frontiers | Research Topics

Molecular, cellular, and ecological processes of haloarchaea

Edited by
Yan Liao, Iain G. Duggin, Maria Ines Gimenez and Zhengshuang Hua

Published in
Frontiers in Microbiology

ISSN 1664-8714
ISBN 978-2-8325-6313-7
DOI 10.3389/978-2-8325-6313-7

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Molecular, cellular, and ecological processes of haloarchaea

Topic editors

Yan Liao — University of Technology Sydney, Australia

Iain G. Duggin — University of Technology Sydney, Australia

Maria Ines Gimenez — National University of Mar del Plata, Argentina

Zhengshuang Hua — University of Science and Technology of China, China

Citation

Liao, Y., Duggin, I. G., Gimenez, M. I., Hua, Z., eds. (2025). *Molecular, cellular, and ecological processes of haloarchaea*. Lausanne: Frontiers Media SA.

doi: 10.3389/978-2-8325-6313-7

Table of contents

frontiers | Frontiers in Microbiology

TYPE Editorial
PUBLISHED 24 April 2025
DOI 10.3389/fmicb.2025.1607148

OPEN ACCESS

EDITED AND REVIEWED BY
Sonja-Verena Albers,
University of Freiburg, Germany

*CORRESPONDENCE
Yan Liao
✉ yan.liao@uts.edu.au

RECEIVED 07 April 2025
ACCEPTED 11 April 2025
PUBLISHED 24 April 2025

CITATION
Liao Y, Duggin IG, Gimenez MI and Hua Z (2025) Editorial: Molecular, cellular, and ecological processes of haloarchaea. *Front. Microbiol.* 16:1607148. doi: 10.3389/fmicb.2025.1607148

Editorial: Molecular, cellular, and ecological processes of haloarchaea

Yan Liao[1]*, Iain G. Duggin[1], Maria Ines Gimenez[2] and Zhengshuang Hua[3]

[1]Australian Institute for Microbiology and Infection, University of Technology Sydney, Ultimo, NSW, Australia, [2]Instituto de Investigaciones Biológicas (IIB-CONICET-UNMDP), Universidad Nacional de Mar del Plata, Mar del Plata, Argentina, [3]Department of Environmental Science and Engineering, University of Science and Technology of China, Hefei, China

KEYWORDS

haloarchaea, molecular, cellular, ecological, interaction

Editorial on the Research Topic
Molecular, cellular, and ecological processes of haloarchaea

Introduction

Halophilic archaea (salt-loving archaea), termed haloarchaea, inhabit some of the most hypersaline environments on Earth. As one of the largest lineages within the domain of Archaea, they serve as invaluable models for studying archaeal cell biology due to their relative ease of cultivation, rapid growth kinetics, and an expanding repertoire of genetic tools. Their remarkable adaptations at the molecular, cellular, and ecological level have driven significant advancements in various aspects of archaeal biology, including gene expression and regulation, protein synthesis and adaptation, cell shape and division, and evolutionary processes. Additionally, their unique biochemical and physiological traits have paved the way for diverse biotechnological applications.

This editorial highlights the key contributions from nine recently published manuscripts in the Research Topic "*Molecular, cellular, and ecological processes of haloarchaea*." These studies explore diverse aspects of haloarchaea, including genetic tool development, genomic adaptations, cellular mechanisms, structural dynamics, and ecological interactions.

Transcriptional adaptation to environmental conditions

The transcriptional behavior of haloarchaea in controlled laboratory conditions may not fully represent their responses in natural environments. Rosselli et al. investigated the metatranscriptomes of the dominant *Haloquadratum walsbyi* from the solar saltern of Santa Pola (Alicante, Spain), comparing its transcription patterns in natural environments to those observed in laboratory-grown strains. This study reveals significant differences in gene expression, likely due to the adaptation of the cultured strain to homogenous laboratory conditions, whereas natural populations respond dynamically to heterogeneous environmental factors such as nutrient competition, viral attack, and other stressors. This study underscores the limitations of extrapolating laboratory-based gene expression data to natural environments and highlights the complex regulatory strategies employed by *H. walsbyi* to thrive in fluctuating hypersaline ecosystems.

Virus-host interactions in haloarchaea

One of the major stressors shaping the evolution of haloarchaea in natural environments is virus infection, which can drive genomic changes and influence cellular defense strategies. Mercier et al. presented an in-depth characterization of an archaeal tailed virus (HRTV-DL11) and its host *Halorubrum lacusprofundi*, a model organism for cold adaptation. They investigated the virus life cycle, host range, transcriptional responses to infection, and virus escape mutants. Notably, the researchers discovered that a laboratory-derived strain (*ACAM34-UNSW*) had lost its megaplasmid and ~38% of its secondary chromosome, resulting in the absence of major virus defense systems and making the strain highly susceptible to HRTV-DL1 infection. In contrast, the type strain (*ACAM34_DSMZ*) retained these defense mechanisms, preventing virus replication. Comparative infection studies between these strains enabled the identification of host response specifically activated upon the loss of virus defense mechanisms, broadening our understanding of virus-host interactions in haloarchaea.

Lipid profiles for potential biomarker discovery

Archaea have a unique lipid composition, characterized by isoprenoid alkyl chains and ether linkages to glycerol-1-phosphate. Yao et al. presented a comprehensive lipidomic study of seven haloarchaeal strains, building an in-house lipid library by using high-performance liquid chromatography coupled with mass spectrometry. They identified a total of 162 lipid features, corresponding to 107 lipids that could be assigned to specific strains. Clustering analyses of both core lipids and total lipid profiles closely aligned with the phylogeny of Halobacteria at the order level. The study also demonstrated how lipidomic features can facilitate the linkage of unknown lipid compounds to phylogeny, refining our understanding of archaeal lipid evolution. This work showcases the power of lipidomics in helping decipher archaeal phylogeny, and offers new avenues for evolutionary studies and biomarker discovery.

Carbohydrate metabolism

Carbohydrate metabolism is essential for haloarchaeal survival in hypersaline environments. Sorokin et al. described the isolation and characterization of a novel haloarchaeal genus and species *Natronoglomus mannanivorans*, capable of utilizing beta-mannans—plant plant-derived polysaccharides—as a primary carbon source. The researchers enriched the cultures with different forms of beta-mannans and isolated four closely related mannan-degrading strains. Phylogenomic analysis placed these isolates within a new genus in the family *Natrialbaceae,* class *Halobacteri*a. These moderate alkaliphiles exhibited extreme halophilicity and aerobic saccharolytic metabolism, enabling them to utilize not only mannans but also cellulose, xylan, and xyloglucan as carbon sources. This study also highlights the metabolic versatility of halophiles in breaking down complex polysaccharides.

Advances in protein expression and purification

Efficient expression and purification of extremophilic proteins remain a challenge due to issues such as misfolding and aggregation in conventional hosts like *Escherichia coli*. Karan et al. systematically investigated the impact of various purification tags and their placements on protein expression and purification in *Haloferax voclanii*. Their results demonstrated that an N-terminal 8× His-tag or Strep-tag®II significantly improves protein production, purity, and yield in *H. volcanii*. The optimal tag positioning varied depending on the protein, with the C-terminal 8X His-tag or Strep-tag®II enhancing mCherry expression, and an N-terminal configuration benefiting halophilic alcohol dehydrogenase. Notably, co-positioning 8× His-tag and Twin-Strep-tag® at the N-terminus led to substantially improved expression across all tested proteins including sfGFP. These findings provide valuable insights into purification tag design, offering a framework for optimizing protein production and purification in haloarchaea.

Nutrient adaptations

Archaeal survival in nutrient-limited environments relies on their ability to regulate essential elements, such as iron, which is crucial for cellular function and microbial physiology. Sailer et al. investigated the proteomic response of *H. volcanii* to iron starvation using data-independent acquisition mass spectrometry. By comparing cells grown under normal and low-iron conditions, this study revealed that iron starvation severely retarded *H. volcanii* growth and altered the abundance of numerous proteins involved in metal transport and cellular homeostasis. They identified the most affected protein, HVO_A0467, a member of RND family permease, is not essential under standard conditions, and may function as a siderophore exporter. This study provides novel insights into siderophore-based metal acquisition strategies in haloarchaea, offering valuable perspectives on microbial survival in iron-limited environments.

Spatial organization and motility regulation

Cellular organization plays a pivotal role in sharing the survival and adaptability of haloarchaea. Proteins of the ParA/MinD family of ATPases play critical roles in the spatial organization of diverse cellular structures in both bacteria and archaea. Two manuscripts in this Research Topic investigated the roles of MinD proteins in *Haloferax voclanii*, revealing their influence on the positioning of motility structures near the poles of the rod shaped cells. Brown and Duggin investigated the interplay between MinD proteins and the cytoskeletal protein CetZ1 (an archaea-specific homolog of FtsZ and tubulin). A key finding of this work was the identification of MinD2 as a critical regulator of CetZ1 localization and motility. Deletion of *minD2* or *minD4* disrupted CetZ1's normal distribution, leading to its mislocalization along the midcell and cell edges while preventing its proper accumulation at the

cell poles. Similarly, MinD4 also influenced CetZ1 positioning, though its effects were weaker compared to MinD2. This study provides the first direct evidence of MinD proteins regulating the positioning of tubulin superfamily proteins in archaea. In a complementary study, Patro et al. identified MinD2 as a determinant of motility and cell shape, influencing the positioning of chemosensory arrays and archaellum motors. Additionally, they found that MinD2 and MinD4 function synergistically to ensure proper cellular organization, highlighting the complexity of MinD-mediated processes in haloarchaea. Together, these two studies underscore the intricate regulation of motility and spatial organization in haloarchaea, improving our understanding on the functional diversity of this large family of ATPases in bacteria and archaea.

Protease regulation

Intramembrane proteases play critical roles in cellular regulation, but their functions in archaea remain poorly understood. Costa et al. investigated the role of rhomboid proteases in *H. volcanii*, focusing on two homologs, rho1 (HVO_1474) and rho2 (HVO_0727), which were not essential for *H. volcanii* viability but significantly influenced motility and biofilm formation. The Δ*rho1* mutant exhibited increased motility and biofilm formation, reduced adhesion to glass surfaces, with profound morphological changes. In contrast, the double mutant (Δ*rho1* Δ*rho2*) showed enhanced adhesion, a mild reduction in motility (similar to Δ*rho2*), and fewer morphological abnormalities. Complementation studies revealed the function of the two proteases only partially overlapped, providing new perspectives on proteolytic regulation of biofilm formation and cell shape in archaea.

Conclusion

These nine manuscripts significantly advance our understanding of the *Molecular, cellular, and ecological processes of haloarchaea*. As research in this field continues to evolve, these studies contribute to a strong foundation for future discoveries, and highlight the importance of haloarchaea as model organisms in archaeal research. We extend our sincere appreciation to all contributing authors, reviewers, and editorial board members for their efforts in assembling this compelling body of Research Topic. We hope that this Research Topic will serve as a valuable resource for researchers in microbiology, molecular biology, and applied biotechnology.

Author contributions

YL: Writing – review & editing, Writing – original draft. ID: Writing – review & editing. MG: Writing – review & editing. ZH: Writing – review & editing.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

frontiers | Frontiers in Microbiology

TYPE Original Research
PUBLISHED 10 November 2022
DOI 10.3389/fmicb.2022.1044446

OPEN ACCESS

EDITED BY
Haiyuan Cai,
University of Oklahoma,
United States

REVIEWED BY
Heng-Lin Cui,
Jiangsu University, China
Shengwei Hou,
Southern University of Science and Technology, China

*CORRESPONDENCE
Henk Bolhuis
henk.bolhuis@nioz.nl

SPECIALTY SECTION
This article was submitted to Extreme Microbiology, a section of the journal Frontiers in Microbiology

RECEIVED 14 September 2022
ACCEPTED 25 October 2022
PUBLISHED 10 November 2022

CITATION
Rosselli R, López-Pérez M, Martin-Cuadrado A-B, Rodriguez-Valera F and Bolhuis H (2022) Differences in gene expression patterns between cultured and natural *Haloquadratum walsbyi* ecotypes. *Front. Microbiol.* 13:1044446. doi: 10.3389/fmicb.2022.1044446

Differences in gene expression patterns between cultured and natural *Haloquadratum walsbyi* ecotypes

Riccardo Rosselli[1,2,3], Mario López-Pérez[4], Ana-Belen Martin-Cuadrado[2,4], Francisco Rodriguez-Valera[3,4] and Henk Bolhuis[1]*

[1]Department of Marine Microbiology and Biogeochemistry, Royal Netherlands Institute for Sea Research (NIOZ), Den Hoorn, Netherlands, [2]Department of Physiology, Genetics and Microbiology, University of Alicante, Alicante, Spain, [3]LABAQUA S.A., Research & Development Department, Las Atalayas, Alicante, Spain, [4]Evolutionary Genomics Group, División de Microbiología, Universidad Miguel Hernández, San Juan de Alicante, Spain

Solar crystallizer ponds are characterized by high population density with a relatively simple community structure in terms of species composition. The microbial community in the solar saltern of Santa Pola (Alicante, Spain), is largely dominated by the hyperhalophilic square archaeon *Haloquadratum walsbyi*. Here we studied metatranscriptomes retrieved from a crystallizer pond during the winter of 2012 and summer of 2014 and compared *Hqr. walsbyi's* transcription patterns with that of the cultured strain *Hqr. walsbyi* HBSQ001. Significant differences were found between natural and the cultured grown strain in the distribution of transcript levels per gene. This likely reflects the adaptation of the cultured strain to the relative homogeneous growth conditions while the natural species, which is represented by multiple ecotypes, is adapted to heterogeneous environmental conditions and challenges of nutrient competition, viral attack, and other stressors. An important consequence of this study is that expression patterns obtained under artificial cultivation conditions cannot be directly extrapolated to gene expression under natural conditions. Moreover, we found 195 significantly differential expressed genes between the seasons, with 140 genes being higher expressed in winter and mainly encode proteins involved in energy and carbon source acquiring processes, and in stress responses.

KEYWORDS

RNA-seq, metatranscriptome, solar saltern, archaea, *Haloquadratum walsbyi*

Introduction

Microbial communities in hypersaline crystallizer ponds have been extensively studied for over 50 years (Oren, 1994). Whereas initial research focused on the physiological characterization of halophilic bacterial and archaeal isolates, developments in high-throughput nucleotide sequencing strongly reduced cultivation efforts. Initial analyzes of

the 16S rRNA gene through Sanger sequencing (Benlloch et al., 1995) provided insight in the taxonomic diversity, while current metagenomic approaches address the genetic potential of halophiles in crystallizers around the world (Legault et al., 2006; Sanchez-Perez et al., 2008; Narasingarao et al., 2011; Plominsky et al., 2014; Naghoni et al., 2017). However, these techniques provide little information about physiological activities of halophiles in their natural environments. A high-resolution approach to deduce active microbial processes in natural habitats is metatranscriptomics as the analysis of transcribed messenger RNA (mRNA) provides information on which genes are actively expressed under extant conditions. Although a one-to-one relationship between gene expression and protein synthesis (or enzymatic activity) is not always evident due to post-transcriptional (Fu et al., 2013; Peterson et al., 2016) and translational modifications (Redon et al., 2005; Li et al., 2014), metatranscriptomics offer a good insight into gene regulation and environmental responses.

In contrast to experiments with monocultures growing under well-defined laboratory conditions, natural communities are exposed not only to numerous environmental fluctuations (light, dark, cold, hot, rain, desiccation, etc.) but also to interactions with other microorganisms and with bacteriophages. Examples of interspecies effects on gene expression patterns are the mutualistic interactions between the marine bacteria *Prochlorococcus* and *Alteromonas macleodii* (Kimes et al., 2014; Aharonovich and Sher, 2016; Biller et al., 2016) between *Synechococcus* sp. PCC7002 and *Shewanella putrefaciens* (Beliaev et al., 2014). Transcriptomics can also reveal antagonistic interactions such as found between the two fish-pathogens, *Moritella viscosa* and *Aliivibrio wodanis* (Hjerde et al., 2015). Interestingly, also within a single species, competition between closely related strains of *Salinibacter ruber* (M8 and M31) revealed a significant change in their transcription patterns when growing in mono- or coculture (Gonzalez-Torres et al., 2015). Transcriptomics was also used to uncover expression of potential mating factors in two near identical mutants of *Haloferax volcanii* (Makkay et al., 2020). Especially genes encoding glycosylation proteins and those with unknown function were overexpressed under mating conditions. Along these lines, several antagonistic interactions were described among extremely halophilic prokaryotes due to the production of antimicrobial metabolites such as halocins (Atanasova et al., 2013; Ghanmi et al., 2016). We therefor need to be aware that gene expression patterns of key species under laboratory conditions may not necessarily be directly translated to the natural behavior of these species.

Here we analyzed and compared gene expression patterns in the cultivated pure culture of *Hqr. walsbyi* strain HBSQ001 with that of natural occurring *Hqr. walsbyi ecotypes* in the crystallizer pond CR30 of the Bras del Port saltern, Santa Pola (Alicante, Spain). *Hqr. walsbyi* is notorious for its peculiar square morphology and largely dominates the crystallizer (Bolhuis et al., 2004) with numbers up to 10^8 cells per ml and constituting up to 80% of the total microbial population (Benlloch et al., 2001; Ghai et al., 2011). This crystallizer basin is the same from which strain HBSQ001 was originally isolated (Bolhuis et al., 2004) and has been the focus of a series of molecular studies over the last 25 years, and substantial information regarding its microbial composition is well documented (Benlloch et al., 1995; Martínez-Murcia et al., 1995; Benlloch et al., 1996; Antón et al., 1999; Benlloch et al., 2002; Øvreås et al., 2003; Papke et al., 2004; Ventosa, 2006; Ghai et al., 2011; Ventosa et al., 2014). A pure culture of *Hqr. walsbyi* strain HBSQ001 has already been the subject of a transcriptome analysis (Bolhuis et al., 2017) and is compared to a metatranscriptome analysis of the CR30 crystallizer inhabiting *Hqr. walsbyi* community that was sampled at two different seasons (summer and winter).

Materials and methods

Sampling

Five liters of water sample were collected in January 2012 (Winter Sample) and July 2014 (Summer Sample) from CR30, a crystallizer pond located in the multi-pond solar saltern Bras del Port (Santa Pola, Alicante, Spain, 38°11′43.0"N 0°35′30.7"W). A hand refractometer was used to measure the salinity which in winter was found to be 31.5% (315 g kg^{-1}) at a brine temperature of 12°C and in summer the salinity was 37.9% (379 g kg^{-1}) at a brine temperature of 27°C. Samples were sequentially filtered through 20, 5 µm pore size polycarbonate filter, and 0.22 µm pore size Sterivex filter (Durapore; polycarbonate filters (Millipore, Billerica, MA, United States) and immediately frozen on dry ice in an RNA-stabilizing agent (RNAlater® Solution; Ambion, United States). Samples were stored at −80°C until processing.

RNA isolation and cDNA synthesis

RNA isolation and cDNA synthesis was largely performed as previously described (Kimes et al., 2014), see details below. Before total RNA extraction with the RNeasy® Mini Kit (Qiagen,) in accordance with the instruction from the manufacturer, filters were treated with 500 µl of TE (Tris–HCl 10 mM, EDTA 1 mM, pH 8.0) containing lysozyme (2 mg/ml) and proteinase K (0.4 mg/ml) for 10 min at room temperature. Residual genomic DNA was removed from the extracted RNA by DNAse I treatment (Sigma-Aldrich) for 30 min at room temperature. Agarose gel electrophoresis and staining confirmed the absence of genomic DNA from the extracted RNA. Total RNA (10 µg) was used to make single-stranded cDNA using the High-Capacity cDNA Reverse Transcription kit (Applied Biosystems) following the manufacturer's instructions. The second strand was synthesized by adding 30 units (U) of *Escherichia coli* Polymerase I (New England Biolabs), 5 U of *E. coli* DNA Ligase (New England Biolabs), 5 U of RNase H (Epicentre), 300 µM of dNTPs (Invitrogen) to the first strand reaction. After 2.5 h at 16°C, 5 U of T4-DNA polymerase (New England Biolabs) was added and

incubated for 40 min at 16°C. Finally, the double-stranded cDNA was cleaned with a QIAquick PCR Purification kit (Qiagen) and quantified using the ND-1000 Spectrophotometer (NanoDrop, Wilmington, United States). The quality of all cDNA samples was determined on an Agilent 2,100 bioanalyzer. Paired-end sequencing of the cDNA was performed at EBI using half a lane of Illumina HiSeq2000 with 100 cycles per run (PE100).

Metatranscriptome analysis

Reads from metatranscriptomes were pre-processed by Trimmomatic (Bolger et al., 2014) in order to remove low-quality bases (minimum Phred-quality score of 20 in a 4-base sliding windows) and reads shorter than 40 bases. The software package EMIRGE (Miller et al., 2011) was used to reconstruct full-length small subunit ribosomal genes and to provide length-normalized estimation the relative abundances. The assembled 16S rRNA sequences were clustered at 97% identity using the software cd-hit (Li and Godzik, 2006) and annotated against the SILVA version 128 reference database (Pruesse et al., 2007). Trimmed sequences were mapped against *Hqr. walsbyi* HBSQ001 (chromosome NC_008282.1, plasmid NC_008213.1) using edge-pro (Magoc et al., 2013). The number of reads aligned on each gene were counted and normalized to give TPM values (transcripts per million) (Wagner et al., 2012). Transcriptomic reads were analyzed and functional annotated using GhostKOALA:KEGG tools (Kanehisa et al., 2016). Transcriptomic reads from cultured *Hqr. walsbyi* HBSQ001 (Bolhuis et al., 2017) were analyzed following the same pipeline. Metatranscriptome reads from the winter and summer sample were deposited at the NCBI SRA sequence read archive (https://www.ncbi.nlm.nih.gov/sra) under bioproject number PRJNA633445.

Results

Metatranscriptome characteristics

More than 27 million and 35 million reads were obtained for the summer and winter samples, respectively (Table 1). Respectively 21% (summer) and 13% (winter) were annotated as 16S rRNA coding reads. Assembly and taxonomic annotation of 16S and 18S ribosomal RNA genes using EMIRGE revealed 91.4% of archaeal ribosomal transcripts sequences in the summer samples and 77.1% in winter samples (Supplementary Dataset-1, Sheet-1). Bacterial derived 16S rRNA reads accounted for 8.4% of the total number in summer transcripts but were nearly 3-fold more abundant during winter (22.7%). The eukaryotic fraction, mainly constituted by the halophilic unicellular algae *Dunaliella* sp., remained stable at 0.2%. Most of the 16S rRNA reads were assigned to the genus of *Haloquadratum*, making up 55% of the total expressed rRNA genes in summer and 47% in winter. Other abundant active genera during the summer period were

TABLE 1 General statistics of the winter and summer metatranscriptomes and reads recruited to *Hqr. walsbyi* HBSQ001.

	Summer	Winter
Total number of reads	27.765.941	35.609.827
% *Hqr. walsbyi* recruited reads	31%	25%
% *Hqr. walsbyi* reads assigned as mRNA	28% (8.7% of total)	40% (10% of total)
Number of genes covered[1]	2.614	2.616
Average read number per gene	884	1.355
Fold coverage per gene	74	109

[1]Only genes were included that recruited at least 10 reads per gene.

Haloarcula (12%) and *Salinibacter* (7%). The winter population revealed a 5-fold lower number of *Haloarcula* assigned reads (2%), while *Salinibacter* accounted for a relative increase up to the 16% of the total 16S-rRNA reads. Other haloarchaeal genera that were significantly present at more than 1% of in the summer community were identified as novel haloarchaeal genera (unclassified reads) together with *Halolamina*, *Natrialba*, *Halobellus*, *Halonotius*, *Halorubrum* and *Halosimplex*. With the exception of the novel haloarchaeal genera, all of these had a lower abundance in the winter sample, especially *Halolamina* and *Natrialba*, which were 18-fold and 11-fold less abundant, respectively.

Difference in gene expression patterns between natural and culture grown *Hqr. walsbyi* HBSQ001

Comparison of the distribution of the normalized number of transcripts per gene (TPM) revealed a clear distinction between the culture transcriptome and environmental metatranscriptome. The TPM distribution in the summer and winter derived metatranscriptomes follow a bell-shape curve with a median of 399 TPM (summer) and 387 TPM (winter), close to the genome average expression of 381 TPM (Figure 1). In the culture transcriptome however, the histogram reveals that most genes are low expression (around 50 to 150 TPM) with a median of 132 TPM whereas a small number of genes (144) have expression levels well above 1,000 TPM. Many genes that are low expressed in culture but significantly expressed in nature are uncharacterized, hypothetical proteins and IS elements. An interesting example is a small, ~10Kb, region of genes mostly encoding hypothetical proteins (HQ_RS07905 - HQ_RS07945) that are expressed in nature but not, or barely, in culture (Supplementary Dataset-1, Sheet-2). This region includes the potential L-lactate permease (*lctP*) that revealed an average expression in both metatranscriptomes (377 TPM) while not a single read was assigned to this gene in the cultured strain. Amongst the genes that are overexpressed

10.3389/fmicb.2022.1044446

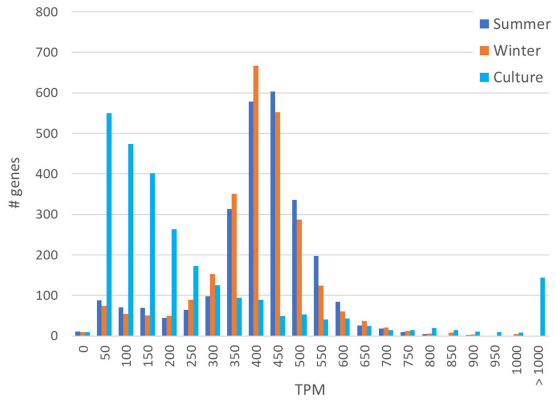


FIGURE 1
Histogram of TPM distribution and number of genes involved in the two metatranscriptomes (summer and winter) in comparison to the transcriptome of the cultivated isolate *Hqr. Walsbyi* HBSQ001.

in the culture but low expressed in nature, there is a 1,626 amino acid cell surface adhesin protein coding gene (HQ_RS01080; 7736 TPM in culture and <100 TPM in nature), two genes encoding uncharacterized proteins, HQ_RS11710 and HQ_RS12980 (respectively 827 and 818 TPM in culture versus <7 TPM in nature) and two genes, *neuA* and *neuB*, involved in sialic acid synthesis (both only expressed in culture but not in nature). Genes that are expressed significantly higher in nature than in culture include ABC-type transporters involved in transport of macrolides, nickel or nitrates and several major facilitator superfamily transport proteins for unknown substrates. The *coxMSL* gene cluster is between 13 and 27-fold higher expressed in nature than in culture.

Seasonal difference in gene expression patterns

Normalized average transcription for 2,625 mRNA coding genes is 381 TPM. The genome of *Hqr. walsbyi* HBSQ001 recruited 31% of the total number of reads from the summer sample and 25% from the winter sample. Messenger RNA genes accounted for 28% (summer) and 40% (winter) of the *Hqr. walsbyi* recruited reads and recovered 2,614 and 2,616 genes of the previously 2,625 identified genes on the genome (Table 1). The average number of reads per gene was 884 (summer) and 1355 (winter) with an estimated 74- and 109-fold coverage per gene.

Comparison of TPM normalized gene expression levels in *Hqr. walsbyi* HBSQ001 between summer and winter derived metatranscriptomes (Figure 2) revealed that out of the 2,625 identified genes, 1,136 genes had a similar expression level (defined as a difference in TPM between summer and winter less than one tenth of the average expression: 38 TPM). In total, 795 genes were higher expressed in summer than in winter and 694 genes were higher expressed in winter relative to the summer (Supplementary Dataset-1, Sheet-3). Significant seasonal different expression was established for genes that had a combined expression (summer + winter) of at least 76 TPM (2 times one tenth of the average) and at least 1.5-fold difference in expression between summer and winter. This resulted in 195 differentially expressed genes, 55 of which were higher expressed in summer and 140 were higher expressed in winter (Supplementary Dataset-1, Sheet-3).

Out of the 55 genes that were higher expressed in summer than in winter, 37 are annotated as hypothetical- or uncharacterized protein coding genes. Summer high expressed genes with predicted functions encode an A-type ATPase subunit H (*atpH*), a Sec-independent translocase (*tatA*), ferredoxin proteins (*fer7* and *fer2*) and three thioredoxin proteins (*trxA6*, *trxA1* and *trxA5*). Out of the 140 genes that were higher expressed in winter than in summer, 14 genes were found to be at least 2.5-fold higher expressed than in summer. Highest differential expression (3.5-fold) was found for the glycerol-kinase encoding gene *glpK*. The gene HQ_RS03760, immediately upstream of *glpK* and well conserved amongst halophilic archaea, was also highly expressed in winter. Genes encoding the dihydroxyacetone PTS system are above average expressed and the dihydroxyacetone kinase subunit-L coding gene (*dhaL)* was significant higher (1.58-fold) expressed in winter than in summer. Other highly differentially expressed genes in winter were *sufB2*, *sufC* and *sufB1* (encoding ferredoxin assembly proteins), *bopI* (encoding the proton translocating bacteriorhodopsin) and *uraA2* (encoding a xanthine/uracil permease). Only 18.5% of the genes encoded for hypothetical or uncharacterized proteins. Furthermore, the winter sample revealed a higher expression of genes encoding the A-type ATPase subunit F (*atpF*), the gas-vesicle protein A (*gvpA*) and the NADH dehydrogenase-like complex subunits A and B (*nuoA* and *nuoB*).

Overall gene expression levels

In addition to seasonal differences in expression, we also looked at the relative expression levels in both samples. Protein coding genes were considered highly expressed if their expression level is at least 1.5 times higher than the average expression (1.5×381 TPM = 571 TPM). In the summer sample, 99 genes were found to be significantly higher expressed and, in winter, 134 genes were expressed above the average. Among these, there are 30 genes that were found highly expressed in both samples. The top twenty-five protein-coding genes in summer and winter are listed in Tables 2, 3 respectively. In summer, the top twenty-five contains five yet uncharacterized proteins and includes gene HQ_RS01075, the overall highest expressed gene in both the summer and winter sample (Table 2). Also highly expressed in summer were the 50S and 30S ribosomal protein coding genes, three IS transposases, genes involved in ferredoxin assembly (*fer7, sufC* and *sufB1*), thiosulfate sulfurtransferase and thioredoxin biosynthesis. In

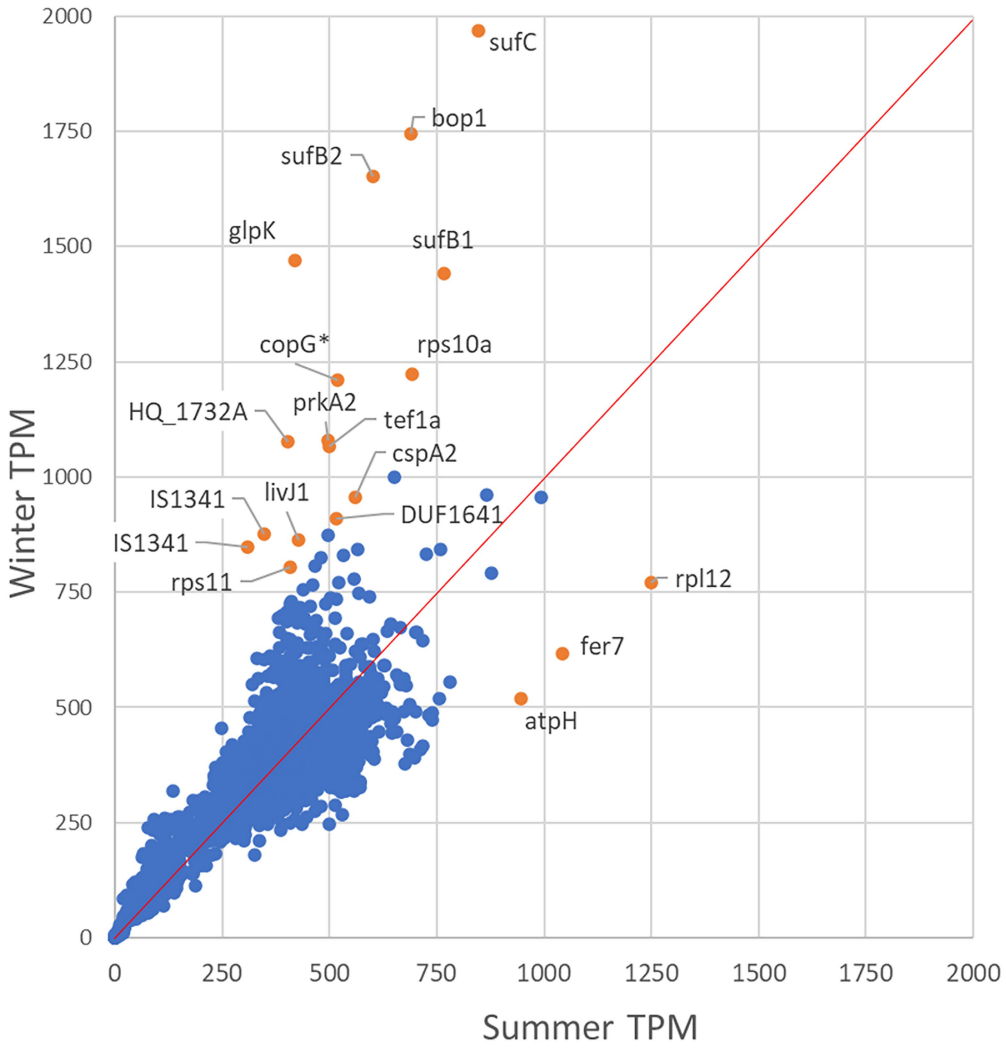


FIGURE 2
Differential expression analysis. The scatter plot shows the gene expression level in the summer (x-axis) versus the winter population (y-axis). Genes below the 1 tot 1 expression line are higher expressed in summer and above the 1 to 1 line are higher expressed in winter. Significantly differentially expressed genes are indicated in orange.

winter, the top twenty-five, headed by HQ_RS01075, also contains an operon with genes involved in ferredoxin assembly (*sufC, sufB1* and *sufB2*), the same two 30S ribosomal protein coding genes as found in summer, IS transposases (two of them were also detected in the summer top 25) and uncharacterized protein coding genes (Table 3). The genes *bop1* and *glpK* were found at the third and fifth most abundant position, respectively. Nine out of the 25 highest expressed genes in summer were also found in the top 25 of the winter sample (Tables 2, 3, bold). Among them, there are several genes involved in stress response; i.e., the thermosome subunits Ths1 and Ths2, the cold shock proteins CspA1 and CspA2, the ATP-dependent protease Lon and the superoxide dismutase Sod.

We also analyzed expression levels of viral genes from the 14 *in-silico* assembled viral genomes assigned to *Hqr. walsbyi* based on the presence of CRISPR spacers (Garcia-Heredia et al., 2012). Curiously, the viral genes were almost one order of magnitude higher expressed in winter than in summer (Figure 3). Exceptions are eHP-9, eHP-41 and eHP-D7 that were barely or not over the full length recruited in the winter and summer samples.

Regarding the low expressed genes, the summer and winter samples contained, respectively, 11 and 9 genes that were not expressed at all, 8 of them were absent from both samples. These unexpressed genes mainly encode uncharacterized and hypothetical proteins, but also encode for acylneuraminate cytidylyltransferase (*neuA*) and sialic acid synthase (*neuB*). Notably, amongst the low expressed genes are several sulfatases, glycosyl transferases and extracellular glycoprotein coding genes including the two short-halomucin coding genes (*hmu2* and *hmu3*). The long-halomucin coding gene *hmu* is expressed at 0.6 (summer) and 0.7 (winter)-fold average.

TABLE 2 Top 25 of highest expressed protein coding genes in the summer population of *Hqr. walsbyi* HBSQ001.

Gene ID	TPM	*gene*	Annotation
HQ_RS01075	**2,380**	-	**Uncharacterized protein**
HQ_RS09700	1,251	*rpl12*	50S ribosomal protein L12
HQ_RS12535	1,043	*fer7*	Ferredoxin (2Fe-2S)
HQ_RS13010	**993**	-	**ISH8-type transposase ISHwa8**
HQ_RS11475	945	*atpH*	A-type ATP synthase subunit H
HQ_RS09370	878	*rpl29*	50S ribosomal protein L29
HQ_RS05300	**866**	-	**ISH9-type transposase ISHwa1**
HQ_RS03630	**847**	***sufC***	**Fe-S cluster assembly ATPase SufC**
HQ_RS08655	780	-	CopG domain protein
HQ_RS03635	**767**	***sufB1***	**SufB domain protein**
HQ_RS09925	**759**	***rps13***	**30S ribosomal protein S13**
HQ_RS13890	755	*tssA2*	Thiosulfate sulfurtransferase
HQ_RS11220	740	-	Uncharacterized protein
HQ_RS01670	738	*trxA1*	Thioredoxin
HQ_RS02195	732	-	Uncharacterized protein
HQ_RS01400	**725**	***cspA1***	**Cold shock protein**
HQ_RS09285	716	-	Uncharacterized protein
HQ_RS12580	716	-	Uncharacterized protein
HQ_RS06655	712	-	Cupin 2 barrel domain protein
HQ_RS09630	704	*rpl8e*	50S ribosomal protein L8e
HQ_RS04415	701	*cspA3*	Cold shock protein
HQ_RS07460	700	-	ISHwa2-type transposase ISHwa2
HQ_RS07640	698	*tatA*	Sec-independent protein translocase subunit TatA
HQ_RS12175	**692**	***rps10a***	**30S ribosomal protein S10a**
HQ_RS00075	**689**	***bop1***	**Bacteriorhodopsin I**

Genes in bold indicated genes that are highly expresses in both summer and winter.

TABLE 3 Top 25 of highest expressed protein coding genes in the winter population of *Hqr. walsbyi* HBSQ001.

Gene ID	TPM	*gene*	Annotation
HQ_RS01075	**2,380**	-	**Uncharacterized protein**
HQ_RS03630	**847**	***sufC***	**Fe-S cluster assembly ATPase SufC**
HQ_RS00075	**689**	***bop1***	**Bacteriorhodopsin I**
HQ_RS03640	601	*sufB2*	SufB domain protein
HQ_RS03765	419	*glpK*	Glycerol kinase
HQ_RS03635	**767**	***sufB1***	**SufB domain protein**
HQ_RS12175	**692**	***rps10a***	**30S ribosomal protein S10a**
HQ_RS02125	519	-	CopG domain protein
HQ_RS10230	496	*prkA2*	Probable PrkA-type serine/threonine protein kinase
HQ_RS03760	404	-	Uncharacterized protein
HQ_RS12180	500	*tef1a*	Translation elongation factor aEF-1 alpha subunit
HQ_RS01950	651	-	Uncharacterized protein
HQ_RS05300	**866**	-	**ISH9-type transposase ISHwa1**
HQ_RS13010	**993**	-	**ISH8-type transposase ISHwa8**
HQ_RS02000	560	*cspA2*	Cold shock protein
HQ_RS00700	515	-	DUF1641 family protein
HQ_RS05510	347	-	IS1341-type transposase
HQ_RS02130	498	*ftsZ2*	Cell division protein FtsZ, type II
HQ_RS03190	428	*livJ1*	ABC-type transport periplasmic substrate-binding protein
HQ_RS01835	310	-	IS1341-type transposase
HQ_RS07465	564	*sod*	Superoxide dismutase (Mn)
HQ_RS09925	**759**	***rps13***	**30S ribosomal protein S13**
HQ_RS01400	**725**	***cspA1***	**Cold shock protein**
HQ_RS05000	533	*aldH2*	Aldehyde dehydrogenase
HQ_RS00695	480	-	Probable anaerobic dehydrogenase alpha subunit

Genes in bold indicated genes that are highly expresses in both summer and winter.

Genome expressed hot spots

Analysis of the genome localization of highly differentially expressed genes revealed three specific regions with a higher than average difference (1.22-fold) in gene expression between summer and winter (Supplementary Dataset-1, Sheet-4). The first region, with an average differential expression of 1.41, ranges from nucleotide position 849,423–900,876, and contains the above-described genes *sufB1*, *sufB2*, *sufC*, *glpK* and HQ_RS03760. The second region has an average differential expression of 1.41 and ranges from nucleotide position 1,270,976–1,345,673. It partly overlaps with one of the six previously identified genomic islands (Cuadros-Orellana et al., 2007), HQ_GI2, and encodes mainly hypothetical and uncharacterized proteins. Three genes encode known proteins and are involved in nucleotide metabolism; i.e. a xanthine/uracil permease family transport gene protein (*uraA2*; 2.48-fold), a purine phosphoribosyltransferase (*apt2*; 1.68-fold), and a signal transduction histidine kinase (HQ_RS14555; 1.90-fold). The third region is found at the end of the genome (2,867,030 to 2,981,278) and partly overlapping with genomic island HQ_GI4. It contains 51 genes, and, with 1.86-fold difference, it was the highest differential expressed region by far. Only one these 51 genes was found to be higher expressed in summer and encodes a hypothetical protein (HQ_RS12985; 1.85-fold difference). High expressed genes in this region consist of

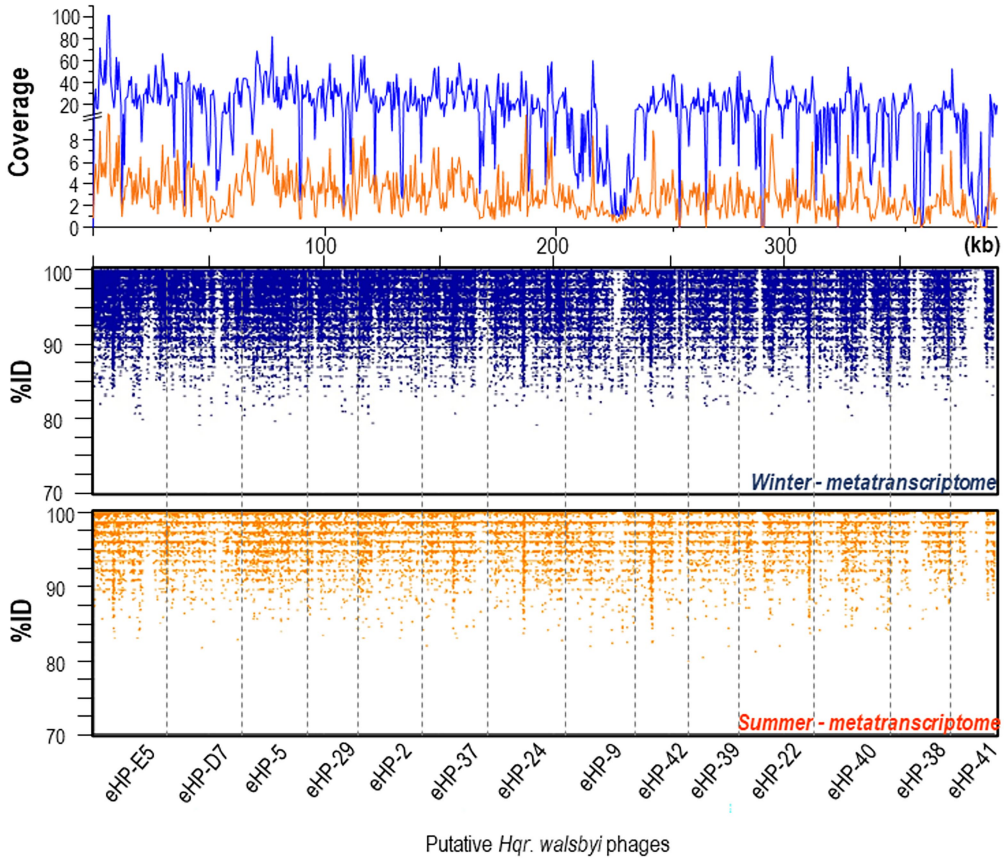


FIGURE 3
Recruitment plots of putative *Hqr. walsbyi* virus genomes (eHP-) against metatranscriptomic reads from summer (bottom panel) and winter (middle panel) samples. (Top panel) Coverage plot of the metatranscriptomic reads mapped onto the viral genes in summer (orange) and winter (blue).

several genes encoding toxin/antitoxin systems such as NUDIX family hydrolase, HicB family protein, AbrB/VapB family protein, and copies of antitoxin VapB and VapC.

