_1176*-myc). The PvuII-digested fragment was ligated to the previously linearized blunt-end pYH10, generating the pYH12 vector containing both *aliB*-His and *hvo_1176*-myc under separate tryptophan-inducible promoters. The same cloning strategy was used to generate pYH13 (pTA963 expressing both *aliB*-His and *hvo_1705*-myc). Demethylated plasmid preparation and *Hfx. volcanii* transformations were performed as described above.

Overexpression vectors carrying both *aliA* and a lipoprotein gene could not be isolated from the *E. coli* DH5α strain, possibly due to the potential toxicity of the plasmids to *E. coli* cells. Therefore, we generated the corresponding Gibson assembly[66] products and directly transformed the products into *Hfx. volcanii*. For example, the p.*tnaA*-*hvo_1176*-myc fragment was amplified by PCR from the pYH3 vector. To facilitate homologous recombination between this fragment and HindIII-digested pYH11 (pTA963-based vector expressing *aliA*), a 22-nucleotide sequence homologous to one blunt end of the linearized pYH11 and a 20-nucleotide sequence homologous to the other blunt end were added to opposite sides of the *p.tnaA*-*hvo_1176*-myc fragment during PCR. The fragment and linearized pYH11 were then added to a homemade Gibson assembly mixture[66] to generate the desired co-expression construct.

Generation of chromosomal deletions in *Hfx. volcanii*

Chromosomal deletions were generated by the pop-in/pop-out method[67] and are briefly outlined below. About 750 nucleotides flanking the corresponding gene were first amplified with PCR (for primers see Supplementary Table 4). For the construction of *ΔaliA*, the upstream flanking region was amplified with oligonucleotides hvo_2859_ups_fwd and hvo_2859_ups_rvr, and the downstream flanking region was amplified with hvo_2859_down_fwd and hvo_2859_down_rvr. The *aliA* upstream and downstream flanking DNA fragments were fused by overlap PCR[68] using oligonucleotides hvo_2859_ups_fwd and hvo_2859_down_rvr followed by cloning into the haloarchaeal

suicide vector pTA131[67] digested with XbaI and XhoI. The resulting construct was verified by sequencing using the same oligonucleotides. The final plasmid construct contained upstream and downstream *aliA* flanking regions and was transformed into the parental strain H53. Pop-out transformants were selected on agar plates containing 5-fluoroorotic acid (FOA) (Toronto Research Chemicals Inc.) at a final concentration of 50 μg mL^{-1}. The equivalent protocol was used for generation of the deletion strains *ΔaliB* and *ΔaliA*/*ΔaliB* (using *ΔaliB* as the parent strain). Primers for the generation of these deletions are listed in Supplementary Table 4. Successful gene deletions were first confirmed by colony PCR. The genomic DNA of the deletion strains were extracted using the GeneJET PCR Purification Kit from Thermo Scientific and quantified using a Qubit 3.0 Fluorometer from Invitrogen. The genomic DNA was then further analyzed by Illumina whole genome sequencing performed by SeqCenter (Pittsburgh, PA, USA) to confirm the complete deletion of the gene as well as search for any secondary genome alterations.

Immunoblotting

A single colony was used to inoculate a 5 mL liquid culture. At an OD$_{600}$ of 0.8, cells from 4 mL liquid culture were harvested by centrifugation at 4300 × g for 10 min at 4 °C. The supernatant was transferred into a new tube and centrifuged again at 10,000 × g for 10 min. The supernatant was then concentrated using the Amicon Ultra-4 3 K centrifugal filter unit at 7500 × g and 4 °C until the final volume was below 250 μL. 2 mL PBS buffer (containing 1 mM AEBSF protease inhibitor from Thermo Scientific) was added to the concentrated supernatant for buffer exchange, and the sample was concentrated again to less than 250 μL. The buffer exchange was repeated once, and the supernatant was concentrated to ~100 μL and transferred to a clean 1.5 mL tube. The cell pellet was resuspended with 1 mL 18% salt water, made by diluting the 30% salt water used in the Hv-Cab medium. The cells were pelleted again at 4300 × g for 10 min at 4 °C to remove any leftover medium. Cells were then resuspended in 250 μL PBS buffer with 1 mM AEBSF and lysed by freezing (with liquid nitrogen) and thawing (at 37 °C) four times. 20 μL RQ1 DNase solution was added to cell lysates, and the mixture was incubated at 37 °C for 30 min to degrade DNA. Unlysed cells were removed by centrifugation at 4300 × g for 10 min at 4 °C, and the supernatant was transferred to a new tube. The protein concentration was measured using the Pierce BCA Protein Assay Kit from Thermo Scientific, and a BioTek Epoch 2 microplate reader with BioTek Gen6 v.1.03.01 (Agilent).