In general, gene expression levels in each island is lower than the genome average. The only exception was represented by HQ_GI1, mainly because the overall highest expressed gene HQ_RS01075 is located on this island (Supplementary Dataset-1, Sheet-5). The average differential gene expression per genomic island between winter and summer was close to the genome average in HQ_G1 (1.25), HQ_G1b (1.27) and HQ_G2 (1.28), and relatively higher in HQ_GI2b (1.47), HQ_GI3 (1.50) and HQ_GI4 (1.58).

Hqr. walsbyi HBSQ001 ecoparalogs

Hqr. walsbyi HBSQ001 contains several potential ecoparalogs, i.e., genes encoding similar proteins (paralogs) that may be differentially expressed under different environmental conditions (Sanchez-Perez et al., 2008). Paralogous genes that differed significantly (>1.5-fold) in expression amongst each other within the same sample qualified as ecoparalogs. We identified 150 genes, grouped in 63 clusters that can be considered as candidate "ecoparalogs" (Supplementary Dataset-1, Sheet-6). In total, 31 clusters with differentially expressed genes were found in summer and 27 in winter. Highest differences in expression between ecoparalogs were found for the uncharacterized proteins HQ_RS05860 and HQ_RS13150 (8.8-fold) in summer and for ISH8 transposases (6.1-fold) in the winter dataset. To see whether seasonality may be a driver of ecoparalog differential expression, we compared their expression between the two samples. Three seasonal-ecoparalogs could be found, two ferredoxin coding genes, *fer2* and *fer7*, and a purine phosphoribosyltransferase coding gene, *apt2*, which had 1.5-fold, 1.7-fold and 1.7-fold higher expression in summer than in winter, respectively. We also calculated the difference between the highest expressed gene in a cluster from one season divided by the lowest expressed gene in that cluster from the other season. In 19 clusters the significant higher expressed gene was found in summer and in 33 clusters in winter. The highest differential expression in summer was found for the above mentioned ISH8 transposase coding genes (6.3-fold difference) and in winter for the uncharacterized proteins HQ_RS05860 and HQ_RS13150 (10-fold) coding genes.

Discussion

The presented metatranscriptome analysis of the crystallizer ponds in summer and winter confirmed the dominance of halophilic Archaea at the highest salinities, and especially of *Hqr. walsbyi*. In summer, Bacteria and Eukarya do not exceed 10% of the total read abundance while in winter, their relative contribution increases to over 20%. At this moment, it is unclear whether this is a consequence of on average lower temperatures, less sun hours and lower sun intensity or whether the nutritional status of the brines dramatically change in winter.

The highest expressed *Hqr. walsbyi* HBSQ001 gene in both summer and winter sample is HQ_RS01075, which encodes a 72 amino acid small protein with a putative DNA binding motive (ribbon-helix–helix) found in the CopG-like protein family IPR002145, suggesting a role in regulating DNA transcription. Interestingly, this gene is absent from the Australian type strain, *Hqr. walsbyi* C23^T (Burns et al., 2007). Moreover, 100% identical homologs of this protein have so far only be found in the salterns from which strain HBSQ001 was isolated and may hint to a local evolved or acquired gene. Among the highly expressed genes in winter, several encode stress response proteins like cold-shock proteins, starvation-induced proteins, and superoxide dismutases. However, only a few of these differed significantly in expression compared to summer, suggesting that winter conditions alone, did not induce this stress response. Apparently, these genes are in a continuous high expression state and may reflect the year-round challenging conditions in the hypersaline ecosystem, such as high salinity, light intensities and UV irradiation, high temperatures and high oxygen radical concentrations (Salin and Brown-Peterson, 1993). Although a connection between these proteins and viral infection is not clear, viruses could be an important factor to explain specific rises of these stress-related proteins in *Hqr. walsbyi*. Two possible interpretations are, (i) that the total *Hqr. walsbyi* population is more susceptible to viruses in a more stressed winter environment, or, alternatively, (ii) that a specific dominant *Hqr. walsbyi* winter-ecotype is more susceptible to these viruses.

The winter high *bopI* expression may reflect a lower energy and nutritional status due to reduced photosynthetic primary production driven by algae like *Dunaliella*. Increased expression of *bopI* may also compensate for less and lower intensity sun hours. High expression of *sufB* and *sufC* in winter may be essential to meet an increased demand for electron translocating ferredoxins or for Fe-S proteins involved in distribution of electrons to various metabolic pathways. Higher expression of the A-type ATPase subunit F coding gene *atpF* and the NADH dehydrogenase-like complex subunits A and B, encoded by *nuoA* and *nuoB*, also agree with an increased need for energy requiring processes. Potential carbon limitation in winter may be deduced from high *glpK* gene expression. A higher glycerol demand in winter is further substantiated by the high expression of the hypothetical protein encoding-gene HQ_RS03760, directly upstream of *glpK*. This gene has a conserved PRK10712 domain, found also in the PTS system fructose-specific transporter subunits IIBC. Potentially, this protein might be involved in the uptake of glycerol rather than fructose and may form a phosphotransferase system with GlpK. The well-established PTS system involved in dihydroxyacetone uptake (Bolhuis et al., 2006; Elevi Bardavid and Oren, 2008; Ouellette et al., 2013) is above average expressed, while the uptake kinase of this system, encoded by *dhaL* is significantly higher expressed in winter.

Evidence for multiple occurring ecotypes of *Hqr. walsbyi*

Potential hotspots for differentially expressed genes relative to the cultured isolate are two of the so-called genomic islands. In contrast to the well conserved core genome, these genomic islands represent regions in the genome that are not conserved amongst all members of the population. Therefore, they might be indicative for the presence of locally adapted ecotypes (Rodriguez-Valera et al., 2016). These islands can be identified by metagenomic recruitment analyzes against the genome of a cultivated strain. The core genome will recruit high numbers of near identical reads whereas the genomic islands will only have a few reads and often with low identity. For *Hqr. Walsbyi* HBSQ001, six of these putative genomic islands were initially identified (Cuadros-Orellana et al., 2007). Later, genome comparison between the two sequenced isolates of *Hqr. walsbyi*, strain HBSQ001 and strain C23^T, suggested the presence of at least 12 of these divergent regions (Cuadros-Orellana et al., 2007; Dyall-Smith et al., 2011). It is believed that the content of these genomic islands can vary per ecotype and may provide specific, selective advantages for their particular ecological niche (López-Pérez et al., 2014; Rodriguez-Valera et al., 2016). The genomic islands in *Hqr. walsbyi* predominantly contained genes involved in biosynthesis of surface layer proteins, genes encoding cell surface glycoproteins and genes involved in the cell envelope formation (Martin-Cuadrado et al., 2015). The overall low gene expression levels within these islands agree with their genes being only present in a smaller subset of the *Hqr. walsbyi* population and hence the existence of different ecotypes that may be more abundant than the isolated strain. Presence of *Hqr. walsbyi* C23^T in the Spanish salterns was also investigated. A distinction between C23^T and HBSQ001 is difficult to make based on the core genome that share over 98% nucleotide identity (Dyall-Smith et al., 2011). However, absence of transcripts mapping to the genomic islands specific to C23^T is in agreement with the absence of this Australian isolate from the Spanish salterns.

In the cultured strain HBSQ001, these islands were also low expressed, but again, with exception of HQ_GI1, where the averaged expression level was 18-fold higher than the genome average. This is mainly caused by the presence of some of the overall highest expressed genes; the potential transcriptional regulator HQ_RS01075 and three cell surface glycoprotein coding genes, HQ_RS01080, HQ_RS01085 and HQ_RS01090. This

suggests a clear difference between growth of *Hqr. walsbyi* in nature or culture. In culture, the cell surface glycoprotein coding genes are over expressed while in nature, the focus appears to be more on energy and carbon acquisition. This may be a consequence of the different geochemical composition of the culture medium compared to the natural brine which may have higher diversity in carbon sources at lower concentrations. Noticeable is also that majority of genes of the Islands HQ_GI2 and HQ_GI4 are higher expressed in nature than in culture and in HQ_GI2b, where 10 out of 12 genes are higher expressed in culture. Moreover, in HQ_GI1, HQ_GI1b, HQ_GI2b and HQ_GI4, more than 70% of the genes that are higher expressed in culture are also highest expressed in winter. Potentially, the isolated strain HBSQ001 or closely related ecotypes are more abundant in the winter season.

Difference in expression patterns between natural and culture grown *Hqr. walsbyi* HBSQ001

One of the discussions in experimental microbial ecology concerns the validity of extrapolating gene expression patterns obtained in lab grown cultures to the natural environment. We therefore compared the metatranscriptome dataset with the transcriptome from *Hqr. walsbyi* HBSQ001 grown in a defined medium (Bolhuis et al., 2017). The most remarkable and noteworthy difference in gene expression pattern is the difference in TPM distribution per gene. The transcript distribution in the natural environment followed a bell-shape curve with a median close to the average TPM suggesting that most genes in nature are expressed close to the average. In culture however the median was much lower than the average and revealed that most genes were low expressed with the exception of a small number of very high expressed genes. This most likely reflects the adaptation of the cultured strain to the relative homogeneous, simple growth conditions and media composition. In nature, *Hqr. walsbyi* is represented by multiple ecotypes, adapted to heterogeneous environmental conditions in terms of temperature, light, UV irradiation and salinity. In addition, they are in continuous competition for the available nutrients and under potential threat of viral attack either with other halophilic species or even amongst their different ecotypes. Under laboratory conditions, several naturally occurring stressors may be absent. A good example is the expression of a gene encoding an L-lactate permease involved in the uptake of lactase that is well expressed in nature but not in culture. Lactate was not present in the cultivation medium but is a natural occurring carbon source in salterns and requires active transported over the membrane for uptake (Oren and Gurevich, 1994). One intriguing observation is the higher expression of the *coxMSL* gene cluster in nature. These genes are annotated as subunits of the aerobic-type carbon monoxide dehydrogenase; however, there is currently no evidence that carbon monoxide can be used as carbon source by *Hqr. walsbyi*. Potentially the true substrate of this molybdopterin binding dehydrogenase may be something else (King and Weber, 2007).

Genes that were overexpressed in culture but low expressed in nature include surface proteins and surface modification proteins that may be expressed in response to the limited movement in the standing batch cultures and may favor formation of biofilms and the much larger cell structures that were observed in the cultures compared to naturally occurring cells (Bolhuis et al., 2004).

Conclusion

Our results clearly show that gene expression patterns in laboratory grown cultures do not necessarily reflect their metabolism in nature and therefore, extrapolating data from laboratory experiments to natural behavior should be done with care. Our results furthermore shows that there are multiple ecotypes of *Hqr. walsbyi* that differ from the HBSQ001 isolate. Low expression of this isolated strain specific genomic islands furthermore suggests that HBSQ001 isolate is not abundantly present in the pond and season from which it was originally isolated. In winter, the system appears more energy and carbon deprived and HBSQ001 is more abundant and hence may also be more virus tolerant, a feature which may have also led to its successful isolation. The differences in gene expression between natural occurring species and cultivated equivalents most likely reflect a higher number of available niches, presence of different nutrients and potential interactions with different species in the natural brine compare to the homogenous glycerol and pyruvate fed monocultures.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: https://www.ncbi.nlm.nih.gov/, PRJNA633445.

Author contributions

ML-P, AM-C, RR, and FR-V: conceived and designed the experiments. ML-P and RR: performed the experiments. RR, ML-P, AM-C, and HB: analyzed the data. RR, ML-P, AM-C, FR-V, and HB: wrote the paper. All authors contributed to the article and approved the submitted version.

Funding

RR received funding from the European Union's Horizon 2020 research and innovation programme under grant

agreement no. 818431 (SIMBA). This output reflects only the author's view, and the European Union cannot be held responsible for any use that may be made of the information contained therein. FRV was supported by grants "VIREVO" CGL2016-76273-P [MCI/AEI/FEDER, EU] (cofounded with FEDER funds) from the Spanish Ministerio de Ciencia e Innovación and "HIDRAS3" PROMETEU/2019/009 from Generalitat Valenciana.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2022.1044446/full#supplementary-material

References

Aharonovich, D., and Sher, D. (2016). Transcriptional response of *Prochlorococcus* to co-culture with a marine Alteromonas: differences between strains and the involvement of putative infochemicals. *ISME J.* 10, 2892–2906. doi: 10.1038/ismej.2016.70

Antón, J., Llobet-Brossa, E., Rodriguez-Valera, F., and Amann, R. (1999). Fluorescence *in situ* hybridization analysis of the prokaryotic community inhabiting crystallizer ponds. *Environ. Microbiol.* 1, 517–523. doi: 10.1046/j.1462-2920.1999.00065.x

Atanasova, N. S., Pietilä, M. K., and Oksanen, H. M. (2013). Diverse antimicrobial interactions of halophilic archaea and bacteria extend over geographical distances and cross the domain barrier. *MicrobiologyOpen* 2, 811–825. doi: 10.1002/mbo3.115

Beliaev, A. S., Romine, M. F., Serres, M., Bernstein, H. C., Linggi, B. E., Markillie, L. M., et al. (2014). Inference of interactions in cyanobacterial–heterotrophic co-cultures via transcriptome sequencing. *ISME J.* 8, 2243–2255. doi: 10.1038/ismej.2014.69

Benlloch, S., Acinas, S. G., Anton, J., Lopez-Lopez, A., Luz, S. P., and Rodriguez-Valera, F. (2001). Archaeal biodiversity in crystallizer ponds from a solar saltern: culture versus pcr. *Microb. Ecol.* 41, 12–19. doi: 10.1007/s002480000069

Benlloch, S., Acinas, S. G., Martínez-Murcia, A. J., and Rodríguez-Valera, F. (1996). Description of prokaryotic biodiversity along the salinity gradient of a multipond solar saltern by direct PCR amplification of 16S rDNA. *Hydrobiologia* 329, 19–31. doi: 10.1007/BF00034544

Benlloch, S., Lopez-Lopez, A., Casamayor, E. O., Øvreås, L., Goddard, V., and Daae, F. L. (2002). Prokaryotic genetic diversity throughout the salinity gradient of a coastal solar saltern. *Environ. Microbiol.* 4, 349–360. doi: 10.1046/j.1462-2920.2002.00306.x

Benlloch, S., Martínez-Murcia, A. J., and Rodríguez-Valera, F. (1995). Sequencing of bacterial and archaeal 16S rRNA genes directly amplified from a hypersaline environment. *Syst. Appl. Microbiol.* 18, 574–581. doi: 10.1016/S0723-2020(11)80418-2

Biller, S. J., Coe, A., and Chisholm, S. W. (2016). Torn apart and reunited: impact of a heterotroph on the transcriptome of Prochlorococcus. *ISME J.* 10, 2831–2843. doi: 10.1038/ismej.2016.82

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170

Bolhuis, H., Martin-Cuadrado, A. B., Rosselli, R., Pasic, L., and Rodriguez-Valera, F. (2017). Transcriptome analysis of *Haloquadratum walsbyi*: vanity is but the surface. *BMC Genomics* 18:510. doi: 10.1186/s12864-017-3892-2

Bolhuis, H., Palm, P., Wende, A., Falb, M., Rampp, M., and Rodriguez-Valera, F. (2006). The genome of the square archaeon Haloquadratum walsbyi: life at the limits of water activity. *BMC Genomics* 7:169. doi: 10.1186/1471-2164-7-169

Bolhuis, H., Poele, E., and Rodriguez-Valera, F. (2004). Isolation and cultivation of Walsby's square archaeon. *Environ. Microbiol.* 6, 1287–1291. doi: 10.1111/j.1462-2920.2004.00692.x

Burns, D. G., Janssen, P. H., Itoh, T., Kamekura, M., Li, Z., Jensen, G., et al. (2007). *Haloquadratum walsbyi* gen. Nov., sp. nov., the square haloarchaeon of Walsby, isolated from saltern crystallizers in Australia and Spain. *Int. J. Syst. Evol. Microbiol.* 57, 387–392. doi: 10.1099/ijs.0.64690-0

Cuadros-Orellana, S., Martin-Cuadrado, A.-B., Legault, B., D'auria, G., Zhaxybayeva, O., Papke, R. T., et al. (2007). Genomic plasticity in prokaryotes: the case of the square haloarchaeon. *ISME J.* 1, 235–245. doi: 10.1038/ismej.2007.35

Dyall-Smith, M. L., Pfeiffer, F., Klee, K., Palm, P., Gross, K., Schuster, S. C., et al. (2011). *Haloquadratum walsbyi*: limited diversity in a global pond. *PLoS One* 6:e20968. doi: 10.1371/journal.pone.0020968

Elevi Bardavid, R., and Oren, A. (2008). Dihydroxyacetone metabolism in *Salinibacter ruber* and in *Haloquadratum walsbyi*. *Extremophiles* 12, 125–131. doi: 10.1007/s00792-007-0114-x

Fu, Y., Deiorio-Haggar, K., Anthony, J., and Meyer, M. M. (2013). Most RNAs regulating ribosomal protein biosynthesis in *Escherichia coli* are narrowly distributed to *Gammaproteobacteria*. *Nucleic Acids Res.* 41, 3491–3503. doi: 10.1093/nar/gkt055

Garcia-Heredia, I., Martin-Cuadrado, A. B., Mojica, F. J., Santos, F., Mira, A., Anton, J., et al. (2012). Reconstructing viral genomes from the environment using fosmid clones: the case of haloviruses. *PLoS One* 7:e33802. doi: 10.1371/journal.pone.0033802

Ghai, R., Pašić, L., Fernández, A. B., Martin-Cuadrado, A.-B., Mizuno, C. M., Mcmahon, K. D., et al. (2011). New abundant microbial groups in aquatic hypersaline environments. *Sci. Rep.* 1:135. doi: 10.1038/srep00135

Ghanmi, F., Carre-Mlouka, A., Vandervennet, M., Boujelben, I., Frikha, D., Ayadi, H., et al. (2016). Antagonistic interactions and production of halocin antimicrobial peptides among extremely halophilic prokaryotes isolated from the solar saltern of Sfax, Tunisia. *Extremophiles* 20, 363–374. doi: 10.1007/s00792-016-0827-9

Gonzalez-Torres, P., Pryszcz, L. P., Santos, F., Martinez-Garcia, M., Gabaldon, T., and Antón, J. (2015). Interactions between closely related bacterial strains are revealed by deep transcriptome sequencing. *Appl. Environ. Microbiol.* 81, 8445–8456. doi: 10.1128/AEM.02690-15

Hjerde, E., Karlsen, C., Sørum, H., Parkhill, J., Willassen, N. P., and Thomson, N. R. (2015). Co-cultivation and transcriptome sequencing of two co-existing fish pathogens *Moritella viscosa* and *Aliivibrio wodanis*. *BMC Genomics* 16:447. doi: 10.1186/s12864-015-1669-z

Kanehisa, M., Sato, Y., and Morishima, K. (2016). BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. *J. Mol. Biol.* 428, 726–731. doi: 10.1016/j.jmb.2015.11.006

Kimes, N. E., López-Pérez, M., Ausó, E., Ghai, R., and Rodriguez-Valera, F. (2014). RNA sequencing provides evidence for functional variability between naturally co-existing *Alteromonas macleodii* lineages. *BMC Genomics* 15:938. doi: 10.1186/1471-2164-15-938

King, G. M., and Weber, C. F. (2007). Distribution, diversity and ecology of aerobic CO-oxidizing bacteria. *Nat. Rev. Microbiol.* 5, 107–118. doi: 10.1038/nrmicro1595

Legault, B. A., Lopez-Lopez, A., Alba-Casado, J. C., Doolittle, W. F., Bolhuis, H., and Rodriguez-Valera, F. (2006). Environmental genomics of "*Haloquadratum walsbyi*" in a saltern crystallizer indicates a large pool of accessory genes in an otherwise coherent species. *BMC Genomics* 7:171. doi: 10.1186/1471-2164-7-171

Li, G.-W., Burkhardt, D., Gross, C., and Weissman, J. S. (2014). Quantifying absolute protein synthesis rates reveals principles underlying allocation of cellular resources. *Cells* 157, 624–635. doi: 10.1016/j.cell.2014.02.033

Li, W., and Godzik, A. (2006). Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics* 22, 1658–1659. doi: 10.1093/bioinformatics/btl158

López-Pérez, M., Martin-Cuadrado, A.-B., and Rodriguez-Valera, F. (2014). Homologous recombination is involved in the diversity of replacement flexible genomic islands in aquatic prokaryotes. *Front. Genet.* 5:147. doi: 10.3389/fgene.2014.00147

Magoc, T., Wood, D., and Salzberg', S. L. (2013). EDGE-pro: estimated degree of gene expression in prokaryotic genomes. *Evol. Bioinforma.* 9, EBO.S11250–EBO.S11136. doi: 10.4137/EBO.S11250

Makkay, A. M., Louyakis, A. S., Ram-Mohan, N., Gophna, U., Gogarten, J. P., and Papke, R. T. (2020). Insights into gene expression changes under conditions that facilitate horizontal gene transfer (mating) of a model archaeon. *Sci. Rep.* 10:22297. doi: 10.1038/s41598-020-79296-w

Martin-Cuadrado, A.-B., Pašić, L., and Rodriguez-Valera, F. (2015). Diversity of the cell-wall associated genomic island of the archaeon *Haloquadratum walsbyi*. *BMC Genomics* 16:603. doi: 10.1186/s12864-015-1794-8

Martínez-Murcia, A. J., Acinas, S. G., and Rodriguez-Valera, F. (1995). Evaluation of prokaryotic diversity by restrictase digestion of 16S rDNA directly amplified from hypersaline environments. *FEMS Microbiol. Ecol.* 17, 247–255. doi: 10.1016/0168-6496(95)00029-A

Miller, C. S., Baker, B. J., Thomas, B. C., Singer, S. W., and Banfield, J. F. (2011). EMIRGE: reconstruction of full-length ribosomal genes from microbial community short read sequencing data. *Genome Biol.* 12:R44. doi: 10.1186/gb-2011-12-5-r44

Naghoni, A., Emtiazi, G., Amoozegar, M. A., Cretoiu, M. S., Stal, L. J., Etemadifar, Z., et al. (2017). Microbial diversity in the hypersaline Lake Meyghan, Iran. *Sci. Rep.* 7:7. doi: 10.1038/s41598-017-11585-3

Narasingarao, P., Podell, S., Ugalde, J. A., Brochier-Armanet, C., Emerson, J. B., Brocks, J. J., et al. (2011). *De novo* metagenomic assembly reveals abundant novel major lineage of Archaea in hypersaline microbial communities. *ISME J.* 6:81.

Oren, A. (1994). The ecology of the extremely halophilic archaea. *FEMS Microbiol. Rev.* 13, 415–439. doi: 10.1111/j.1574-6976.1994.tb00060.x

Oren, A., and Gurevich, P. (1994). Production of d-lactate, acetate, and pyruvate from glycerol in communities of halophilic archaea in the Dead Sea and in saltern crystallizer ponds. *FEMS Microbiol. Ecol.* 14, 147–155.

Ouellette, M., Makkay, A. M., and Papke, R. T. (2013). Dihydroxyacetone metabolism in *Haloferax volcanii*. *Front. Microbiol.* 4:376. doi: 10.3389/fmicb.2013.00376

Øvreås, L., Bourne, D., Sandaa, R.-A., Casamayor, E. O., Benlloch, S., Goddard, V., et al. (2003). Response of bacterial and viral communities to nutrient manipulations in seawater mesocosms. *Aquat. Microb. Ecol.* 31, 109–121. doi: 10.3354/ame031109

Papke, R. T., Koenig, J. E., Rodríguez-Valera, F., and Doolittle, W. F. (2004). Frequent recombination in a saltern population of *Halorubrum*. *Science* 306, 1928–1929. doi: 10.1126/science.1103289

Peterson, J. R., Thor, S., Kohler, L., Kohler, P. R. A., Metcalf, W. W., and Luthey-Schulten, Z. (2016). Genome-wide gene expression and RNA half-life measurements allow predictions of regulation and metabolic behavior in *Methanosarcina acetivorans*. *BMC Genomics* 17:924. doi: 10.1186/s12864-016-3219-8

Plominsky, A. M., Delherbe, N., Ugalde, J. A., Allen, E. E., Blanchet, M., Ikeda, P., et al. (2014). Metagenome sequencing of the microbial community of a solar saltern crystallizer pond at Cáhuil lagoon, Chile. *Genome Announc.* 2, e01172–e01114. doi: 10.1128/genomeA.01172-14

Pruesse, E., Quast, C., Knittel, K., Fuchs, B. M., Ludwig, W., Peplies, J., et al. (2007). SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. *Nucleic Acids Res.* 35, 7188–7196. doi: 10.1093/nar/gkm864

Redon, E., Loubière, P., and Cocaign-Bousquet, M. (2005). Role of mRNA stability during genome-wide adaptation of *Lactococcus lactis* to carbon starvation. *J. Biol. Chem.* 280, 36380–36385. doi: 10.1074/jbc.M506006200

Rodriguez-Valera, F., Martin-Cuadrado, A.-B., and López-Pérez, M. (2016). Flexible genomic islands as drivers of genome evolution. *Curr. Opin. Microbiol.* 31, 154–160. doi: 10.1016/j.mib.2016.03.014

Salin, M. L., and Brown-Peterson, N. J. (1993). Dealing with active oxygen intermediates: A halophilic perspective. *Experientia* 49, 523–529. doi: 10.1007/BF01955155

Sanchez-Perez, G., Mira, A., Nyirő, G., Pašić, L., and Rodriguez-Valera, F. (2008). Adapting to environmental changes using specialized paralogs. *Trends Genet.* 24, 154–158. doi: 10.1016/j.tig.2008.01.002

Ventosa, A.. (2006). *Unusual Micro-Organisms from Unusual Habitats: Hypersaline Environments*. Cambridge: Cambridge University Press.

Ventosa, A., Fernández, A. B., León, M. J., Sánchez-Porro, C., and Rodriguez-Valera, F. (2014). The Santa Pola saltern as a model for studying the microbiota of hypersaline environments. *Extremophiles* 18, 811–824. doi: 10.1007/s00792-014-0681-6

Wagner, G. P., Kin, K., and Lynch, V. J. (2012). Measurement of mRNA abundance using RNA-seq data: RPKM measure is inconsistent among samples. *Theory Biosci.* 131, 281–285. doi: 10.1007/s12064-012-0162-3

frontiers | Frontiers in Microbiology

TYPE Original Research
PUBLISHED 18 September 2023
DOI 10.3389/fmicb.2023.1274068

OPEN ACCESS

EDITED BY
Qunxin She,
Shandong University (Qingdao), China

REVIEWED BY
Hanna M. Oksanen,
University of Helsinki, Finland
Zhirui Zeng,
Southern University of Science and Technology, China

*CORRESPONDENCE
Susanne Erdmann
✉ serdmann@mpi-bremen.de

RECEIVED 07 August 2023
ACCEPTED 04 September 2023
PUBLISHED 18 September 2023

CITATION
Mercier C, Thies D, Zhong L, Raftery MJ and Erdmann S (2023) Characterization of an archaeal virus-host system reveals massive genomic rearrangements in a laboratory strain. *Front. Microbiol.* 14:1274068. doi: 10.3389/fmicb.2023.1274068

Characterization of an archaeal virus-host system reveals massive genomic rearrangements in a laboratory strain

Coraline Mercier[1], Daniela Thies[1], Ling Zhong[2], Mark J. Raftery[2] and Susanne Erdmann[1,3]*

[1]Max Planck Institute for Marine Microbiology, Archaeal Virology, Bremen, Germany, [2]Bioanalytical Mass Spectrometry Facility, The University of New South Wales, Sydney, NSW, Australia, [3]School of Biotechnology and Biomolecular Sciences, The University of New South Wales, Sydney, NSW, Australia

Halophilic archaea (haloarchaea) are known to exhibit multiple chromosomes, with one main chromosome and one or several smaller secondary chromosomes or megaplasmids. *Halorubrum lacusprofundi*, a model organism for studying cold adaptation, exhibits one secondary chromosome and one megaplasmid that include a large arsenal of virus defense mechanisms. We isolated a virus (Halorubrum tailed virus DL1, HRTV-DL1) infecting *Hrr. lacusprofundi*, and present an in-depth characterization of the virus and its interactions with *Hrr. lacusprofundi*. While studying virus-host interactions between *Hrr. lacusprofundi* and HRTV-DL1, we uncover that the strain in use (ACAM34_UNSW) lost the entire megaplasmid and about 38% of the secondary chromosome. The loss included the majority of virus defense mechanisms, making the strain sensitive to HRTV-DL1 infection, while the type strain (ACAM34_DSMZ) appears to prevent virus replication. Comparing infection of the type strain ACAM34_DSMZ with infection of the laboratory derived strain ACAM34_UNSW allowed us to identify host responses to virus infection that were only activated in ACAM34_UNSW upon the loss of virus defense mechanisms. We identify one of two S-layer proteins as primary receptor for HRTV-DL1 and conclude that the presence of two different S-layer proteins in one strain provides a strong advantage in the arms race with viruses. Additionally, we identify archaeal homologs to eukaryotic proteins potentially being involved in the defense against virus infection.

KEYWORDS

***Halorubrum lacusprofundi*, secondary chromosome, haloarchaea, archaeal virus, virus defense, genome plasticity**

Introduction

Secondary chromosomes and megaplasmids are additional chromosomes that coexist alongside the primary circular chromosome, and are found in bacteria and archaea (Harrison et al., 2010; Hall et al., 2022). Amongst archaea, halophilc archaea belonging to the Euryarchaeota (haloarchaea) are particularly rich in additional chromosomes (DasSarma et al., 2009). While the exact number of additional chromosomes can vary between different species, it is not uncommon for haloarchaea to possess two or more of these extra genetic elements (DasSarma et al., 2009; Wang et al., 2015). Secondary chromosomes, are distinguished from megaplasmid by carrying essential housekeeping genes (Hall et al., 2022). While some of the additional

chromosomes in haloarchaea strongly interact with the main chromosomes and exhibit some essential genes (Harrison et al., 2010; Wang et al., 2015), others have been lost and shown to be non-essential (Hawkins et al., 2013). *Hrr. lacusprofundi* exhibits three replicons, the main chromosome CHR1, a secondary chromosome CHR2 and the megaplasmid pHLAC01. A metagenomics study revealed that the main chromosome of *Hrr. lacusprofundi* is highly conserved between strains, but the secondary chromosomes and megaplasmids are very diverse and responsible for the genetic diversity of the species (Tschitschko et al., 2018).

Haloarchaea are infected by diverse viruses including head-tailed viruses, pleomorphic viruses, spherical viruses, and spindle shaped viruses (Luk et al., 2014). A number of haloarchaeal viruses exhibit lysogenic and chronic life cycles, however, the majority of isolated viruses exhibit lytic life cycles, reflecting a methodical bias caused by using plaque assays for isolation, selecting mostly for lytic viruses (Alarcón-Schumacher et al., 2022). A recent study described a number of new haloarchaeal tailed viruses and proposed a new classification of archaeal tailed viruses (arTV) (Liu Y. et al., 2021). The majority of haloarchaeal tailed viruses, exhibiting major capsid proteins (MCPs) with the HK97 fold, belong to the "myovirus" or "siphovirus" morphotype, only one isolate exhibits "podovirus" morphology (Pietilä et al., 2013b). While all package a linear genome with size ranging from 26 kb to 143 kb, a few exhibit circular permuted ends and the majority exhibit direct terminal repeats ranging in size from 229 bp up to 739 bp (Liu Y. et al., 2021). The host range of haloarchaeal tailed viruses appears to be very variable, with some exhibiting a very broad host range and some being very host specific (Atanasova et al., 2012). While it is still unclear how arTVs exit their host cells, the S-layer has recently been identified as possible receptor for a haloarchaeal tailed virus (Schwarzer et al., 2023).

Even though arTV appear morphological similar to head-tailed bacteriophages, they encounter cells that are dramatically different, in particular with respect to the cell surface, and replication, transcription and translation mechanisms. While the differences between bacterial and archaeal head-tailed viruses is clearly reflected by genomic differences, the effect of these differences on virus-host interactions remains severely understudied. To date only very few arTV infecting haloarchaea have been characterized in more depth (Luk et al., 2014; Liu et al., 2021).

Archaea use primarily the CRISPR-Cas (Clustered Regularly Interspaced Short Palindromic Repeats and CRISPR-associated genes) system, Toxin-Antitoxin (TA) systems, and Restriction Modification (RM) systems, additional to changes of cell surface proteins, to defend against viruses (Koonin et al., 2017). The majority of haloarchaea exhibit CRISPR-Cas systems, RM systems and TA systems, and CRISPR systems have been experimentally shown to be active against extrachromosomal elements in haloarchaea (Li et al., 2014; Maier et al., 2019). Argonaut proteins, known to be important for virus defense in eukaryotes (Obbard et al., 2009), have also been identified on archaeal genomes (Willkomm et al., 2018). However, while DNA-targeting activity has been shown experimentally for an archaeal Argonaut protein (Zander et al., 2017), evidence for exclusion of extrachromosomal elements by archaeal Argonauts is pending. Recent studies, using bioinformatics approaches, identified an enormous diversity of previously unrecognized bacterial virus exclusion mechanisms (Bernheim and Sorek, 2020; Millman et al., 2022), of which some were also detected in archaea. For example, CBASS (Cyclic oligonucleotide-based antiphage signaling system) systems were detected in a number of archaeal genomes and were shown to block virus propagation in bacteria by inducing cell death of the infected cell (Duncan-Lowey and Kranzusch, 2022). The BREX (Bacteriophage Exclusion) system, also discovered on archaeal genomes, was shown to block virus replication (Goldfarb et al., 2015). While a Dnd defense system has recently been experimentally investigated in archaea (Xiong et al., 2019), the majority of the newly discovered defense systems that are not as widespread amongst archaea as CRISPR-Cas, TA or RM systems, remain uncharacterised.

Here we present an in-depth characterization of the interactions of an archaeal tailed virus and its host, including the characterization of the virus life cycle, determination of the virus host range, analysis of the host transcriptional response to virus infection and the characterization of virus escape mutants. We uncover that the host strain in use lost an entire megaplasmid and a part of its secondary chromosome while being grown under laboratory conditions. This loss included the majority of virus defense systems, including a CRISPR-Cas system, TA and RM systems and a BREX system, which allows us to discover new putative virus defense mechanisms and the host receptor for virus attachment.

Results and discussion

Isolation of a lytic virus that infects *Halorubrum lacusprofundi* ACAM34 (DSM 5036)

A virus with head-tailed morphology (Figure 1A) was obtained from the culture supernatant of a *Halorubrum lacusprofundi* strain, that was isolated from a sample taken in 2014 from Deep Lake, Antarctica (Gibson, 1999). Analytic restriction digest of DNA isolated from the viral lysate revealed a dsDNA genome of approximately 60 kb (Figure 1B). The lysate produced plaques on *Hrr. lacusprofundi* strain ACAM34 (Franzmann et al., 1988) that was isolated from a Deep Lake sample in 1988, while it remains elusive when the sample was actually obtained from Deep Lake. Lysis was also observed in liquid cultures of this strain. The viral lysate was formerly describe as DLHTHV (Tschitschko et al., 2018), however, we rename it hereby as HRTV-DL (Halorubrum Tailed Virus-Deep Lake), according to the nomenclature that was previously used for archaeal tailed viruses (arTV) (Liu Y. et al., 2021). HRTV-DL is the first virus infecting *Hrr. lacusprofundi* that has been isolated from its original environment, Deep Lake. *Hrr. lacusprofundi* ACAM34 was used to purify the virus through several rounds of plaque assays. When analyzing genomic DNA of virus particles propagated from single plaques, we observed a remarkable high genomic variability (Supplementary Figure S1) and a reduction of the genome size. Such genomic variation, upon isolation, has previously been observed for *Halobacterium salinarium* virus Phi H (Schnabel et al., 1982) and suggests the presence of several variants in the initial HRTV-DL preparation. After two rounds of plaque purification, we chose one variant (P2V1, Supplementary Figure S1C), HRTV-DL1, for further characterization. Virus particles analyzed by electron microscopy exhibit a head-tailed morphology, with a head diameter of about 50 nm and a non-contractile tail of approximately 80 nm-100 nm in length (Figure 1B and Supplementary Figure S2).

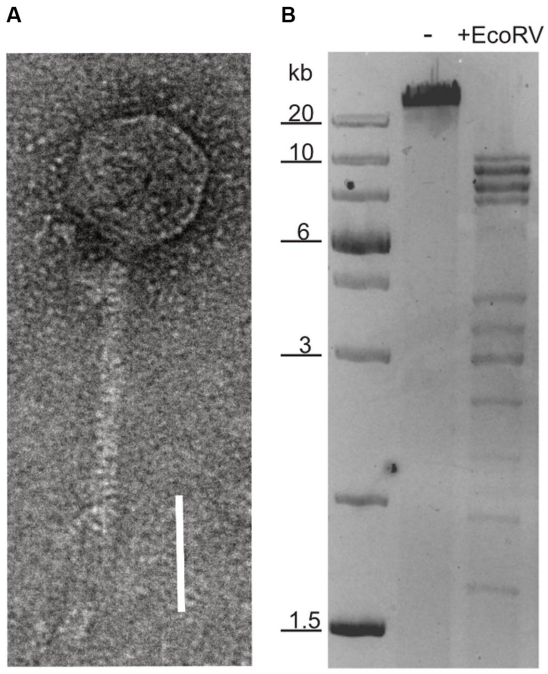


FIGURE 1
Nucleic acid content and particle of HRTV-DL. **(A)** Electron micrograph of HRTV-DL particle. Sample was negatively stained with 2% uranyl acetate. Size bar: 50nm. **(B)** Total DNA of HRTV-DL undigested (-) and digested with EcoRV (+EcoRV). MW size marker is shown to the left of the gel (GeneRuler 1 kb Plus DNA Ladder, Thermo Fisher Scientific). DNA was separated on 1% agarose gels and stained with SYBR™ Safe DNA stain. Size bar: 50 nm. Original images have been modified by cropping to improve visual presentation.

Genomic features of HRTV-DL1 and proteins associated with the virus particles

DNA isolated directly from HTRV-DL1 particles appeared sensitive to Exonuclease treatment, indicating a linear genome (Supplementary Figure S3). However, the majority of viral DNA isolated out of infected cells was shown to be insensitive to Exonuclease. Therefore, we suggest that the genome is replicated as a circular genome within the host and packaged as linear genome into virus particles. When comparing the EcoRV digested genome of the intracellular virus after Exonuclease treatment with the EcoRV digested genome that is packaged in virus particles, we observed that a band of app. 2,500 bp was missing (Supplementary Figure S3). We conclude that this fragment only appears when digesting the linear genome and could represent one end of the linear virus genome. Analytic digest of virus and host DNA with *Dpn*I, showed that both the host and the virus genome are dam methylated (Supplementary Figure S3).

Sequencing and assembly of HRTV-DL1 genome revealed a single circular contig of 37.7 kb, with no indication in the sequence data for discrete ends. The genome of HRTV-DL1 contains 49 putative open reading frames (ORFs) (Figure 2 and Supplementary Table S1). Using Hidden Markov models (HMM) (Söding, 2004) and domain prediction (Blum et al., 2020) we can predict the function for the products of about 40% of these ORFs (Supplementary Table S1). Two genes were identified that we propose to be involved in virus genome replication, both are also encoded by arTV CGphi46 (Figure 2). ORF48 product is predicted as primase-helicase, probably generating a short primer that is then elongated by the host DNA polymerase. ORF41 product is predicted as DNA polymerase sliding clamp, binding the DNA polymerase and preventing its dissociation from the template DNA. Two genes could be involved in transcription. ORF43 product, a predicted transcription initiation factor IIB, might be responsible for recruiting the host RNA Polymerase and ORF32, with weak hits to antitermination protein NusA, could associate with the host RNA polymerase elongation complex. ORF29 product is a predicted adenine-specific methyltransferase and 100% identical to a host protein, probably recruited by the virus to protect its genome from host RM-systems. ORF7 and ORF8 products were identified as small and large subunit of the terminase, responsible for genome packaging into the virus particle.

Proteins building the virus particle were identified by mass spectrometry of purified virus particles and confirm our predictions (Supplementary Tables S1, S2). The major capsid protein exhibits the HK97 fold and is most similar to that of arTV BJ1 (45% identity). Additionally, we could identify the phage portal protein, two capsid proteins with similarity to *Natrialba* virus PhiCh1 capsid proteins, the putative head-tail adaptor, the tail tape measure protein, and the tail terminator protein. ORF18 and ORF25 products, identified in the virus particle, did not show any similarities to other viral proteins, therefore, their role within the virus particle remains enigmatic. Two proteins that could be involved in host attachment were identified. ORF24 product, that is present in virus particle, exhibits a immunglobulin-like fold (fibronectin type III domain) at the N-terminus. This domain is commonly found in bacterial phage tails and has been shown to exhibit membrane binding activity (Fraser et al., 2006; Kageyama et al., 2009; Pell et al., 2010). In the C-terminus we identified a sialidase domain. Sialidase (neuraminidase) is known to cleave glycosidic linkages and is essential for some eukaryotic viruses (von Itzstein, 2007). It could be involved in destabilizing the glycosylated S-layer and facilitating access to the host membrane. ORF23 product, also detected in virus preparations, shows similarities to a phage receptor binding protein with a high probability ($p = 99.06$). ORF4 product, with significant similarities to DNA double-strand break repair protein MRE11, and ORF43 product, annotated as transcription initiation factor, were also detected in virus particles. Both could be associated with the virus genome. ORF4 product might facilitate circularization of the virus genome within the host, while ORF43 product could recruit the host RNA polymerase ensuring timely transcription of the viral genome after injection.