The desired amounts of protein were supplemented with 50 mM dithiothreitol and NuPAGE LDS Sample Buffer (1X). Samples were incubated at 70 °C for 10 min before being loaded onto NuPAGE 10%, Bis-Tris gels (1.0 mm × 10 wells) with NuPAGE MOPS SDS Running Buffer. After electrophoresis, proteins were transferred to a polyvinylidene difluoride (PVDF) membrane (Millipore) using a semi-dry transfer apparatus (BioRad) at 15 V for 30 min. Subsequently, the membrane was stained with Ponceau S Staining Solution (Cell Signaling Technology) to verify the successful protein transfer. The membrane was then washed for 10 min twice in PBS buffer and blocked for 1 h in 5% non-fat milk (LabScientific) in PBST buffer (PBS with 1% Tween-20). After blocking, the membrane was washed twice in PBST and once in PBS. For detection of the myc tag, anti-Myc antibody (9E10, UPenn Cell Center Service with a catalog number 3207) was diluted 1000 times with PBS buffer containing 3% bovine serum albumin, which was then used to incubate the membrane overnight at 4 °C. Subsequently, the membrane was washed twice in PBST and once in PBS, followed by a 45 min incubation at room temperature with the secondary antibody solution: Amersham ECL anti-mouse IgG (horseradish peroxidase-linked whole antibody from sheep, Cytiva; catalog number NA931-1 mL, lot number 17376630), diluted 10,000 times in PBS containing 10% non-fat milk. After incubation, the membrane was washed three times in PBST and once in PBS. HRP activity was assessed

using the Amersham ECL Prime Western Blotting Detection Reagent (GE). Membranes were imaged using an Amersham Imager 600, with image analysis conducted using the Amersham Imager 600 Analysis software.

Triton X-114 extraction

Triton X-114 extraction was performed as previously described[41] with some modification. Briefly, 100 µL ice-cold 20% Triton X-114 and 900 µL *Hfx. volcanii* protein sample containing ~500 µg proteins were added to a chilled glass container. The container was put into a glass jar filled with ice to maintain the mixture temperature at 0–4 °C. The mixture was stirred on ice at 340 rpm for 2 h before being transferred to a chilled microtube. The mixture was then centrifuged at 4 °C for 10 min at 15,000 × g to remove any precipitation. The supernatant was incubated at 37 °C for 10 min and centrifuged at room temperature for 10 min at 10,000 × g. After centrifugation, the mixture was separated into the upper aqueous (AQ) phase and lower detergent (TX) phase. 88.8 µL 20% Triton X-114 was added to the AQ phase, while 800 µL of ice-cold PBS buffer was added to the TX phase. After a brief vortex, the mixture was incubated on ice for 10 min and then at 37 °C for 10 min, followed by centrifugation at room temperature for 10 min at 10,000 × g. The new TX phase of the AQ phase sample was discarded, and 9 mL of ice-cold methanol was added to the AQ phase. For the TX phase sample, the new AQ phase was discarded, and 900 µL ice-cold methanol was added to the TX phase. Both AQ and TX phase samples were stored at −80 °C overnight to precipitate the proteins. The next day, the samples were centrifuged at 4 °C for 10 min (15,000 g for microtubes and 12,000 g for 15 mL Falcon tubes). The sediments of the AQ and TX samples containing the desired proteins were resuspended in the same volume of PBS buffer with 1 mM AEBSF and stored at −20 °C. The same volume of AQ and TX samples were used for western blots.