Genome comparison with 63 arTV (Liu Y. et al., 2021) and phylogenetic tree construction (Supplementary Figure S4), suggests HRTV-DL1 to be closest related to arTV BJ1 (Pagaling et al., 2007) and CGphi46 (NC_021537) (Figure 2) that have been classified into the *Flexireviridae* family. However, HRTV-DL1 only shares nine genes with CGphi46, but none with a sequence identity above 60%, and shares sixteen genes with BJ1 of which only three have a sequence identity above 60% (ORF6, ORF33, ORF34). Therefore, HRTV-DL1 should be classified into a different genus, that we propose to name *Deelavirus* in accordance with its origin (Deep Lake).

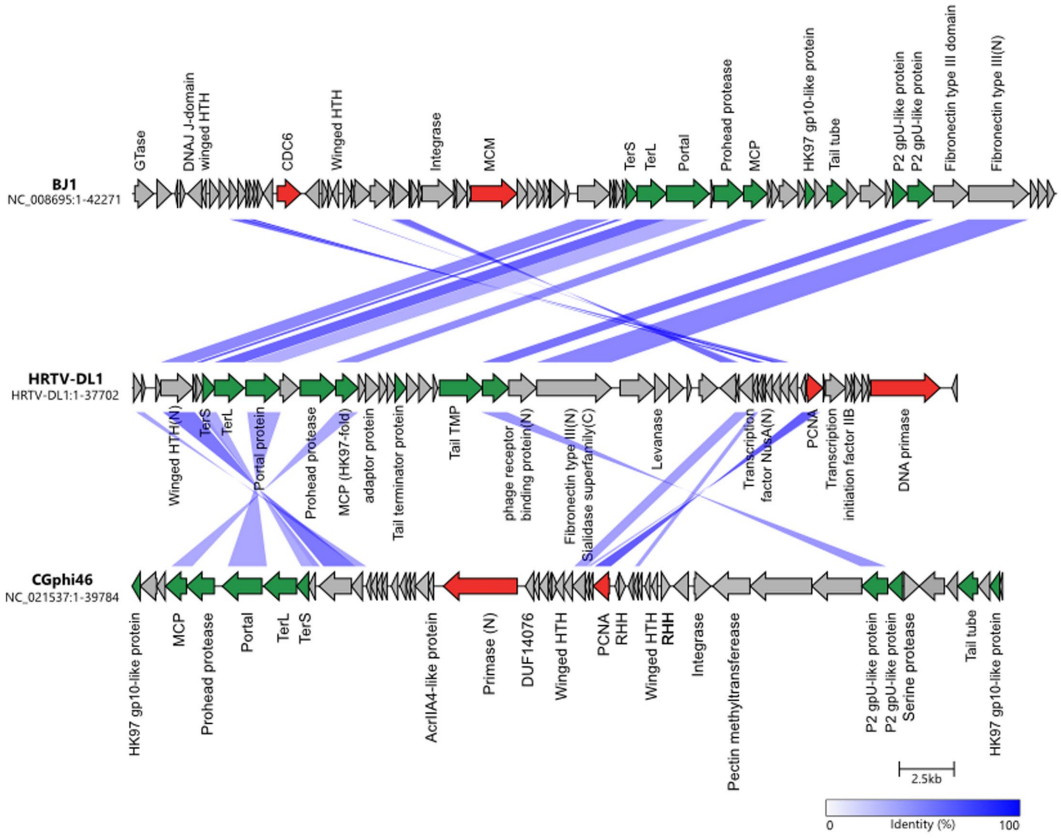


FIGURE 2
Comparisons of genomes HRTV-DL1 (OP630575) with BJ1 (NC_008695) and CGphi46 (NC_021537) belonging to the *Flexireviridae*. Protein functions are indicated above or below the corresponding ORFs. Genes encoding virus morphogenesis–related proteins are colored in green, whereas replication-related genes are colored in red. Homologous genes shared between viruses are connected by shadings of different degrees of blue based on the amino acid sequence identity. Figure was generated with clinker & clustermap.js (Gilchrist and Chooi, 2021).

Life cycle of HRTV-DL1

To gain insights into the initial stages of HRTV-DL1 entry, we followed the adsorption of HRTV-DL1 to ACAM34 host cells. The adsorption was very rapid, with ~75% of virions being bound to cells within the first 30 s of infection. Further adsorption plateaus after 15 min with 99% of particles bound to host cells (Figure 3A). Such a rapid adsorption is rather uncommon for haloarchaeal arTV. *Haloarcula hispanica* tailed virus 1 (HHTV-1), binds extremely slowly with virions adsorbed only after 3 h and binding of *Halorubrum* virus HRTV-1 was only detected after 25 min (Kukkaro and Bamford, 2009). Such diversity in adsorption rates indicates differences in adsorption mechanisms and receptors. The rapid adsorption of HRTV-DL1 suggests that the receptor is easily accessible and highly abundant on the cell surface. After successful infection, host cell lysis occurred between 56 and 68 h post infection (p.i.) at 28°C, which appears to be a long intracellular phase prior to lysis when compared to other haloarchael tailed viruses (Pietilä et al., 2013a). The virus-to-host ratio (VHR) was determined by qPCR as 95 viral genome copies per host main chromosome copy number a few hours prior lysis (Figure 3B). Interestingly, we observed changes to the cell morphology of infected *Hrr. lacusprofundi* at the onset of cell lysis. While the majority of the uninfected cells retained their rod-shaped morphology, infected cells tended to round up in the late stages of infection (Supplementary Figure S5). However, a significant increase in cell size, as shown for another archaeal virus (Liu J. et al., 2021), was not observed.

Both *Hrr. lacusprofundi* and HRTV-DL1 were isolated from a lake that experiences very low temperatures (Gibson, 1999). To determine whether low temperatures have an influence on the virus life cycle, we performed growth experiments at different temperature between 4°C and 30°C. Host growth rates change at different temperatures, consequently, lysis was observed at different time points post infection. However, lysis always occurred in early exponential phase at all temperatures (Supplementary Figure S6). Therefore, we conclude that temperature does not influence the life cycle of HRTV-DL1 under the laboratory growth conditions tested.

HRTV-DL1 exhibits a narrow host range that uncovers genetic changes in the laboratory strain serving as host organisms

HRTV-DL1 was isolated from Deep Lake, that is dominated by 4 different haloarchaea (*Halohasta litchfieldiae* tADL, *Hrr. lacusprofundi*, halophilic archaeon DL31, and *Halobacterium* DL1) (DeMaere et al., 2013). CRISPR spacers against HRTV-DL1 were detected in the CRISPR loci of all except 1 species (*Halobacterium* DL1) (Supplementary Table S3). CRISPR spacers of all three

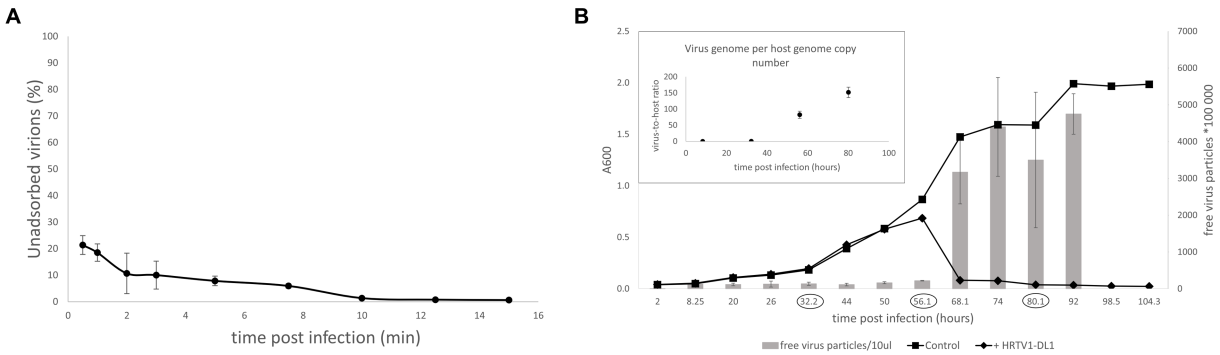


FIGURE 3
Life cycle of HRTV-DL1. **(A)** Adsorption of HRTV-DL1 to cells of *Hrr. lacusprofundi ACAM34_UNSW*. Cells were incubated with HRTV-DL1 at room temperature (20°C). The number of unbound virus particles was determined at different time points post infection by plaque assay. Graph represents one of three biological replicates. Error bars represent standard deviation from three independent experiments. **(B)** Growth curve of uninfected control and HRTV-DL1 infected *Hrr. lacusprofundi* ACAM34_UNSW. Free virus particles in 10ul culture supernatant were determined by plaque assay. Inset: Virus-host ratio determined by comparing copy numbers of host main chromosome and virus genome. Copy numbers were determined per ml cell culture by qPCR and samples were taken at time points encircled in the growth curve in the main figure. Graphs represent one of three biological replicates. Error bars represent standard deviation from three independent experiments.

different Deep Lake haloarchaea target different positions on the virus genome, indicating that they were acquired independently and not by gene transfer. Therefore, DL31 and tADL could be potential host organisms for HRTV-DL1. We tested all 4 species, but only *Hrr. lacusprofundi* ACAM34 was susceptible to infection. Adsorption assays revealed that attachment of the virus to all tested strains is impaired (Supplementary Figure S7), showing that infection is already compromised at the adsorption stage. We assume that the attachment site of HRTV-DL1 is highly diverse. Modifications of cell surface proteins, in particular the S-layer that was found to be highly distinct, has already been proposed as a virus exclusion mechanisms of Deep Lake haloarchaea (Tschitschko et al., 2015, 2018).

When continuing characterization of HRTV-DL1 in a new laboratory, we used a fresh stock of *Hrr. lacusprofundi* ACAM34 provided by the DSMZ. Surprisingly, while the virus successfully adsorbed to the host cell and injected its genome, HRTV-DL1 was not able to complete a lytic life cycle in the DSMZ strain. No lysis was observed in cultures and virus DNA could not be detected within the host at 56 h p.i. (Figure 4). We subsequently assumed that the HRTV-DL1 susceptible strain, from now on referred to as ACAM34_UNSW (for University of New South Wales, Australia, the location of the laboratory the strain originated) had experienced genomic changes while being grown over a long time in the laboratory.

Genome comparison of the ACAM34 laboratory strain with the ACAM34 strain from a culture collection reveals the loss of the majority of virus exclusion mechanisms in the laboratory strain

To determine genetic differences between the DSMZ strain (ACAM34_DSMZ) and the HRTV-DL1 sensitive strain (ACAM34_UNSW), we re-sequenced the genome of both strains (genomes provided in Supplementary Material).

ACAM34_DSMZ assembled into 4 contigs. Contig 1 is circular with a size of 2,735,247 bp, (coverage 1,080) representing the main chromosome (CHR1) with 99.99% identity to the published genome. Contig 2 is circular with a size of 525,899 bp (coverage 1,167), representing the secondary chromosome (CHR2) with 99.99% identity to the published sequence. Contig 3 is circular with a size of 431,344 bp (coverage 1,137), representing the plasmid pHLAC01 with 99.99% identity to the published sequence. Contig 4 is linear with the size of 6,176 bp matching a short region on CHR2 covering Hlac_3232 to Hlac_3234 with an insertion of a transposon within Hlac_3234. This contig has a lower coverage (432) than contig 2 and likely represents a variant of this region in the CH2 of the strain. In conclusion, the genome of ACAM34_DSMZ is to 99.99% identical with the published sequence.

Surprisingly, ACAM34_UNSW assembled into only one single circular contig of 3,058,421 bp. Genome comparison with the published genome (Figure 5A) revealed that the entire plasmid pHLAC01 and about 38% (173 kb) of CHR2 is not present, while the remaining 61.11% of CHR2 (Hlac_2782 – Hlac_3106) is integrated into CHR1 between Hlac_1757 and Hlac_1759, two copies of ISH3 family transposase ISHla1 (Siguier et al., 2006) with 100% sequence identity. Hlac_2781 and Hlac_3106 flanking the CHR2 region integrated into CHR1 are also ISHla1 transposases that are 100% identical to Hlac_1757 and Hlac_1759. We suggest that the integration of CHR2 occurred by homologs recombination between the copies of ISHla1 and at the same time caused the loss of the remaining CHR2. The insertion of a large genomic fragment into the main chromosome due to transposition activity has already been reported for a *Halobacterium salinarium* laboratory strain (Pfeiffer et al., 2008), and the insertion of a large plasmid into the main chromosome has been reported for *Haloferax volcanii* (Hawkins et al., 2013). Apart from the insertion of CHR2, CHR1 remains highly conserved with an identity of 99.999%. One minor change is found in an intergenic region between Hlac_0599 and Hlac_0600 that is missing a stretch of 23 nucleotides. The integrated remains of CHR2 have three significant changes. First, a transposase was

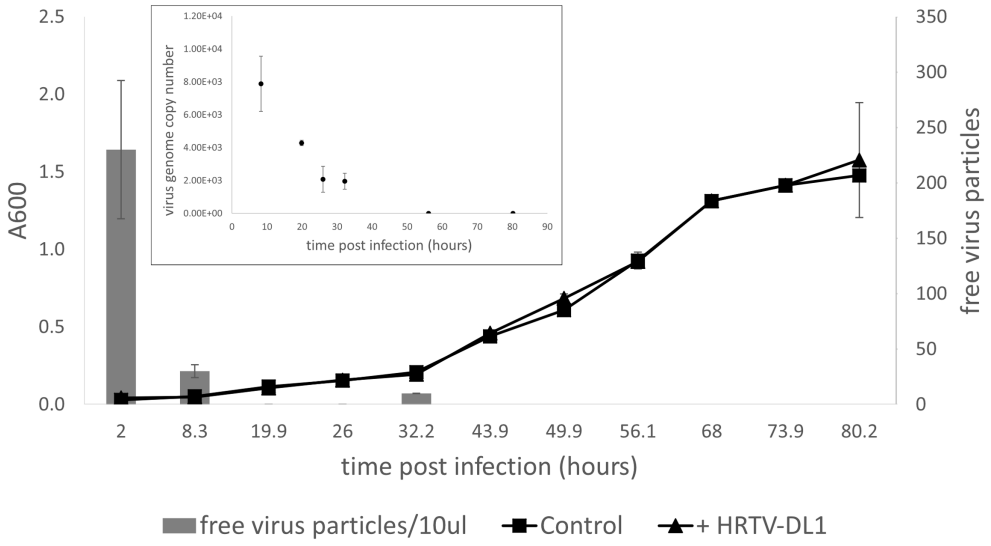


FIGURE 4
Infection of *Hrr. lacusprofundi* ACAM34_DSMZ with HRTV-DL1. Growth curve of uninfected control and HRTV-DL1 infected *Hrr. lacusprofundi* ACAM34_DSMZ. Free virus particles in 10ul culture supernatant were determined by plaque assay. Inset: Virus genome copy number determined per ml cell culture by qPCR. Samples were taken from the growth curve shown in the main figure. Samples in which virus genomes could not be detected by qPCR were set to 0. Graphs represent one of three biological replicates. Error bars represent standard deviation from three independent experiments.

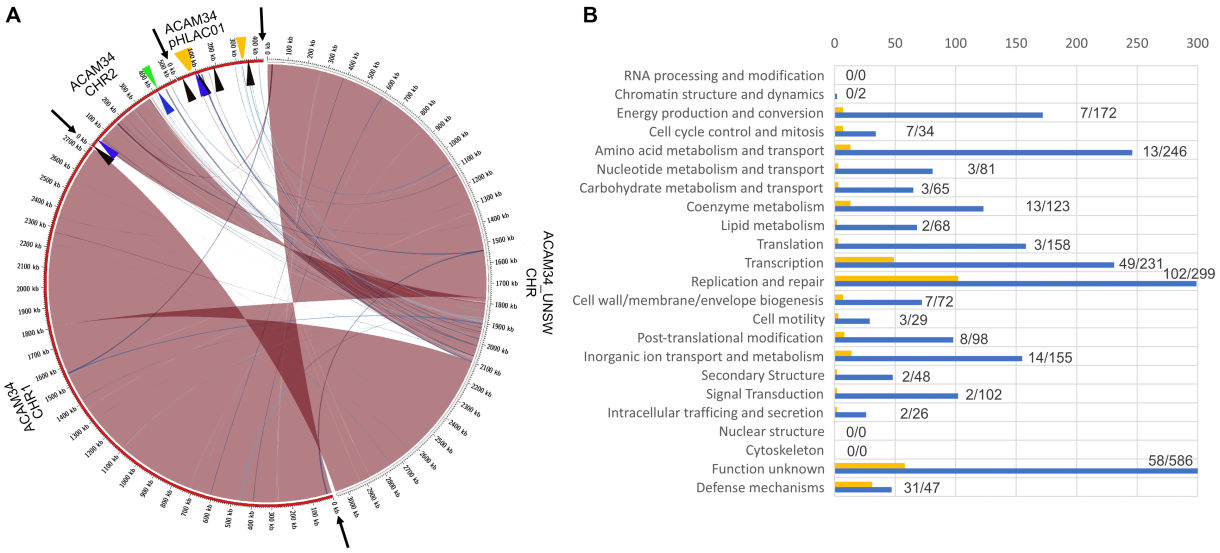


FIGURE 5
(A) Genome comparison of ACAM34_UNSW with the three replicons CHR1, CHR2 and pHLAC01 (NC_012029.1, NC_012028.1 and NC_012030.1) of the reference genome from NCBI (ATCC 49239). Interconnecting lines highlight regions present in both genomes (red/blue colors indicates same/reverse orientation on the two genomes). The figure highlights that the ACAM34_UNSW genome (on the right side, white outer ring) consists of a single replicon comprising the full length primary replicon (NC_012029.1) and parts of one secondary replicon (NC_012028.1) of the reference genome while pHLAC01 (NC_012030.1) of the reference genome (on the right side, red outer ring) is lost. Genomic regions shared between the two genomes were identified using NUCmer (Kurtz et al., 2004) and visualized with Circos (Krzywinski et al., 2009) wrapped within the script mummer2circos.py (https://github.com/metagenlab/mummer2circos). Black arrows indicate the borders between the three replicons of the reference genome. Colored arrows mark the position of the virus defense systems predicted with PADLOC (Payne et al., 2022). Black: DNA modification system, Blue: Argonaut-like protein, Green: BREX system, Orange: CRISPR system. **(B)** Functional profile of genes lost in ACAM34_UNSW. Blue bars represent the number of genes assigned to each particular functional category of all genes present on the ACAM34_DSMZ genome (X). Yellow bars represent the number of genes assigned to each particular functional category of all genes lost in ACAM34_UNSW (Y). Numbers behind the bars indicated genes lost / total amount of genes assigned to the category (Y/X). Functional classification of the *Hrr. lacusprofundi* genomes was performed using the cluster of orthologous genes (COG) database.

inserted into Hlac_2793, a predicted ADP-ribosylglycohydrolase, and disabled the gene. Second, Hlac_2835, with a MarR-like HTH domain within the N-terminus, has several changes on amino acid level within the C-terminus. And third, one of the two 16 s ribosomal RNA genes has a number of nucleotide changes that make up 3% of the sequence.

It remains unknown how the plasmid pHLAC01 was lost in ACAM34_UNSW. Analysis of the functional potential of the genes lost in ACAM34_UNSW revealed that the majority of them belong to the functional categories "Defense" (31 out of 47) and "Replication and repair" (102 out of 299), that are likely essential under virus infection or other stress condition influencing genome stability (Figure 5B). However, when grown under defined laboratory conditions these genes do not seem to be essential and are therefore likely too costly to be maintained. pHLAC01 exhibits the only CRISPR locus on the ACAM34 genome. This CRISPR locus contains 10 spacers that give a 100% match to the HRTV-DL1 genome (Supplementary Table S3). We therefore assume, that the resistance of ACAM34_DSMZ against HRTV_DL1 infection could be caused by the intact CRISPR system on pHLAC01, that is lost in ACAM34_UNSW. Nevertheless, there are also a number of other putative virus defense systems encoded on pHLAC01 and the region of CHR2 that is lost in ACAM34_UNSW, such as restriction modification systems (RM-systems), Argonaut proteins and a BREX system [Figure 5 and Supplementary Table S4, determined with PADLOC (Payne et al., 2022)]. These could also cause ACAM34_DSMZ resistance to HRTV-DL1 infection. Only one putative defense mechanism, an Argonaut protein (ACAM34UNSW_01791, Hlac_2785) (Willkomm et al., 2018), is still present in ACAM34_UNSW, leaving the strain basically unprotected against viral infection without any known virus exclusion mechanism.

HRTV-DL1 infection induces the re-mobilization of CHR2 in ACAM34_UNSW

We analyzed the host response to HRTV-DL1 infection by determining transcriptional changes at two time points 32 h p.i. (middle phase, after infection has been established, time point 1 = T1) and 56 h p.i. (late phase, immediately before lysis, time point 2 = T2) for both strains ACAM34_DSMZ and ACAM34_UNSW.

The ACAM34_DSMZ transcriptome recovered no reads for the virus genome in both samples, which is consistent with our results that the virus genome copy number is already very low at 32 h p.i. (time point 1 = T1) and not detectable at 56 h p.i. (time point 2 = T2) (Figure 4). None of the predicted virus defense mechanism (Supplementary Table S4), including the CRISPR system, showed significant upregulation ($\log2 > 1$) (Supplementary Table S5). We therefore concluded that the time points chosen were too late post infection to determine the defense system active against HRTV-DL1 in ACAM34_UNSW. However, when comparing uninfected controls with infected samples, we still detect some changes to the host transcription. At T1 (32 h p.i.) 468 genes are upregulated more than 2 fold ($\log2 > 1$) and none are downregulated, and at T2 (56 h p.i.) we detect 98 genes up- and 5 downregulated (Supplementary Table S5). When analyzing the functional potential of the upregulated genes at both time points, the majority falls into the categories transcription, replication and repair mechanisms, posttranslational modification, nucleotide metabolism and energy production and conversion (Figure 6).

The ACAM34_UNSW transcriptome recovered 0.2% of all mapped reads 32 h p.i. and 41% of all mapped reads at 56 h p.i. for the virus genome, consistent with the virus genome copy numbers increasing over time (Figure 3).

The transcriptional profile at time point 1 is very similar to that of T1 in ACAM34_DSMZ with 371 genes upregulated ($\log2 > 1$) and none downregulated, and functional categories transcription, replication and repair mechanisms, posttranslational modification, nucleotide metabolism and energy production and conversion being the most upregulated (Figure 6). Additionally, we find the same genes among the 30 most upregulated genes, encoding for cold-shock proteins, ribosomal proteins, the ferredoxin Hlac_2176 and transcriptional regulators, being upregulated in ACAM34_UNSW and ACAM34_DSMZ, indicating that both strains exhibit a very similar response to HRTV-DL1 infection 32 h p.i. (Supplementary Table S6).

However, T2 shows a different profile, with 201 genes upregulated and 333 genes downregulated. The most upregulated functional category is 'replication and repair', consistent with a takeover of HRTV-DL1 and the degradation of the host genome (Supplementary Figure S8). Genes involved in metabolic processes seem to be subjected to downregulation. Amongst the 30 most upregulated genes, only two were also detected as upregulated in ACAM34_DSMZ 56 h p.i. (Hlac_0148, Hlac_0919). Surprisingly, the majority of genes on former CHR2 inserted in CHR1 (ACAM34UNSW_01788 to ACAM34UNSW_02100) was upregulated (Supplementary Table S6). Therefore, we assumed that this region might have been re-mobilized during virus infection and that the upregulation is due to an increased copy number of the secondary replicon. Indeed when comparing copy numbers of CHR1 with former CHR2 by qPCR in the same samples, we detect a slightly higher copy number for CHR2 in the late stage of infected samples (1.3x the copy number of the main chromosome) (Supplementary Figure S8). We conclude that CHR2 is mobilized only in a portion of cells in the population, because CHR2 usually has a copy number of 1.5–2 copies per CHR1 in ACAM34_DSMZ. When normalizing the expression values to the copy numbers, only 30 of 109 genes are still upregulated ($\log2 > 1$), correcting the total number of upregulated genes from 201 to 122.

Within the 30 most downregulated genes at T2 we find genes encoding for ribosomal proteins, for proteins involved in amino acid metabolism, proteins involved in energy production and conversion and proteins involved in oxidative stress response (Supplementary Table S6), being a typical signature for a virus driven takeover of the host cell metabolism (Karlsson et al., 2005; Howard-Varona et al., 2020).

The loss of virus defense mechanisms in ACAM34_UNSW leads to the activation of potential alternative virus exclusion mechanisms upon infection with HRTV-DL1

Only one gene, predicted to be potentially involved in virus defense (ACAM34UNSW_01791), an Argonaut related nuclease [PADLOC (Payne et al., 2022)] that is still present in ACAM34_UNSW, is slightly upregulated ($\log2 = 0.9$) in ACAM34_UNSW, but not in ACAM34_DSMZ (Hlac_2785) at any time point. We suggest that ACAM34UNSW_01791 is activated in ACAM34_UNSW due to

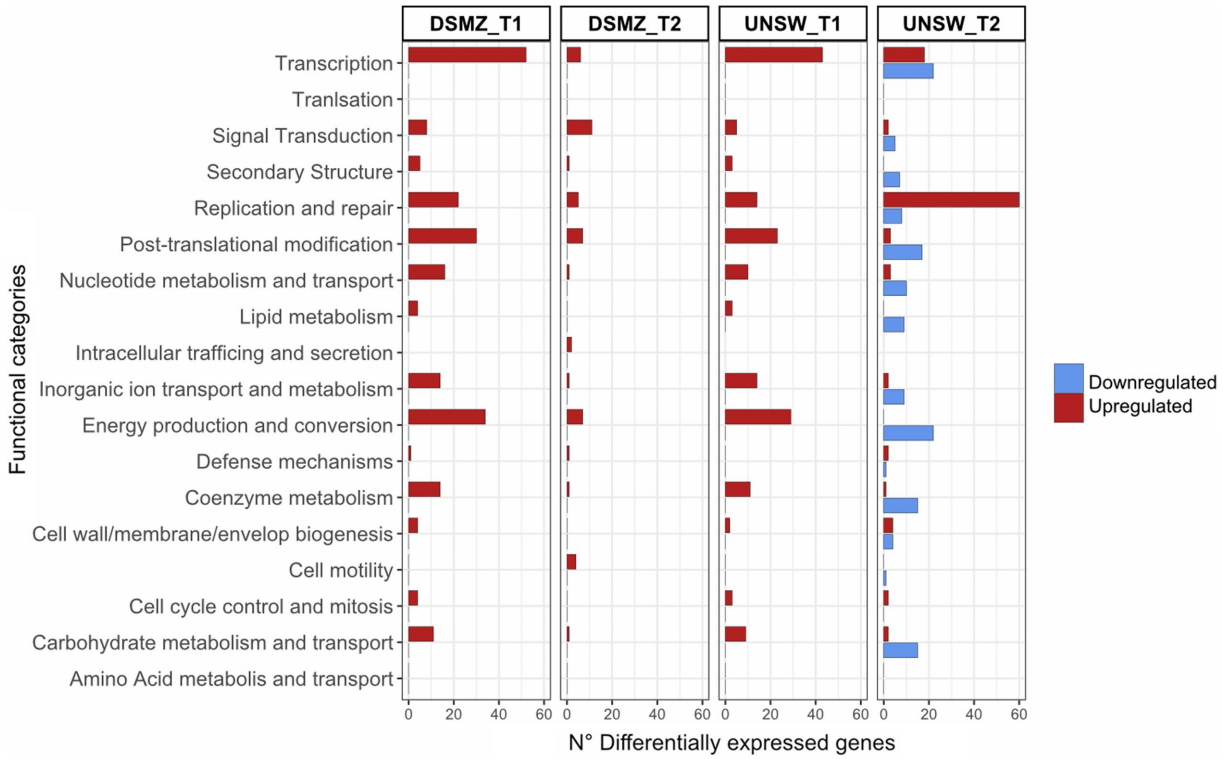


FIGURE 6
Functional profile of differentially expressed genes in ACAM34_DSMZ and ACAM34_UNSW under infection with HRTV-DL1. Bars represent the number of differentially expressed genes assigned to each particular functional category at a given time point (T1 = 32 h post infection, T2 = 56 h post infection). Functional classification of the *Hrr. lacusprofundi* genomes was performed using the cluster of orthologous genes (COG) database. The colors of the bars indicate if genes are upregulated (red) or downregulated (blue) respectively. Genes that were not assigned to any functional category are not displayed on the plot (for detailed information see Supplementary Tables S5, S6).

the lack of other defense mechanisms, however, its activity does not seem to interrupt the lytic cycle of HRTV-DL1. Whether it has an impact on the efficiency of infection or virus production remains to be determined. To identify other potential virus exclusion mechanisms we analyzed the 30 most upregulated genes. We propose that some of these genes might be implicated in virus exclusion of ACAM34_UNSW, that has lost all computational identifiable virus defense systems. The predicted functions of the upregulated genes are summarized in Supplementary Table S6 and some of particular interest, are discussed below.

ACAM34UNSW_01107 (Hlac_1086), shows significant similarities (HMM, Probability 99.95, E-value: 8.2e-26) to Transposon-associated TnpB, that is guided by an RNA derived from a sequence upstream of the TnpB (Karvelis et al., 2021) and also shows significant similarities (HMM, Probability 99.95, E-value: 8.2e-26) to ACAM34UNSW_01107. Interestingly, we do find a short (118 amino acids) ORF (ACAM34UNSW_01107a) upstream of ACAM34UNSW_01107 that is highly similar to the N-terminus of Hlac_2960, also showing high similarity to RNA-guided DNA endonuclease TnpB. ACAM34UNSW_01107a might represent TnpA that is also very small in size (140 amino acids) and responsible for the mobilization of TnpB. In absence of any other virus defense mechanism in ACAM34_UNSW, this gene might play an important role as a virus exclusion mechanism.

Two hypothetical proteins (ACAM34UNSW_01618 and ACAM34UNSW_01620) within the 30 most upregulated genes, enclose a predicted chromosome maintenance protein (SMC), ACM34UNSW_01619 (log2 = 1.1, T2). Additionally, another predicted SMC, ACM34UNSW_01796, is also slightly upregulated (log2 = 1.1, T2). SMC complexes were found to interfere with virus replication in humans (Gibson and Androphy, 2020; Han et al., 2022). Additionally, SMC-like proteins were found to be associated with the recently discovered Wadjet system, that has been shown to exclude circular foreign DNA in bacteria (Doron et al., 2018; Liu et al., 2022). Since both ACAM34UNSW_01618 and ACAM34UNSW_01620 do not show any similarity to other proteins involved virus defense, an antiviral activity of this gene cluster requires experimental verification.

Finally, another interesting candidate for a virus exclusion mechanism is ACAM34UNSW_02085 (Hlac_3088). The gene is a passenger gene of a transposon with transposase ACAM34UNSW_02084 (Hlac_3078), and could be mobilized similarly to other virus exclusion mechanisms (Broecker and Moelling, 2019) or antibiotic resistance genes (Babakhani and Oloomi, 2018). Hlac_3088 exhibits a PGF-TERM domain, an archaeal protein-sorting motif recognized by an archaeosortase (Abdul Halim et al., 2013), and is predicted to be a non-cytoplasmic protein with a TM domain. HMM search showed significant similarities with the S-layer protein of *H. volcanii*, however, Hlac_3088 is too small to be an S-layer protein. When comparing Hlac_3088 expression levels with those of the S-layer protein (Hlac_2976), Hlac_3088 has usually lower expression levels compared to the S-layer. However, at T2 it actually exceeds the expression levels of the S-layer gene

(Supplementary Table S7). We suggest that Hlac_3088 is located at the surface of the cell and could protect the cell by hiding the receptor and interfering with the attachment of the virus, similar to TraT, encoded by the F plasmid of *Escherichia coli*, that interacts with OmpA and blocks phage adsorption (Achtman et al., 1977). Alternatively, it could stabilize the cell and prevent or at least delay cell lysis.

While none of the proposed mechanisms are fully effective against HRTV-DL1, because the virus is completing its lytic cycle, they might prevent the lysis of the entire population and allow virus escape mutants to evolve, and could be responsible for the long intracellular phase of the virus life cycle.

HRTV-DL1 gene expression during infection reveals presence of a PATE with strong activity

Transcriptomic data from HRTV-DL1 infected ACAM34_UNSW cultures revealed that all annotated HRTV-DL1 ORFs were expressed, while ORF3 was, by far, the most transcribed ORF (Supplementary Table S8). When examining the reads mapping ORF3, we discovered that there is a peak that expands upstream of the annotated gene covering a region of approximately 300 bp and the majority of reads mapped antisense to the predicted coding sequence. Upon blasting the sequence against public databases, we discovered that it matched 99% (1 mismatch, 1 gap) to a region within the hosts genome and a region that was previously described in plasmid pR1SE infecting the same host (Erdmann et al., 2017). Further inspection revealed that this region represents a PATE (Palindrome-Associated Transposable Element), with a length of 489 bp that expands over the gap between ORF2 and ORF3. Only the C-terminal end of the PATE shows high expression, which includes the N-terminus of predicted ORF3, and expression occurs almost exclusively in reverse direction. We assume that ORF3 is a pseudogene and expression within this ORF is due to the PATE activity. In *Halobacterium salinarium* a PATE has been shown to be active as ncRNA (Gomes-Filho et al., 2015) and we imagine a similar function for this PATE. However, it is encoded by both the host and the virus, so it is difficult to predict whether its activity is beneficial for the host or the virus, and whether the virus acquired it from the host or vice versa.

ACAM34_UNSW can develop resistance to HRTV-DL1 infection

To observe adaptation to HRTV-DL1 infection, we established long-term infected cultures, in triplicates, for ACAM34_UNSW with an MOI of only 1 to allow a recovery of infected cultures. For ACAM34_UNSW lysis was observed as expected between 45 and 70 h post infection and the cultures were maintained to allow the growth of HRTV-DL1 resistant clones. The cultures slowly recovered over a time frame of app. 200 h. After transfer of the recovered cells in fresh media, the cultures experienced another long lag phase of about 120 h before establishing a normal growth (Supplementary Figure S9). Cells from the recovered cultures were plated on agar plates to obtain single colonies and for each biological replicate we tested 24 colonies for infection with HRTV-DL1 by PCR. In 14 of the 72 tested colonies we could still detect the virus by PCR (Supplementary Figure S10), indicating that the majority of the cells were able to exclude the virus. We chose two escape mutants for further characterization.

A large genomic deletion in escape mutant ACAM34_UNSW_2.14 identifies the S-layer as HRTV-DL1 receptor

Escape mutant ACAM34_UNSW_2.14 was found to be fully resistant to HRTV-DL1 infection (Figure 7A). Infection is already impaired at the adsorption stage (Supplementary Figure S11), indicating that the receptor for HRTV-DL1 might have undergone changes in this mutant. Genome analyzes of UNSW_2.14 revealed a large gap covering the entire integrated CHR2 with two interruptions, indicating a massive rearrangement in several independent events. The region between ACAM34_UNSW_01838 (Hlac_2826) and UNSW_01867 (Hlac_2855), both transposases, is still present, as well as the region between the two transposases UNSW_01829 (Hlac_2819) and UNSW_01802 (transposase inserted into Hlac_2793). The coverage of these two regions is identical with the remaining chromosome, indicating that they are still integrated. We detect a few other changes, including twelve SNPs in intergenic regions, two silent SNPs in ORFs, three SNPs that lead to an amino acid substitution (Supplementary Table S9). None of these changes affect genes that would have an influence on virus-host interactions. However, the large deletion includes one of two S-layer genes (Hlac_2976) encoded by *Hrr. lacusprofundi*. Additionally, we also detected a deletion of two amino acids (D45 and S46) in the N-terminus of the second S-layer gene (Hlac_0412, 41% identity with Hlac_2976). Since adsorption of HRTV-DL1 to UNSW_2.14 is fully abolished, we conclude that the S-layer represents the receptor for HRTV-DL1, as it has been proposed for another arTV (HFTV1) (Schwarzer et al., 2023). Variability of S-layer genes of ACAM34 was previously detected by two 'omics'-based studies of Deep Lake, and was suggested to be driven by an arms race between viruses and host (Tschitschko et al., 2015, 2018). Our results strongly support this hypothesis. We also propose that the deletion of Hlac_2976 causes the slightly reduced growth rate of UNSW_2.14 (Figure 7A). Cell morphology, as observed by light microscopy, does not seem to be significantly impacted (Supplementary Figure S12), suggesting that Hlac_0421 alone is able to maintain the S-layer.

Escape mutant ACAM34_UNSW_3.3 confirms the S-layer as HRTV-DL1 receptor and suggest a Cdc6 protein to be involved in virus defense

Adsorption of HRTV-DL1 to escape mutant ACAM34_UNSW_3.3 is greatly reduced, but not fully impaired (Supplementary Figure S11). The virus genome could be detected in host cells, however, it is unclear whether it is actively replicated, and no cell lysis was observed and no virus particles were detected in culture supernatants (Figure 7B and Supplementary Table S11). We conclude that adsorption is partially impaired, and either virus genome replication, virus particle production or cell lysis is inhibited. Genome sequencing revealed 12 SNPs in intergenic regions, two silent SNPs, two aa substitutions observed previously in UNSW_2.14 and a

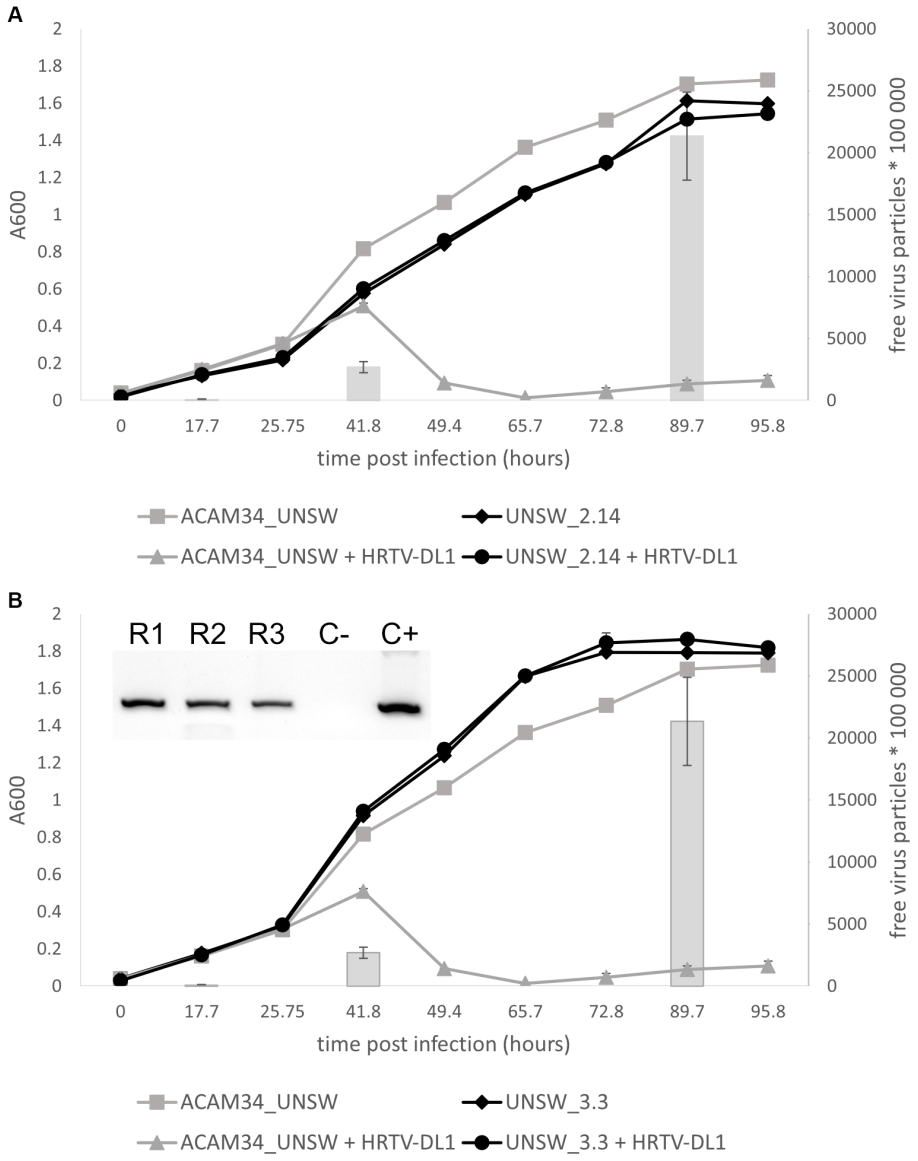


FIGURE 7
Life cycle of HRTV-DL1 in ACAM34_UNSW escape mutants. Growth curves of uninfected control and HRTV-DL1 infected *Hrr. lacusprofundi* ACAM34_UNSW, and escape mutants UNSW_2.14 **(A)** and UNSW_3.3 **(B)**. Free virus particles in 10 ul culture supernatant were determined by plaque assay and are presented by gray bars for the parental strain. For number of virus particles in supernatants of escape mutant cultures please refer to Supplementary Table S11. Graphs represent one replicate for the uninfected controls and one of three biological replicates for infected cultures. Error bars represent standard deviation from three independent experiments. Inset in **(B)** PCR confirming infection of three replicates of UNSW_3.3 with HRTV-DL1 (R1-R3 replicate 1–3; C– negative control; C+ positive control).

frame shift in ACAM34UNSW_01960, a hypothetical protein. We do not expect that any of these mutations has an influence on the virus life cycle, since the majority of them are also present in UNSW_2.14 that has a different phenotype. The major change influencing virus-host interactions is an insertion of 12 aa (TPPTVSRLCFDT) between amino acids 209 and 210 of ACAM34UNSW_01982 (Hlac_2976), the S-layer gene that is also affected in UNSW_2.14. However, the mutation only accounts for 90% of the population, suggesting that the drastically reduced adsorption is caused by this mutation. Some virus particles can adsorb to the remaining 10% of the population, or can possibly attach to cells via the second S-layer protein (Hlac_0412). Hlac_0412 did not experience a mutation in UNSW_3.3, but we still see a dramatically reduced adsorption, indicating that either Hlac_2976 is the preferred receptor of HRTV-DL1, or the Hlac_2976 is preferably used by the cell to build the S-layer.