Core lipid extraction

A single colony of *Hfx. volcanii* was inoculated into 5 mL of liquid medium. When the culture reached an OD_{600} of 0.3–0.8, it was used to inoculate 25 mL of secondary culture (starting $OD_{600} = 0.01$) in the same medium. Cells were harvested at $OD_{600} = 0.65$–0.75 by centrifugation at 7000 × g for 20 min at 4 °C. Cell pellets were resuspended in 2 mL of methanol (MeOH), transferred to glass vials, and the solvent was dried off under N_2 stream. The samples were then base hydrolyzed in 2 mL of 1 M KOH in MeOH at 75 °C for 3 h. The reaction was neutralized by the addition of 1 mL of 2 M HCl in MeOH and diluted with 5 mL of deionized water. Core lipids were extracted three times by the addition of 5 mL dichloromethane (DCM). The three DCM extracts were pooled and evaporated under N_2 stream. The resultant core lipids were resuspended in 1 mL of 9:1 MeOH:DCM and filtered through 0.45 µm polytetrafluoroethylene (PTFE) filters for analysis via reverse phase (RP) liquid chromatography-mass spectrometry (LC-MS).

Methyl iodide treatment of lipoprotein extracts

Methyl iodide treatment was performed as previously described by Sagami et al. and by Maltese and Ermand, with minor modifications[45,46]. Briefly, Triton X-114 lipoproteins extracted from 80 mg total proteins were resuspended in 400 µL of 50 mM Tris-HCl buffer (pH 8.0) and transferred to 4 mL glass vials. Then, 200 µL 3% formic acid was added to bring the solution to a final concentration of 1% formic acid, followed by the addition of 600 µL methyl iodide. Samples were capped, vortexed, and then shaken in the dark at 200 rpm at 37 °C for 24 h. Then, methyl iodide was evaporated under N_2 stream, samples were brought to a final volume of 1 mL by the addition of ~400 µL 50 mM Tris-HCl buffer, and the pH of the solution was adjusted to 12 by the addition of 10 N NaOH. The incubation was continued for 24 h. Afterwards, the released lipids were extracted by the addition of 1 mL 9:1 chloroform:MeOH, vortexed, and centrifuged at 4200 x g for 20 min at

4 °C to achieve separation. This was repeated twice more, and the chloroform:MeOH extracts were pooled and evaporated under N_2 stream. The extracts were resuspended in 500 µL of 2:1:1 MeOH:iso-propanol:hexane and filtered through 0.45 µm PTFE filters for analysis via RP LC-MS.

Lipid analysis

Lipids were analyzed on an Agilent 1260 Infinity II series high-performance liquid chromatography instrument coupled to an Agilent G6125B single quadrupole mass spectrometer using electrospray ionization (ESI). Samples were analyzed in positive mode with the following parameters: 8.0 L/min drying gas flow rate, 35 psi nebulizer pressure, 3000 V capillary voltage, and either a 300 °C drying gas temperature (for core lipids) or 250 °C drying gas temperature (for methylthio-archaeol analysis).

Core lipids were separated on a Kinetex 1.7 µm XB-C18 100 Å LC column (150 × 2.1 mm) maintained at 45 °C with a mobile phase A of methanol containing 0.04% formic acid and 0.03% NH_3 and a mobile phase B of isopropanol containing 0.04% formic acid and 0.03% NH_3 by a method modified from Rattray and Smittenberg 2020[69]. Lipids were eluted at a flow rate of 0.2 mL/min with an initial mixture of 60% A:40%B held for 1 min, then linearly ramped to 50%A:50%B over 19 min, and returned to 60%A:40%B over 5 min, for a total runtime of 25 min. Samples were analyzed in full scan mode with a m/z range of 600 to 750. For quantification of the relative of abundance of each core lipid species, the m/z peak areas corresponding to the $[M + H]^+$, $[M + NH_4]^+$, $[M+Na]^+$, $[M + K]^+$, and $[M+Isopropanol+H]^+$ adduct ions were calculated and summed using Agilent MassHunter B.08.00 and verified manually. For archaeol, the adduct ions m/z values analyzed were as follows: $[M + H]^+ = 653.7$, $[M + NH_4]^+ = 670.8$, $[M+Na]^+ = 675.7$, $[M + K]^+ = 691.7$, and $[M+Isopropanol+H]^+ = 712.8$. For the unsaturated archaeols, the adduct ion m/z values analyzed were equal to the values for archaeol – (2 x number of double bonds).