Despite the reduced infection efficiency, we detect virus genomes within cells, however, infected cells do not produce virus particles (Figure 7B). Only one other genomic change was detected in UNSW_3.3, the deletion of an origin of replication (ORI), that could possibly be responsible for preventing virus replication, virus particle production or cell lysis. ORIs are binding sites for Orc1/Cdc6 proteins and the deletion of an ORI could redirect Cdc6 causing the same effect as the upregulation of Cdc6 proteins. Indeed, the transcriptomics data show that while we only detected one (Hlac_3217) of fifteen annotated *orc1/cdc6* genes being upregulated in resistant ACAM34_DSMZ, five of seven remaining *orc1/cdc6*

genes are differentially regulated in sensitive ACAM34_UNSW (Supplementary Table S5), including the *orc1/cdc6* gene adjacent to the destroyed origin (ACAM34UNSW_01845, Hlac_2833). Studies have shown that Cdc6 can bind dsDNA without sequence specificity (Feng et al., 2000), that dysregulation of Cdc6 expression can lead to inhibition of replication (Kundu et al., 2010) and that Cdc6 is also able to recruit the RNA polymerase I for rDNA transcription initiation (Huang et al., 2016). Additionally, a recent study showed that an Orc1/Cdc6 homolog, encoded by the archaeal virus SNJ2, is involved in regulating the lysogenic-lytic life cycle of SNJ2 (Chen et al., 2023). Indeed, we find an integrase encoded four genes upstream of the Cdc6 that is affected by the ORI deletion in ACAM34_UNSW, indicating that this Cdc6 might have been acquired from a virus genome. Therefore, we suggest that Cdc6 does affect the life cycle of HRTV-DL1, however, the mechanism needs to be investigated experimentally.

Conclusion

In this work, we characterize a new head-tailed archaeal virus (HRTV-DL1) isolated from Deep Lake, Antarctica. HRTV-DL1 belongs to the *Flexireviridae* family, and we propose to classify HRTV-DL1 into the new genera *Deelavirus*. HRTV-DL1 exhibits a linear genome of 37.7 kb in size, that is replicated as circular genome within host cells, and is most likely terminally redundant similar to the genome of a number of other archaeal tailed viruses (Liu Y. et al., 2021).

HRTV-DL1 exhibits a lytic life cycle in a *Hrr. lacusprofundi* strain (ACAM34_UNSW) grown in the laboratory for a time that is not retraceable anymore. However, the type strain of *Hrr. lacusprofundi* obtained from the DSMZ did not show susceptibility to HRTV-DL1 infection. Genome sequencing uncovered that ACAM34_UNSW had undergone a massive genomic rearrangement and lost about 18% of its genome, a phenomenon that has also been described for other haloarchaea (Pfeiffer et al., 2008; Hawkins et al., 2013). While probably essential in the natural environment, this proportion of the genome, known to be highly variable in *Hrr. lacusprofundi* (Tschitschko et al., 2018), appears to be redundant when the strain is grown isolated and under laboratory conditions. Interestingly, the loss of genetic information included the majority of virus defense mechanisms present on the genome of the type strain, indicating that they could be very costly for the strain to maintain.

We were not able to determine with certainty the mechanisms that is responsible for HRTV-DL1 exclusion from ACAM34_DSMZ. However, the fact that ACAM34_UNSW lost the majority of virus defense mechanisms allowed us to uncover the receptor for virus binding and a number of undescribed putative virus exclusion mechanisms. Analysis of the HRTV-DL1 infected transcriptome of ACAM34_UNSW revealed the activity of potential TnpB-like RNA-guided DNA endonuclease, and a transposon passenger gene that could potentially prevent the virus from attaching to the host cell. Most interestingly, we discover the possible involvement of proteins shared between archaea and eukaryotes, a SMC-like protein and a CDC6-like protein, in virus defense, though antiviral activity of these proteins will require experimental verification. Infection of ACAM34_UNSW with HRTV-DL1 induces further rearrangements of the genome, in particular of the integrated CHR2, and allowed the isolation of escape mutants. We uncovered that one of two S-layer proteins is the preferred virus receptor, as recently also discovered for an arTV infecting *Haloferax gibbonsii* (Schwarzer et al., 2023), indicating that exhibiting two S-layer proteins could be a strategy of virus exclusion.

Our study provides important insights into the genome plasticity of *Hrr. lacusprofundi,* and highlights that genome plasticity has important implication for virus-host interactions. *Hrr. lacusprofundi* can be genetically modified (Gebhard et al., 2023), making *Hrr. lacusprofundi* and HRTV-DL1 to a great model system for studying virus-host interactions of archaeal tailed viruses. The new putative virus defense mechanisms discovered in this study can be investigated in detail, and possibly be identified in other organisms. ACAM34_UNSW itself can be used as model to study virus defense mechanisms. It is susceptible to a variety of viruses, including HRTV-DL1 and others (Nuttall and Dyall-Smith, 1993; Porter et al., 2005; Erdmann et al., 2017; Dyall-Smith, 2021; Alarcón-Schumacher et al., 2022), and individual virus defense mechanisms can be re-introduced to test their activity against diverse viruses.

Materials and methods

Isolation of HRTV-DL, strains and culture conditions

HRTV-DL, formerly described as DLHTHV (Tschitschko et al., 2018), was isolated from the supernatant of a colony that lysed upon propagation in liquid culture. Briefly, a sample, taken from Deep Lake (Antarctica) (Gibson, 1999) in 2014, was used to generate enrichment cultures in liquid medium (culture medium and conditions see below). Enrichment cultures were plated on solid media and a single colonie picked into liquid media. The colony was identified as *Halorubrum lacusprofundi* by 16 s rDNA sequencing using universal primers [21F and 1510R (Takahashi et al., 2014)]. After lysis was observed in the culture, cells and cell debris were removed by centrifugation (4,500 × *g*, 30 min), the supernatant of the culture (lysate) was filtered through 0.2 µm filters and used for infection of *Halorubrum lacusprofundi* ACAM34. *Halorubrum lacusprofundi* ACAM34 (DSM 5036), hereafter named ACAM34 _DSMZ, was obtained from the German Collection of Microorganisms and Cell Cultures (DSMZ). The parental strain of *Halorubrum lacusprofundi* ACAM34_UNSW was provided by either Peter Franzmann or John Bowman (not traceable anymore) and subsequently cultured in the laboratory (Ricardo Cavicchioli, School of Biotechnology and Biomolecular Sciences, The University of New South Wales, Sydney) from glycerol stocks stored at −80°C. The strain went through several rounds of culturing, plating and − 80°C storage in three different laboratories (UNSW and University of Technology, Sydney, Australia; Max Planck Institute for Marine Microbiology, Bremen, Germany) prior to re-sequencing. The number of generations that the strain had undergone during this time is not traceable. *Halohasta litchfieldiae* tADL (DSM 22187) was provided by the DSMZ. Halophilic archaeon DL31 and *Halobacterium* DL were isolated previously (DeMaere et al., 2013). All strains were grown in DBCM2 media (Dyall-Smith, 2021), with 5 g peptone and 1 g yeast extract added per liter. Incubation temperature was 28°C unless stated otherwise. Cultures were incubated in glass flasks aerobically at 120 rpm. Solid agar plates contained 16 g agar per liter and top layer

agar contained 6 g agar per liter. Plaque assays were performed by mixing virus and host with 10 mL of top layer agar that was subsequently poured on solid plates and incubated at 28°C until growth was visible.

Isolation and purification of HRTV-DL1 particles

For virus production, *Hrr. lacusprofundi* was grown in liquid culture to mid exponential phase, with an optical density at 600 nm (OD_{600}) of 0.5. Cells were harvested (4,500 x *g* for 45 min), mixed with viral suspension with a multiplicity of infection (MOI) of 5–10 and incubated for 2 h at room temperature to allow viral adsorption. Treated cells were inoculated into 500 mL liquid cultures that were monitored by measuring optical density changes (OD_{600}). After lysis occurred, cultures were centrifuged at 4,500 x *g* for 45 min to pellet the cells. The supernatant was recovered and viruses were subsequently precipitated with polyethylene glycol (PEG) 6,000 (10% w/v final concentration) and incubation at 4°C overnight. Then, viral preparations were collected by centrifugation (30,000 x *g*, 45 min, 4°C). Pellets were resuspended in DBCM2 salt solution (DBCM2 media without nutrient sources pyruvate, trypton and yeast extract), sterile filtered (pore size 0.2 µm) one time when used for plaque assay and twice when used in liquid culture infection assays. Virus solutions were stored at 4°C and were active for a minimum of 1 year. For downstream analyzes such as genome sequencing and mass spectrometry the virus particles were further purified by Cesium chloride (CsCl) density gradient centrifugation. The virus solution was treated with 200 U of DNase I and 50 µg/mL of RNase A to reduce host genomic contamination prior to loading on a CsCl density gradient (0.45 g CsCl /ml in DBCM2 salt solution) and centrifuged at 38,000 rpm for 22 h, 4°C (SW41 Ti Swinging-Bucket rotor, Beckman & Coulter). Bands containing virus particles were extracted with a syringe, diluted in three volumes of DBCM2 salt solution and re-precipitated with PEG (final concentration 10%, 4°C, overnight). After centrifugation at 30,000 x *g* for 30 min the resulting pellet was washed twice with DBCM2 salt solution and stored at −20°C before further processing.

Imaging

For Transmission electron microscopy (TEM), virus containing solution was adsorbed for 5 min to carbon coated copper grids and stained for 1 min with 2% uranyl acetate (w/v in water). Electron micrographs were generated using JEM2100 Plus at 200 kV acceleration voltage. For light microcopy, cells were fixed with 1% glutaraldehyde for 1 h at room temperature and then stored at 4°C until imaging with a Zeiss AxioPhot microscope with AxioCam MRm.

DNA extraction, manipulation and genome sequencing

Genomic DNA of all samples, if not otherwise stated, was extracted using genomic DNA extraction kit (Bioline, London, United Kingdom) according to the manufacturer's instructions. Plasmid extraction was performed with ISOLATE II Plasmid Mini Kit (Bioline, London, United Kingdom) according to the manufacturer's instructions. Nuclease digestions using specified restriction enzymes (New England Biolabs [NEB], 20 units) or exonuclease III (NEB, 5 units) were performed on 2–3 µg of DNA extracted from virus particles for 1 h at 37°C. PCR reactions were performed with the Q5® High-Fidelity DNA Polymerase (NEB) and contained 0.02 units/µL of DNA polymerase, primer concentration of 0.1 µM for both forward and reverse, 1X of Q5 Reaction Buffer and 1X Q5 High GC Enhancer. The following program was used: 5 min at 95°C, followed by 35 cycles of 30 s at 95°C, 30 s at 68°C for annealing and 30 s at 72°C for elongation. Digested DNA and PCR products were separated on 1% agarose gels in Tris-borate-EDTA buffer and stained with SYBR™ Safe DNA stain (Invitrogen). For HRTV-DL1, UNSW_2.14 and UNSW_3.3 library preparation (FS DNA Library, NEBNext® Ultra™) and sequencing (Illumina HiSeq3000, 2 × 150 bp, 1 Gigabase per sample) was performed at the Max Planck-Genome-Center Cologne (Cologne, Germany). For PacBio sequencing of ACAM34_UNSW and ACAM34_DSMZ, DNA extraction, library preparation, sequencing (Pacific Biosciences Sequel, 4 samples on one SMRT cell) and assembly ('Microbial Assembly' function in PacBio SMRT® Link v8.0 for ACAM34_DSMZ and flye assembly tool v2.6 (Kolmogorov et al., 2019) for ACAM34_UNSW was performed at the Max Planck-Genome-Center Cologne (Cologne, Germany).

Genome assembly, annotation and phylogenetic analysis

After assembly with SPAdes (Bankevich et al., 2012), the HRTV-DL genome was manually closed to one contig using primers listed in Supplementary Table S10, followed by Sanger sequencing (MICROMON DNA Sequencing Facility, Monash University, Victoria, Australia) of PCR products. The HRTV-DL1 genome was assembled using metaviralSPAdes (Antipov et al., 2020). Genome annotation for all genomes was done using Prokka (Seemann, 2014) followed by manual corrections using conserved domain based searches (Blum et al., 2020) or hidden Markov model (HMM) based searches (Söding, 2004). Phylogenetic tree reconstructions from protein sequences was done with VICTOR under optimal settings (formula VICTOR d6), as implemented at the DSMZ webserver https://victor.dsmz.de. Prediction of termini was done using PhageTerm (Garneau et al., 2017).

Protein content analysis

Protein content analysis of HRTV-DL1 particles was performed by mass spectrometry on purified virus pellets as described previously for plasmid vesicles and membrane vesicles (Erdmann et al., 2017). Peak lists were generated using Mascot Daemon/Mascot Distiller (Matrix Science, London, England) and initially submitted to the database search program Mascot (Matrix Science). Search parameters were: Precursor tolerance 4 ppm and product ion tolerances ±0.05 Da; Met(O) carboxyamidomethyl-Cys specified as variable modification; enzyme specificity was trypsin; 1 missed cleavage was possible; customized databases searched: *Hrr. lacusprofundi* ACAM34 and HRTV-DL. After genome sequencing of ACAM34_UNSW, results presented in Supplementary Table S2 were obtained using the

SEQUEST search algorithm with Thermo Proteome Discoverer™ 2.5.0.400 (Thermo Fisher Scientific) with the same settings searching customized databases *Hrr. lacusprofundi* ACAM34_UNSW and HRTV-DL1.

Virus infectivity and kinetics

To study the life cycle, cultures of *Hrr. lacusprofundi* (all strains) were synchronized using an adaptation of the "Stationary phase method" (Cutler and Evans, 1966). The strain was scratched from a -80°C glycerol stock into liquid culture and grown up to an $OD_{600} \sim 1$. Then a 20-fold dilution step in fresh media was performed (final $OD_{600} = 0.05$) and cultures were then regrown up to $OD_{600} \sim 1$. Iterative dilution and growth of the culture were repeated twice before considering a culture synchronized. For infection with HRTV-DL1, cells from synchronized cultures ($OD_{600} \sim 1$) were collected by centrifugation, resuspended in 1 mL of fresh media and infected with HRTV-DL1 virus with a MOI of 5. After incubation (2 h, room temperature), cells were transferred into liquid cultures and growth was monitored by optical density (OD_{600}). Long-term cultures were established by continuously diluting infected cultures at $OD_{600} \sim 1$ to $OD_{600} \sim 0.05$. Infection was confirmed by PCR using primers HRTV-DLF and HRTV-DLR (Supplementary Table S10). Viral titer in culture supernatants was quantified by plaque assay after removal of cells by centrifugation (11,000 × *g*, 10 min, room temperature [RT]). Intracellular virus titers and host chromosome copy numbers were quantified by qPCR. Briefly, samples of 2 mL culture in biological replicates were collected and pelleted (11,000 × *g*, 10 min, RT). Cell pellets were washed two times with 1 mL of fresh media and stored at −20°C upon DNA or RNA extraction. Quantification of *Hrr. lacusprofundi* and HRTV-DL1 genome copy number were carried out using a CFX96 Touch Real-Time PCR (Bio-Rad Laboratories, Inc., Hercules, CA, United States) and the software CFX Manager™ Software. Primers are listed in Supplementary Table S10. Each reaction (10 µL) contained 1X SsoAdvanced Universal SYBR™ Green Supermix (Bio-Rad) and primer concentrations as stated in Supplementary Table S10. The following amplification thermal cycling program was used for both primer sets: 5 min at 95°C, followed by 40 cycles of 30 s at 95°C and 30 s at annealing temperature stated in Supplementary Table S10, with readings taken between each cycle. Efficiencies of the assays were 95–102%, with R^2 values ≥0.99 for all assays. The specificity of the qPCR was confirmed by unique signals in melting curves and gel electrophoresis of PCR products.

Adsorption assays and host range assessment of HRTV-DL1

For adsorption assays, 5 mL of cells at $OD_{600} = 0.5$ were harvested by centrifugation (4,500 × *g*, 30 min, RT) and resuspended in 1 mL fresh medium. The cells were subsequently infected using a MOI of 5. At defined time intervals the adsorption was stopped by immediate centrifugation (10,000 × *g*, 5 min, RT). The number of remaining free virus particles in the supernatant was determined by plaque assay. A cell-free control (only media) served as control. Three different strains of haloarchaea [*Halohasta litchfieldiae* tADL, halophilic archaeon DL31, and *Halobacterium* DL1 (DeMaere et al., 2013)] were tested to determine the host range of HRTV-DL1. CRISPR matches to HRTV-DL1 were identified by searching for CRISPR loci using CRISPRs finder (Grissa et al., 2007) and blasting (BlastN) identified spacer against the HRTV-DL1 genome. Adsorption of HRTV-DL1 to the different strains was determined by adsorption assay as described above and the absence of HRTV-DL1 DNA within cells was confirmed by PCR on cell pellets 1 day and 5 days post infection.

Transcriptomic analyzes

RNA extraction of frozen cell pellets with 3 biological replicates for the infected and 2 biological replicates for the controls, was performed with the Direct-zol™ RNA miniprep Kit (R2051, Zymo Research). RNA concentration and integrity were assessed using Nanodrop DS-11 Spectrophotometer (DeNovix) according to the manufacturer's instructions. Library preparation and sequencing was done at the Max Planck-Genome-Center Cologne (Cologne, Germany). Briefly, ribosomal RNA were depleted prior to sequencing using the rRNA depletion Kit riboPOOL™, for *Haloferax volcanii* (88.36% identity to 16 s rDNA sequences of *Hrr. lacusprofundi*), siTOOLs Biotech®. Libraries were prepared with library kit NEBNext® Ultra™ II RNA Library Prep Kit for Illumina and sequencing was performed on an Illumina HiSeq3000 sequencer, following a 1 × 150 run. Read trimming and mapping was performed with the "Map to reference" function (Mapper 'Geneious RNA') with medium-low sensitivity within Geneious Prime® 2022.2.1. Expression values (FPKM values) were calculated using standard settings and comparison of expression levels were performed using DeSeq2 within Geneious Prime® 2022.2.1 using default settings. Genes with *p*-values <0.01 and a fold change of at least two times ($\log2FC \geq 1$ or ≤ -1) were considered to be differentially expressed (DE).

Analysis of ACAM34_UNSW escape mutants

For the isolation of ACAM34_UNSW HRTV-DL1 escape mutants we established long-term infected cultures as described above (*Virus infectivity and kinetics*), in triplicates. Cells were only infected with an MOI of ~1 to allow a recovery of infected cultures. After lysis of ACAM34_UNSW the cultures were maintained to allow the growth of HRTV-DL1 resistant clones. Cells from the recovered cultures were plated on agar plates to obtain single colonies. HRTV-DL1 infection was determined by PCR as described above. HRTV-DL1 life cycle in escape mutants and adsorption of HRTV-DL1 to escape mutants was determined as described above. Isolation of genomic DNA and genome sequencing is described above. For genomic analysis of escape mutants, reads were mapped against ACAM34_UNSW and HRTV-DL1 using 'geneious mapper' with medium-low sensitivity and default settings and variants were called using the Geneious function 'Find Variations/SNPs' with default settings in Geneious Prime® 2022.2.1. Variants with a coverage below 150 and a variant frequency below 75% were excluded from the analysis.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at: https://www.ncbi.nlm.nih.gov/, OP630574, https://www.ebi.ac.uk/ena, PRJEB56294, and https://www.ncbi.nlm.nih.gov/, PRJNA887929.

Author contributions

CM: Investigation, Methodology, Writing – review & editing, Data curation. DT: Investigation, Writing – review & editing. LZ: Investigation, Data curation, Formal analysis, Methodology, Writing – review & editing. MR: Data curation, Writing – review & editing, Funding acquisition, Supervision. SE: Funding acquisition, Supervision, Writing – review & editing, Conceptualization, Investigation, Methodology, Visualization, Writing – original draft.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This work was supported by the Australian Research Council [DP170102576] Max Planck Society [Max Planck Research Group Archaeal Virology].

Acknowledgments

We thank Ricardo Cavicchioli for valuable discussions and help with the acquisition of funding. We thank the Max Planck-Genome-Center Cologne (Cologne, Germany) for assistance with DNA and RNA sequencing. We thank Ingrid Kunze (MPI for Marine Microbiology, Bremen, Germany) for assistance with some of the experiments. We thank Tomas Alarcón-Schumacher and Bernhard Tschitschko for help with some of the figures. We thank the entire Archaeal Virology group at the Max-Planck-Institute for Marine Microbiology and Friedhelm Pfeiffer for critically reviewing the manuscript. We also acknowledge Friedhelm Pfeiffer and Mike Dyall-Smith for identifying the PATE and helpful discussions. Mass spectrometry results were obtained at the Bioanalytical Mass Spectrometry Facility (BMSF) within the Analytical Center of the University of New South Wales. Subsidized access to the BMSF is gratefully acknowledged. Finally, we want to thank the Max-Planck-Institute for Marine Microbiology and the Max-Planck-Society for continuous support.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2023.1274068/full#supplementary-material

References

Abdul Halim, M. F., Pfeiffer, F., Zou, J., Frisch, A., Haft, D., Wu, S., et al. (2013). *Haloferax volcanii* archaeosortase is required for motility, mating, and C-terminal processing of the S-layer glycoprotein. *Mol. Microbiol.* 88, 1164–1175. doi: 10.1111/mmi.12248

Achtman, M., Kennedy, N., and Skurray, R. (1977). Cell--cell interactions in conjugating *Escherichia coli*: role of traT protein in surface exclusion. *Proc. Natl. Acad. Sci. U. S. A.* 74, 5104–5108. doi: 10.1073/pnas.74.11.5104

Alarcón-Schumacher, T., Naor, A., Gophna, U., and Erdmann, S. (2022). Isolation of a virus causing a chronic infection in the archaeal model organism *Haloferax volcanii* reveals antiviral activities of a provirus. *Proc. Natl. Acad. Sci. U. S. A.* 119:e2205037119. doi: 10.1073/pnas.2205037119

Antipov, D., Raiko, M., Lapidus, A., and Pevzner, P. A. (2020). MetaviralSPAdes: assembly of viruses from metagenomic data. *Bioinformatics* 36, 4126–4129. doi: 10.1093/bioinformatics/btaa490

Atanasova, N. S., Roine, E., Oren, A., Bamford, D. H., and Oksanen, H. M. (2012). Global network of specific virus–host interactions in hypersaline environments. *Environ. Microbiol.* 14, 426–440. doi: 10.1111/j.1462-2920.2011.02603.x

Babakhani, S., and Oloomi, M. (2018). Transposons: the agents of antibiotic resistance in bacteria. *J. Basic Microbiol.* 58, 905–917. doi: 10.1002/jobm.201800204

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., et al. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* 19, 455–477. doi: 10.1089/cmb.2012.0021

Bernheim, A., and Sorek, R. (2020). The pan-immune system of bacteria: antiviral defense as a community resource. *Nat. Rev. Microbiol.* 18, 113–119. doi: 10.1038/s41579-019-0278-2

Blum, M., Chang, H. Y., Chuguransky, S., Grego, T., Kandasaamy, S., Mitchell, A., et al. (2020). The InterPro protein families and domains database: 20 years on. *Nucleic Acids Res.* 49, D344–D354. doi: 10.1093/nar/gkaa977

Broecker, F., and Moelling, K. (2019). Evolution of immune systems from viruses and transposable elements. *Front. Microbiol.* 10:51. doi: 10.3389/fmicb.2019.00051

Chen, Z., Liu, Y., Wang, Y., du, X., Deng, X., Xiang, J., et al. (2023). A virus-borne DNA damage signaling pathway controls the lysogeny-induction switch in a group of temperate pleolipoviruses. *Nucleic Acids Res.* 51, 3270–3287. doi: 10.1093/nar/gkad125

Cutler, R. G., and Evans, J. E. (1966). Synchronization of bacteria by a stationary-phase method. *J. Bacteriol.* 91, 469–476. doi: 10.1128/jb.91.2.469-476.1966

DasSarma, S., Capes, M., and DasSarma, P. (2009) in *Haloarchaeal Megaplasmids, in microbial Megaplasmids*. ed. E. Schwartz (Berlin, Heidelberg: Springer Berlin Heidelberg), 3–30.

DeMaere, M. Z., Williams, T. J., Allen, M. A., Brown, M. V., Gibson, J. A. E., Rich, J., et al. (2013). High level of intergenera gene exchange shapes the evolution of haloarchaea in an isolated Antarctic lake. *Proc. Natl. Acad. Sci. U. S. A.* 110, 16939–16944. doi: 10.1073/pnas.1307090110

Doron, S., Melamed, S., Ofir, G., Leavitt, A., Lopatina, A., Keren, M., et al. (2018). Systematic discovery of antiphage defense systems in the microbial pangenome. *Science* 359:eaar4120. doi: 10.1126/science.aar4120

Duncan-Lowey, B., and Kranzusch, P. J. (2022). CBASS phage defense and evolution of antiviral nucleotide signaling. *Curr. Opin. Immunol.* 74, 156–163. doi: 10.1016/j.coi.2022.01.002

Dyall-Smith, M.. (2021). The Halohandbook: protocols for Halobacterial genetics. Available at: https://haloarchaea.com/wp-content/uploads/2018/10/Halohandbook_2009_v7.3mds.pdf

Erdmann, S., Tschitschko, B., Zhong, L., Raftery, M. J., and Cavicchioli, R. (2017). A plasmid from an Antarctic haloarchaeon uses specialized membrane vesicles to disseminate and infect plasmid-free cells. *Nat. Microbiol.* 2, 1446–1455. doi: 10.1038/s41564-017-0009-2

Feng, L., Wang, B., Driscoll, B., and Jong, A. (2000). Identification and characterization of *Saccharomyces cerevisiae* Cdc6 DNA-binding properties. *Mol. Biol. Cell* 11, 1673–1685. doi: 10.1091/mbc.11.5.1673

Franzmann, P. D., Stackebrandt, E., Sanderson, K., Volkman, J. K., Cameron, D. E., Stevenson, P. L., et al. (1988). *Halobacterium lacusprofundi* sp. nov., a halophilic bacterium isolated from deep Lake, Antarctica. *Syst. Appl. Microbiol.* 11, 20–27. doi: 10.1016/S0723-2020(88)80044-4

Fraser, J. S., Yu, Z., Maxwell, K. L., and Davidson, A. R. (2006). Ig-like domains on bacteriophages: a tale of promiscuity and deceit. *J. Mol. Biol.* 359, 496–507. doi: 10.1016/j.jmb.2006.03.043

Garneau, J. R., Depardieu, F., Fortier, L. C., Bikard, D., and Monot, M. (2017). PhageTerm: a tool for fast and accurate determination of phage termini and packaging mechanism using next-generation sequencing data. *Sci. Rep.* 7:8292. doi: 10.1038/s41598-017-07910-5

Gebhard, L. J., Duggin, I. G., and Erdmann, S. (2023). Improving the genetic system for *Halorubrum lacusprofundi* to allow in-frame deletions. *Front. Microbiol.* 14:1095621. doi: 10.3389/fmicb.2023.1095621

Gibson, J. A. E. (1999). The meromictic lakes and stratified marine basins of the Vestfold Hills, East Antarctica. *Antarctic Sci.* 11, 175–192. doi: 10.1017/S0954102099000243

Gibson, R. T., and Androphy, E. J. (2020). The SMC5/6 complex represses the replicative program of high-risk human papillomavirus type 31. *Pathogens* 9:786. doi: 10.3390/pathogens9100786

Gilchrist, C. L. M., and Chooi, Y. H. (2021). Clinker & clustermap.Js: automatic generation of gene cluster comparison figures. *Bioinformatics* 37, 2473–2475. doi: 10.1093/bioinformatics/btab007

Goldfarb, T., Sberro, H., Weinstock, E., Cohen, O., Doron, S., Charpak-Amikam, Y., et al. (2015). BREX is a novel phage resistance system widespread in microbial genomes. *EMBO J.* 34, 169–183. doi: 10.15252/embj.201489455

Gomes-Filho, J. V., Zaramela, L. S., Italiani, V. C. S., Baliga, N. S., Vêncio, R. Z. N., and Koide, T. (2015). Sense overlapping transcripts in IS1341-type transposase genes are functional non-coding RNAs in archaea. *RNA Biol.* 12, 490–500. doi: 10.1080/15476286.2015.1019998

Grissa, I., Vergnaud, G., and Pourcel, C. (2007). CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res.* 35, W52–W57. doi: 10.1093/nar/gkm360

Hall, J. P. J., Botelho, J., Cazares, A., and Baltrus, D. A. (2022). What makes a megaplasmid? *Philos. Trans. R Soc. Lond. B Biol. Sci.* 377:20200472. doi: 10.1098/rstb.2020.0472

Han, C., Zhang, D., Gui, C., Huang, L., Chang, S., Dong, L., et al. (2022). KSHV RTA antagonizes SMC5/6 complex-induced viral chromatin compaction by hijacking the ubiquitin-proteasome system. *PLoS Pathog.* 18:e1010744. doi: 10.1371/journal.ppat.1010744

Harrison, P. W., Lower, R. P. J., Kim, N. K. D., and Young, J. P. W. (2010). Introducing the bacterial 'chromid': not a chromosome, not a plasmid. *Trends Microbiol.* 18, 141–148. doi: 10.1016/j.tim.2009.12.010

Hawkins, M., Malla, S., Blythe, M. J., Nieduszynski, C. A., and Allers, T. (2013). Accelerated growth in the absence of DNA replication origins. *Nature* 503, 544–547. doi: 10.1038/nature12650

Howard-Varona, C., Lindback, M. M., Bastien, G. E., Solonenko, N., Zayed, A. A., Jang, H. B., et al. (2020). Phage-specific metabolic reprogramming of virocells. *ISME J.* 14, 881–895. doi: 10.1038/s41396-019-0580-z

Huang, S., Xu, X., Wang, G., Lu, G., Xie, W., Tao, W., et al. (2016). DNA replication initiator Cdc6 also regulates ribosomal DNA transcription initiation. *J. Cell Sci.* 129, 1429–1440. doi: 10.1242/jcs.178723

Kageyama, Y., Murayama, M., Onodera, T., Yamada, S., Fukada, H., Kudou, M., et al. (2009). Observation of the membrane binding activity and domain structure of gpV, which comprises the tail spike of bacteriophage P2. *Biochemistry* 48, 10129–10135. doi: 10.1021/bi900928n

Karlsson, F., Malmborg-Hager, A. C., Albrekt, A. S., and Borrebaeck, C. A. K. (2005). Genome-wide comparison of phage M13-infected vs. uninfected *Escherichia coli*. *Can. J. Microbiol.* 51, 29–35. doi: 10.1139/w04-113

Karvelis, T., Druteika, G., Bigelyte, G., Budre, K., Zedaveinyte, R., Silanskas, A., et al. (2021). Transposon-associated TnpB is a programmable RNA-guided DNA endonuclease. *Nature* 599, 692–696. doi: 10.1038/s41586-021-04058-1

Kolmogorov, M., Yuan, J., Lin, Y., and Pevzner, P. A. (2019). Assembly of long, error-prone reads using repeat graphs. *Nat. Biotechnol.* 37, 540–546. doi: 10.1038/s41587-019-0072-8

Koonin, E. V., Makarova, K. S., and Wolf, Y. I. (2017). Evolutionary genomics of defense Systems in Archaea and Bacteria. *Annu. Rev. Microbiol.* 71, 233–261. doi: 10.1146/annurev-micro-090816-093830

Krzywinski, M., Schein, J., Birol, İ., Connors, J., Gascoyne, R., Horsman, D., et al. (2009). Circos: an information aesthetic for comparative genomics. *Genome Res.* 19, 1639–1645. doi: 10.1101/gr.092759.109

Kukkaro, P., and Bamford, D. H. (2009). Virus-host interactions in environments with a wide range of ionic strengths. *Environ. Microbiol. Rep.* 1, 71–77. doi: 10.1111/j.1758-2229.2008.00007.x

Kundu, L. R., Kumata, Y., Kakusho, N., Watanabe, S., Furukohri, A., Waga, S., et al. (2010). Deregulated Cdc6 inhibits DNA replication and suppresses Cdc7-mediated phosphorylation of Mcm2-7 complex. *Nucleic Acids Res.* 38, 5409–5418. doi: 10.1093/nar/gkq262

Kurtz, S., Phillippy, A., Delcher, A. L., Smoot, M., Shumway, M., Antonescu, C., et al. (2004). Versatile and open software for comparing large genomes. *Genome Biol.* 5:R12. doi: 10.1186/gb-2004-5-2-r12

Li, M., Wang, R., Zhao, D., and Xiang, H. (2014). Adaptation of the *Haloarcula hispanica* CRISPR-Cas system to a purified virus strictly requires a priming process. *Nucleic Acids Res.* 42, 2483–2492. doi: 10.1093/nar/gkt1154

Liu, J., Cvirkaite-Krupovic, V., Baquero, D. P., Yang, Y., Zhang, Q., Shen, Y., et al. (2021). Virus-induced cell gigantism and asymmetric cell division in archaea. *Proc. Natl. Acad. Sci.* 118:e2022578118. doi: 10.1073/pnas.2022578118

Liu, Y., Demina, T. A., Roux, S., Aiewsakun, P., Kazlauskas, D., Simmonds, P., et al. (2021). Diversity, taxonomy, and evolution of archaeal viruses of the class Caudoviricetes. *PLoS Biol.* 19:e3001442. doi: 10.1371/journal.pbio.3001442

Liu, H. W., Roisné-Hamelin, F., Beckert, B., Li, Y., Myasnikov, A., and Gruber, S. (2022). DNA-measuring Wadjet SMC ATPases restrict smaller circular plasmids by DNA cleavage. *Mol. cell* 82, 4727–4740.e6. doi: 10.1016/j.molcel.2022.11.015

Luk, A. W., Williams, T., Erdmann, S., Papke, R., and Cavicchioli, R. (2014). Viruses of haloarchaea. *Life* 4, 681–715. doi: 10.3390/life4040681

Maier, L. K., Stachler, A. E., Brendel, J., Stoll, B., Fischer, S., Haas, K. A., et al. (2019). The nuts and bolts of the Haloferax CRISPR-Cas system I-B. *RNA Biol.* 16, 469–480. doi: 10.1080/15476286.2018.1460994

Millman, A., Melamed, S., Leavitt, A., Doron, S., Bernheim, A., Hör, J., et al. (2022). An expanded arsenal of immune systems that protect bacteria from phages. *Cell Host Microbe* 30, 1556–1569.e5. doi: 10.1016/j.chom.2022.09.017

Nuttall, S. D., and Dyall-Smith, M. L. (1993). HF1 and HF2: novel bacteriophages of halophilic archaea. *Virology* 197, 678–684. doi: 10.1006/viro.1993.1643

Obbard, D. J., Gordon, K. H. J., Buck, A. H., and Jiggins, F. M. (2009). The evolution of RNAi as a defence against viruses and transposable elements. *Philos. Trans. R. Soc. Lond. Ser. B Biol. Sci.* 364, 99–115. doi: 10.1098/rstb.2008.0168

Pagaling, E., Haigh, R. D., Grant, W. D., Cowan, D. A., Jones, B. E., Ma, Y., et al. (2007). Sequence analysis of an archaeal virus isolated from a hypersaline lake in Inner Mongolia, China. *BMC Genom.* 8:410. doi: 10.1186/1471-2164-8-410

Payne, L. J., Meaden, S., Mestre, M. R., Palmer, C., Toro, N., Fineran, P. C., et al. (2022). PADLOC: a web server for the identification of antiviral defence systems in microbial genomes. *Nucleic Acids Res.* 50, W541–W550. doi: 10.1093/nar/gkac400

Pell, L. G., Gasmi-Seabrook, G. M. C., Morais, M., Neudecker, P., Kanelis, V., Bona, D., et al. (2010). The solution structure of the C-terminal Ig-like domain of the bacteriophage λ tail tube protein. *J. Mol. Biol.* 403, 468–479. doi: 10.1016/j.jmb.2010.08.044

Pfeiffer, F., Schuster, S. C., Broicher, A., Falb, M., Palm, P., Rodewald, K., et al. (2008). Evolution in the laboratory: the genome of *Halobacterium salinarum* strain R1 compared to that of strain NRC-1. *Genomics* 91, 335–346. doi: 10.1016/j.ygeno.2008.01.001

Pietilä, M. K., Laurinmäki, P., Russell, D. A., Ko, C. C., Jacobs-Sera, D., Butcher, S. J., et al. (2013a). Insights into head-tailed viruses infecting extremely halophilic Archaea. *J. Virol.* 87, 3248–3260. doi: 10.1128/JVI.03397-12

Pietilä, M. K., Laurinmäki, P., Russell, D. A., Ko, C. C., Jacobs-Sera, D., Hendrix, R. W., et al. (2013b). Structure of the archaeal head-tailed virus HSTV-1 completes the HK97 fold story. *Proc. Natl. Acad. Sci. U. S. A.* 110, 10604–10609. doi: 10.1073/pnas.1303047110

Porter, K., Kukkaro, P., Bamford, J. K. H., Bath, C., Kivelä, H. M., Dyall-Smith, M. L., et al. (2005). SH1: a novel, spherical halovirus isolated from an Australian hypersaline lake. *Virology* 335, 22–33. doi: 10.1016/j.virol.2005.01.043

Schnabel, H., Zillig, W., Pfäffle, M., Schnabel, R., Michel, H., and Delius, H. (1982). *Halobacterium halobium* phage øH. *EMBO J.* 1, 87–92. doi: 10.1002/j.1460-2075.1982.tb01129.x

Schwarzer, S., Hackl, T., Oksanen, H. M., and Quax, T. E. F. (2023). Archaeal host cell recognition and viral binding of HFTV1 to its Haloferax host. *MBio* 14:e0183322. doi: 10.1128/mbio.01833-22

Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30, 2068–2069. doi: 10.1093/bioinformatics/btu153

Siguier, P., Perochon, J., Lestrade, L., Mahillon, J., and Chandler, M. (2006). ISfinder: the reference Centre for bacterial insertion sequences. *Nucleic Acids Res.* 34, D32–D36. doi: 10.1093/nar/gkj014

Söding, J. (2004). Protein homology detection by HMM–HMM comparison. *Bioinformatics* 21, 951–960. doi: 10.1093/bioinformatics/bti125

Takahashi, S., Tomita, J., Nishioka, K., Hisada, T., and Nishijima, M. (2014). Development of a prokaryotic universal primer for simultaneous analysis of Bacteria and Archaea using next-generation sequencing. *PLoS One* 9:e105592. doi: 10.1371/journal.pone.0105592

Tschitschko, B., Erdmann, S., DeMaere, M. Z., Roux, S., Panwar, P., Allen, M. A., et al. (2018). Genomic variation and biogeography of Antarctic haloarchaea. *Microbiome* 6:113. doi: 10.1186/s40168-018-0495-3

Tschitschko, B., Williams, T. J., Allen, M. A., Páez-Espino, D., Kyrpides, N., Zhong, L., et al. (2015). Antarctic archaea–virus interactions: metaproteome-led analysis of invasion, evasion and adaptation. *ISME J.* 9, 2094–2107. doi: 10.1038/ismej.2015.110

von Itzstein, M. (2007). The war against influenza: discovery and development of sialidase inhibitors. *Nat. Rev. Drug Discov.* 6, 967–974. doi: 10.1038/nrd2400

Wang, H., Peng, N., Shah, S. A., Huang, L., and She, Q. (2015). Archaeal extrachromosomal genetic elements. *Microbiol. Mol. Biol. Rev.* 79, 117–152. doi: 10.1128/MMBR.00042-14

Willkomm, S., Makarova, K. S., and Grohmann, D. (2018). DNA silencing by prokaryotic Argonaute proteins adds a new layer of defense against invading nucleic acids. *FEMS Microbiol. Rev.* 42, 376–387. doi: 10.1093/femsre/fuy010

Xiong, L., Liu, S., Chen, S., Xiao, Y., Zhu, B., Gao, Y., et al. (2019). A new type of DNA phosphorothioation-based antiviral system in archaea. *Nat. Commun.* 10:1688. doi: 10.1038/s41467-019-09390-9

Zander, A., Willkomm, S., Ofer, S., van Wolferen, M., Egert, L., Buchmeier, S., et al. (2017). Guide-independent DNA cleavage by archaeal Argonaute from *Methanocaldococcus jannaschii*. *Nat. Microbiol.* 2:17034. doi: 10.1038/nmicrobiol.2017.34

frontiers | Frontiers in Microbiology

TYPE Original Research
PUBLISHED 24 November 2023
DOI 10.3389/fmicb.2023.1297600

OPEN ACCESS

EDITED BY
Zhengshuang Hua,
University of Science and Technology of China, China

REVIEWED BY
Yinzhao Wang,
Shanghai Jiao Tong University, China
Nicole Jane Bale,
Royal Netherlands Institute for Sea Research (NIOZ), Netherlands

*CORRESPONDENCE
Wenjie Xiao
✉ wjxiaocug@126.com
Fengfeng Zheng
✉ zhengff@sustech.edu.cn

RECEIVED 20 September 2023
ACCEPTED 08 November 2023
PUBLISHED 24 November 2023

CITATION
Yao W, Zhang W, He W, Xiao W, Chen Y, Zhu Y, Zheng F and Zhang C (2023) Lipidomic chemotaxonomy aligned with phylogeny of Halobacteria.
Front. Microbiol. 14:1297600.
doi: 10.3389/fmicb.2023.1297600

Lipidomic chemotaxonomy aligned with phylogeny of Halobacteria

Wenyong Yao[1], Wan Zhang[1,2], Wei He[1], Wenjie Xiao[1,3]*, Yufei Chen[1], Yuanqing Zhu[1,4], Fengfeng Zheng[1]* and Chuanlun Zhang[1,4]

[1]Shenzhen Key Laboratory of Marine Archaea Geo-Omics, Department of Ocean Science and Engineering, Southern University of Science and Technology, Shenzhen, China, [2]GEOMAR Centre for Marine Biotechnology (GEOMAR-Biotech), Research Unit Marine Natural Products Chemistry, GEOMAR Helmholtz Centre for Ocean Research Kiel, Kiel, Germany, [3]Department of Biology, Hadal, Nordcee & DIAS, University of Southern Denmark, Odense, Denmark, [4]Shanghai Sheshan National Geophysical Observatory, Shanghai, China

Archaea play an important role in global biogeochemical cycles and are considered ancestral to eukaryotes. The unique lipid composition of archaea, characterized by isoprenoid alkyl chains and ether linkage to glycerol-1-phosphate, offers valuable insights into archaeal phylogeny and evolution. However, comprehensive studies focusing on archaeal lipidomes, especially at the intact polar lipid level, are currently limited. Here, we built an in-house library of archaeal lipids by using high-performance liquid chromatography coupled with mass-spectrometry, which was integrated with bioinformatics and molecular network analyses. Seven halobacterial strains, representing three distinct orders, were cultured under identical conditions to investigate their lipidomes. A total of 162 features were identified, corresponding to 107 lipids that could be assigned to different strains. Clustering analyses of both core lipids and total lipids matched the phylogeny of Halobacteria at the order level. Notably, lipids such as triglycosyl diether-phosphatidyl acid and bis-sulfate glycosyl lipids were specific to particular groups and could serve as diagnostic intact lipid biomarkers for Halobacteria. Furthermore, the analysis of network-coordinated features facilitated the linkage of unknown lipid compounds to phylogeny, which promotes a lipidome to phylogeny matchup among three *Haloferax* strains, thereby expanding the knowledge of the halobacterial lipidome. Our study provides a comprehensive view of the lipidomes of the seven strains of Halobacteria and highlights the potential of lipidomics for studying archaeal phylogeny.