Lipids released from methyl iodide treated lipoproteins were separated on an Agilent Poroshell 120 EC-C18 column (1.9 µm, 2.1 × 150 mm) maintained at 40 °C with a mobile phase A of 95:5 MeOH:water containing 0.04% formic acid and 0.03% NH_3 and mobile phase B of 50:50 isopropanol:hexane containing 0.04% formic acid and 0.03% NH_3 by a method modified from Connock et al. [70]. Lipids were eluted at a flow rate of 0.2 mL/min with an initial mixture of 70% A:30%B linearly ramped to 50%A:50%B over 35 min, then immediately back to 70%A:30%B and held for 15 min, for a total runtime of 50 min. Samples were analyzed in full scan mode with a m/z range of 650 to 750. For quantification of methylthio-archaeol, the peak areas of m/z = 683.7, 700.8, and 705.7, (corresponding to the $[M + H]^+$, $[M + NH_4]^+$, $[M+Na]^+$ adducts of methylthio-archaeol) were calculated and summed using Agilent MassHunter B.08.00 and verified manually. The areas were converted to masses by the use of a 10-point standard curve ($R^2 = 0.999$) generated by injecting 0.064 ng, 0.128 ng, 0.256 ng, 0.512 ng, 1.024 ng, 2.048 ng, 4.096 ng, 8.192 ng, 11.2 ng, and 19.2 ng of the synthesized methylthio-archaeol standard into the LC-MS with the same program used to analyze the samples. The methylthio-archaeol (diphytanylglyceryl methyl thioether) standard was synthesized by Pharmaron, Inc. via a synthesis scheme outlined in the Supplementary Methods and was confirmed with NMR and LC-MS analysis.

Hfx. volcanii growth curve

Colonies for each strain were inoculated in 5 ml Hv-Cab liquid medium and grown until they reached mid-log phase (OD_{600} between 0.3 and 0.8). 1 ml cultures were then harvested via centrifugation (15,800 g, 1 min at room temperature) and washed twice with 18% salt water. Cultures were then diluted to an OD_{600} of 0.01 in a flatbottom polystyrene 96-well plate (Corning) with 200 µl Hv-Cab medium or fresh CDM-based minimal medium supplemented with 20 mM glucose and 25 mM alanine. 200 µl of medium was also aliquoted into each well for

two rows of perimeter wells of the plate to help prevent culture evaporation. Growth curves were measured using a BioTek Epoch 2 microplate reader with BioTek Gen6 v.1.03.01 (Agilent). Readings were taken every 30 min with double orbital, continuous fast shaking (355 cycles per min) in between. Readings were taken at a wavelength of 600 nm for about 90 h and then plotted using GraphPad Prism version 10.0.1 for Windows 64-bit.

Motility assays and motility halo quantification

Motility was assessed on 0.35% agar Hv-Cab medium plates with supplements as required. A toothpick was used to stab-inoculate the agar followed by incubation at 45 °C. Motility assay plates were removed from the incubator after three days and imaged after one day of room temperature incubation, using an iPhone 13. Motility halos were quantified using ImageJ version 1.54 h 15[71]. Images were uploaded to ImageJ, and the scale was set based on the 100 mm Petri dish diameter. The statistical significance of halo diameters was assessed with a one-way ANOVA test using GraphPad Prism version 10.0.1 for Windows 64-bit.