KEYWORDS

Halobacteria, lipidomics, phylogeny, chemotaxonomy, network-coordinated features, bioinformatics

1 Introduction

The identification of archaea has widely broadened our knowledge of life forms on Earth and motived the proposal of three domain theory (Woese et al., 1990). Archaea are ubiquitously found in a wide variety of geological settings, from extreme conditions such as salt lakes, hot springs, and hydrothermal vents, to more commonplace settings such as soils, lakes, rivers, marine water column and sediments (Baker et al., 2020). Although only a limited number of archaeal taxa have been isolated and cultured, advances in sequence-based

approaches such as 16S rRNA and metagenomic sequencing have recovered a high phylogenetic diversity and versatile metabolic potentials of archaea, suggesting they might play an important role in multiple elemental cycling processes (Baker et al., 2020). Recently, the newly discovered Asgard archaea showed a close phylogenetic relationship with eukaryotes, and their genomes encode an expanded repertoire of eukaryotic signature proteins (Spang et al., 2015; Baker et al., 2020), pointing to an archaeal origin of eukaryotes (Zaremba-Niedzwiedzka et al., 2017; MacLeod et al., 2019). These new evidences largely challenged the three domains theory, highlighting the need for further research on archaeal phylogeny and evolution.

Archaea are characterized for their cell membrane lipids consisting of isoprenoid alkyl chains that are ether-linked to glycerol-1-phosphate, whereas bacterial and eukaryotic membranes are mostly comprised of fatty acids that are ester-linked to glycerol-3-phosphate (Valentine, 2007; Koga, 2011). The membrane of archaea is composed of dialkyl glycerol diethers (e.g., archaeol) or glycerol dialkyl glycerol tetraethers (e.g., GDGTs) or both. The structure of ether bound lipids in archaea is more stable than bacterial and eukaryotic counterparts, which is considered as one of the key strategies for archaea to adapt to extreme conditions and prolongs the preservation of fossil archaeal biomarkers (Summons et al., 2022). Compared with DNA, RNA and proteins, lipids do not contain phylogenetic signals, but are suitable biomarkers with longer lifespan and diverse structures to indicate ecological occurrence and activity of archaea (Briggs and Summons, 2014). The distribution of lipids is suggested to be regulated by both genetic constraints and physiological modification responding to environmental stresses. Therefore, archaeal lipids are widely used as biomarkers to reconstruct paleoenvironment and ancient biological activities (reviewed by Schouten et al., 2013). They also serve as useful chemotaxonomic biomarkers as each taxon may produce a distinct lipid profile (e.g., Elling et al., 2017).

Halobacteria (also known as Haloarchaea) is a subgroup of *Euryarchaeota*, consisting of three main orders, namely *Halobacteriales*, *Haloferacales*, and *Natrialbales* (Gupta et al., 2015, 2016). These microorganisms can resist saline stress and survive in diverse hypersaline environments such as salt lakes and salt marshes (Oren, 2014). The membrane lipids of Halobacteria primarily consist of bilayer diethers, with the absence of monolayer tetraether lipids (Villanueva et al., 2014; Bale et al., 2019). The structural diversity of halobacterial archaeol cores can be extended by replacing a C_{20} phytanyl chain with an elongated C_{25} chain and/or by forming several double bonds (uns) on the phytanyl chain (Dawson et al., 2012; Bale et al., 2019; Vandier et al., 2021). This structural flexibility enables them to adapt and survive in hypersaline environments (Dawson et al., 2012). The diether core structures in Halobacteria can be further linked to a phospho- and/or glycosyl-head group, forming intact polar lipids (IPLs). Phosphatidyl glycerol (PG) and methylated phosphatidyl glycerol phosphate (Me-PGP) are the dominant IPL components identified in Halobacteria (Kamekura and Kates, 1999; Sprott et al., 2004; Kellermann et al., 2016). PG and Me-PGP are negatively charged phospholipids that can interact with cations, which helps the organisms to resist osmotic shock in hypersaline environments (Tenchov et al., 2006). The halobacterial lipids also include the cardiolipin analogs (archaeal cardiolipin) that have two core structure of diethers sharing one polar head group (Sturt et al., 2004; Corcelli, 2009). Menaquinone (MK, also known as vitamin K2) that resides inside the bilayer as an electron carrier in the respiratory chains (Kellermann et al., 2016), and carotenoid (mainly bacterioruberin) which is responsible for the red appearance of Halobacteria (Hartmann et al., 1980), are important non-bilayer forming lipids in the Halobacteria.

Hitherto, hundreds of diverse lipid components have been identified in Halobacteria and other archaeal categories using a variety of analytical techniques, such as NMR, GC-MS, LC-MS (Bale et al., 2019; Law and Zhang, 2019). Recent advances in lipidomics with high-resolution mass spectrometry (HRMS) have significantly improved the detection coverage, sensitivity, and throughput. This enables the analysis of a larger number of lipids in a single run (Tsugawa et al., 2020), which have continuously enhanced our knowledge of archaeal lipidome (Knappy et al., 2009; Law et al., 2021). Remarkably, the structural diversity and chemotaxonomic potential have been carefully examined with HRMS-based lipidomic approach for Thaumarchaeota (Elling et al., 2017), halo(alkali)philic methanogens and Halobacteria (Bale et al., 2019). However, the data processing of HRMS is still time-consuming and the identifications largely relied on manual interpretation, which may substantially affect the efficiency and accuracy of archaeal lipids analysis.

The development of bioinformatic tools has facilitated the analysis of the expanded dataset, which can provide reliable lipid identification by integrating multidimensional mass spectrometry information, for instance, MS^1, retention time (RT) and the MS^2 spectra (Tsugawa et al., 2020; Schmid et al., 2021). Bioinformatic approaches such as feature-based molecular networking have been widely used in HRMS-based lipidomic or metabolic research, allowing further mass spectra interpretation by linking the features with similar MS^2 spectra to generate a molecular network (Guthals et al., 2012; Watrous et al., 2012). This approach has increasingly been used in uncovering the lipidomes from environmental and pure culture samples, which enables the visualization of the relationships between different lipid features, the categorization of unknown features, and the structural prediction of unknown analogs (Ding et al., 2021, 2022).

In a specific environmental niche, different archaeal species will encounter the same environmental/growth conditions. This underscores the necessity of conducting a lipidome investigation under controlled conditions. In this study, we hypothesize that the lipid composition and abundance of halobacterial strains are in alignment with the phylogeny, which may manifest the underlying mechanisms governing phylogenetic control over lipid biosynthesis. We cultured seven strains from three prominent halobacterial orders under identical conditions, minimizing the potential confounding effects of environmental fluctuations. We employed and assessed the application of ultra-high performance liquid chromatography-high resolution mass spectrometry (UPLC-HRMS) coupled with bioinformatics tools. This encompassed an *in silico* spectral library-based lipid identification and a feature-based molecular network approach. The streamlined procedure of lipidomic analysis using our in house library allowed us to comprehensively characterize the lipidomic composition of archaea and evaluate their chemotaxonomic potential. This study expands our understanding of the archaeal lipidome, which may provide a new insight for exploring the relevance between lipidome and phylogeny.

2 Materials and methods

2.1 Cultivation of halobacterial strains

Seven pure cultured strains of Halobacteria were used in this study, including *Haloarcula* (*Ha.*) *argentinensis*; *Haloferax* (*Hf.*) *larsenii*; *Haloferax mediterranei*; *Haloferax volcanii*; *Haloterrigena* (*Ht.*) *turkmenica*; *Natrialba asiatica (Na.)* and *Natrinema (Nt.) gari* (Table 1). *Hf. larsenii*, *Hf. mediterranei* and *Hf. volcanii* are from the order of *Haloferacales*; *Nn. gari*, *Ht. turkmenica* and *Na. asiatica* are from the order of *Natrialbales*, with *Na. asiatica* being distantly related to *Ht. turkmenica* and *Nn. gari*. *Ha. argentinensis* is the only species in the order *Halobacteriales* (Figure 1). All strains were grown at 37°C and 200 rpm in triplicate using 300 mL of DSMZ 589 liquid medium, which was originally designed for cultivating *Halobacterium lacusprofundii*. The medium contained NaCl 180 g/L; $MgCl_2$•$6H_2O$ 75 g/L; $MgSO_4$•$7H_2O$ 7.4 g/L; KCl 7.4 g/L; $CaCl_2$•$2H_2O$ 1 g/L; yeast extract 1 g/L; sodium succinate 10 g/L; vitamin solution 1 mL/L. The growth pH was adjusted to 7.4. The vitamin solution was a mixture of biotin 0.1 g/L; vitamin B_{12} 0.1 g/L; and thiamine-HCl•$2H_2O$ 0.1 g/L. Cell growths were monitored at a wavelength of 600 nm (OD600) and harvested at the stationary phase. 45 mL of the cultures were harvested by centrifuging at 10000 g for 12 min; the cell pellets were collected, freezing-dried, and stored at −80°C for lipid extraction.

2.2 Lipid extraction

The cell pellets were lyophilized and extracted using a modified Bligh and Dyer method (Findlay, 2004). Briefly, the samples were resuspended with 3.2 mL of 50 mM phosphate buffer (PB, dissolve 8.7 g K_2HPO_4 in 1 L distilled H_2O, pH was adjusted to 7.4 using HCl) and transferred to Teflon centrifuge tubes (Thermo Scientific). The solvent mixture was adjusted to a ratio of 2:1:0.8 (methanol (MeOH): dichloromethane (DCM): PB, v/v/v) by adding 8 mL methanol, and 4 mL dichloromethane. 1-O-hexadecyl-2-acetyl-sn-glycerol-3-phosphocholine (C_{16}-PAF, 0.1 μg/μL) and 1,2-dinonadecanoyl-sn-glycerol-3-phosphocholine (C_{19}-PC, 50 pmol/μL) were spiked as internal standards. The tubes were vortexed and then ultrasonicated for 2 min (twice with 1-min intervals). Subsequently, the samples were incubated in dark at the 4°C overnight (Findlay, 2004). The solvent ratio was adjusted to 1:1:0.9 (MeOH: DCM: H_2O, v/v/v) by adding 4 mL ddH_2O and 4 mL DCM. The lipid extracts were left to stand briefly for phase separation. The DCM layer was collected, and the remaining aqueous phase was re-extracted twice with 5 mL of DCM. Samples from all DCM layer collection were pooled as total lipid extracts (TLEs), which were dried under a gentle nitrogen flow and stored at −80°C until lipid analysis.

2.3 Lipid analyses

The TLEs were dissolved in 1 mL MeOH and 10 μL of each sample was pooled as a QC sample for peak alignment. An aliquot of the TLEs was analyzed on a Waters ACQUITY I-class ultra-performance liquid chromatography (UPLC) coupled to Waters SYNAPT G2-S*i* quadrupole time-of-flight mass spectrometer (qTOF); electrospray ionization (ESI) in positive mode was equipped as the ion source.

Separation of archaeal lipids was achieved on a Waters Acquity BEH C_{18} column maintained at 45°C (Wörmer et al., 2013; Kellermann et al., 2016). The samples were kept at 7°C and the flow rate was kept at 0.3 mL/min during the run. Solvent A was 85% MeOH (Optima™ LC/MS Grade, Fisher Chemical) with 15% H_2O (Optima™ LC/MS Grade, Fisher Chemical) and solvent B was 50% isopropanol (Optima™ LC/MS Grade, Fisher Chemical) with 50% MeOH, both amended with 0.1% NH_4OH (25–30% NH_3 basis, Sigma-Aldrich) and 0.04% formic acid (>99.0%, Optima™ LC/MS Grade, Fisher Chemical). The lipids were eluted with the following gradient: 100% A was maintained for 2.5 min. B was linearly increased to 15% at 3 min, 85% at 27 min, and 100% at 28 min and kept the fraction until 37 min. The flow was reversed to 100% A at 38 min and re-equilibrated for 7 min.

The MS setting was identical to Chen et al. (2022): capillary 2.5 kV, source temperature 120°C, sampling cone 45, source offset 80, desolvation gas flow 800 L/h at 350°C, cone gas flow 50 L/h, nebulizer gas flow 6.5 bar. The TOF analyzer was operated at Resolution mode and the data acquisition mode was FAST-DDA. The scan time was 0.2 s and the mass range for MS^1 and MS^2 was m/z 100–2,000 and m/z 50–2,000, respectively. Ions with the top 5 intensity and >20,000 total ion chromatogram counts were fragmented by collision-induced dissolution (CID) to obtain product ion spectra. For lower-intensity ions, a real-time dynamic exclusion was enabled. Ramped collision energy from 10 V to 55 V for low mass and 15 V to 65 V for high mass was used to obtain the MS^2 spectra. The MS was calibrated before the analysis using sodium iodide solution (m/z 50–2,000; residual mass error < 0.5 ppm). A leucine enkephalin solution (m/z = 556.2771 for $[M+H]^+$) was used for real-time calibration (scan time 0.2 s, 20 s interval).

TABLE 1 Strains used in this study and their optimum growth conditions.

Species	Accession No.	Temperature °C	NaCl M	pH	GC %	References
Haloarcula argentinensis	DSM 12282	40	2.5–3.0	7.5	62.0	Ihara et al. (1997)
Haloferax larsenii	JCM 13917	43–45	2.2–3.4	6.5–7.0	62.2	Xu et al. (2007)
Haloferax mediterranei	ATCC 33500	35–37	3.4	6.5	60.0	Rodriguez-Valera et al. (1983)
Haloferax volcanii	DSM 3757	45	1.7–2.5	7.0	63.5	Mullakhanbhai and Larsen (1975)
Haloterrigena turkmenica	DSM 5511	45–51	2.6–3.4	7.4	59.8	Minegishi and Kamekura (2009) and Zvyagintseva and Tarasov (1988)
Natrialba asiatica	DSM 12278	30–40	4.0	6.6–7.0	62.3	Hezayen et al. (2001)
Natrinema gari	JCM 14663	37–40	2.6–3.4	6.0–6.5	65.4	Tapingkae et al. (2008)

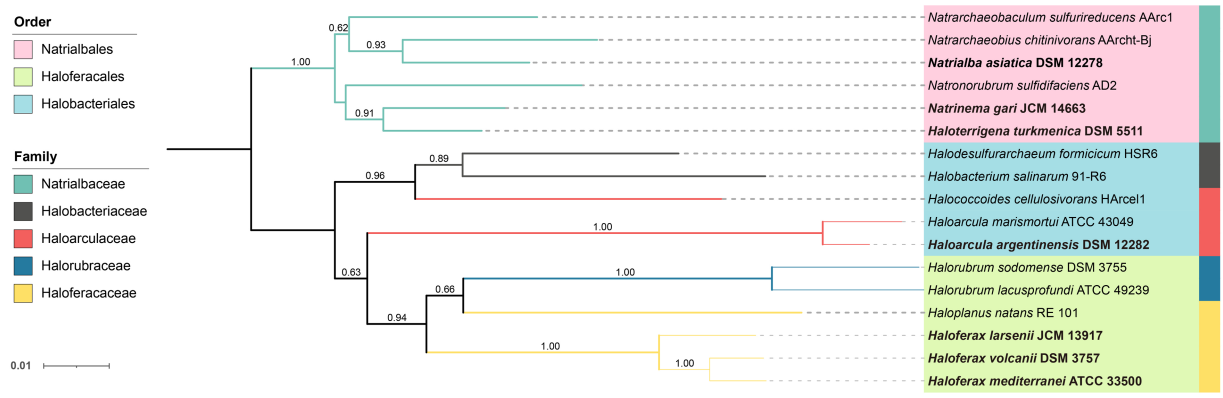


FIGURE 1
Neighbor-joining phylogenetic tree of Halobacteria based on the 16S rRNA gene. The strains studied are depicted in bold font. 16S rRNA genes of other stains were obtained from NCBI database. The tree was bootstrapped 1,000 times and values above 0.5 are marked. The branch length represents the evolutionary distance between the taxa.

The continuum raw data was centroid using MassLynx v 4.1 software and then transformed into mzML format using Proteowizard v 3.0.20247. The mzML files were then imported into MS-DIAL v4.9 software for further lipidomic analysis (Tsugawa et al., 2020). The retention time (RT) drifts between the cohort samples were correlated using internal standards with the RT deviations <0.8 min. Peak detection and integration were performed in MS-DIAL v4.9 with a minimum peak height of 3,000 and mass slice width of 0.1 Da. The detected peaks were aligned against a QC sample as reference with RT tolerance 0.7 min and MS^1 tolerance 0.05 Da. The mass features detected in samples lower than 10 folds of that present in blanks were filtered.

An in-house *in silico* archaeal lipid MS^2 library was used for feature annotation and lipid identification in MS-DIAL v4.9. The *in silico* library was made based on and modified from the LipidBlast Templates (Kind et al., 2014), with fragmentation rules obtained from the published MS^2 data of archaeal lipids (Sprott, 1992; Yoshinaga et al., 2012; Elling et al., 2016; Bale et al., 2019; Kropp et al., 2022) and the data obtained under our equipment setting. Lipids were annotated with mass error within 0.01 Da for MS^1 and 0.05 Da for MS^2 and total scores over 70% matching to the library were considered as identified compounds. The detailed MS^2 of representative compounds were shown in Supplementary Figures S1–S3. The annotation results were further manually confirmed, and our results indicated that the total score of 70% as threshold was an optimized balance of the annotation coverage and accuracy.

The $[M+H]^+$ and $[M+NH_4]^+$ ions were included in the peak integration. Lipid abundance was semi-quantified by the internal standard 1,2-dinonadecanoyl-sn-glycerol-3-phosphocholine (Kellermann et al., 2016), while the response factors of different compounds were not determined due to the lack of commercial standards for most compounds. Data of logarithmic relative abundance of lipids were scaled by samples and then used for cluster analysis by pheatmap package v 1.0.12 (Kolde, 2012) in *R* v 4.21 (Team, 2013). Complete linkage method was used for cluster analysis, while single linkage, Ward. D linkage and average linkage methods were also tested to confirm the consistency of the cluster results (Supplementary Figure S7). The principal component analysis (PCA) and partial least squares discrimination analysis (PLS-DA) analysis were performed in MetaboAnalyst v5.0.[1] The data were normalized by sum and log transformed during the submission, and the default setting was applied for each analysis (Pang et al., 2022).

2.4 Feature-based molecular network construction

MS^2 spectral summary and feature quantification were exported from MS-DIAL v4.9 and then imported for the Feature-Based Molecular Networking analysis on Global Natural Products Social Molecular Networking (GNPS) following the detailed documentation (Wang et al., 2016; Nothias et al., 2020; Ding et al., 2021). Mass tolerance was set to 0.05 Da for precursor and MS^2 fragment ions. The MS^2 spectra were filtered by collecting the top 6 fragment ions within a 50 Da window and discarding fragments within 17 Da of the precursor. Spectrum pairs with a cosine score above 0.7 and sharing 6 peaks were considered linked pairs. Spectra were further annotated in the GNPS library with the same filter (Horai et al., 2010; Wang et al., 2016; Mohimani et al., 2018). Network visualization was performed in Cytoscape v 3.9.1 software (Shannon et al., 2003). Nodes involved in blank samples or with $m/z < 500$ (to collect features that are more likely to be lipids), and networks containing less than 4 nodes were filtered. After the filtration, only one annotated features of G-1,2-Didecanoyl PC ($m/z = 566.381$) was remained. SIRIUS v 5.6.2 was used for assistance in compound formula indication (Dührkop et al., 2019).

2.5 Phylogenetic analysis

An aliquot of the collected biomass was used for DNA analysis. DNA were extracted using Fast DNA SPIN KIT (MP Biomedicals, OH, United States) following the manufacturer's protocol. The extracted DNA were sent to Sangon Biotech for Sanger sequencing. Primer pair of Arch 21F: TTCCGGTTGATCCYGCCGGA and

1 https://metaboanalyst.ca

1492R: TACGGYTACCTTGTTACGACTT was used to acquire full-length 16S rRNA gene sequences (Oyama et al., 2018). SeqMan v7.1.0 was used for sequence assembly (Swindell and Plasterer, 1997). Ten extra 16S rRNA gene sequences were obtained from NCBI as reference sequences (Supplementary Table S1). All sequences were imported into MEGA 11 for further analysis (Tamura et al., 2021). The sequences were aligned using the built-in MUSCLE algorithm and the gaps were manually removed (Edgar, 2004). A neighborhood joining phylogenetic tree was constructed with 1,000 bootstrap tests (Saitou and Nei, 1987).

3 Results

3.1 Lipid identification of Halobacteria

A total of 192 peaks were annotated against an in-house *in silico* archaeal lipid database after peak alignment using MS-DIAL v4.9. The annotated peaks were manually curated to reduce false positive annotations and interference peaks from in-source fragmentation, finally resulting in 162 correctly annotated features. Regardless of different isomer and adduct forms, we finally had 107 unique lipids (Figure 2, Table 2). These lipids included a large range of lipid diversity, composing of 16 core structure of diethers, 30 phospholipids (2 phosphatidyl glycerolhexose, Gly-PG; 9 Me-PGP; 13 PG; 2 phosphatidyl glycerolsulphate, PGS; 1 phosphatidyl acid, PA; 2 phosphatidyl ethanolamine, PE; 1 tentative phosphatidyl inositol, $PI^{(a)}$), 27 glycolipids (8 sulphated/disulphated glycosyl aminohexanehexaol, S-Gly-AHH / 2S-Gly-AHH; 10 non-sulphated glycolipids; 9 sulphated-glycolipids), 22 archaeal cardiolipins (7 bisphosphateglycerol, BPG; 5 non-sulphated glycocardiolipins; 10 sulphated-glycocardiolipins), 8 quinones of MK, and 4 bacterioruberin (Figure 2). Representative MS^2 spectra and the matches with reference spectra of the *in silico* archaeal lipid library are provided in the Supplemental material (Supplementary Figures S1–S3, Supplementary Table S2). One feature ($m/z = 895.711$) was tentatively identified as PI, which showed characteristic ions of phospholipids but lacked a fragment ion resembling the polar head group ($m/z = 261.038$).

3.2 Comparative analysis of Halobacteria lipidome

3.2.1 Composition of diether core lipids

For a complete investigation of the core membrane lipid composition of Halobacteria, the detected core lipids were manually combined with the IPLs that share the same core lipids. The core lipid (including archaeol that are identified directly or involved in the IPLs) of the seven Halobacteria strains primarily consisted of diether core structures with two fully saturated C_{20} chains (AR_C_{20}:C_{20}, short for AR, 47.71–99.62%), while diether variant structures of unsaturated and extended chain length (C_{25}) were also identified in some strains as expected. The degree of unsaturation in archaeol ranged from 0 to 8, and the extended archaeol were composed of a C_{20} chain and a C_{25} chain (AR_C_{20}:C_{25}, short for Ext-AR), whereas di-extended archaeol of two C_{25} chains (AR_C_{25}:C_{25}, diExt-AR) was not detected in any halobacterial strains in this study.

The distribution of unsaturated diether lipids varied among the three orders of Halobacteria (Figure 3B). Particularly, *Haloferacales*, which comprises three strains belonging to the genus *Haloferax*, exhibited the highest fraction of unsaturated diether lipids (11.17–20.72%) compared to *Natrialbales* (2.27–5.01%) and *Ha. argentinensis* (the sole strain in the order *Halobacteriales*, 0.32%). *Hf. mediterranei* produced the most abundant unsaturated lipids (20.72%) in *Haloferax* with predominantly low degrees of unsaturation (unsaturation <4, 97.44%). Additionally, the unsaturated diether lipids synthesized by *Hf. volcanii* were primarily low-unsaturated (99.61%), while *Hf. larsenii* exhibited significantly higher production of polyunsaturated lipids compared to its counterparts (unsaturation ≥4, 21.27%). Among the *Natrialbales* order, *Na. asiatica* and *Nn. gari* were characterized by the highest production of polyunsaturated lipids (unsaturation ≥4, 33.93 and 44.43% respectively). Although less capable of highly unsaturated lipid production among *Natrialbales*, *Ht. turkmenica* still yielded more polyunsaturated lipids (9.05%) than *Hf. mediterranei* and *Hf. volcanii*.

The Ext-AR was detected abundantly in the *Natrialbales*, accounting for 13.44–26.16% of total archaeol chains. This corresponded to an Ext-AR content in the range of 26.88–52.33% of the total core lipids (Figure 3A). While traces of Ext-AR were also detected in other strains, its abundance was less than 0.3%. These results are consistent with previous studies emphasizing the predominant occurrence of Ext-AR in *Natrialbales* (Vandier et al., 2021). Remarkably, under the same culture conditions, the core lipid compositions exhibited significant variations among the seven strains. Cluster analysis of these lipids successfully differentiated the strains at the order level, indicating a strong connection between phylogeny and core lipid composition (Figure 3C). Notably, the strains with the highest optimal salt demand, *Hf. mediterranei* and *Na. asiatica,* produced the largest amounts of unsaturated lipids or Ext-AR, leading to independent branches of both strains in their respective groups (Figure 3C).

3.2.2 Composition of intact polar lipids and non-bilayer forming lipids

The intact polar lipids (phospholipids, cardiolipins and glycolipids) and non-bilayer forming lipids (quinone and bacterioruberin) were found in the strains but with varying fractions. Phospholipids comprised the largest fraction of the total lipids in most strains (45.70–69.17%, Figure 4A), except for *Hf. mediterranei* (32.76%). Among all strains, Me-PGP and PG were the predominant phospholipids (34.03–74.35% and 24.81–63.41%, respectively), which aligned with the observation of previous studies (Kellermann et al., 2016). PA, a biosynthetic intermediate for other phospholipids (Corcelli, 2009), was detected in all strains but only present in minor fractions (0.13–0.91%). The distribution of other phospholipids, such as PE, $PI^{(a)}$, and PGS, showed taxonomic specificity. PE was found exclusively in strains of *Haloferax* and *Natrialbales* (0.06–0.65%), while $PI^{(a)}$ was specific to *Haloferax* strains (0.03–0.12%). PGS was detected in *Ha. argentinensis*, *Hf. mediterranei* and *Nn. gari* with contents of 2.44, 0.14 and 3.54%, respectively (Figure 4B, Table 2).

The production of cardiolipins was higher in three *Haloferax* species and *Ha. argentinensis*. In *Natrialbales*, only a small amount of glycocardiolipins (sulfate diglycosyl diether-phosphatidyl acid, S-DGD-PA) was detected (0.44‰) in *Nn. gari*. Comparing to *Haloferax* (1.42% in *Hf. mediterranei*, 5.59% in *Hf. larsenii* and 7.61%

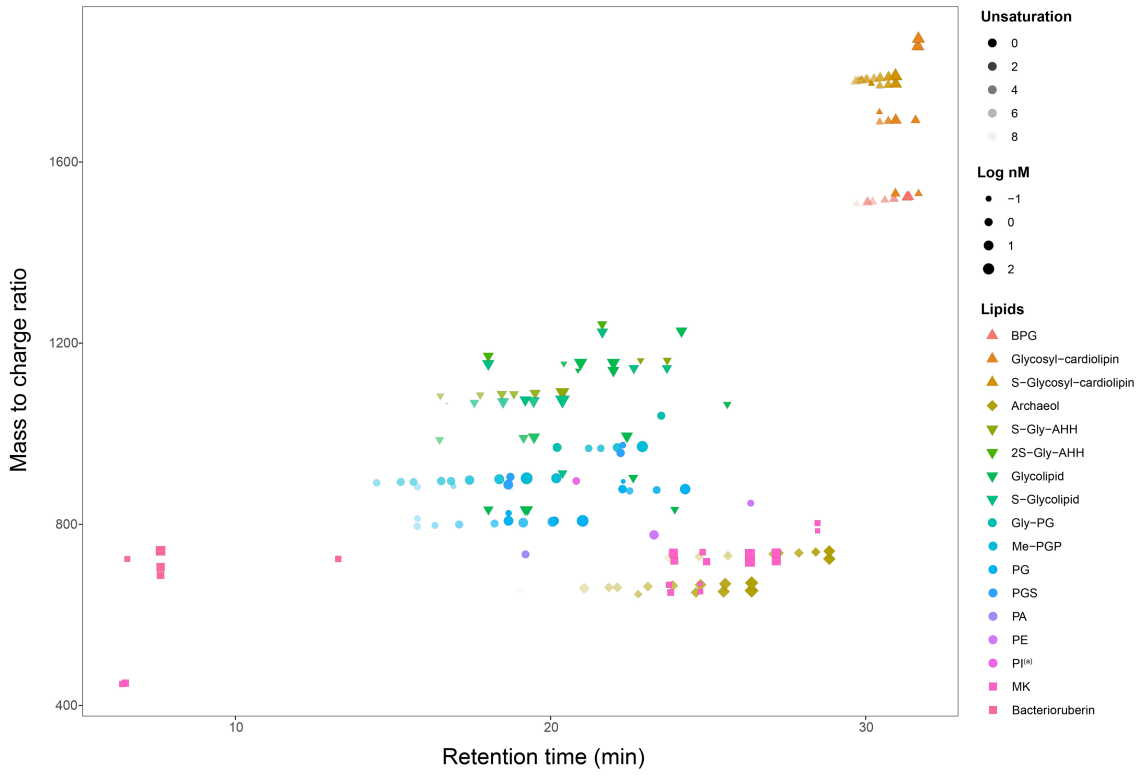


FIGURE 2
A map of all 162 lipids features (107 unique lipids in 17 classes, including AR, IPLs, cardiolipin, quinone, and bacterioruberin) from seven cultured Halobacteria. (a) Tentative identification as the MS^2 spectrum lack some of the characteristic fragments.

in *Hf. volcanii*), *Ha. argentinensis* exhibited a significantly higher cardiolipin content (11.60%). In all cases of stains producing cardiolipin, glycocardiolipins were predominant over BPGs (1.41–10.67% comparing to 0.02–0.61%, Figure 4A).

Glycolipids were found to be less abundant than phospholipids and made up 2.50–12.74% of the total lipids. The compositions of glycolipids also exhibited taxonomic specificity in these seven strains. The three *Haloferax* strains displayed similar glycolipid profiles, with sulfate diglycosyl diether (S-DGD) being the dominant glycolipid (69.30–77.18%), followed by S-Gly-AHH (17.53–24.70%), diglycosyl diether (DGD) (4.80–7.03%), and minor sulfate monoglycosyl diether (S-MGD) (0.21–0.43%, Figure 4C, Table 2). *Ht. turkmenica* and *Na. asiatica* also produced diglycosyl-based glycolipids as major components but were also capable of producing bis-sulfate glycosyl moiety, forming disulfate diglycosyl diether (2S-DGD) (28.10 and 56.29%) or 2S-Gly-AHH (5.82 and 15.59%). *Na. asiatica* showed a higher abundance of bis-sulfate glycosyl lipids than *Ht. turkmenica*. Both strains produced more monoglycosyl diether (MGD) compared to *Haloferax*. However, *Nn. gari*, despite belonging to the same order, displayed distinct glycolipids. It lacked the bis-sulfate glycosyl moiety, and S-Gly-AHH was only present in small amounts. Instead, triglycosyl diether (TGD, 64.83%) was the primary glycolipid in *Nn. gari*, along with MGD (17.12%), S-DGD (15.34%), and DGD (2.53%). In *Ha. argentinensis*, TGD was the predominant glycolipid, accounting for almost all the glycolipids detected (99.36%) (Figure 4C).

Respiratory quinones were also identified in all halobacterial strains. The dominating compounds were polyunsaturated MKs with 8 isoprenoid units ($MK_{8:8}$-$MK_{8:7}$). The MKs were abundant in all cultures with a maximal abundance being detected in *Hf. larsenii* (27.08%) and lowest in *Hf. mediterranei* (8.62%). Notably, bacterioruberin, a red carotenoid pigment, is a well-known biomarker for Halobacteria (Oren, 2002); however, our results showed this compound was not universally produced in Halobacteria. Five strains (except for *Hf. mediterranei* and *Na. asiatica*) showed high abundance of bacterioruberin, ranging from 1.20 to 8.62%. In contrast, *Hf. mediterranei* exhibited only a minor amount of bacterioruberin at 0.04% (Figure 4A, Table 2). Notably, bacterioruberin was entirely absent in the culture of *Na. asiatica*, which has been reported as a pigment-free species (Hezayen et al., 2001). Thereby, the detection of bacterioruberin was in line with the observation of a pink or red color exhibited in the cultures of the other five strains.

The PCA analysis using the grouped identified lipids (Supplementary Figure S8A) revealed that most strains could be differentiated based on the variation in lipid groups, except *Hf. larsenii* and *Hf. volcanii*, which exhibited similar patterns as expected (Figure 4). The three *Natrialbales* strains showed great divergence in the PCA plot compared to *Haloferax* strains. To investigate the relationship between lipid groups and DNA-based phylogeny, we further performed cluster analysis using lipid groups as elements. The results demonstrated a clear separation of three groups at the order level, with three *Haloferacales*, three *Natrialbales*, and the single *Halobacteriales* forming distinct clusters (Figure 4D). Compared to the cluster of core lipid variation alone (Figure 3C), there was a shift in the position of the three groups. In the grouped identified lipid cluster analysis, *Ha. argentinensis* appeared as a distinct branch (Figure 4D), whereas it was more closely related to *Natrialbales* in the

10.3389/fmicb.2023.1297600

TABLE 2 Lipid profiles of seven Halobacteria strains.

Order	*Halobacteriales*	*Haloferacales*			*Natrialbales*		
Genus/ species	*Haloarcula argentinensis*	*Haloferax mediterranei*	*Haloferax larsenii*	*Haloferax volcanii*	*Haloterrigena turkmenica*	*Natrialba asiatica*	*Natrinema gari*
Unsaturation		>10%			1 ~ 10%		
C_{25} chain					>10%		
MK	***	***	***	***	***	***	***
Archaeol	***	***	***	***	***	***	***
Me-PGP	***	***	***	***	***	***	***
PG	***	***	***	***	***	***	***
Bacterioruberin	***	*	***	***	***		***
PGS	***	***					***
TGD	***						***
BPG	***	*	***	*			
TGD-PA	***						
DGD-PA	***	***	***	***			
S-DGD-PA		***	***	***			*
S-DGD		***	***	***	***	***	***
DGD	*	***	***	***	***	*	***
S-Gly-AHH		***	***	***	***	***	*
PE		***	***	*	***	***	***
MGD			*		***	***	***
PI[a]		***	*	*			
2S-DGD					***	***	
2S-Gly-AHH					***	***	
Gly-PG				*	*		***
PA	*	*	*	*	*	*	*
MGD-PA	*	*	*	***			
S-MGD		*	*	*	*		

***Relative abundance > 1/1,000; *relative abundance > 1/10,000. Identification of content less than 1/10,000 is suspectable due to the background noises. [a]Tentative identification as the MS^2 spectrum lacks some of the characteristic fragments.

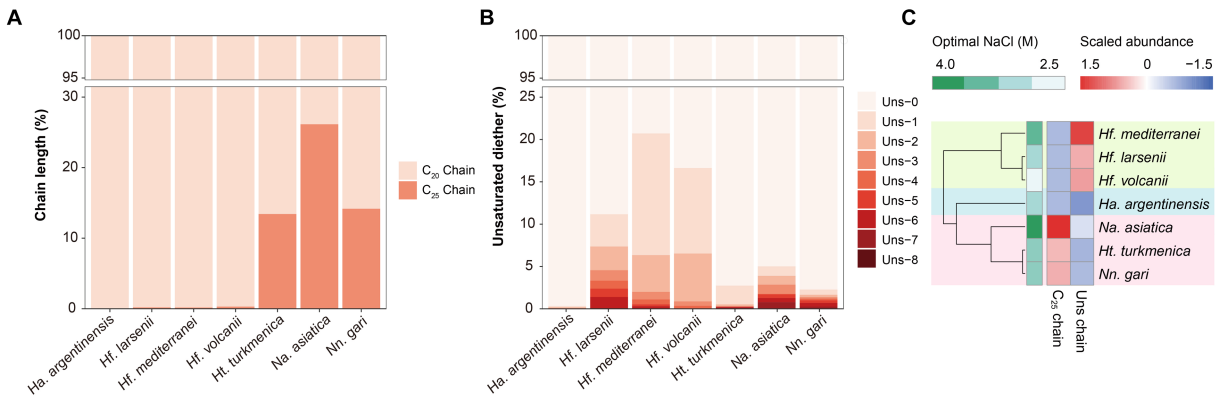


FIGURE 3
Different variables of core lipids, expressed in percent, across the seven Halobacteria cultures, including **(A)** unsaturation in diether molecules and **(B)** isoprenoid chain length. **(C)** Complete linkage cluster analysis of the variation in core lipids. Proportion of unsaturated lipids in total lipids and proportion of C_{25} chains in total chains are used in this analysis.

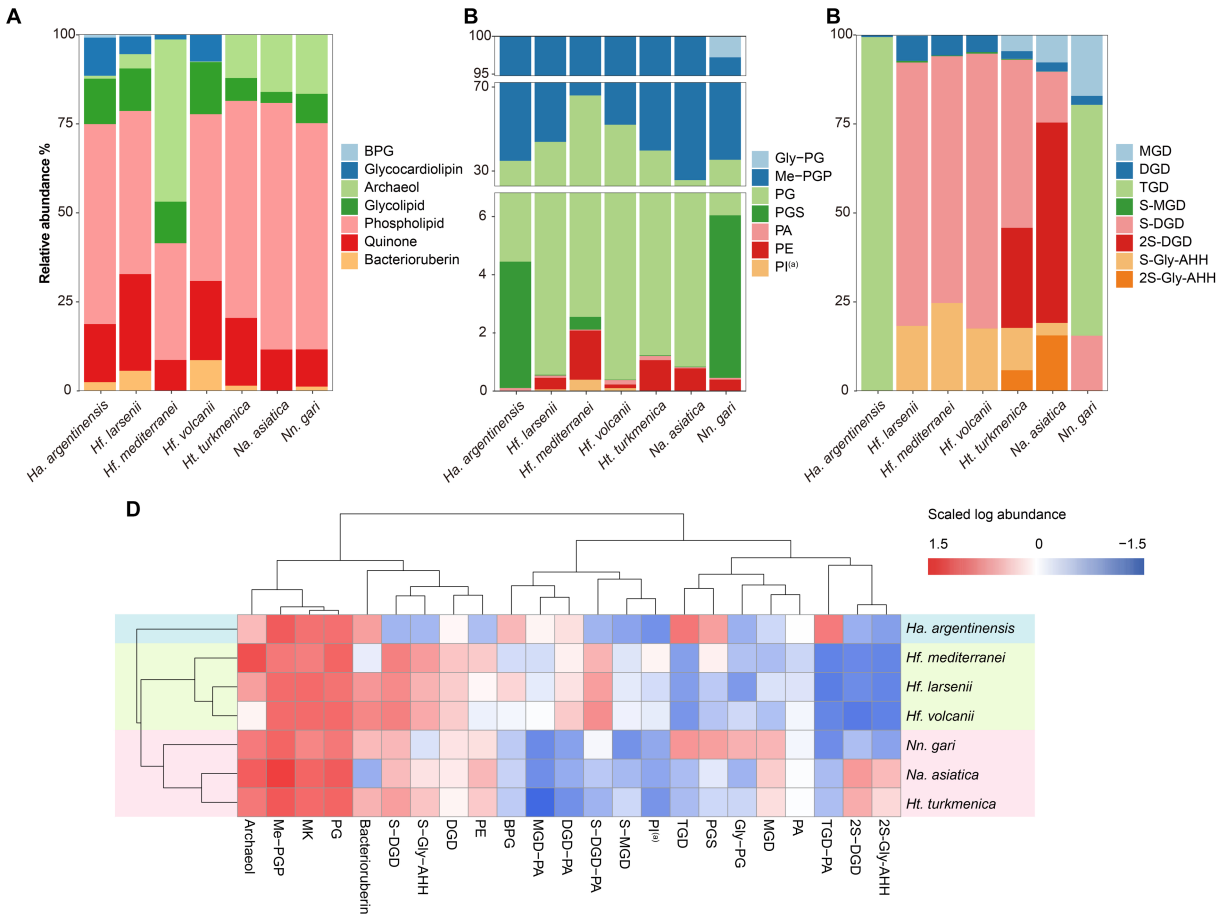


FIGURE 4
Lipidome composition of seven cultured Halobacteria, including **(A)** composition of total lipids, **(B)** composition of glycolipids (including lipids with a reported structure of S-$C_{12}H_{23}NO_{10}$ (Bale et al., 2019), which is named as sulphated Gly-AHH in this study) and **(C)** composition of phospholipids. **(D)** Complete cluster analysis of the lipidome using grouped lipids as an element. The lipid classes examined include BPG: bisphosphatidylglycerol; MGD: monoglycosyl diether; DGD: diglycosyl diether; TGD: triglycosyl diether; S-MGD: sulfo-monoglycosyl diether; S-DGD: sulfo-diglycosyl diether; 2S-DGD: disulfo-diglycosyl diether; Gly-PG: phosphatidylglycerolhexose; Me-PGP: methylated phosphatidylglycerolphosphate; PG: phosphatidylglycerol; PGS: phosphatidylglycerolsulfate; PA: phosphatidic acid; PE: phosphatidylethanolamine; PI: phosphatidyl inositol; MK: menaquinone. (a): Tentative identification, the MS^2 lack some of the characteristic fragments.

core lipid cluster analysis (Figure 3C). In *Natrialbales*, *Nn. gari* formed a separated branch, replacing *Na. asiatica*. Obviously, both the phospholipid and glycolipid compositions supported a distinct lipidome of *Nn. gari* among *Natrialbales*. However, the distant relationship of *Nn. gari* contradicted the phylogenetic relation, which suggested an independent branch for *Na. asiatica* (Figures 1, 4D).

3.3 Extended identification of lipids using feature-based molecular network analysis

3.3.1 Lipids involved in the networks

The feature-based molecular network comprised 444 lipid features after construction and filtration. Of these, 157 had been described previously and were thus identified using our in-house library (Figure 5A). The remaining 287 unknown features were either analogs of the identified ones (122) or new structures (165) that formed different networks on their own (Supplementary Figure S4). All networks were then manually assigned into 10 categories based on the major constituent of known lipids and occurrence in different strains.