Live-cell imaging and analysis

Hfx. volcanii strains were inoculated from a single colony into 5 ml of Hv-Cab liquid medium with appropriate supplementation. Cultures were grown until they reached an OD_{600} of 0.07. Three biological replicates of each strain were grown as described above, and for each culture, 1 ml aliquots were centrifuged at 4900 g for 6 min. Pellets were resuspended in ~10 µl of medium. 1.5 µl of the resuspension was placed on glass slides (Fisher Scientific), and a glass coverslip (Globe Scientific Inc.) was placed on top. Slides were visualized using a Leica Dmi8 inverted microscope attached to a Leica DFC9000 GT camera with Leica Application Suite X (version 3.6.0.20104) software, and both brightfield and differential interference contrast (DIC) images were captured at 100x magnification. Brightfield images were quantified using CellProfiler (version 4.2.1)[72]. A published CellProfiler pipeline[48] was used for image analysis, available at https://doi.org/10.5281/zenodo.8404691. Briefly, images were uploaded to CellProfiler, and parameters were set for each image set (strain and replicate) based on specific image conditions to eliminate noise and maximize the number of identified cells. Cell-specific data for each image set were exported, and aspect ratio was calculated. The statistical significance of aspect ratio comparisons between strains was assessed with a one-way ANOVA test using GraphPad Prism 10.0.1 for Windows 64-bit.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Whole genome sequencing data can be accessed through the NCBI Sequence Read Archive under the accession number PRJNA1182751. The PDB files of the AlphaFold models, variant calling results from whole genome sequencing, as well as raw data from growth curves, motility assays, and live-cell imaging experiments are available on Zenodo, with the identifier https://doi.org/10.5281/zenodo.12730247 [https://doi.org/10.5281/zenodo.12730247]. The LC-MS raw data and quantification can be accessed through the Stanford Digital Repository [https://doi.org/10.25740/vq241jx8767]. All other data supporting the findings of this study are provided within the paper and its supplementary information files. Source data are provided with this paper.

References

1. Chaban, B., Ng, S. Y. M. & Jarrell, K. F. Archaeal habitats—from the extreme to the ordinary. *Can. J. Microbiol.* **52**, 73–116 (2006).
2. Baker, B. J. et al. Diversity, ecology and evolution of Archaea. *Nat. Microbiol.* **5**, 887–900 (2020).
3. Pfeifer, K. et al. Archaea biotechnology. *Biotechnol. Adv.* **47**, 107668 (2021).
4. Haque, R. U., Paradisi, F. & Allers, T. *Haloferax volcanii* for biotechnology applications: challenges, current state and perspectives. *Appl. Microbiol. Biotechnol.* **104**, 1371–1382 (2020).
5. Chibani, C. M. et al. A catalogue of 1167 genomes from the human gut archaeome. *Nat. Microbiol.* **7**, 48–61 (2022).
6. Coker, O. O., Wu, W. K. K., Wong, S. H., Sung, J. J. & Yu, J. Altered gut archaea composition and interaction with bacteria are associated with colorectal cancer. *Gastroenterology* **159**, 1459–1470 (2020).
7. Siliakus, M. F., van der Oost, J. & Kengen, S. W. M. Adaptations of archaeal and bacterial membranes to variations in temperature, pH and pressure. *Extremophiles* **21**, 651–670 (2017).
8. Villanueva, L. et al. Bridging the membrane lipid divide: bacteria of the FCB group superphylum have the potential to synthesize archaeal ether lipids. *ISME J.* **15**, 168–182 (2021).
9. Gattinger, A., Schloter, M. & Munch, J. C. Phospholipid etherlipid and phospholipid fatty acid fingerprints in selected euryarchaeotal monocultures for taxonomic profiling. *FEMS Microbiol. Lett.* **213**, 133–139 (2002).
10. Lombard, J., López-García, P. & Moreira, D. Phylogenomic investigation of phospholipid synthesis in archaea. *Archaea* **2012**, 630910 (2012).
11. Abdul Halim, M. F. et al. Permuting the PGF signature motif blocks both archaeosortase-dependent C-terminal cleavage and prenyl lipid attachment for the *Haloferax volcanii* S-layer glycoprotein. *J. Bacteriol.* **198**, 808–815 (2016).
12. Braun, V. & Hantke, K. Lipoproteins: structure, function, biosynthesis. in *Bacterial Cell Walls and Membranes* 39–77 (Springer, 2019).
13. El Rayes, J., Rodríguez-Alonso, R. & Collet, J.-F. Lipoproteins in Gram-negative bacteria: new insights into their biogenesis, subcellular targeting and functional roles. *Curr. Opin. Microbiol.* **61**, 25–34 (2021).
14. Nguyen, M.-T., Matsuo, M., Niemann, S., Herrmann, M. & Götz, F. Lipoproteins in Gram-positive bacteria: abundance, function, fitness. *Front. Microbiol.* **11**, 582582 (2020).
15. Babu, M. M. et al. A database of bacterial lipoproteins (DOLOP) with functional assignments to predicted lipoproteins. *J. Bacteriol.* **188**, 2761–2773 (2006).
16. Thompson, B. J. et al. Investigating lipoprotein biogenesis and function in the model Gram-positive bacterium *Streptomyces coelicolor*. *Mol. Microbiol.* **77**, 943–957 (2010).
17. Gan, K., Gupta, S. D., Sankaran, K., Schmid, M. B. & Wu, H. C. Isolation and characterization of a temperature-sensitive mutant of *Salmonella typhimurium* defective in prolipoprotein modification. *J. Biol. Chem.* **268**, 16544–16550 (1993).
18. Tokunaga, M., Loranger, J. M. & Wu, H. C. Prolipoprotein modification and processing enzymes in *Escherichia coli*. *J. Biol. Chem.* **259**, 3825–3830 (1984).
19. Gupta, S. D. & Wu, H. C. Identification and subcellular localization of apolipoprotein *N*-acyltransferase in *Escherichia coli*. *FEMS Microbiol. Lett.* **78**, 37–42 (1991).
20. Gardiner, J. H. I. V. et al. Lipoprotein *N*-acylation in *Staphylococcus aureus* is catalyzed by a two-component acyl transferase system. *mBio* **11**, e01619 (2020.
21. Armbruster, K. M. & Meredith, T. C. Identification of the lyso-form *N*-acyl intramolecular transferase in low-GC firmicutes. *J. Bacteriol.* **199**, https://doi.org/10.1128/jb.00099-17 (2017).
22. Pohlschroder, M., Pfeiffer, F., Schulze, S. & Halim, M. F. A. Archaeal cell surface biogenesis. *FEMS Microbiol. Rev.* **42**, 694–717 (2018).
23. Yerushalmi, G., Zusman, T. & Segal, G. Additive effect on intracellular growth by *Legionella pneumophila* Icm/Dot proteins containing a lipobox motif. *Infect. Immun.* **73**, 7578–7587 (2005).
24. Storf, S. et al. Mutational and bioinformatic analysis of haloarchaeal lipobox-containing proteins. *Archaea* **2010**, 410975 (2010).

25. Giménez, M. I., Dilks, K. & Pohlschröder, M. *Haloferax volcanii* twin-arginine translocation substates include secreted soluble, C-terminally anchored and lipoproteins. *Mol. Microbiol.* **66**, 1597–1606 (2007).

26. Mattar, S. et al. The primary structure of halocyanin, an archaeal blue copper protein, predicts a lipid anchor for membrane fixation. *J. Biol. Chem.* **269**, 14939–14945 (1994).

27. Pohlschroder, M., Schulze, S., Pfeiffer, F. & Hong, Y. *Haloferax volcanii*: a versatile model for studying archaeal biology. *J. Bacteriol.* **0**, e00062–25 (2025).

28. Göker, M. & Oren, A. Valid publication of four additional phylum names. *Int. J. Syst. Evol. Microbiol.* **73**, 006024 (2023).

29. Makarova, K. S., Wolf, Y. I. & Koonin, E. V. Archaeal clusters of orthologous genes (arCOGs): an update and application for analysis of shared features between Thermococcales, Methanococcales, and Methanobacteriales. *Life* **5**, 818–840 (2015).

30. Makarova, K. S., Wolf, Y. I. & Koonin, E. V. Towards functional characterization of archaeal genomic dark matter. *Biochem. Soc. Trans.* **47**, 389–398 (2019).

31. Teufel, F. et al. SignalP 6.0 predicts all five types of signal peptides using protein language models. *Nat. Biotechnol.* **40**, 1023–1025 (2022).

32. Camacho, C. et al. BLAST+: architecture and applications. *BMC Bioinforma.* **10**, 421 (2009).

33. Holm, L., Laiho, A., Törönen, P. & Salgado, M. DALI shines a light on remote homologs: one hundred discoveries. *Protein Sci. Publ. Protein Soc.* **32**, e4519 (2023).