3.3.2 Resolution of features within networks

By analyzing features within the network, the phospholipids-related molecular networks involved 111 lipid features, and 64 of them remain unidentified. The molecular networking and further analysis of their MS^2 spectra revealed that most of these features were closely related to the identified phospholipids and contained the characteristic fragments of the known phospholipids. These features might result from new adduct forms of PG, Me-PGP, and Gly-PG in the analyzer, or in-source fragments. We tentatively identified these features as phospholipids that have lost one isoprenoid chain, highly unsaturated lipids attracting an extra chain or even another lipid, and novel adduct forms with short-chain alkylamine (Supplementary Figure S5).

We further identified networks consisting of unknown features (Figure 5B, Supplementary Figure S4). These features can be assigned as compounds of PG, PGP-Me, and Gly-PG with sodium adduct ions ($[M+Na]^+$) based on manual examination of MS^1 and RT. Previous

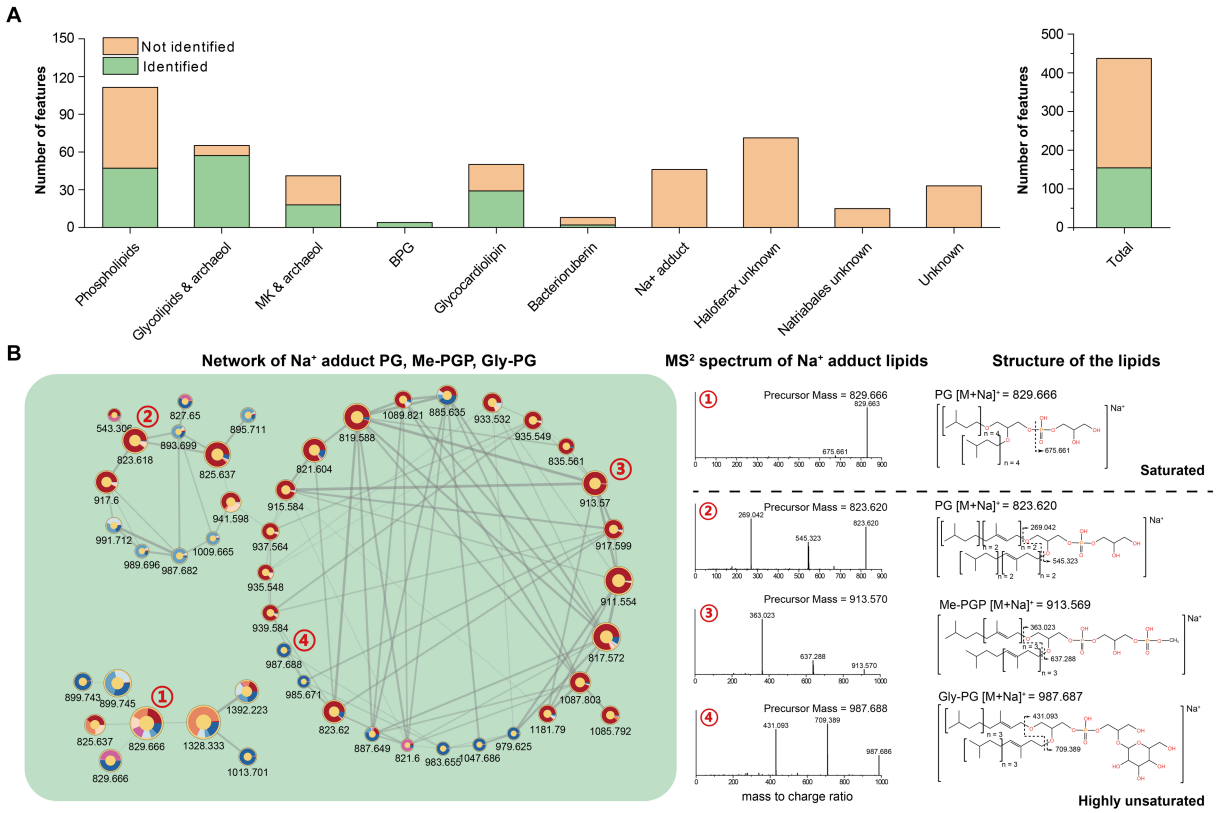


FIGURE 5
(A) Quantification of lipid features within each network; **(B)** Network of PG, Me-PGP, and Gly-PG with Na^+ adduct.

studies showed that archaeal lipids with sodium adduct ions ($[M+Na]^+$) tended to resist fragmentation under CID, making it difficult to provide structural information. Therefore, sodium adducts were not included in our in-house *in silico* library for their poor resolvability, which may lead to false annotation. However, we observed that the MS^2 spectral of unsaturated archaeal diethers with sodium adduct ions ($[M+Na]^+$) could produce fragments ions associated with the intact unsaturated chain loss (Figure 5B), which may provide additional structural information for these lipids.

Finally, some sub-networks were entirely composed of unknown lipids. They presented a strong taxa specificity and could be defined as *Haloferax*-dominant unknown lipids or *Natrialbales*-dominant unknown lipids. This indicated that Halobacteria of *Haloferax* and *Natrialbales* were likely to biosynthesize unique lipids with highly chemotaxonomic potential, which should give traces for halobacterial chemotaxonomy; however, their structures still require elucidation and identification (Supplementary Figure S6).

3.3.3 Enhancing the lipidomic capability of chemotaxonomy by network-coordinated features

Comparing to the PCA analysis using lipid groups, the analysis using network-involved features resulted in a better separation between and within the three groups (Supplementary Figures S8B,C). The PLS-DA analysis revealed that several features involved in network showed significant capability in distinguishing the three halobacterial orders. Among these features, two features in unknown network showed the highest variable importance in the projection (VIP) scores, and 7 unidentified features, two of which were located in the *Natrialbales*-dominant unknown sub-network, ranked among the top 15 VIP scores (Supplementary Figure S9H). This highlighted the importance of incorporating new features to uncover the chemical distinction between strains. To further investigate the chemotaxonomic potential of the lipidomic approach, we performed cluster analysis using individual lipid features (Figure 6A). Compared to the lipid group cluster analysis, the individual identified lipids cluster showed minimal differences except for the positioning of *Ha. argentinensis* (Figures 4D, 6A), which was consistent with core lipids cluster (Figure 3C). The individual lipid cluster provided additional information regarding the core lipids that was not apparent in the grouped cluster analysis, revealing a less diverse lipid profile in *Ha. argentinensis* and contributing to its distinct chemotaxonomic status.

In previous cluster analyses, *Hf. mediterranei* appeared as a distant relative within the *Haloferax* group, possibly due to its high content of archaeol in total lipids and the presence of large amounts of unidentified lipids. When lipid features involved in the identified networks were used, which encompassed all possible analogs of known lipids, there was no significant variation observed in the cluster topology (Figure 6B). However, when all features from the networks including those in unknown networks were introduced, changes occurred in the relationship between *Haloferax* strains. *Hf. larsenii* emerged as the distant relative, which better aligned with the phylogenetic relation (Figures 1, 6C).

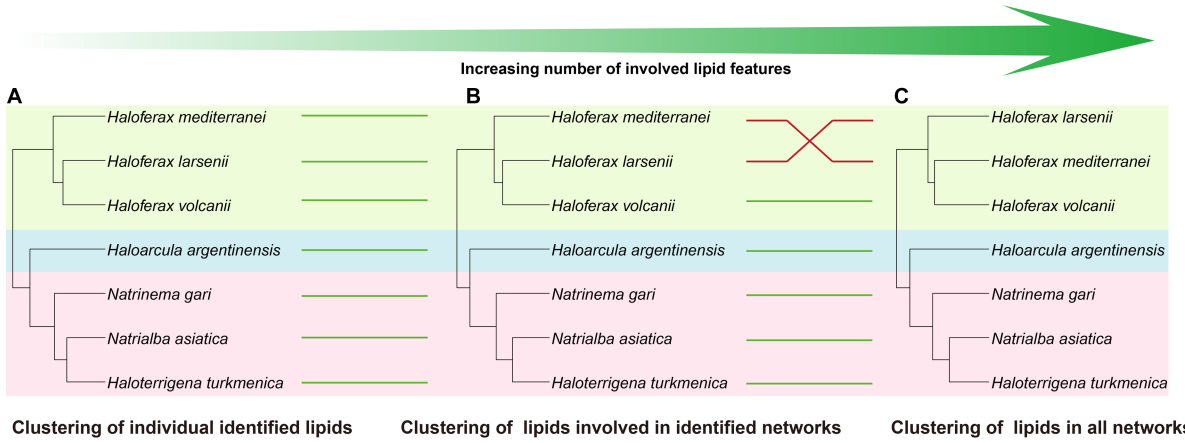


FIGURE 6
Cluster analysis using different levels of features, including **(A)** all identified lipids, **(B)** lipids involved in identified networks, and **(C)** lipids in all networks. Four cluster methods (complete linkage; single linkage; Ward. D linkage; average linkage) are used for each analysis to ensure consistency (Supplementary Figure S7). Strains in different clusters are linked using green or red lines. Green line refers to consistent cluster result while the red line refers to inconsistent cluster result.

4 Discussion

4.1 Lipidome profiling strengthened by high-resolution MS and informatic data processing

Considerable efforts have been devoted to reveal archaeal lipid diversity with the developments of analytical approaches. For example, Elling et al. (2017) successfully identified 118 structurally different lipids in 10 thaumarchaeal cultures, which provided an elaborate database of their membrane lipid composition, enabling us to compare the distribution of lipids in different lineages. Bale et al. (2019) identified a wide range of known and novel lipid compounds in 13 strains of halo(alkali)philic methanogens and Halobacteria, which revealed structural diversity and chemotaxonomic potential of archaeal lipids. However, the traditional manual identification of archaeal lipids analyzed by HRMS is still challenging given the demand for processing and identifying large amounts of lipid features, which may substantially affect the efficiency and accuracy of the analysis.

Combining the untargeted lipidomic approaches and the *in silico* mass spectral library, we were able to simultaneously determine a wide range of structurally different lipids. After peak alignment, a total of 6,470 mass features were detected, with 2,347 features carrying MS^2 spectra, of which 162 features were annotated with our *in silico* library, resulting in an annotation coverage of 6.90%. The annotation results referred to 107 unique lipids, encompassing the majority of lipids previously reported in Halobacteria. These included unsaturated and extended diether core structures; IPLs of PG, Me-PGP, DGD and S-DGD; cardiolipins of both BPG and glycocardiolipins; non-bilayer forming lipids of MK and bacterioruberin (Kamekura and Kates, 1999; Ventosa et al., 1999; Hezayen et al., 2001; Sprott et al., 2003; Xu et al., 2007; Tapingkae et al., 2008; Minegishi and Kamekura, 2009; Ventosa et al., 2009; Elling et al., 2016).

The application of high-resolution mass spectrometry and bioinformatical approaches has allowed us to identify lipids that may be a minor component but are still important in archaea. For example, we were able to detect diether IPLs with PE and PA headgroups in halobacterial strains. PA was a universal intermediate compound in the biosynthesis of phospholipid but has often been overlooked in earlier research (Mullakhanbhai and Larsen, 1975; Ihara et al., 1997). Additionally, PE has been suggested to be absent in Halobacteria (Tenchov et al., 2006), possibly due to the low sensitivity of instrumentation used in previous studies. Recent studies, however, have detected PE in several strains of Halobacteria facilitated by HRMS (Kellermann et al., 2016; Sorokin et al., 2019). These findings have expanded our knowledge of the halobacterial lipidome and highlight the power of bioinformatic tools supervised by manual corrections in lipidome analysis.

4.2 The prospect of feature-based molecular network analysis

Feature-based molecular network techniques are an efficient tool in elucidating lipidomics data by linking features with similar MS^2 spectra. This approach enables the visualization and manually annotation of molecular networks, presenting the distribution of features in grouped samples, m/z value, and structural similarity between different lipid feature or lipid groups (Ding et al., 2021). Our results suggest that FBMN analysis also allows for the characterization of specific unknown features that are potentially crucial in some Halobacteria. Although the detailed structure of those features cannot be directly determined, their structural characteristics can be inferred to some extent according to the relationships in network or using different analytical methods. For example, despite not being identified previously, the feature with an m/z value of 852.745, involved in the network of phospholipids, is likely a novel phospholipid compound (Supplementary Figure S5A). Additionally, some unassigned lipid groups can form identical networks with their known analogs, shedding light on the identification of new lipid types.

In this study, the resulting network comprises 444 features and accounts for 18.92% of the total MS^2 features, far more than the identified features (162 features, 6.90%). This means that lipidome

data can be better elucidated by introducing network analysis. On the basis of network analysis, a fragmentation rule for unsaturated IPLs with Na^+ adduct could be concluded, i.e., the sequential losses of two intact unsaturated isoprenoid chains (Figure 5B). Additionally, the phospholipid network revealed the features of short-chain alkylamine adducts and unsaturated phospholipids that exhibit a propensity to attract extra unsaturated chains or lipids (Supplementary Figure S5). Although these features do not correspond to new lipids, it provides new perspectives to track known lipids in different adduct forms.

In addition, several unannotated networks were found, which contained features with MS^2 characteristics indicating the loss of one or two isoprenoid chains. This demonstrated the presence of novel IPLs or other lipids containing isoprenoid chains. Moreover, based on their distributions among different species, some of these unknown features are also taxa-specific and potentially contribute to chemotaxonomy (Supplementary Figures S4, S6). Despite the structural analysis of these features are not available, formula predictions can be performed using the MS^2 information in SIRIUS v5.6.2 (Supplementary Figure S6), which allowed for a primary prediction of element composition of the new lipid compounds. By introducing network features into the cluster analysis, especially those unknown components, a change in the cluster topology occurred, making the chemotaxonomy to better fit the phylogeny inferred from the 16S rRNA genes (Figures 1, 6). This highlights the improved effectiveness of lipid chemotaxonomy when incorporating unknown lipid features. Overall, the application of molecular networks has the potential to drive a new paradigm in lipidomic research, opening new avenues for exploration on archaea (see below).

4.3 Lipidomic potential for phylogenetic classification of Halobacteria

Attempts for chemotaxonomy analysis of microorganisms have been made for a long period (Fenselau, 1993; Vandamme and Sutcliffe, 2021). A recent Raman spectroscopy application on cell classification reveals distinct chemotaxonomic pattern among halophilic archaea, thermophilic archaea, Thaumarchaeota and bacteria (Wang et al., 2021). Lipids are one of the most important chemotaxonomic biomarkers and have been widely used in identifying new isolated bacteria and archaea strains (Elling et al., 2017; Villanueva and Coolen, 2022; Edwards, 2023). In previous studies, several taxa-specific lipids of archaea have been identified and lipid component-based phylogenetic inference was used to indicate the presence of certain taxa. For example: PGS head group was suggested to be biomarkers for neutrophilic haloarchaea and *Methanothermaceae* was unique for having inositol as the sole phospholipids head group (Koga et al., 1993; Bale et al., 2019).

In this study, we illustrated that seven halobacterial strains, including three strains from *Haloferax* genus, could be chemotaxonomically differentiated through a comprehensive analysis of their lipidomes (Supplementary Figure S8). Archaeol, PG, Me-PGP and MK are common lipid components identified in all 7 cultures of Halobacteria, while other lipid constitutions, such as diverse polar head groups and core lipid structures, exhibit selective distribution among the phylogenetic taxonomy (Table 2). Both lipid modification of unsaturation and extended length on carbon chains in diether core structures are suggested to be adaptive strategies under hyper-salinity stress. Our results revealed that different halobacterial groups, may choose distinct membrane adaptation strategy to salinity stress. For instance, *Haloferax* chooses the strategy by regulating the unsaturation of carbon chains, while *Natrialbales* by extending the length of carbon chains, which can be used to distinguish the three Halobacteria groups. The phospholipid composition in Halobacteria is relatively consistent, but a few minor components such as PE and PI are present in specific groups. Whereas the glycolipids are quite diverse, varying in the number of glycosyl moiety and sulfate moiety among different strains. Bacterioruberin is absent in *Na. asiatica* and has a low amount in *Hf. mediterranei*, which are extreme salt-adapted strains (Merino et al., 2019) that have the highest optimum NaCl concentrations (3.4 M and 4.0 M, respectively) among the 7 strains (Figure 3C) (Hezayen et al., 2001; Ventosa et al., 2009). Accordingly, we inferred that strains thriving under such extreme salt conditions may not favor the production of bacterioruberin.

Among the seven studied strains, *Ha. argentinensis* has the simplest lipidomic diversity, with only a few lipids containing unsaturation or extended chain detected in its core lipids. Interestingly, two novel phospholipids (PE and PI) that were identified in other strains were absent in *Ha. argentinensis*. Instead, we observed that the dominance of triglycosyl diether-phosphatidylglycerol (TGD) is a unique feature of *Ha. argentinensis*. TGD is also detected in *Nn. gari*, but the incapability of *Nn. gari* to produce cardiolipin makes triglycosyl diether-phosphatidyl acid (TGD-PA) a characteristic of *Ha. argentinensis*. Another report of TGD-related cardiolipin is sulfate triglycosyl diether-phosphatidyl acid (S-TGD-PA) in *Halobacterium salinarum* (Corcelli et al., 2000), which belongs to the same order as *Ha. Argentinensis*, but a different family from *Ha. Argentinensis* (Figure 1). Based on these findings, TGD-PA was likely to be a specific biomarker for order *Halobacteriales*.

The lipidome of *Haloferax* genus indicated their ability to produce the unsaturated core lipids, with PI headgroup serving as a featured intact polar lipid. All three strains within this genus also have the ability to produce minor PE, while PGS were only present in *Hf. mediterranei*. The composition of glycolipids in the three strains was similar, with S-DGD being the major component, followed by S-Gly-AHH and DGD. The main cardiolipin lipids in the three *Haloferax* strains were glycocardiolipins, which was similar to *Ha. argentinensis* and strains in *Haloferax* were capable of synthesizing glycocardiolipins in sulphated form.

The three *Natrialbales* strains shared the ability to produce abundant lipids with C_{25} chain, as well as moderate amounts of unsaturated lipids, PE, and sulfated glycosyl moiety. However, they lacked the ability to produce $PI^{(a)}$. The main difference among these stains were the biosynthesis of diglycosyl-based glycolipids as their main glycolipids in *Ht. turkmenica* and *Na. asiatica* (which is also a specific feature in *Natrialbales*), while *Nn. gari* produced TGD instead. The absence of bis-sulfate glycosyl moieties and sulphated-Gly-AHH groups in *Nn. gari* further contributed to the differentiation of their lipidome. For phospholipids, highly abundant Gly-PG and PGS were exclusively found in *Nn. gari*. While minor cardiolipin of S-DGD-PA was present in *Nn. gari*, it represented the only detectable trace of cardiolipin within order *Natrialbales*.

The lipidome differentiation among the 7 strains grown under identical cultivation conditions was in alignment with their phylogenetic relationship at the order level. Such alignment

underscores the mechanism of genetic control on lipid biosynthesis, uncovering the potential for a lipidomic approach in revealing microbial community dynamics in ancient environment when DNA, RNA or proteins were no longer available. However, this study only examined a limited number of strains and further investigations should expand into other archaeal clades.

5 Summary and conclusion

Our study provides a comprehensive view of the lipidome of Halobacteria. At the order level, the clustering results of both CLs and TLs are consistent with the phylogeny of Halobacteria, showing the potential of using lipids as biomarkers for phylogeny. In addition, some lipids are specific to certain taxa and can serve as diagnostic intact lipid biomarkers for specific groups within the Halobacteria, for example: TGD-PA in *Ha. argentinensis*, sulfate glycocardiolipins and abundant production of unsaturated archaeol in *Haloferax*, and bis-sulfate glycosyl lipids in *Natrialbales*. The integration of molecular network analysis in lipidomics expands the comprehensibility of lipidomic data and improves the accuracy of phylogenetic inference, aligning it more closely with the actual relationships among *Haloferax* species. Overall, our study highlights an application of lipidomics as a valuable tool for investigating archaeal phylogeny. Future research is warranted to deepen our understanding of the functional significance of lipidomics in Halobacteria and other archaeal lineages.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found at: https://www.ncbi.nlm.nih.gov/nuccore/NR_176502.1, https://www.ncbi.nlm.nih.gov/nuccore/KT247970.1, https://www.ncbi.nlm.nih.gov/nuccore/NR_029142.2, https://www.ncbi.nlm.nih.gov/nuccore/NR_149760.1, https://www.ncbi.nlm.nih.gov/nuccore/NR_025555.1, https://www.ncbi.nlm.nih.gov/nuccore/NR_177325.1, https://www.ncbi.nlm.nih.gov/nuccore/NR_121590.1, https://www.ncbi.nlm.nih.gov/nuccore/NR_043389.1, https://www.ncbi.nlm.nih.gov/nuccore/NR_028244.1, https://www.ncbi.nlm.nih.gov/nuccore/NR_043803.1.

Author contributions

WY: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Visualization, Writing – original draft. WZ: Investigation, Methodology, Writing – review & editing. WH: Methodology, Writing – review & editing. WX: Funding acquisition, Supervision, Writing – review & editing. YC: Writing – review & editing. YZ: Supervision, Writing – review & editing. CZ: Conceptualization, Funding acquisition, Project administration, Resources, Supervision, Writing – review & editing. FZ: Methodology, Supervision, Project administration, Writing – review & editing, Conceptualization, Funding acquisition.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. This study was supported by the National Natural Science Foundation of China (42141003, 42206040, 42003063), the Stable Support Plan Program of Shenzhen Natural Science Fund (Program Contract No. 20200925173954005), the Guangdong Basic and Applied Basic Research Foundation (2021B1515120080), the Shenzhen Key Laboratory of Marine Archaea Geo-Omics (ZDSYS201802081843490), and the Shanghai Sheshan National Geophysical Observatory (2020Z01). Computation in this study was supported by the Centre for Computational Science and Engineering at the Southern University of Science and Technology.

Acknowledgments

We thank Zhirui Zeng and Wei Yang from Southern University of Science and Technology for providing pure culture strains.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2023.1297600/full#supplementary-material

References

Baker, B. J., De Anda, V., Seitz, K. W., Dombrowski, N., Santoro, A. E., and Lloyd, K. G. (2020). Diversity, ecology and evolution of Archaea. *Nat. Microbiol.* 5, 887–900. doi: 10.1038/s41564-020-0715-z

Bale, N. J., Sorokin, D. Y., Hopmans, E. C., Koenen, M., Rijpstra, W. I. C., Villanueva, L., et al. (2019). New insights into the polar lipid composition of extremely halo(alkali) philic Euryarchaea from Hypersaline Lakes [original research]. *Front. Microbiol.* 10:377. doi: 10.3389/fmicb.2019.00377

Briggs, D. E., and Summons, R. E. (2014). Ancient biomolecules: their origins, fossilization, and role in revealing the history of life. *BioEssays* 36, 482–490. doi: 10.1002/bies.201400010

Chen, Y., Zheng, F., Yang, H., Yang, W., Wu, R., Liu, X., et al. (2022). The production of diverse brGDGTs by an Acidobacterium providing a physiological basis for paleoclimate proxies. *Geochimica et Cosmochimica Acta*. doi: 10.1016/j.gca.2022.08.033

Corcelli, A. (2009). The cardiolipin analogues of Archaea. Biochimica et Biophysica Acta (BBA). *Biomembranes* 1788, 2101–2106. doi: 10.1016/j.bbamem.2009.05.010

Corcelli, A., Colella, M., Mascolo, G., Fanizzi, F. P., and Kates, M. (2000). A novel glycolipid and phospholipid in the purple membrane. *Biochemistry* 39, 3318–3326. doi: 10.1021/bi992462z

Dawson, K. S., Freeman, K. H., and Macalady, J. L. (2012). Molecular characterization of core lipids from halophilic archaea grown under different salinity conditions. *Org. Geochem.* 48, 1–8. doi: 10.1016/j.orggeochem.2012.04.003

Ding, S., Bale, N. J., Hopmans, E. C., Villanueva, L., Arts, M. G., Schouten, S., et al. (2021). Lipidomics of environmental microbial communities. II: characterization using molecular networking and information theory. *Front. Microbiol.* 12:659315. doi: 10.3389/fmicb.2021.659315

Ding, S., Henkel, J. V., Hopmans, E. C., Bale, N. J., Koenen, M., Villanueva, L., et al. (2022). Changes in the membrane lipid composition of a Sulfurimonas species depend on the electron acceptor used for sulfur oxidation. *ISME Commun.* 2:121. doi: 10.1038/s43705-022-00207-3

Dührkop, K., Fleischauer, M., Ludwig, M., Aksenov, A. A., Melnik, A. V., Meusel, M., et al. (2019). SIRIUS 4: a rapid tool for turning tandem mass spectra into metabolite structure information. *Nat. Methods* 16, 299–302. doi: 10.1038/s41592-019-0344-8

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 32, 1792–1797. doi: 10.1093/nar/gkh340

Edwards, B. R. (2023). Lipid biogeochemistry and modern Lipidomic techniques. *Annu. Rev. Mar. Sci.* 15, 485–508. doi: 10.1146/annurev-marine-040422-094104

Elling, F. J., Becker, K. W., Könneke, M., Schröder, J. M., Kellermann, M. Y., Thomm, M., et al. (2016). Respiratory quinones in a rchaea: phylogenetic distribution and application as biomarkers in the marine environment. *Environ. Microbiol.* 18, 692–707. doi: 10.1111/1462-2920.13086

Elling, F. J., Könneke, M., Nicol, G. W., Stieglmeier, M., Bayer, B., Spieck, E., et al. (2017). Chemotaxonomic characterisation of the thaumarchaeal lipidome [doi:10.1111/1462-2920.13759]. *Environ. Microbiol.* 19, 2681–2700. doi: 10.1111/1462-2920.13759

Fenselau, C. (1993). *Mass spectrometry for the characterization of Microorganisms*. American Chemical Society, 541, 1–7.

Findlay, R. H. (2004). Determination of microbial community structure using phospholipid fatty acid profiles. *Mol. Microb. Ecol. Manual* 4:10.1007/978-1-4020-2177-0_408, 983–1004.

Gupta, R. S., Naushad, S., and Baker, S. (2015). Phylogenomic analyses and molecular signatures for the class Halobacteria and its two major clades: a proposal for division of the class Halobacteria into an emended order Halobacteriales and two new orders, Haloferacales Ord. Nov. and Natrialbales Ord. Nov., containing the novel families Haloferacaceae fam. Nov. and Natrialbaceae fam. Nov. *Int. J. Syst. Evol. Microbiol.* 65, 1050–1069. doi: 10.1099/ijs.0.070136-0

Gupta, R. S., Naushad, S., Fabros, R., and Adeolu, M. (2016). A phylogenomic reappraisal of family-level divisions within the class Halobacteria: proposal to divide the order Halobacteriales into the families Halobacteriaceae, Haloarculaceae fam. Nov., and Halococcaceae fam. Nov., and the order Haloferacales into the families, Haloferacaceae and Halorubraceae fam nov. *Antonie Van Leeuwenhoek* 109, 565–587. doi: 10.1007/s10482-016-0660-2

Guthals, A., Watrous, J. D., Dorrestein, P. C., and Bandeira, N. (2012). The spectral networks paradigm in high throughput mass spectrometry. *Mol. BioSyst.* 8, 2535–2544. doi: 10.1039/c2mb25085c

Hartmann, R., Sickinger, H.-D., and Oesterhelt, D. (1980). Anaerobic growth of halobacteria. *Proc. Natl. Acad. Sci.* 77, 3821–3825. doi: 10.1073/pnas.77.7.3821

Hezayen, F. F., Rehm, B. H., Tindall, B. J., and Steinbüchel, A. (2001). Transfer of *Natrialba asiatica* B1T to *Natrialba taiwanensis* sp. nov. and description of Natrialba aegyptiaca sp. nov., a novel extremely halophilic, aerobic, non-pigmented member of the Archaea from Egypt that produces extracellular poly(glutamic acid). *Int. J. Syst. Evol. Microbiol.* 51, 1133–1142. doi: 10.1099/00207713-51-3-1133

Horai, H., Arita, M., Kanaya, S., Nihei, Y., Ikeda, T., Suwa, K., et al. (2010). MassBank: a public repository for sharing mass spectral data for life sciences. *J. Mass Spectrom.* 45, 703–714. doi: 10.1002/jms.1777

Ihara, K., Watanabe, S., and Tamura, T. (1997). *Haloarcula argentinensis* sp. nov. and *Haloarcula mukohataei* sp. nov., two new extremely halophilic Archaea collected in Argentina. *Int. J. Syst. Bacteriol.* 47, 73–77. doi: 10.1099/00207713-47-1-73

Kamekura, M., and Kates, M. (1999). Structural diversity of membrane lipids in members of Halobacteriaceae. *Biosci. Biotechnol. Biochem.* 63, 969–972. doi: 10.1271/bbb.63.969

Kellermann, M. Y., Yoshinaga, M. Y., Valentine, R. C., Wörmer, L., and Valentine, D. L. (2016). Important roles for membrane lipids in haloarchaeal bioenergetics. *Biochim. Biophys. Acta* 1858, 2940–2956. doi: 10.1016/j.bbamem.2016.08.010

Kind, T., Okazaki, Y., Saito, K., and Fiehn, O. (2014). LipidBlast templates as flexible tools for creating new *in-silico* tandem mass spectral libraries. *Anal. Chem.* 86, 11024–11027. doi: 10.1021/ac502511a

Knappy, C. S., Chong, J. P. J., and Keely, B. J. (2009). Rapid discrimination of archaeal tetraether lipid cores by liquid chromatography-tandem mass spectrometry. *J. Am. Soc. Mass Spectrom.* 20, 51–59. doi: 10.1016/j.jasms.2008.09.015

Koga, Y. (2011). Early evolution of membrane lipids: how did the lipid divide occur? *J. Mol. Evol.* 72, 274–282. doi: 10.1007/s00239-011-9428-5

Koga, Y., Akagawa-Matsushita, M., Ohga, M., and Nishihara, M. (1993). Taxonomic significance of the distribution of component parts of polar ether lipids in methanogens. *Syst. Appl. Microbiol.* 16, 342–351. doi: 10.1016/S0723-2020(11)80264-X

Kolde, R. (2012). Pheatmap: pretty heatmaps. *R package version* 1:726.

Kropp, C., Lipp, J., Schmidt, A. L., Seisenberger, C., Linde, M., Hinrichs, K.-U., et al. (2022). Identification of acetylated diether lipids in halophilic Archaea. *MicrobiologyOpen* 11:e1299. doi: 10.1002/mbo3.1299

Law, K. P., He, W., Tao, J., and Zhang, C. (2021). A novel approach to characterize the Lipidome of marine archaeon Nitrosopumilus maritimus by ion mobility mass spectrometry. *Front. Microbiol.* 12:735878. doi: 10.3389/fmicb.2021.735878

Law, K. P., and Zhang, C. L. (2019). Current progress and future trends in mass spectrometry-based archaeal lipidomics. *Org. Geochem.* 134, 45–61. doi: 10.1016/j.orggeochem.2019.04.001

MacLeod, F., Kindler, G. S., Wong, H. L., Chen, R., and Burns, B. P. (2019). Asgard archaea: diversity, function, and evolutionary implications in a range of microbiomes. *AIMS Microbiol.* 5, 48–61. doi: 10.3934/microbiol.2019.1.48

Merino, N., Aronson, H. S., Bojanova, D. P., Feyhl-Buska, J., Wong, M. L., Zhang, S., et al. (2019). Living at the extremes: Extremophiles and the limits of life in a planetary context [Review]. *Front. Microbiol.* 10. doi: 10.3389/fmicb.2019.00780

Minegishi, H., and Kamekura, M. (2009). "Haloterrigena" in *Bergey's Manual of Systematics of Archaea and Bacteria*, 1–11.

Mohimani, H., Gurevich, A., Shlemov, A., Mikheenko, A., Korobeynikov, A., Cao, L., et al. (2018). Dereplication of microbial metabolites through database search of mass spectra. *Nat. Commun.* 9:4035. doi: 10.1038/s41467-018-06082-8

Mullakhanbhai, M. F., and Larsen, H. (1975). *Halobacterium volcanii* spec. Nov., a Dead Sea halobacterium with a moderate salt requirement. *Arch. Microbiol.* 104, 207–214. doi: 10.1007/BF00447326

Nothias, L.-F., Petras, D., Schmid, R., Dührkop, K., Rainer, J., Sarvepalli, A., et al. (2020). Feature-based molecular networking in the GNPS analysis environment. *Nat. Methods* 17, 905–908. doi: 10.1038/s41592-020-0933-6

Oren, A. (2002). Halophilic microorganisms and their environments. Dordrecht: Springer Science & Business Media.

Oren, A. (2014). Taxonomy of halophilic Archaea: current status and future challenges. *Extremophiles* 18, 825–834. doi: 10.1007/s00792-014-0654-9

Oyama, K., Shimada, K., Ishibashi, J.-I., Miki, H., and Okibe, N. (2018). Silver-catalyzed bioleaching of enargite concentrate using moderately thermophilic microorganisms. *Hydrometallurgy* 177, 197–204. doi: 10.1016/j.hydromet.2018.03.014

Pang, Z., Zhou, G., Ewald, J., Chang, L., Hacariz, O., Basu, N., et al. (2022). Using MetaboAnalyst 5.0 for LC–HRMS spectra processing, multi-omics integration and covariate adjustment of global metabolomics data. *Nat. Protoc.* 17, 1735–1761. doi: 10.1038/s41596-022-00710-w

Rodriguez-Valera, F., Juez, G., and Kushner, D. J. (1983). *Halobacterium mediterranei* spec, nov., a new carbohydrate-utilizing extreme halophile. *Syst. Appl. Microbiol.* 4, 369–381. doi: 10.1016/S0723-2020(83)80021-6

Saitou, N., and Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4, 406–425.

Schmid, R., Petras, D., Nothias, L.-F., Wang, M., Aron, A. T., Jagels, A., et al. (2021). Ion identity molecular networking for mass spectrometry-based metabolomics in the GNPS environment. *Nat Commun.* 12, 3832.

Schouten, S., Hopmans, E. C., and Sinninghe Damsté, J. S. (2013). The organic geochemistry of glycerol dialkyl glycerol tetraether lipids: a review. *Org. Geochem.* 54, 19–61. doi: 10.1016/j.orggeochem.2012.09.006

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., et al. (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.* 13, 2498–2504. doi: 10.1101/gr.1239303

Sorokin, D. Y., Elcheninov, A. G., Toshchakov, S. V., Bale, N. J., Sinninghe Damsté, J. S., Khijniak, T. V., et al. (2019). Natrarchaeobius chitinivorans gen. Nov., sp. nov., and Natrarchaeobius halalkaliphilus sp. nov., alkaliphilic, chitin-utilizing haloarchaea from hypersaline alkaline lakes. *Syst. Appl. Microbiol.* 42, 309–318. doi: 10.1016/j.syapm.2019.01.001

Spang, A., Saw, J. H., Jørgensen, S. L., Zaremba-Niedzwiedzka, K., Martijn, J., Lind, A. E., et al. (2015). Complex archaea that bridge the gap between prokaryotes and eukaryotes. *Nature* 521, 173–179. doi: 10.1038/nature14447

Sprott, G. D. (1992). Structures of archaebacterial membrane lipids. *J. Bioenerg. Biomembr.* 24, 555–566. doi: 10.1007/BF00762348

Sprott, G. D., Dicaire, C. J., Gurnani, K., Deschatelets, L. A., and Krishnan, L. (2004). Liposome adjuvants prepared from the total polar lipids of *Haloferax volcanii*, Planococcus spp. and *Bacillus firmus* differ in ability to elicit and sustain immune responses. *Vaccine* 22, 2154–2162. doi: 10.1016/j.vaccine.2003.11.054

Sprott, G. D., Larocque, S., Cadotte, N., Dicaire, C. J., McGee, M., and Brisson, J. R. (2003). Novel polar lipids of halophilic eubacterium Planococcus H8 and archaeon *Haloferax volcanii*. Biochimica et Biophysica Acta (BBA) - molecular and cell biology of. *Lipids* 1633, 179–188. doi: 10.1016/j.bbalip.2003.08.001

Sturt, H. F., Summons, R. E., Smith, K., Elvert, M., and Hinrichs, K.-U. (2004). Intact polar membrane lipids in prokaryotes and sediments deciphered by high-performance liquid chromatography/electrospray ionization multistage mass spectrometry—new biomarkers for biogeochemistry and microbial ecology. *Rapid Commun. Mass Spectrom.* 18, 617–628. doi: 10.1002/rcm.1378

Summons, R. E., Welander, P. V., and Gold, D. A. (2022). Lipid biomarkers: molecular tools for illuminating the history of microbial life. *Nat. Rev. Microbiol.* 20, 174–185. doi: 10.1038/s41579-021-00636-2

Swindell, S. R., and Plasterer, T. N. (1997). "SEQMAN: Contig assembly" in *Sequence data analysis guidebook*, 75–89.

Tamura, K., Stecher, G., and Kumar, S. (2021). MEGA11: molecular evolutionary genetics analysis version 11. *Mol. Biol. Evol.* 38, 3022–3027. doi: 10.1093/molbev/msab120

Tapingkae, W., Tanasupawat, S., Itoh, T., Parkin, K., Benjakul, S., Visessanguan, W., et al. (2008). *Natrinema gari* sp. nov., a halophilic archaeon isolated from fish sauce in Thailand. *Int. J. Syst. Evol. Microbiol.* 58, 2378–2383. doi: 10.1099/ijs.0.65644-0

Team, R. C. (2013). *R: A language and environment for statistical computing.*

Tenchov, B., Vescio, E. M., Sprott, G. D., Zeidel, M. L., and Mathai, J. C. (2006). Salt tolerance of archaeal extremely halophilic lipid membranes *. *J. Biol. Chem.* 281, 10016–10023. doi: 10.1074/jbc.M600369200

Tsugawa, H., Ikeda, K., Takahashi, M., Satoh, A., Mori, Y., Uchino, H., et al. (2020). A lipidome atlas in MS-DIAL 4. *Nat. Biotechnol.* 38, 1159–1163. doi: 10.1038/s41587-020-0531-2

Valentine, D. L. (2007). Adaptations to energy stress dictate the ecology and evolution of the Archaea. *Nat. Rev. Microbiol.* 5, 316–323. doi: 10.1038/nrmicro1619

Vandamme, P., and Sutcliffe, I. (2021). Out with the old and in with the new: time to rethink twentieth century chemotaxonomic practices in bacterial taxonomy. *Int. J. Syst. Evol. Microbiol.* 71:005127. doi: 10.1099/ijsem.0.005127

Vandier, F., Tourte, M., Doumbe-Kingue, C., Plancq, J., Schaeffer, P., Oger, P., et al. (2021). Reappraisal of archaeal C20-C25 diether lipid (extended archaeol) origin and use as a biomarker of hypersalinity. *Org. Geochem.* 159:104276. doi: 10.1016/j.orggeochem.2021.104276

Ventosa, A., de la Haba, R. R., and Sánchez-Porro, C. (2009). "Haloferax" in *Bergey's Manual of Systematics of Archaea and Bacteria*, 1–16.

Ventosa, A., Gutiérrez, M. C., Kamekura, M., and Dyall-Smith, M. L. (1999). Proposal to transfer *Halococcus turkmenicus*, *Halobacterium trapanicum* JCM 9743 and strain GSL-11 to *Haloterrigena turkmenica* gen. Nov., comb. nov. *Int. J. Syst. Evol. Microbiol.* 49, 131–136. doi: 10.1099/00207713-49-1-131

Villanueva, L., and Coolen, M. J. L. (2022). Contributions of genomics to lipid biomarker research: from paleoclimatology to evolution. *Elements* 18, 87–92. doi: 10.2138/gselements.18.2.87

Villanueva, L., Damsté, J. S. S., and Schouten, S. (2014). A re-evaluation of the archaeal membrane lipid biosynthetic pathway. *Nat. Rev. Microbiol.* 12, 438–448. doi: 10.1038/nrmicro3260

Wang, M., Carver, J. J., Phelan, V. V., Sanchez, L. M., Garg, N., Peng, Y., et al. (2016). Sharing and community curation of mass spectrometry data with global natural products social molecular networking. *Nat. Biotechnol.* 34, 828–837. doi: 10.1038/nbt.3597

Wang, Y., Xu, J., Cui, D., Kong, L., Chen, S., Xie, W., et al. (2021). Classification and identification of Archaea using single-cell Raman ejection and artificial intelligence: implications for investigating uncultivated microorganisms. *Anal. Chem.* 93, 17012–17019. doi: 10.1021/acs.analchem.1c03495

Watrous, J., Roach, P., Alexandrov, T., Heath, B. S., Yang, J. Y., Kersten, R. D., et al. (2012). Mass spectral molecular networking of living microbial colonies. *Proc. Natl. Acad. Sci.* 109, E1743–E1752. doi: 10.1073/pnas.1203689109

Woese, C. R., Kandler, O., and Wheelis, M. L. (1990). Towards a natural system of organisms: proposal for the domains Archaea, Bacteria, and Eucarya. *Proc. Natl. Acad. Sci.* 87, 4576–4579. doi: 10.1073/pnas.87.12.4576

Wörmer, L., Lipp, J. S., Schröder, J. M., and Hinrichs, K.-U. (2013). Application of two new LCESIMS methods for improved detection of intact polar lipids (IPLs) in environmental samples. *Organic Geochem.* 59, 10–21. doi: 10.1016/j.orggeochem.2013.03.004

Xu, X.-W., Wu, Y.-H., Wang, C., Oren, A., Zhou, P.-J., and Wu, M. (2007). *Haloferax larsenii* sp nov, an extremely halophilic archaeon from a solar saltern. *Int. J. Syst. Evol. Microbiol.* 57, 717–720. doi: 10.1099/ijs.0.64573-0

Yoshinaga, M. Y., Wörmer, L., Elvert, M., and Hinrichs, K.-U. (2012). Novel cardiolipins from uncultured methane-metabolizing archaea. *Archaea* 2012:832097. doi: 10.1155/2012/832097

Zaremba-Niedzwiedzka, K., Caceres, E. F., Saw, J. H., Bäckström, D., Juzokaite, L., Vancaester, E., et al. (2017). Asgard archaea illuminate the origin of eukaryotic cellular complexity. *Nature* 541, 353–358. doi: 10.1038/nature21031

Zvyagintseva, I., and Tarasov, A. (1988). Extreme halophilic bacteria from saline soils. *Microbiology (USA)* 56, 664–669.

frontiers | Frontiers in Microbiology

TYPE Original Research
PUBLISHED 12 March 2024
DOI 10.3389/fmicb.2024.1364606

OPEN ACCESS

EDITED BY
Mohammad Ali Amoozegar,
University of Tehran, Iran

REVIEWED BY
Heng-Lin Cui,
Jiangsu University, China
Doug Bartlett,
University of California, San Diego,
United States

*CORRESPONDENCE
Dimitry Y. Sorokin
soroc@inmi.ru;
d.sorokin@tudelft.nl

RECEIVED 02 January 2024
ACCEPTED 14 February 2024
PUBLISHED 12 March 2024

CITATION
Sorokin DY, Elcheninov AG, Bale NJ, Sinninghe Damsté JS and Kublanov IV (2024) *Natronoglomus mannanivorans* gen. nov., sp. nov., beta-1,4-mannan utilizing natronoarchaea from hypersaline soda lakes. *Front. Microbiol.* 15:1364606. doi: 10.3389/fmicb.2024.1364606

Natronoglomus mannanivorans gen. nov., sp. nov., beta-1,4-mannan utilizing natronoarchaea from hypersaline soda lakes

Dimitry Y. Sorokin[1,2]*, Alexander G. Elcheninov[1], Nicole J. Bale[3], Jaap S. Sinninghe Damsté[3] and Ilya V. Kublanov[1]

[1]Winogradsky Institute of Microbiology, Research Centre of Biotechnology, Russian Academy of Sciences, Moscow, Russia, [2]Department of Biotechnology, Delft University of Technology, Delft, Netherlands, [3]NIOZ Royal Netherlands Institute for Sea Research, Texel, Netherlands

Beta-mannans are insoluble plant polysaccharides with beta-1,4-linked mannose as the backbone. We used three forms of this polysaccharide, namely, pure mannan, glucomannan, and galactomannan, to enrich haloarchaea, which have the ability to utilize mannans for growth. Four mannan-utilizing strains obtained in pure cultures were closely related to each other on the level of the same species. Furthermore, another strain selected from the same habitats with a soluble beta-1,4-glucan (xyloglucan) was also able to grow with mannan. The phylogenomic analysis placed the isolates into a separate lineage of the new genus level within the family *Natrialbaceae* of the class *Halobacteria*. The strains are moderate alkaliphiles, extremely halophilic, and aerobic saccharolytics. In addition to the three beta-mannan forms, they can also grow with cellulose, xylan, and xyloglucan. Functional genome analysis of two representative strains demonstrated the presence of several genes coding for extracellular endo-beta-1,4-mannanase from the GH5_7 and 5_8 subfamilies and the GH26 family of glycosyl hydrolases. Furthermore, a large spectrum of genes encoding other glycoside hydrolases that were potentially involved in the hydrolysis of cellulose and xylan were also identified in the genomes. A comparative genomics analysis also showed the presence of similar endo-beta-1,4-mannanase homologs in the cellulotrophic genera *Natronobiforma* and *Halococcoides*. Based on the unique physiological properties and the results of phylogenomic analysis, the novel mannan-utilizing halolarchaea are proposed to be classified into a new genus and species *Natronoglomus mannanivorans* gen. nov., sp. nov. with the type strain AArc-m2/3/4 (=JCM 34861=UQM 41565).

KEYWORDS

hypersaline lakes, haloarchaea, glucomannan, galactomannan, beta-1,4-mannan

Introduction

Recent extensive studies on the functional diversity of extremely halophilic archaea belonging to the class *Halobacteria* resulted in a significant shift in a longstanding perception of this unique group of extremophiles as copiotrophic dissipotrophs utilizing either rich amino acid mixtures or simple soluble monomeric compounds, such as sugars or organic acids. In particular, it is becoming evidently clear that many haloarchaeal species possess extensive hydrolytic potential against various groups of recalcitrant insoluble

polysaccharides, such as chitin, cellulose, and a range of other neutral glucans (Sorokin et al., 2015, 2018, 2019a,b, 2022a, 2023). In our previous study, we used selective enrichments with a range of soluble and insoluble glucans to obtain pure cultures of haloarchaea utilizing various glucans with different monomers and back-bond structures as growth substrates (Sorokin et al., 2022a). One of the selected groups included five isolates of alkaliphilic haloarchaea that have the ability to grow with beta-mannans, including pure beta-1,4-mannan and its two modifications, glucomannan (with glucose and mannose in the backbone) and galactomannan (with galactose in the side chains). To date, the beta-mannans with backbone have been proven to serve as a growth substrate for three characterized species of haloarchaea: *Natronoarchaeum mannanilyticum* (Shimane et al., 2010) and *Haloarcula mannanilytica* (Enomoto et al., 2020) use galactomannan as a substrate and cellulotrophic *Natronobiforma cellulositropha* uses pure beta-1,4-mannan as a substrate (Sorokin et al., 2018). The latter and the heteropolymeric glucomannan have never been tested before to enrich mannan-utilizing haloarchaea. According to the CAZy database (http://www.cazy.org), the only enzymatically characterized mannan-specific glycoside hydrolases (several subfamilies of the GH family 5 and the GH family 26) are known in bacteria and eukarya, mostly in fungi. This finding certainly indicates a substantial gap in the knowledge on the potential presence of functionally active archaeal beta-mannanases and on the identity of archaea that are able to utilize such polymers as growth substrates.

In this study, we provide taxonomic evaluation and functional genome analysis focusing on glycosyl hydrolase families for several strains of haloarchaea, which were previously enriched from neutral and alkaline hypersaline lakes in southwestern Siberia on insoluble beta-1,4-mannans. All mannan-utilizing isolates are closely related, forming a separate lineage within the family *Natrialbaceae*, and are proposed to be classified as *Natronoglomus mannanivorans* gen. nov., sp. nov.

Materials and methods

Media and cultivation conditions

Mix surface sediments and brines collected from several hypersaline salt and soda lakes in the Kulunda Steppe (Altai, Russia) were used as inoculum for enrichment cultures. The brine pH (measured with the field pH meter, WWR) in neutral lakes of the chloride-sulfate type ranged from 7.5 to 8.2, and in soda lakes, the pH values ranged from 9.8 to 11. The total salt concentration (measured in the field with a refractometer and verified by the gravimetry in the laboratory) was found to be ranging from 200 to 400 g l^{-1}. Before cultivation, the sediment suspension (one part of sediment:nine parts of brines, v:v) was preincubated for 3 days at 28°C on a rotary shaker with the addition of 200 mg l^{-1} each of ampicillin and streptomycin to suppress the growth of bacteria. The suspended solids were precipitated by centrifugation and resuspended into two types of base mineral medium containing either 4 M total NaCl at pH 7 or a 3:1 (v:v) mix of 4 M NaCl and sodium-carbonate base with 4 M total Na^+ (the final pH was 9.5). The cultures were supplemented with 1 g l^{-1} each of three types of beta-1,4-mannan (Megazyme, Ireland): pure mannan; glucomannan (beta-1,4 heteropolymer of mannose and glucose with a Gl:Man ratio of 4:1); and galactomannan (beta-1,4-mannan with a side chain decoration of alpha-1,6-linked galactose). The polymers were prepared as 10% suspensions in sterile distilled water and sterilized at 110°C for 30 min, and the enrichments were incubated on a rotary shaker in closed bottles at 37°C until visible substrate degradation (accessed by light microscopy) and the appearance of pink turbidity in the liquid phase occurs. Pure culture isolation was achieved after the plating of final positive serial dilution from single colonies forming the clearance zones on the mannan plates (the mannan suspension was sonicated to reduce the particle size) and showing stable growth in the liquid medium with the beta-1,4-mannan as the sole carbon and energy source.

Phenotypic characterization

Phase contrast microscopy was performed using Zeiss Axioplane Imaging 2 (Germany). The growth in liquid cultures was monitored OD_{600} after 30 min of gravity sedimentation of insoluble particles or directly in the case of soluble substrates. Thin-section electron microscopy was performed for the cells of type strain AArc-m2/3/4 grown either with beta-mannan or xylan, using a JEOL100 instrument (Japan) as described previously (Sorokin et al., 2018). The ability for anaerobic growth was tested in serum bottles sealed with butyl rubber stoppers after the removal of dissolved oxygen and subjected first to "cold boiling" under a vacuum followed by flashing three times with sterile argon gas. Aerobic substrate utilization, pH-salt profiling, and other standard phenotypic testing were performed as described by Sorokin et al. (2022a). Membrane polar lipids and respiratory menaquinones were extracted from freeze-dried biomass of strains AArc-m2/3/4 and AArc-xg1-1 that were grown with cellobiose at 37°C, 4 M total Na^+, and pH 9.5 until the late exponential growth phase and were resolved by Ultra High Pressure Liquid Chromatography-High Resolution Mass Spectrometry (UHPLCHRMS) using an Agilent 1290 Infinity I UHPLC (ThermoFisher Scientific), as described previously (Bale et al., 2021; Sorokin et al., 2022b).

Growth tests were performed in a basal alkaline medium containing 4 M total Na^+ (1 M as sodium carbonates and 3 M as NaCl) at pH 9.5. Substrates were added at a final concentration of 1 g l^{-1} after sterilization from 10% concentrated stocks. In total, 10 ml cultures in 30 ml bottles with screw-capped rubber septum (to prevent evaporation) were incubated on a rotary shaker at 37°C. The growth was monitored by measuring OD_{600} against a control without substrate, directly in the case of soluble substrate, and after vigorous homogenization and 10 min stasis to allow particles to sediment in the case of insoluble polysaccharides.

Genome sequencing and phylogenomic and functional genomic analyses

The genomes of AArc-m2/3/4 and AArc-xg1-1 were sequenced using the MiSeq Illumina platform and assembled as described previously (Sorokin et al., 2022a). Their draft assemblies were

deposited in the GenBank under the numbers GCA_025517485 and GCA_025517495, respectively.

For phylogenomic analysis, 122 conserved single copy archaeal protein markers (Rinke et al., 2021) from the *in silico* translated genomes of two mannan-utilizing natronoarchaea, the closest related strains KZCA124 and TS33, and all described species of the family *Natrialbaceae*, were identified and aligned using the GTDB-tk v.1.7.0 with reference data v.202 (Chaumeil et al., 2019). The resulting alignment was treated using trimAL v1.4.1 with –gt 1 option for complete gap elimination (Capella-Gutiérrez et al., 2009). The phylogenomic tree was constructed and drawn using the RAxML v.8.2.12 (the PROTGAMMAILG model, 1,000 rapid bootstrap replications) and iTOL v.6.5.2, respectively (Stamatakis, 2014; Letunic and Bork, 2019). The whole genome comparisons (ANI, AAI, and POPC) were calculated as described earlier (Sorokin et al., 2023).

Functional genome analysis of the glycoside hydrolases (GH) with particular focus on the mannan-hydrolyzing families including the GH5_7, GH5_8, and GH26 families in AArc-m2/3/4 and AArc-xg1-1 (as well as the genomes of related strains KZCA124 and TS33) was performed using the dbCAN v.4 (Zhang et al., 2018) with HMMER (Mistry et al., 2013) and Diamond tools (Buchfink et al., 2015) to detect the target GH domains. Additional domains were searched using the InterPro database (Paysan-Lafosse et al., 2023). Gene clusters of AArc-m2/3/4, AArc-xg1-1, and several other species containing genes encoding endo-beta-mannanases were visualized using the gggenes package in R (https://cran.r-project.org/web/packages/gggenes). The GH5 and GH26 enzymes of strains AArc-m2/3/4 and AArc-xg1-1 and related proteins found by the NCBI BLAST (the organism list was limited to *Archaea*; the e-value threshold were 1e-10 and 1e-5 for GH5 and GH26, respectively) were used for the phylogenetic reconstruction of the haloarchaeal beta-mannanases. Furthermore, all proteins were analyzed using dbCAN v.4 with HMMER tool to detect target domains; only proteins with target catalytic domains (minimal coverage is 0.4) and some additional carbohydrate-binding domains (minimal coverage is 0.5) were kept in the final protein set. Both sets of sequences (the GH5 and GH26 families) were aligned using MAFFT v.7 with the E-INS-i method (Katoh et al., 2017). The alignments obtained were trimmed using trimAL v1.4.1 with -gt 0.75 (Capella-Gutiérrez et al., 2009). The phylogenetic trees for GH5 and GH26 glycosidases were constructed and decorated as described above for the "ar122"-based phylogenomic tree with a decreased number of bootstrap replications (500).

Results and discussion

Enrichment and isolation of pure cultures of mannan-utilizing haloarchaea

Five enrichment cultures (one from neutral salt lakes and four from soda lakes) showed stable growth with four different polymers. The neutral and soda lake enrichments with beta-1,4-mannan yielded strains HArc-m1 and AArc-m2/3/4^T; the soda lake enrichments with galactomannan yielded strain AArc-glctm5 and with glucomannan yielded strain AArc-gm3/4/5-2. An additional enrichment with xyloglucan provided an additional beta-1,4-mannan-utilizing isolate AArc-xg1-1. It was confirmed that the first four isolates selected with different mannan variations could grow with beta-1,4-mannan. The xyloglucan-yielded isolate AArc-xg1-1 was first identified as a potential mannan utilizer based on its phylogenomic proximity to the mannan-enriched strains and functional genome analysis, and it was later confirmed by a growth test. An interesting fact to note is that the neutral salt lake isolate, HArc-m1, turned out to be identical to its counterpart mannan-selected soda lake natronarchaeon, AArc-m2/3/4, and was confirmed to be able to grow optimally at moderately alkaline conditions used for the enrichment of the soda lake isolates, which is a rare example (at least in our experience). Evidently, two factors might have played a role in such a coincidence: the first is the close proximity of the chloride-sulfate and soda lakes in the southern Kulunda steppe (a few kilometers range) and the second is the unique substrate specialization overruling the significance of the alkali-related adaptation.

Phenotypic properties

The cells of all mannan-utilizing strains were mostly true non-motile cocci that were ~1 μm in diameter, becoming especially refractive when grown with mannan and cellulose (Figure 1). Thin sectioning electron microscopy revealed a major difference in the cell ultrastructure between the mannan- and xylan-grown cells: the former had a much thicker cell wall consisting of several layers and an extended pseudoperiplasmic space (Figure 2), which resembled the cellulose-growing cells of *Natronobiforma* (Sorokin et al., 2018). It might be due to the necessity of these haloarchaea to interact physically with the insoluble glucans.

Membrane phospholipid profiling of strains AArc-m2/3/4 and AArc-xg1-1 indicated domination of C_{20}-C_{20} archaeol and C_{20}-C_{25} extended archaeol as the core, with phosphatidylglycerolphosphate methyl ester (PGP-Me) and phosphatidylglycerol (PG) as the polar heads in equal proportion (Supplementary Table S1). The glyco- and sulfo-lipids were not detected. The majority of respiratory lipoquinones were represented by the saturated MK-8:8 (75 and 95% in AArc-m2/3/4 and AArc-xg1-1, respectively), and the rest was monounsaturated MK-8:7.

The range of polysaccharides supporting the growth of all isolates included beta-1,4 mannan, glucomannan, amorphous cellulose, xyloglucan, xylan, and arabinoxylan, while galactomannan supported only a weak growth with incomplete degradation. Alpha-glucans from the starch family, beta-glucans with the 1,3- and 1,6-backbone, and the beta-fructans did not support the growth of isolates. Active hydrolysis of the beta-mannan and amorphous cellulose can be observed as the formation of clearance zones around the colonies (Supplementary Figure S1). The range of utilized sugars investigated in the strain AArc-m2/3/4^T included galactose, lactose, trehalose, raffinose, rhamnose, sucrose, cellobiose, maltose, melezitose, and melibiose. Interestingly, mannose and glucose, the monomers of two groups of growth-supporting beta-1,4-glucans, were not among the utilized sugars. Most likely, only oligomers were transported inside the cells.

10.3389/fmicb.2024.1364606

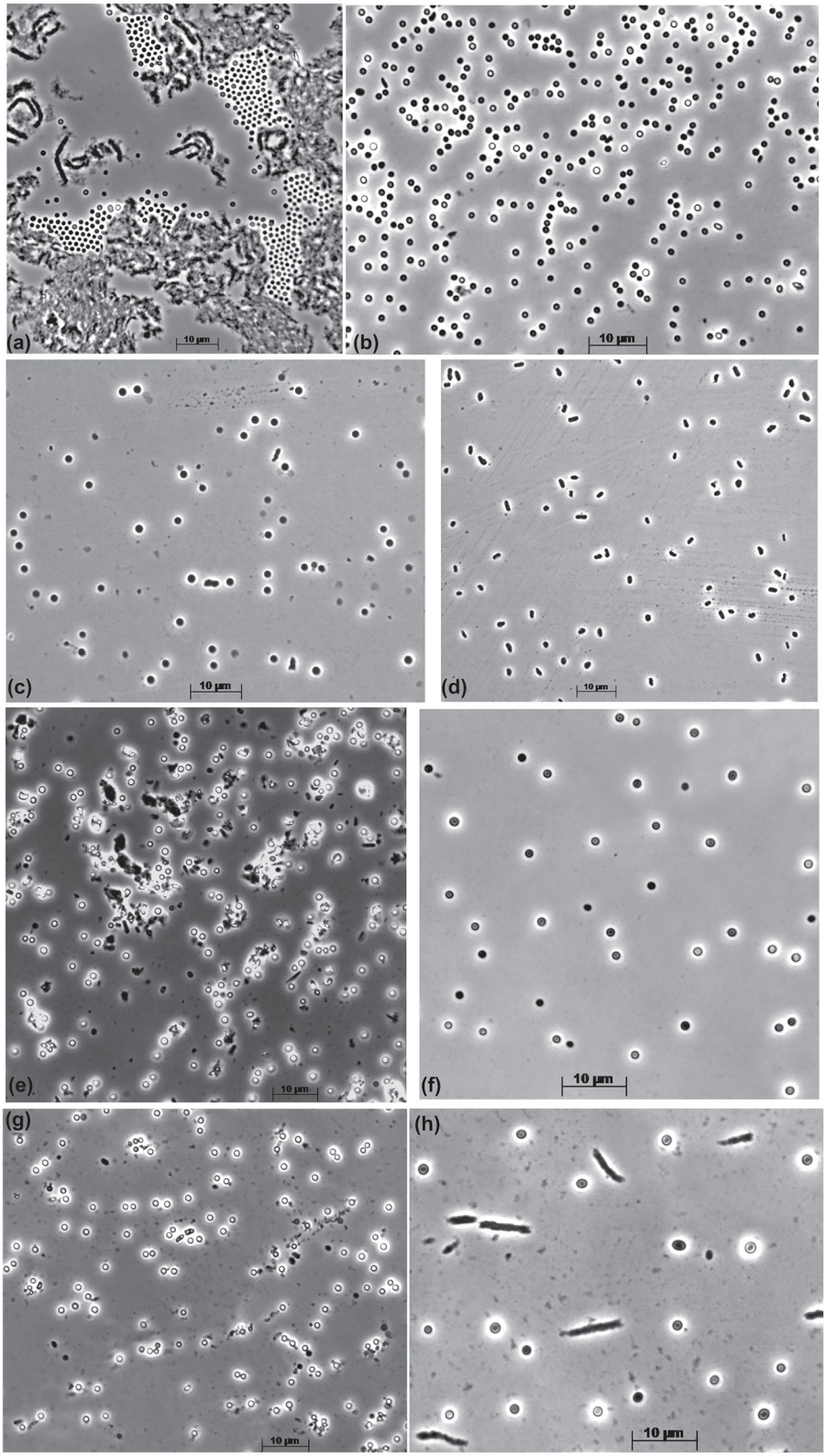


FIGURE 1
Cell morphology of the beta-1,4-mannan-utilizing haloarchaea growing at 4 M Na^+, pH 9.5, and 37°C. **(a, b)** Strain AArc-m2/3/4 on mannan [**(a)** aggregated cells in the solid phase with polymer, **(b)** free cells]; **(c, d)** strain AArc-xg1-1 [**(c)** on mannan, **(d)** on xyloglucan]; **(e)** strain HArc-m1 on mannan; **(f)** strain AArc-glctm5 on mannan; **(g, h)** strain AArc-gm3/4/5-2 [**(g)** on mannan, **(h)** on glucomannan].

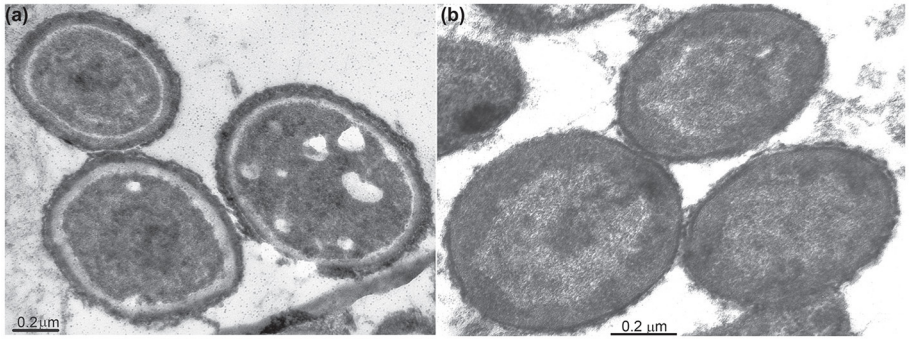


FIGURE 2
Thin-section electron microscopy microphotographs showing an ultrastructural difference of the AArc-m2/3/4 cells grown with beta-1,4-mannan **(a)** and xylan **(b)**. The mannan-grown cells have an extra external cell layer and extended pseudoperiplasm in comparison to the cells grown on xylan.

Anaerobic growth with cellobiose as a substrate was not observed both at fermentative conditions and in the presence of various electron acceptors, including nitrate, nitrite, N_2O, sulfur, thiosulfate, DMSO, and fumarate. Ammonium and urea, but not nitrate, served as the nitrogen source. The cellobiose-dependent growth at pH 9.5 was not inhibited by streptomycin, ampicillin, kanamycin, vancomycin, and tetracycline at concentrations of up to 100 mg l^{-1}.

AArc-m2/3/4^T grew optimally (with cellobiose at pH 9) at 3.5 M total Na^+ and with the range from 2.0 to 4.5 M. The pH profiling (with cellobiose at 4 M total Na^+) was performed for the soda lake isolate, AArc-m2/3/4^T, and for the salt lake isolate, HArc-m1. The former showed a weak growth starting from pH 7.2 and grew up to pH 9.7, while the latter had a lower growth pH limit starting from a minimum of pH 6.8 and was able to grow up to pH 9.5. Both strains grew optimally around pH 9. Thus, they can be qualified as facultatively alkaliphilic extreme halophiles.

Phylogenetic analysis

The genomes of AArc-m2/3/4 and AArc-xg1-1 contain a single *rrn* operon with the 16S rRNA gene most closely related to the species of the genera *Halovivax* and *Saliphagus* (~95% sequence identity). However, according to the results of the phylogenomic analysis based on 122 archaeal conserved single-copy protein markers, mannan-utilizing isolates formed a distinct clade within the *Natrialbaceae* family neighboring the cellulotrophic genus *Natronobiforma* (Figure 3, Supplementary Figure S2). Furthermore, the whole-genome-based comparative analyses indicated that the novel isolates form an independent lineage of a new genus level within the *Natrialbaceae* (Supplementary Table S2). Average nucleotide identities (ANI) between the AArc-m2/3/4^T and AArc-xg1-1 strains was 98.5%, which means that they belong to the same species; ANI values between novel isolates and the representatives of the closest genera, *Natronobiforma* and *Saliphagus*, were 77.9 and 77.1%, respectively, and the range for other type species of the *Natrialbaceae* varied from 74.1 to 78.6%. Average amino acid identity (AAI) values between the AArc-m2/3/4^T and AArc-xg1-1 strains and the type species of the *Natrialbaceae* ranged from 58.9 to 70.3%, with the highest value being *N. cellulositropha*. These values are significantly below the threshold for the demarcation of genera within the *Natrialbaceae* (76%; de la Haba et al., 2021). The percentage of conserved proteins (POCP) between novel strains and the type species of the *Natrialbaceae* varied from 49.4 to 66.2% with the highest values being *N. cellulositropha* and *Saliphagus infecundisoli*. The highest POCP values were slightly higher than the proposed genus-level threshold of 50% (Qin et al., 2014). However, many authors often indicated that a strict boundary of 50% is not suitable for the demarcation of genera within some taxa (Aliyu et al., 2016; Orata et al., 2018; Wirth and Whitman, 2018). Thus, it is likely that, in the case of the *Natrialbaceae*, the POCP threshold needs to be revised.

Functional genome analysis

Beta-mannan utilization potential

The main functional property of the new isolates is the potential to utilize insoluble beta-1,4-mannans and beta-1,4-glucans (cellulose) as a growth substrate. The two sequenced genomes contained a very similar array of genes encoding the key glycoside hydrolases that were potentially involved in the unique substrate specificity of the novel natronoarchaea. A most interesting group of such enzymes included beta-1,4-mannanases belonging to the families GH26 (one enzyme), GH5_7 (four enzymes, two of them with a carbohydrate binding domain, CBM), and GH5_8 (one enzyme with a CBM) (Table 1). All six proteins have TAT translocation signals (i.e., are extracellular) and are co-located in two large genomic islands together with the genes coding for multiple endo-beta-1,4-glucanases from the GH5 superfamily, which is most likely responsible for cellulose degradation and/or endo-beta-1,4-xylanases from the GH10 family.

The first cluster contains genes of three endo-beta-1,4-mannosidases from the GH5_7 and GH5_8 subfamilies and the GH26 family and a beta-mannosidase from the GH2 family (Figure 4A). Moreover, endo-alpha-1,5-L-arabinanase from the GH43 family, a putative enzyme from the GH109 family (homologous to various sugar dehydrogenases), two

10.3389/fmicb.2024.1364606

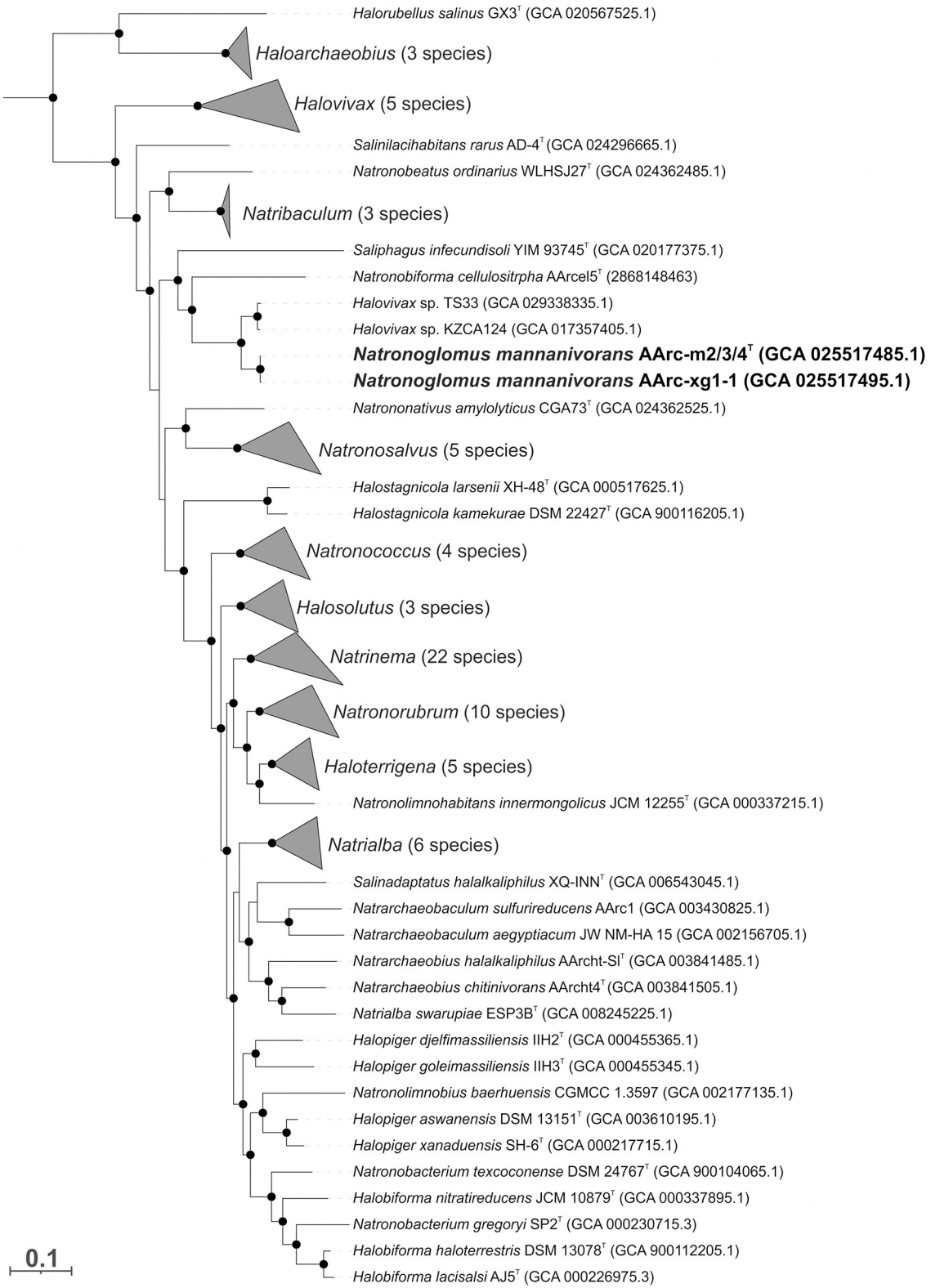


FIGURE 3
The maximum likelihood phylogenetic tree based on 122 conserved archaeal proteins demonstrating the position of two mannan-utilizing natronoarchaea (in bold) within the family *Natrialbaceae*. The branch lengths correspond to the number of substitutions per site with corrections associated with the models. The black circles at nodes indicate that the percentage of corresponding support values was above 50. The species of some genera are collapsed, and the numbers of species are given in brackets. *Archaeoglobus fulgidus* VC-16^T, *Methanocella paludicola* SANAE^T, and *Methanothermobacter thermautotrophicus* Delta H^T were used as an outgroup (not shown).

10.3389/fmicb.2024.1364606

TABLE 1 Polysaccharide hydrolases potentially involved in the beta-mannan, cellulose, and xylan hydrolysis encoded in the genome of mannan-utilizing strain AArc-m2/3/4.

Locus tag MCU4972+	Protein family	Putative function	Protein size (aa)	Localization (signal peptide)
0716	GH5	Endo-beta-1,4-glucanase	523	(E) TAT/SPI
2295	GH11	Endo-beta-1,4-xylanase	428	(E) TAT/SPI
2297	GH10/2 × CBM85	Endo-beta-1,4-xylanase	1,035	(E) TAT/SPI-SPII
2716	GH5/PKD	Endo-beta-1,4-glucanase	637	(E) TAT/SPI
2717	GH5	Endo-beta-1,4-glucanase	610	(E) TAT/SPI
2718	GH5/CBM6	Endo-beta-1,4-glucanase	1,755	(E) TAT/SPI
2719	–	S-layer glycoprotein	258	(M) 2 TMH
2720	**GH5_7/CBM5**	**Endo-beta-1,4-mannanase**	759	(E) TAT/SPI
2721	GH5	Endo-beta-1,4-glucanase	646	(E) Sec/SPI
2722	GH10	Endo-beta-1,4-xylanase	589	(E) TAT/SPI
2723	GH10	Endo-beta-1,4-xylanase	741	(E) TAT/SPI
2724	GH10	Endo-beta-1,4-xylanase	577	(E) TAT/SPI
2725	GH5	Endo-beta-1,4-glucanase	600	(E) TAT/SPI
2726	–	Dockerin/PKD family	815	(E) TAT/SPI
2727	CBM6	Cellulose-binding	592	(E) TAT/SPI
2728	GH5	Endo-beta-1,4-glucanase	598	(E) TAT/SPI
2729	GH5/CBM6	Endo-beta-1,4-glucanase	835	(E) TAT/SPI
2730	GH5	Endo-beta-1,4-glucanase	490	(E) TAT/SPI
2732	CBM3	Chitin/cellulose-binding	578	(E) TAT/SPI
2733	**GH5_7**	**Endo-beta-1,4-mannanase**	469	(E) Sec/SPI
2734	**GH5_7**	**Endo-beta-1,4-mannanase**	779	(E) TAT/SPI
2735	GH5/CBM6	Endo-beta-1,4-glucanase	827	(E) TAT/SPI
3710	GH9	cellulase	821	(E) TAT/SPI
4111	GH5/CBM9	Endo-beta-1,4-glucanase	653	(E) TAT/SPI
5514	**GH5_7**	**Endo-beta-1,4-mannanase**	417	(E) TAT/SPI
5518	**GH26**	**Endo-beta-1,4-mannanase**	466	(E) TAT/SPI
5520	**GH5_8**/CBM6	**Endo-beta-1,4-mannanase**	407	(E) TAT/SPI
5524	GH2	Beta-mannosidase	844	(C)

(E), extracellular; (C), cytoplasmic; CBM, carbohydrate-binding domain. Mannan-specific enzymes are in bold.

mannonate dehydratases, putative UDP-glucose 4-epimerase, and a transcriptional regulator of the IclR family were encoded in this cluster. In the genome of AArc-xg1-1, there was an additional CRISPR-associated endonuclease. Very similar clusters were found in the genomes of undescribed strains KZCA124 and TS33, which were closely related to our mannan isolates, but they possessed an additional GH5_7 gene, while genes of GH43 glycosidase were absent. Homologous cluster in *N. cellulositropha* AArcel5 lacked genes of GH5_8 enzyme and one mannonate dehydratase gene, while *Halococcoides cellulosivorans* HArcel1 had only a single endo-beta-1,4-mannosidase from GH26.

The second cluster was much larger and encoded up to 21 proteins (Figure 4B). Identical clusters encoding 11 glycosides from the GH5 family (both endo-beta-1,4-mannanases and endoglucanases), three endo-beta-1,4-xylanases from the GH10 family, putative polysaccharide lyase (PL)14 with an unknown function, and several proteins containing putative CBM domains (galactose-binding-like domain, fibronectin type III, and type I dockerin) were observed in the genomes of AArc-m2/3/4 and AArc-xg1-1. Clusters in the genomes of KZCA124 and TS33 did not encode proteins of the PL14 family and fibronectin type III domain-containing proteins. Moreover, there were fewer genes of endo-glucanases from GH5 and GH10 enzymes. The largest endo-glucanase encoding cluster was found in the genome of *N. cellulositropha* AArcel5 with 10 x GH5 and GH9 endo-beta-1,4-glucanases, GH81 endo-beta-1,3-1,4-glucanase but lack the GH10 xylanases. Other clusters of the second type consisted of four to nine genes (mostly genes of glycosidases from GH5 and GH5_7). Both *H. cellulosivorans* and *Halomicrobium salinisoli* had a single gene encoding endo-beta-1,4-mannanases from the GH5_7 subfamily.

10.3389/fmicb.2024.1364606

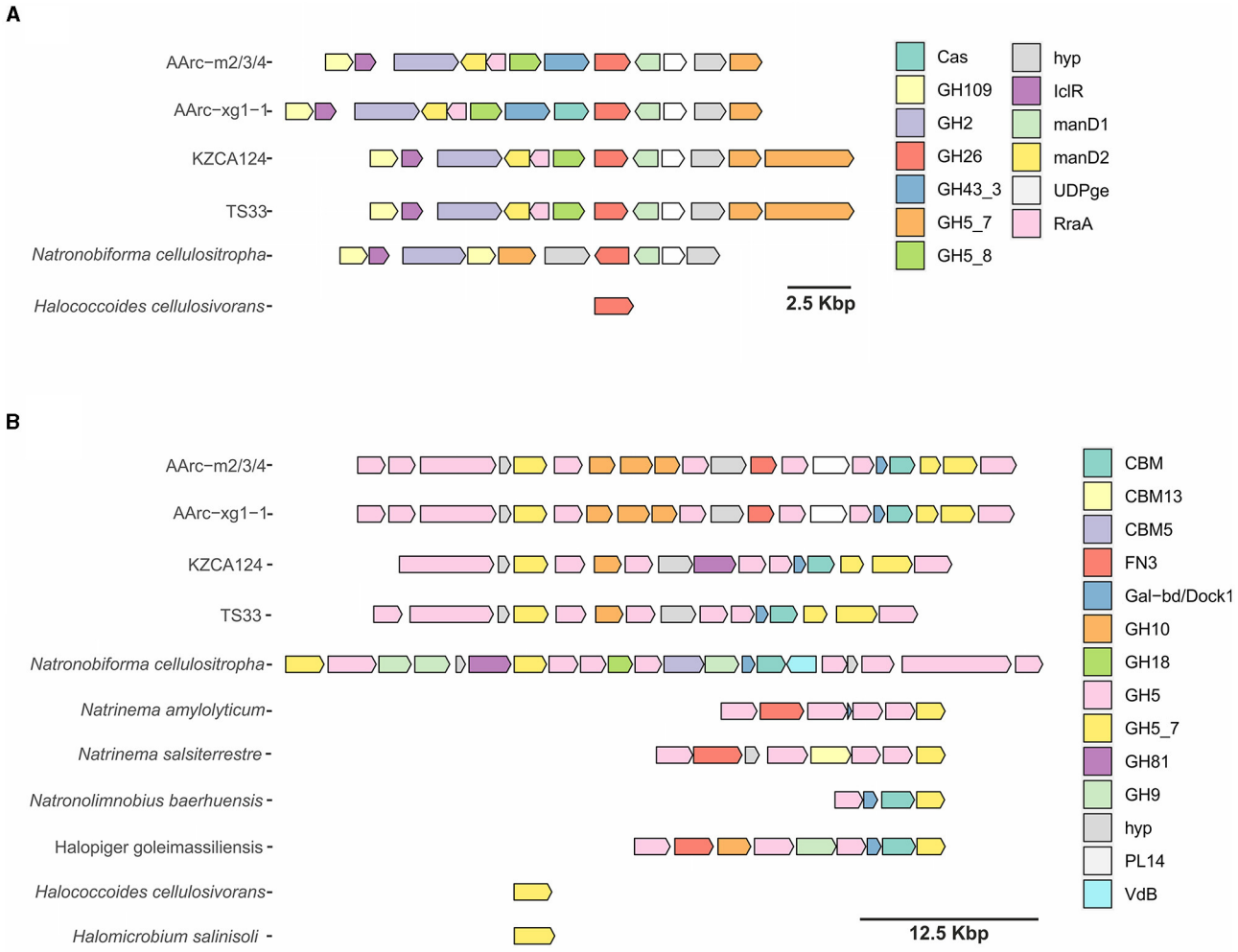


FIGURE 4
Clusters containing endo-beta-mannanase genes of AArc-m2/3/4^T and AArc-xg1-1 strains and similar gene clusters of other haloarchaeae encoding GH26, GH5_7, and GH5_8 families **(A)** or enzymes of the GH5_7 family only **(B)**: Cas, CRISPR-associated nuclease; GHx, glycoside hydrolase of "x" family; hyp, hypothetical protein; IclR, HTH-type transcriptional regulator of IclR family; manD, mannonate dehydratase; UDPge, UDP-glucose 4-epimerase; RraA, putative oxaloacetate decarboxylase of RraA family; CBM, carbohydrate-binding module; FN3, fibronectin type III domain-containing protein; Gal-db/Dock1, galactose-binding-like domain- and/or type I dockerin domain-containing protein; PL14, polysaccharide lyase of PL14 family; and VbB, vanadium-dependent bromoperoxidase.

To predict the functions of the GH26 and GH5_7/8 subfamilies more reliably, a phylogenetic analysis of these families was performed. This analysis demonstrated that the GH26 encoded in the genomes of the AArc-m2/3/4^T and AArc-xg1-1 strains are indeed endo-beta-1,4-mannanases belonging to the clade containing biochemically characterized beta-mannanases from bacteria and fungi (Figure 5). In addition to the catalytic GH26 domain, these enzymes contain the dockerin type I domain participating in the binding of substrate. Enzymes of KZCA124 and TS33, as well as *N. cellulositropha* and *H. cellulosivorans*, also belong to this clade (II). It is remarkable that there were more than 100 proteins of the GH26 family from halophilic archaea (and other archaea) belonging to another clade (I), which have a very low level of homology with the confirmed beta-1,4-mannanases. It is likely that these proteins together with the characterized beta-1,3-xylanases should be separated into a different subfamily of the GH26 or even into a separate family. Due to the complexity of the GH5 family, recently including 57 subfamilies (Supplementary Figure S3), analysis of these glycosidases was focused on clades containing endo-beta-mannosidases (GH5_7 and GH5_8 subfamilies). Both AArc-m2/3/4^T and AArc-xg1-1 strains possess four enzymes from the GH5_7 subfamily (Figure 6), of which two of them contained additional domains (carbohydrate-binding module family 5/12, fibronectin type III or PKD, and concanavalin A-like lectin domains), which may enhance the hydrolytic activity against insoluble polysaccharides. In addition to the abovementioned enzymes, KZCA124 and TS33 had another protein of GH5_7 containing three fibronectin type III domains. Moreover, one enzyme of the GH5_8 subfamily was encoded in each genome (Figure 7). These proteins consisted of a catalytic domain and the CBM5/12. Both GH5_7 and GH5_8 enzymes from the studied haloarchaea were clustered together with biochemically characterized endo-beta-1,4-mannonases from bacteria, fungi, and

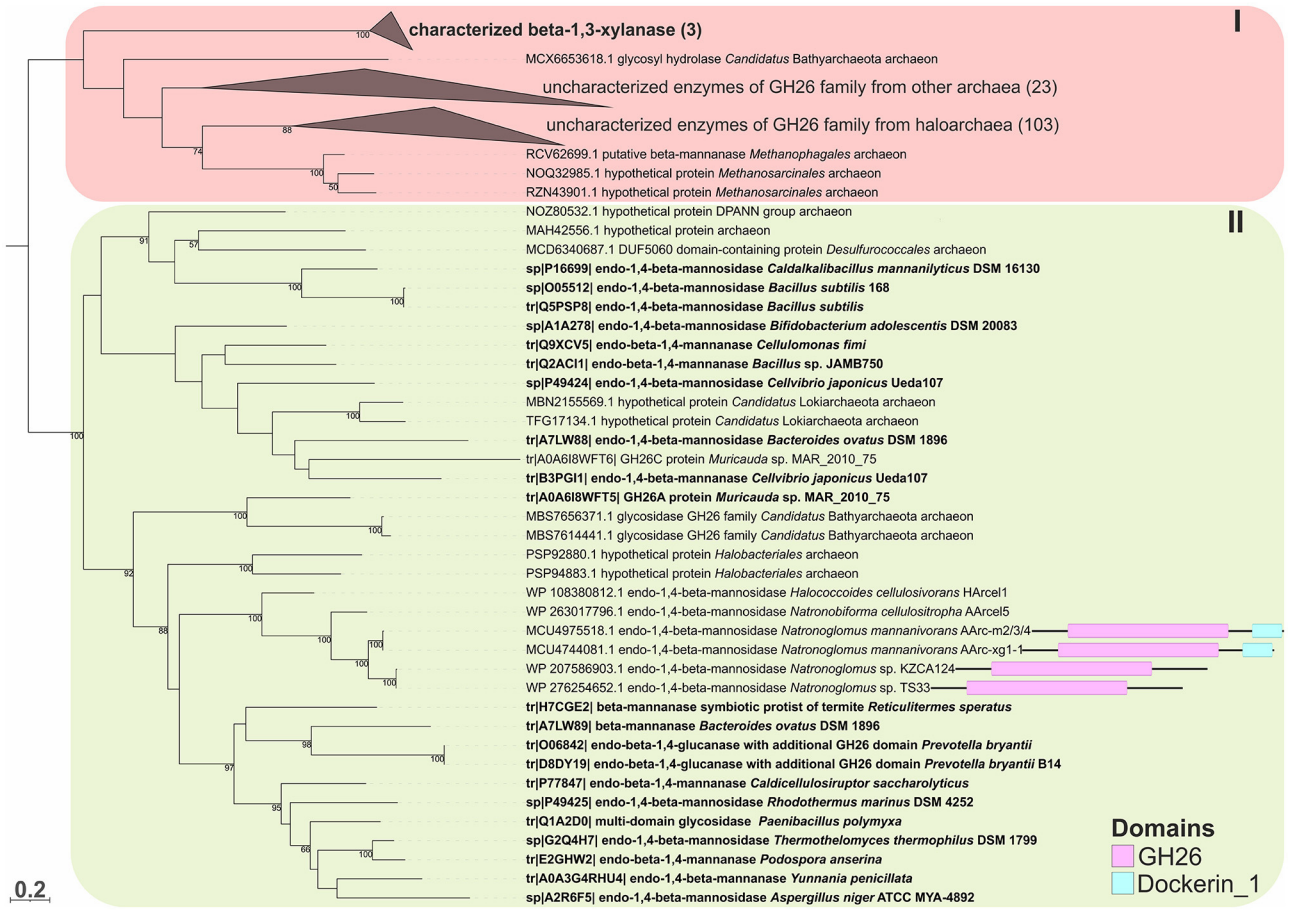


FIGURE 5
The phylogenetic analysis of the GH26 family endo-beta-1,4-mannanases (with domain architecture) of strains AArc-m2/3/4^T and AArc-xg1-1 and closely related strains KZCA124 and TS33. Biochemically characterized enzymes are highlighted in bold. Domain architectures for endo-beta-1,4-mannosidases of AArc-m2/3/4, AArc-xg1-1 strains, and its closest relative are given. The numbers at nodes indicate the percentage of corresponding support values (only values higher than 50% are shown).

plants indicating that haloarcheal proteins should also possess endo-mannonolytic activity.