34. Bazan, J. F., MacDonald, B. T. & He, X. The TIKI/TraB/PrgY family: a common protease fold for cell signaling from bacteria to metazoa? *Dev. Cell* **25**, 225 (2013).

35. Zimmermann, L. et al. A completely reimplemented MPI bioinformatics toolkit with a new HHpred server at its core. *J. Mol. Biol.* **430**, 2237–2243 (2018).

36. Mao, G. et al. Crystal structure of *E. coli* lipoprotein diacylglyceryl transferase. *Nat. Commun.* **7**, 10198 (2016).

37. Sankaran, K. et al. Roles of histidine-103 and tyrosine-235 in the function of the prolipoprotein diacylglyceryl transferase of *Escherichia coli*. *J. Bacteriol.* **179**, 2944–2948 (1997).

38. Jain, R., Rivera, M. C. & Lake, J. A. Horizontal gene transfer among genomes: the complexity hypothesis. *Proc. Natl. Acad. Sci. USA* **96**, 3801–3806 (1999).

39. Makarova, K. S., Galperin, M. Y. & Koonin, E. V. Comparative genomic analysis of evolutionarily conserved but functionally uncharacterized membrane proteins in archaea: Prediction of novel components of secretion, membrane remodeling and glycosylation systems. *Biochimie* **118**, 302–312 (2015).

40. Bordier, C. Phase separation of integral membrane proteins in Triton X-114 solution. *J. Biol. Chem.* **256**, 1604–1607 (1981).

41. Hasebe, A., Saeki, A. & Shibata, K. Lipoprotein extraction from microbial membrane and lipoprotein/lipopeptide transfection into mammalian cells. in *Periodontal Pathogens* (eds. Nagano, K. & Hasegawa, Y.) vol. 2210 195–204 (Springer US, 2021).

42. Gavitt, T. D. et al. B cells oppose *Mycoplasma pneumoniae* vaccine enhanced disease and limit bacterial colonization of the lungs. *Npj Vaccines* **7**, 1–13 (2022).

43. Müller, G. A., Tschöp, M. H. & Müller, T. D. Upregulated phospholipase D activity toward glycosylphosphatidylinositol-anchored proteins in micelle-like serum complexes in metabolically deranged rats and humans. *Am. J. Physiol. Endocrinol. Metab.* **318**, E462–E479 (2020).

44. Sagami, H., Kikuchi, A. & Ogura, K. A novel type of protein modification by isoprenoid-derived materials. Diphytanylglycerylated proteins In Halobacteria. *J. Biol. Chem.* **270**, 14851–14854 (1995).

45. Maltese, W. A. & Erdman, R. A. Characterization of isoprenoid involved in the post-translational modification of mammalian cell proteins. *J. Biol. Chem.* **264**, 18168–18172 (1989).

46. Sagami, H., Kikuchi, A., Ogura, K., Fushihara, K. & Nishino, T. Novel isoprenoid modified proteins in halobacteria. *Biochem. Biophys. Res. Commun.* **203**, 972–978 (1994).

47. Albers, S.-V., Koning, S. M., Konings, W. N. & Driessen, A. J. Insights into ABC transport in archaea. *J. Bioenerg. Biomembr.* **36**, 5–15 (2004).

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

49. de Silva, R. T. et al. Improved growth and morphological plasticity of *Haloferax volcanii*. *Microbiology* **162**, 001012 (2021).

50. Duggin, I. G. et al. CetZ tubulin-like proteins control archaeal cell shape. *Nature* **519**, 362–365 (2015).

51. Schiller, H. et al. Identification of structural and regulatory cell-shape determinants in *Haloferax volcanii*. *Nat. Commun.* **15**, 1414 (2024).

52. Wolf, Y. I., Makarova, K. S., Yutin, N. & Koonin, E. V. Updated clusters of orthologous genes for Archaea: a complex ancestor of the Archaea and the byways of horizontal gene transfer. *Biol. Direct* **7**, 46 (2012).

53. Bergo, M. O. et al. Inactivation of *Icmt* inhibits transformation by oncogenic K-Ras and B-Raf. *J. Clin. Invest.* **113**, 539–550 (2004).

54. Backlund, P. S. Post-translational processing of RhoA: carboxyl methylation of the carboxyl-terminal prenylcysteine increases the half-life of RhoA. *J. Biol. Chem.* **272**, 33175–33180 (1997).

55. Tschumi, A. et al. Functional analyses of mycobacterial lipoprotein diacylglyceryl transferase and comparative secretome analysis of a mycobacterial lgt mutant. *J. Bacteriol.* **194**, 3938–3949 (2012).

56. Dautin, N. et al. Role of the unique, non-essential phosphatidylglycerol::prolipoprotein diacylglyceryl transferase (Lgt) in *Corynebacterium glutamicum*. *Microbiology* **166**, 759–776 (2020).

57. Edgar, R. C. Muscle5: High-accuracy alignment ensembles enable unbiased assessments of sequence homology and phylogeny. *Nat. Commun.* **13**, 6968 (2022).

58. Esterman, E. S., Wolf, Y. I., Kogay, R., Koonin, E. V. & Zhaxybayeva, O. Evolution of DNA packaging in gene transfer agents. *Virus Evol.* **7**, veab015 (2021).

59. Steinegger, M. et al. HH-suite3 for fast remote homology detection and deep protein annotation. *BMC Bioinforma.* **20**, 473 (2019).

60. Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree 2-approximately maximum-likelihood trees for large alignments. *PloS One* **5**, e9490 (2010).

61. Hallgren, J. et al. DeepTMHMM predicts alpha and beta transmembrane proteins using deep neural networks. 2022.04.08.487609 Preprint at https://doi.org/10.1101/2022.04.08.487609 (2022).

62. Mirdita, M. et al. ColabFold: making protein folding accessible to all. *Nat. Methods* **19**, 679–682 (2022).

63. Meng, E. C. et al. UCSF ChimeraX: Tools for structure building and analysis. *Protein Sci. Publ. Protein Soc.* **32**, e4792 (2023).

64. Kauri, T., Wallace, R. & Kushner, D. J. Nutrition of the halophilic archaebacterium, *Haloferax volcanii*. *Syst. Appl. Microbiol.* **13**, 14–18 (1990).

65. Dyall-Smith, M. The Halohandbook: protocols for haloarchaeal genetics. *Haloarchaeal Genet. Lab. Melb.* **14** (2008).

66. Gibson, D. G. et al. Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nat. Methods* **6**, 343–345 (2009).

67. Allers, T., Ngo, H.-P., Mevarech, M. & Lloyd, R. G. Development of additional selectable markers for the halophilic archaeon *Haloferax volcanii* based on the leuB and trpA genes. *Appl. Environ. Microbiol.* **70**, 943–953 (2004).

68. Tripepi, M., Imam, S. & Pohlschröder, M. *Haloferax volcanii* flagella are required for motility but are not involved in PibD-dependent surface adhesion. *J. Bacteriol.* **192**, 3093–3102 (2010).

69. Rattray, J. E. & Smittenberg, R. H. Separation of branched and isoprenoid glycerol dialkyl glycerol tetraether (GDGT) isomers in peat soils and marine sediments using reverse phase chromatography. *Front. Mar. Sci.* **7** (2020).

70. Connock, G. T. Owens, J. D. & Liu, X.-L. Biotic induction and microbial ecological dynamics of Oceanic Anoxic Event 2. *Commun. Earth Environ.* **3**, 1–10 (2022).

71. Schneider, C. A., Rasband, W. S. & Eliceiri, K. W. NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* **9**, 671–675 (2012).

72. Stirling, D. R. et al. CellProfiler 4: improvements in speed, utility and usability. *BMC Bioinforma.* **22**, 433 (2021).

Acknowledgements

Author contributions

Y.H., K.S.M., A.A.G, R.X., F.P., P.V.W., and M.P. designed research; Y.H., K.S.M., A.A.G, R.X., F.P., P.V.W., and M.P. performed research; Y.H., K.S.M., A.A.G, R.X., F.P., P.V.W., and M.P. analyzed data; Y.H., K.S.M., A.A.G, R.X., F.P., P.V.W., and M.P. wrote the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41467-025-63625-6.

Correspondence and requests for materials should be addressed to Mechthild Pohlschroder.

Peer review information *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.