Other repertoire of polysaccharide hydrolases encoded in the genomes

In addition to the mannanase-cellulase-xylanase GH families, both sequenced genomes also encoded a large repertoire of other polysaccharide-active enzymes including putative PL families listed in Supplementary Tables S3 and S4. The major difference between these groups of hydrolases is cell localization: the former are generally all extracellular, while many proteins from the latter group are cytoplasmic. The putative substrate for such enzymes includes cellulose/xyloglucan (GH3, GH5, and GH9 families), xylan/arabinoxylan (GH3, GH10, GH11, GH51, GH67, and GH115 families) arabinogalactan (GH2, GH30, GH42, GH43, GH51, and GH154 families), starch (GH13 family), acidic polysaccharides, such as polygalacturonan (PL1, PL22, and GH28), rhamnogalacturonan (PL11, PL26, PL42, GH4, and GH105), and alginate, and their oligomers. However, growth and hydrolytic activity were only observed with arabinoxylan and xyloglucan.

Central sugar metabolism

Taking into account that the GH2 family beta-mannosidases are intracellular and the strains are unable to grow on mannose, it is plausible to assume that mannan is hydrolyzed to mannooligosaccharides outside the cells, and then, the oligomers are imported into the cells. The phosphotransferase system (PTS) transporter genes were not found in the genomes of AArc-m2/3/4^T and AArc-xg1-1 similar to cellulotrophic haloarchaea (Elcheninov et al., 2023). On the other hand, there were multiple genes of putative carbohydrate-specific ABC transporters (e.g., MCU497+ 1351–1354; 1378–1381; 3422–3426; 4795.1–4797.1; 4083–4081; and 5056–5061) in AArc-m2/3/4^T, some of which likely import mannooligosaccharides into the cells, and furthermore, they are hydrolyzed by the GH2 beta-mannosidases. It is likely that mannose is oxidized during the side activity of glucose dehydrogenase or putative oxidoreductase (MCU4975802) distantly homologous to mannose-specific aldohexose dehydrogenase from *Thermoplasma acidophilum* (Nishiya et al., 2004). Under the action of mannonate dehydratases ManD (the genes of two of them, 5517 and 5522, were co-located in the gene cluster with mannanases), mannose is converted

10.3389/fmicb.2024.1364606

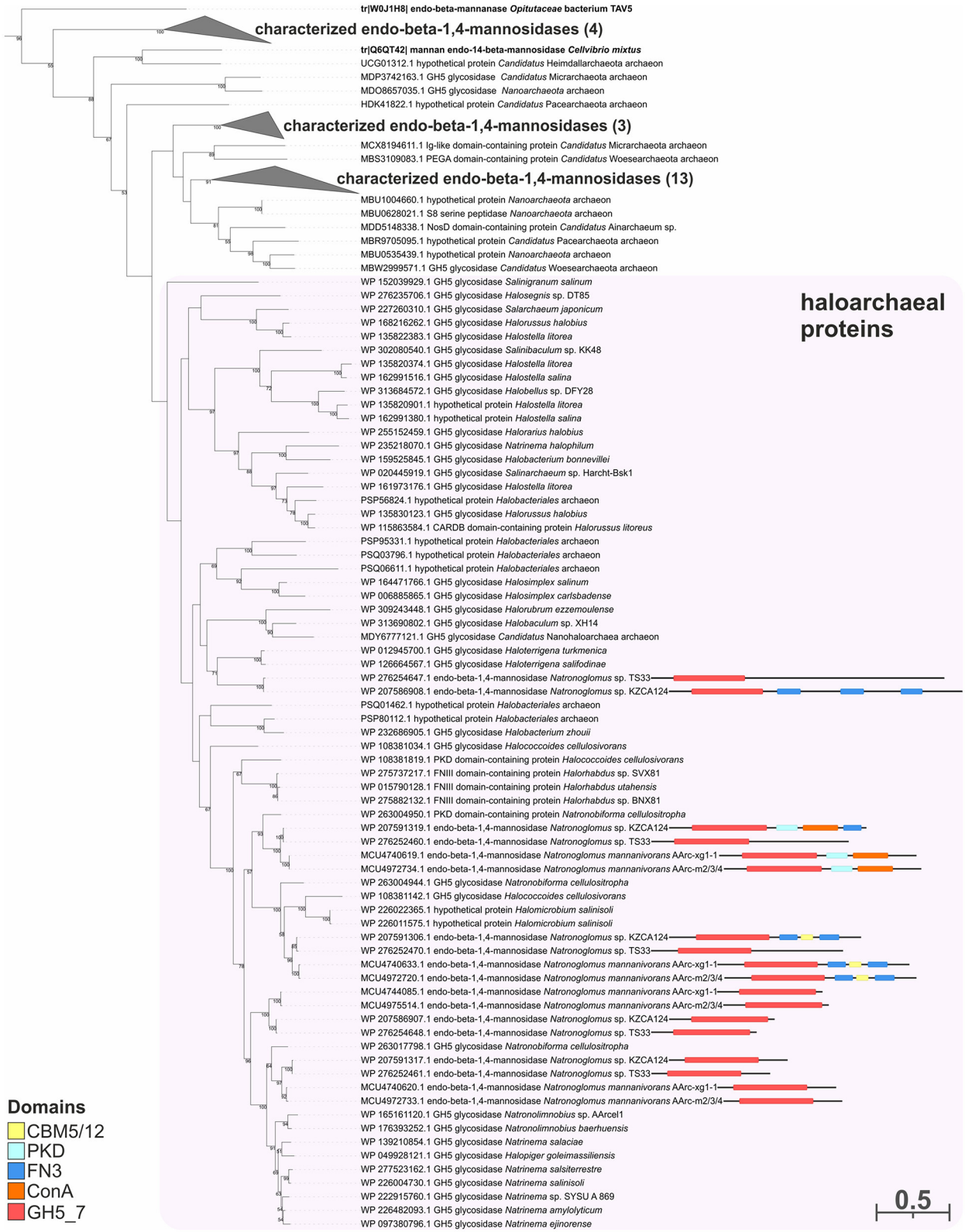


FIGURE 6
The phylogenetic analysis of enzymes of strains AArc-m2/3/4 and AArc-xg1-1 belonging to the GH5 family (the GH5_7 subfamily). Biochemically characterized enzymes are highlighted in bold. Domain architectures for endo-beta-1,4-mannosidases of AArc-m2/3/4, AArc-xg1-1 strains, and its closest relative (KZCA124 and TS33) are given. The numbers at nodes indicate the percentage of corresponding support values (only values higher than 50% are shown).

10.3389/fmicb.2024.1364606

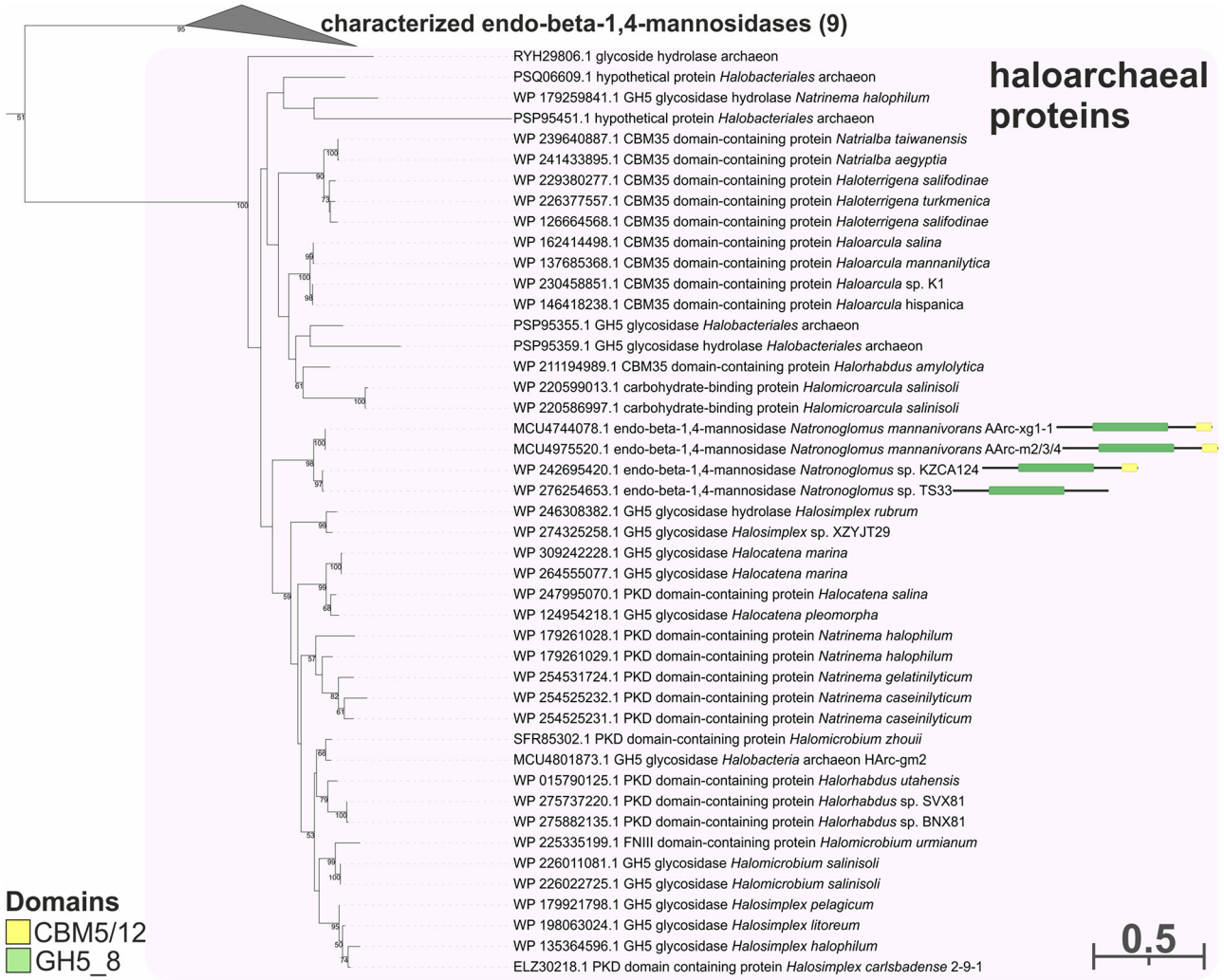


FIGURE 7
The phylogenetic analysis of enzymes of strains AArc-m2/3/4 and AArc-xg1-1 belonging to the GH5 family (the GH5_8 subfamily). Domain architectures for endo-beta-1,4-mannosidases of AArc-m2/3/4, AArc-xg1-1 strains, and its closest relative (KZCA124 and TS33) are given. The numbers at nodes indicate the percentage of corresponding support values (only values higher than 50% are shown).

to 2-dehydro-3-deoxy-D-gluconate. Finally, it enters into the semi-phosphorylative Entner–Doudoroff (ED) pathway with the generation of glyceraldehyde-3-phosphate and pyruvate (Figure 8). Genes encoding enzymes of terminal steps in the ED pathway were also present: 2-dehydro-3-deoxy-D-gluconate kinase (3928 and 5533) and 2-dehydro-deoxy-6-phosphogluconate aldolase (2034).

Other metabolically important traits

Other functionally important proteins from the genome of AArc-m2/3/4^T are listed in Supplementary Table S5, including the ion-pH homeostasis, energy-related respiratory complexes, and some others. In particular, an incomplete denitrification pathway is encoded, including Cu-nitrite reductase NirK, archaeal type of NO-reductase qNor, and N_2O reductase, but lacks nitrate reductase Nar. Despite this, the organisms did not grow anaerobically with either nitrite or N_2O as acceptors (with cellobiose as substrate).

Taxonomy conclusion

Comparative properties of the novel isolates in relation to the type species of the most related genera are shown in Table 2. The most similar in properties from the three related genera is obviously *N. cellulositropha*, but it was unable to grow with glucomannan and galactomannans, and its sugar utilization spectrum is much more restricted. The other two species of the related genera (*Natronosalvus* and *Saliphagus*) lack the beta-mannanase genes in their genomes and, thus, most likely would be able to grow with this polysaccharide type. In addition, *S. infecundisoli* is a neutraphilic haloarchaeon unable to grow at high pH. Furthermore, the mannan-utilizing natronoarchaea nearly lacked glycolipids, while in *Natronosalvus* and *Saliphagus*, several glycolipids constituted a large fraction of the membrane polar lipids, and the latter also has sulfolipid PGS not usually present in the soda lake natronoarchaea (Bale et al., 2021).

10.3389/fmicb.2024.1364606

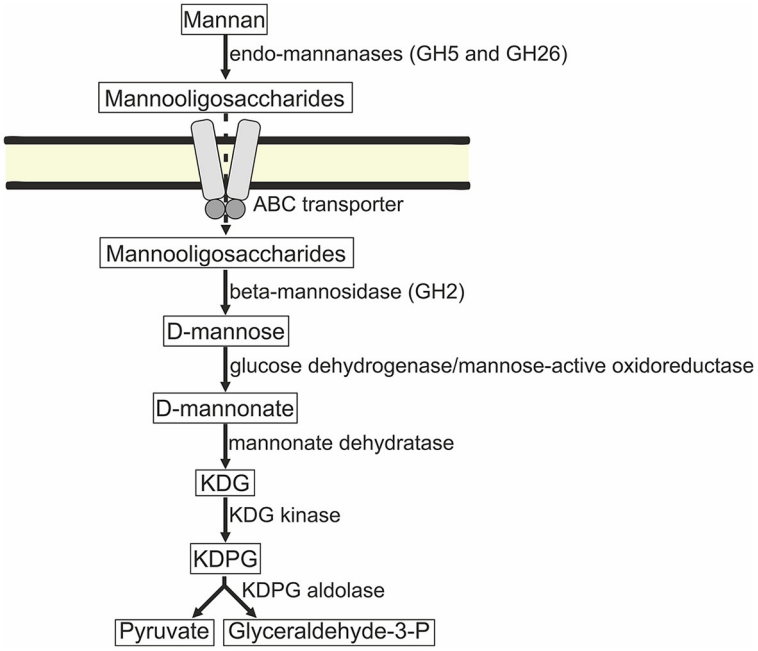


FIGURE 8
Predicted mannose catabolism pathway in strains AArc-m2/3/4 and AArc-xg1-1. ABC, ATP-binding cassette; KDG, 2-dehydro-3-deoxy-D-gluconate; KDPG, 2-dehydro-deoxy-6-phosphogluconate.

In conclusion, taking into account the unique functional specialization and distant phylogeny, the beta-mannan-utilizing natronoarchaea from Siberian hypersaline soda lakes are proposed to form a new genus and species *Natronoglomus mannanivorans.*

Description of *Natronoglomus mannanivorans* gen. nov., sp. nov.

Natronoglomus gen. nov.

Na.tro.no.glo'mus. Gr. neut. n. *natron* arbitrarily derived from Arabic n. *natrun* or *natron,* soda; L. neut. n. *glomus*, a ball; N.L. neut. n. *Natronoglomus,* a coccoid natronoarchaeon.

Natrononoarchaea from hypersaline lakes with mostly coccoid cells. Aerobic organoheterotrophs with the ability to utilize insoluble beta-mannans and cellulose as growth substrates. Extremely halophilic and moderately alkaliphilic. The dominant polar membrane lipids are phosphatidylglycerophosphate methyl ether (PGP-Me) and phosphatidylglycerol (PG) with C_{20}-C_{20} and C_{20}-C_{25} archaeol cores. The genus belongs to the *Natrialbaceae* family, class *Halobacteria*. A three-letter abbreviation is *Ngm*.

Natronoglomus mannanivorans sp. nov.

man.na.ni.vo'rans. N.L. neut. n. *mannanum*, mannan; L. pres. part. *vorans*, devouring; N.L. part. adj. *mannanivorans*, mannan devouring.

The cells are mostly non-motile cocci of 0.8–1.2 μm producing red pigments. The cells grown on mannan have a thick cell wall and extended pseudoperiplasm. The cells lyze in distilled water. The colony morphology varied depending on the substrate. On insoluble polysaccharides, the colonies were mostly pale-pink, thin, spreading, and irregular, up to 4 mm; on soluble sugars, the colonies were pink, more regular, and convex, up to 3 mm. The core membrane diether lipids are dominated by C_{20}-C_{20} DGE (archaeol) and C_{20}-C_{25} DGE (extended archaeol). The polar lipid head groups include phosphatidylglycerolphosphate methyl ester (PGP-Me) and phosphatidylglycerol (PG). The dominant respiratory menaquinone is MK-8:8, with the MK-8:7 second in abundance. The species are obligatory aerobic saccharolytic heterotrophs that are able to grow with insoluble beta-1,4-mannan, glucomannan, and cellulose. Sugars supporting the growth include galactose, rhamnose, raffinose, lactose, sucrose, cellobiose, maltose, trehalose, melezitose, and melibiose. Weak growth was observed on glycerol and pyruvate. Organic compounds tested (in the type strain) but not utilized included fructose, arabinose, ribose, xylose, sugar alcohols, acetate, ethanol, lactate, succinate, fumarate, malate, glycine, aspartate, glutamate, arginine, yeast extract, and peptone from casein. Ammonium and urea serve as the nitrogen source. Oxidase and catalase are positive. Mesophilic, with a maximum growth temperature of 48°C, is a low Mg-demanding, extreme halophile, with a range of Na^+ for growth from 2.5 to 4.5 M (optimum at 3.5 M), and a facultative alkaliphile, with a pH range from 6.8 to 9.6 for growth (optimum at 9–9.2). The G + C content of the DNA is 62% (two genomes), and the habitat is hypersaline salt lakes. The type strain (AArc-m2/3/4^T = JCM 34861 = UQM 41565) was isolated from sediments of hypersaline soda lakes in the Kulunda Steppe (Altai, Russia). The species also includes other four closely related isolates, one of which is deposited in JCM (AArc-xg1-1 = JCM 34866). The draft genome assemblies' accession numbers of strains AArc-m2/3/4^T and AArc-xg1-1 in the GenBank are GCA_025517485 and GCA_025517495.

TABLE 2 Comparative properties of the β-mannan utilizing isolates with the closest related genera from the family *Natrialbaceae* (Sorokin et al., 2018; Yin et al., 2018; Tan et al., 2023).

Property	*"Natronoglomus mannanivorans"*	*Natronobiforma cellulositropha*	*Natronosalvus amylolyticus*	*Saliphagus infecundisoli*
The number of isolates	5	4	1	2
Cell morphology	Non-motile cocci	Motile flat rods or non-motile cocci	Non-motile rods or angular coccoids	Non-motile cocci
Pigmentation	Red	Pink	Red	Pink-yellowish
Anaerobic growth	–	–	–	–
Substrates for aerobic growth				
Polysaccharides	Beta-mannans, cellulose, xylan, xyloglucan, and arabinoxylan	Cellulose, xylan, and beta-mannan	Starch	Starch
Sugars	Galactose, lactose, rhamnose, raffinose, sucrose, cellobiose, maltose, trehalose, melezitose, and melibiose	Cellobiose and maltose	Mannose, galactose, and sucrose	Glucose, mannose, raffinose, sucrose, maltose, and atrehalose
Others	–	–	Pyruvate, alanine, and ornithine	Pyruvate, succinate, glutamate, aspartate, lysine, and ornithine
Beta-mannanase genes	+ (6)	+(4)	–	–
Amylase	–	–	+	+
Esterase/lipase	– (tributyrin)	– (tributyrin)	– (Tween-80)	+ (Tween-20)
Protease	– (casein)	– (casein)	+ (casein)	+ (casein)
Catalase/oxidase	+/+	+/+	nd	+/+
Indole from tryptophane	+	–	–	–
Salinity range (optimum) M Na^+	2.0–4.5 (3.5)	2.5–4.8 (4.0)	0.9–4.8 (3.4)	2.0–6.0 (2.5–3.0)
pH range (optimum)	7.2–9.7 (9.0–9.2)	7.5–9.9 (8.5–9.0)	6.0–9.5 (8.0–8.5)	6.5–8.5 (7.0–7.5)
Temperature max (°C)	48 (at pH 8.5)	53 (at pH 8.5)	55*	50 (at optimum pH)
Core lipids (archaeols)	C_{20}-C_{20}, C_{20}-C_{25} DGE	C_{20}-C_{20}, C_{20}-C_{25} DGE	C_{20}-C_{20}, C_{20}-C_{25} DGE	nd
Intact polar lipids				
Phospholipids	PG, PGP-Me	PG, PGP-Me	PG, PGP-Me, PA	PG, PGP-Me, PGS
Glycolipids	MG (trace amount)	MG, DG	DGD-1, S-DGD-1	S-DGD-1
			S-TGD-1	
Respiratory lipoquinones	MK-8:8 (major)	nd	nd	nd
	MK-8:7 (minor)			
	MK-7:7 (trace amount)			
DNA G + C (%, genomic)	62.0 (2 strains)	65.5 (1 strain)	63.7	64.0
Isolation source	Inland hypersaline salt and soda lakes	Inland hypersaline soda lakes	Hypersaline alkaline lake	Saline soil

nd, not reported.

*The pH is not reported.

Lipids: PA, phosphatidic acid; PG, phosphatidylglycerol; PGP, phosphatidylglycerophosphate; PGP-Me, phosphatidyl-glycerophosphate methyl ester; PGS, phosphatidylglycerol sulfate; MG, phosphatidyl monoglycosyl diether; DG, phosphatidyl diglycosyl diether; DGD-1, mannosyl glycosyl diether; S-DGD-1, sulfated mannosyl glycosyl diether; S-TGD-1, sulfated galactosyl mannosyl glycosyl diether; and DGE, dialkyl glycerol ether.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary material.

Author contributions

DS: Conceptualization, Investigation, Writing—original draft. AE: Investigation, Writing—original draft. NB: Investigation, Writing—original draft. JS: Writing—original draft. IK: Writing—original draft.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. All Russian authors were supported by the Russian Ministry of Higher Education and Science. DS was also partially supported by the Gravitation-SIAM Program of the Dutch Ministry of Education and Sciences (grant no. 24002002). NB and JS-D acknowledge support from the ERC Horizon 2020 Program (grant agreement no. 694569 – MICROLIPIDS).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2024.1364606/full#supplementary-material

References

Aliyu, H., Lebre, P., Blom, J., Cowan, D., and De Maayer, P. (2016). Phylogenomic re-assessment of the thermophilic genus *Geobacillus*. *Syst. Appl. Microbiol.* 39, 527–533. doi: 10.1016/j.syapm.2016.09.004

Bale, N. J., Ding, S., Hopmans, E. C., Villanueva, L., and Boschman, R. C. (2021). Lipidomics of environmental microbial communities. I: visualization of specific niches using untargeted analysis of high-resolution mass spectrometry data. *Front. Microbiol.* 12:659302. doi: 10.3389/fmicb.2021.659302

Buchfink, B., Xie, C., and Huson, D. H. (2015). Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* 12, 59–60. doi: 10.1038/nmeth.3176

Capella-Gutiérrez, S., Silla-Martínez, J. M., and Gabaldón, T. (2009). trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. *Bioinformatics* 25, 1972–1973. doi: 10.1093/bioinformatics/btp348

Chaumeil, P. A., Mussig, A. J., Hugenholtz, P., and Parks, D. H. (2019). GTDB-Tk: a toolkit to classify genomes with the genome taxonomy. *Bioinformatics* 36, 1925–1927. doi: 10.1093/bioinformatics/btz848

de la Haba, R.R., Minegishi, H., Kamekura, M., Shimane, Y., Ventosa, A. (2021). Phylogenomics of *Haloarchaea*: the controversy of the genera *Natrinema-Haloterrigena*. *Front. Microbiol.* 12:740909. doi: 10.3389/fmicb.2021.740909

Elcheninov, A. G., Ugolkov, Y. A., Elizarov, I. M., Klyukina, A. A., Kublanov, I. V., Sorokin, D. Y., et al. (2023). Cellulose metabolism in halo(natrono)archaea: a comparative genomics study. *Front. Microbiol.* 14:1112247. doi: 10.3389/fmicb.2023.1112247

Enomoto, S., Shimane, Y., Ihara, K., Kamekura, M., Itoh, T., Ohkuma, M., et al. (2020). *Haloarcula mannanilytica* sp. nov., a galactomannan-degrading haloarchaeon isolated from commercial salt. *Int. J. Syst. Evol. Microbiol.* 70, 6331–6337. doi: 10.1099/ijsem.0.004535

Katoh, K., Rozewicki, J., and Yamada, K. D. (2017). MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization. *Brief. Bioinform.* 20, 1160–1166. doi: 10.1093/bib/bbx108

Letunic, I., and Bork, P. (2019). Interactive Tree Of Life (iTOL) v4: recent updates and new developments. *Nucleic Acids Res.* 47, W256–W259. doi: 10.1093/nar/gkz239

Mistry, J., Finn, R. D., Eddy, S. R., Bateman, A., and Punta, M. (2013). Challenges in homology search: HMMER3 and convergent evolution of coiled-coil regions. *Nucleic Acids Res.* 41:e121. doi: 10.1093/nar/gkt263

Nishiya, Y., Tamura, N., and Tamura, T. (2004). Analysis of bacterial glucose dehydrogenase homologs from thermoacidophilic archaeon *Thermoplasma acidophilum*: finding and characterization of aldohexose dehydrogenase. *Biosci. Biotechnol. Biochem.* 68, 2451–2456. doi: 10.1271/bbb.68.2451

Orata, F. D., Meier-Kolthoff, J. P., Sauvageau, D., and Stein, L. Y. (2018). Phylogenomic analysis of the gammaproteobacterial methanotrophs (order *Methylococcales*) calls for the reclassification of members at the genus and species levels. *Front. Microbiol.* 9:3162. doi: 10.3389/fmicb.2018.03162

Paysan-Lafosse, T., Blum, M., Chuguransky, S., Grego, T., Pinto, B. L., Salazar, G. A., et al. (2023). InterPro in 2022. *Nucleic Acids Res.* 51, D418–D427. doi: 10.1093/nar/gkac993

Qin, Q.-L., Xie, B.-B., Zhang, X.-Y., Chen, X.-L., Zhou, B.-C., Zhou, J., et al. (2014). A proposed genus boundary for the prokaryotes based on genomic insights. *J. Bacteriol.* 196, 2210–2215. doi: 10.1128/JB.01688-14

Rinke, C., Chuvochina, M., Mussig, A. J., Chaumeil, P.-A., Davín, A. A., Waite, D. W., et al. (2021). A standardized archaeal taxonomy for the Genome Taxonomy Database. *Nat. Microbiol.* 6, 946–959. doi: 10.1038/s41564-021-00918-8

Shimane, Y., Hatada, Y., Minegishi, H., Mizuki, T., Echigo, A., Miyazaki, M., et al. (2010). *Natronoarchaeum mannanilyticum* gen. nov., sp. nov., an aerobic, extremely halophilic archaeon isolated from commercial salt. *Int. J. Syst. Evol. Microbiol.* 60, 2529–2534. doi: 10.1099/ijs.0.016600-0

Sorokin, D. Y., Elcheninov, A. G., Khijniak, T. V., Kolganova, T. V., and Kublanov, I. V. (2022a). Selective enrichment on a wide polylsaccharide spectrum allowed isolation of novel metabolic and taxonomic groups of haloarchaea from hypersaline lakes. *Front. Microbiol.* 13:1059347. doi: 10.3389/fmicb.2022.1059347

Sorokin, D. Y., Elcheninov, A. G., Toshchakov, S. V., Bale, N. J., Sinninghe Damsté, J. S., Khijniak, T. V., et al. (2019b). *Natrarchaeobius chitinivorans* gen. nov., sp. nov., *Natrarchaeobius halalkaliphilus* sp. nov., alkaliphilic, chitin-utilizing haloarchaea from hypersaline alkaline lakes. *Syst. Appl. Microbiol.* 42, 309–318. doi: 10.1016/j.syapm.2019.01.001

Sorokin, D. Y., Elcheninov, A. S., Merkel, A. Y., Bale, N. J., Sinninghe Damsté, J. S., Kublanov, I. V., et al. (2023). *Halapricum hydrolyticum*, a beta-1,3-glucan utilizing haloarchaeon from hypersaline lakes. *Syst. Appl. Microbiol.* 46:126471. doi: 10.1016/j.syapm.2023.126471

Sorokin, D. Y., Khijniak, T. V., Kostrikina, N. A., Elcheninov, A. G., Toshchakov, S. V., Bale, N. J., et al. (2018). *Natronobiforma cellulositropha* gen. nov., sp. nov., a novel haloalkaliphilic member of the family Natrialbaceae (class Halobacteria) from hypersaline alkaline lakes. *Syst. Appl. Microbiol.* 41, 355–362. doi: 10.1016/j.syapm.2018.04.002

Sorokin, D. Y., Khijniak, T. V., Kostrikina, N. A., Elcheninov, A. G., Toshchakov, S. V., Bale, N. J., et al. (2019a). *Halococcoides cellulosivorans* gen. nov., sp. nov., an extremely halophilic cellulose-utilizing haloarchaeon from hypersaline lakes. *Int. J. Syst. Evol. Microbiol.* 69, 1327–1335. doi: 10.1099/ijsem.0.003312

Sorokin, D. Y., Toshchakov, S. V., Kolganova, T. V., and Kublanov, I. V. (2015). Halo(natrono)archaea isolated from hypersaline lakes utilize cellulose and chitin as growth substrates. *Front. Microbiol.* 6:942. doi: 10.3389/fmicb.2015.00942

Sorokin, D. Y., Yakimov, M. M., Messina, E., Merkel, A. Y., Koenen, M., Bale, N. J., et al. (2022b). *Natranaeroarchaeum sulfidigenes* gen. nov., sp. nov., carbohydrate-utilizing sulfur-respiring haloarchaeon from hypersaline soda lakes, a member of a new family Natronoarchaeaceae fam. nov. in the order Halobacteriales. *Syst. Appl. Microbiol.* 45:126356. doi: 10.1016/j.syapm.2022.126356

Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* 30, 1312–1313. doi: 10.1093/bioinformatics/btu033

Tan, S., Cheng, M., Li, X.-X., Hu, Y., Ma, X., Hou. J., et al. (2023). *Natronosalvus halobius* gen. nov., sp. nov., *Natronosalvus caseinilyticus* sp. nov., *Natronosalvus vescus* sp. nov., *Natronosalvus rutilus* sp. nov., *Natronosalvus amylolyticus* sp. nov., halophilic archaea isolated from salt lakes and soda lakes. *Int. J. Syst. Evol. Microbiol.* 73:006036. doi: 10.1099/ijsem.0.006036

Wirth, J. S., and Whitman, W. B. (2018). Phylogenomic analyses of a clade within the *Roseobacter* group suggest taxonomic reassignments of species of the genera *Aestuariivita, Citreicella, Loktanella, Nautella, Pelagibaca, Ruegeria, Thalassobius, Thiobacimonas and Tropicibacter*, and the proposal. *Int. J. Syst. Evol. Microbiol.* 68, 2393–2411. doi: 10.1099/ijsem.0.002833

Yin, X.-Q., Liu, B.-B., Chu, X., Salam, N., Li, X., Wang, Z.-W., et al. (2018). *Saliphagus infecundisoli* gen. nov., sp. nov., an extremely halophilic archaeon isolated from a saline soil. Int. J. Syst. Evol. Microbiol. 67, 4154–4160. doi: 10.1099/ijsem.0.002270

Zhang, H., Yohe, T., Huang, L., Entwistle, S., Wu, P., Yang, Z., et al. (2018). dbCAN2: a meta server for automated carbohydrate-active enzyme annotation. *Nucleic Acids Res.* 46, W95–W101. doi: 10.1093/nar/gky418

frontiers | Frontiers in Microbiology

TYPE Original Research
PUBLISHED 30 May 2024
DOI 10.3389/fmicb.2024.1403623

OPEN ACCESS

EDITED BY
Mohammad Ali Amoozegar,
University of Tehran, Iran

REVIEWED BY
Satya P. Singh,
Saurashtra University, India
Adele Williamson,
University of Waikato, New Zealand

*CORRESPONDENCE
Ram Karan
✉ ramkaran@south.du.ac.in
Dominik Renn
✉ dominik.renn@kaust.edu.sa

RECEIVED 19 March 2024
ACCEPTED 15 May 2024
PUBLISHED 30 May 2024

CITATION
Karan R, Renn D, Allers T and Rueping M (2024) A systematic analysis of affinity tags in the haloarchaeal expression system, *Haloferax volcanii* for protein purification.
Front. Microbiol. 15:1403623.
doi: 10.3389/fmicb.2024.1403623

A systematic analysis of affinity tags in the haloarchaeal expression system, *Haloferax volcanii* for protein purification

Ram Karan[1,2]*, Dominik Renn[2]*, Thorsten Allers[3] and Magnus Rueping[2,4]

[1]Department of Microbiology, University of Delhi, South Campus, New Delhi, India, [2]King Abdullah University of Science and Technology (KAUST), KAUST Catalysis Center, Thuwal, Makkah, Saudi Arabia, [3]School of Life Sciences, University of Nottingham, Queen's Medical Centre, Nottingham, United Kingdom, [4]Institute for Experimental Molecular Imaging, University Clinic, RWTH Aachen University, Aachen, Germany

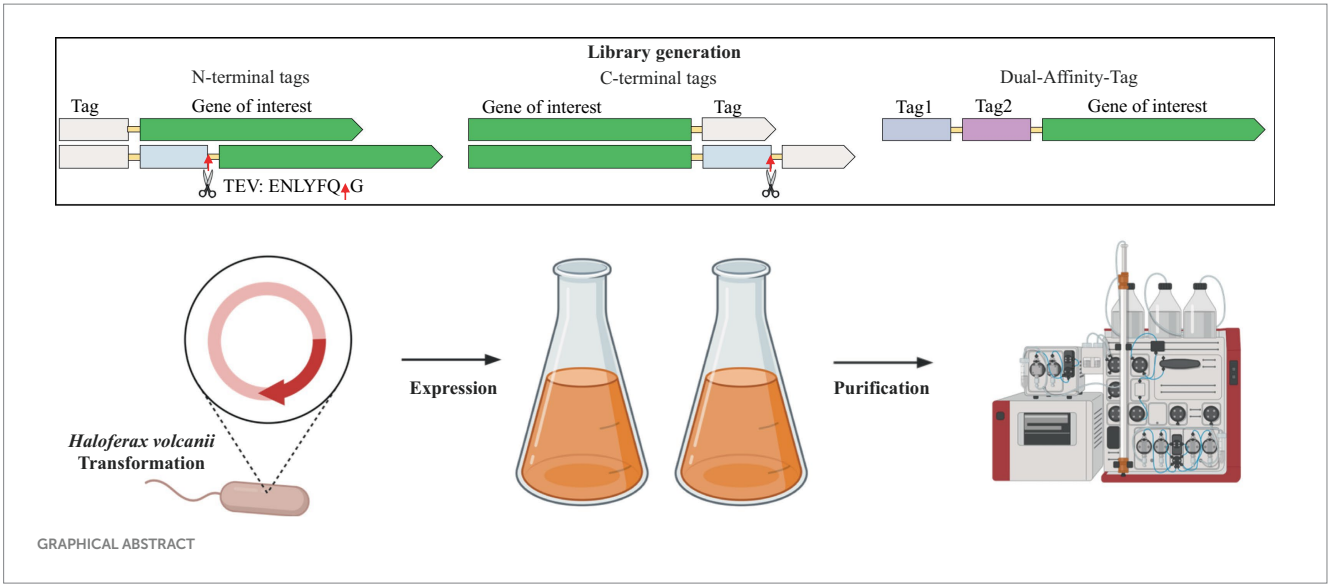
Extremophilic proteins are valuable in various fields, but their expression can be challenging in traditional hosts like *Escherichia coli* due to misfolding and aggregation. *Haloferax volcanii* (*H. volcanii*), a halophilic expression system, offers a solution. This study examined cleavable and non-cleavable purification tags at both the N- and C-termini when fused with the superfolder green fluorescent protein (sfGFP) in *H. volcanii*. Our findings reveal that an N-terminal 8xHis-tag or Strep-tag®II significantly enhances protein production, purity, and yield in *H. volcanii*. Further experiments with mCherry and halophilic alcohol dehydrogenase (ADH) showed improved expression and purification yields when the 8xHis-tag or Strep-tag®II was positioned at the C-terminus for mCherry and at the N-terminus for ADH. Co-positioning 8xHis-tag and Twin-Strep-tag® at the N-terminus of sfGFP, mCherry, and ADH yielded significantly enhanced results. These findings highlight the importance of thoughtful purification tag design and selection in *H. volcanii*, providing valuable insights for improving protein production and purification with the potential to advance biotechnological applications.

KEYWORDS

halophiles, archaea, extremophiles, purification tags, His-tag, Strep-tag®II, *Haloferax volcanii*, bioengineering

Introduction

Extremophiles thrive in extreme physical or geochemical conditions that are inhospitable to most life forms. These environments encompass extreme temperatures, significant variations in salinity, high pressures, the presence of heavy metals, and drastic pH levels (Sysoev et al., 2021). Extremophilic proteins hold a unique position in biotechnology due to their exceptional stability under harsh environmental conditions (Wang et al., 2024). They contribute to sustainable biotechnology on Earth and offer valuable insights into the potential existence of life on other planets (Allers, 2010; Schultz et al., 2023; Wang et al., 2024). The efficient expression and cost-effective purification of recombinant proteins are paramount in today's pharmaceutical and biotechnology industries (Puetz and Wurm, 2019). While *Escherichia coli* (*E. coli*) expression systems are widely adopted for producing heterologous recombinant



proteins, they come with inherent limitations, including a limited capacity for forming disulfide bonds, the inability to perform posttranslational modifications, and the absence of an efficient secretion system (Baeshen et al., 2015; Ibrahim et al., 2023). Extremophilic proteins, which are adapted to extremes of temperature, pH, or salinity, often elude successful expression in mesophilic systems like *E. coli* due to issues of protein misfolding, conformational stress, and the formation of inclusion bodies (Elleuche et al., 2014; Kruglikov et al., 2022). Consequently, there is a pressing demand for alternative expression systems that can effectively accommodate the expression of extremophilic proteins in their native state.

Haloferax volcanii (*H. volcanii*) offers a solution to these challenges as an extremophilic protein expression host. *H. volcanii* is an obligate halophilic archaeon, originally isolated from the Dead Sea, which naturally thrives in extreme conditions (Strillinger et al., 2016; Grötzinger et al., 2018; Akal et al., 2019; Sysoev et al., 2021). *H. volcanii* has been equipped with a suite of microbiological and molecular genetics techniques, including an efficient DNA transformation system, shuttle plasmids, and selectable markers, making it a versatile platform for genetic studies, proteomics, and biotechnology research (Allers, 2010; Strillinger et al., 2016; Haque et al., 2019; Pérez-Arnaiz et al., 2020; Karan et al., 2023). Moreover, *H. volcanii* grows, compared to other extreme halophiles, under laboratory culture conditions at high density and in a wide range of temperatures (37–55°C) and salt concentrations (1.8–3.5 M NaCl) (Haque et al., 2019, 2020; Sysoev et al., 2021). Due to the fragile nature of the outer S-layer cell walls, *H. volcanii cells easily* lyse in hypotonic conditions, such as low salt buffer or water, releasing all cellular proteins (Dyall-Smith, 2008). This allows a simplified downstream protein purification with minimal labor and cost, making *H. volcanii* attractive for biotechnology use (Grötzinger et al., 2018; Akal et al., 2019; Karan et al., 2020, 2023; Alshehri et al., 2022).

The *H. volcanii* strain H1895, coupled with the pTA1992 plasmid, is a dedicated host with genetic modifications for enhanced protein overexpression and purification (Haque et al., 2020). Several genetic modifications have been made to *H. volcanii* strain H1895 to optimize it as an expression host for protein overexpression. Notably, biofilm-forming genes HVO_1033 and HVO_1034 have been intentionally deleted. Additionally, naturally histidine-rich proteins PitA and Cdc48d have been replaced with orthologous genes containing fewer histidine residues (Allers et al., 2010; Strillinger et al., 2016; Haque et al., 2019, 2020). These changes serve a dual purpose: they reduce potential contaminants and enable polyhistidine-tag (His-tag) purification, a common method in protein purification. Moreover, H1895 has proven to be an excellent platform for expressing various enzymes and larger molecular structures (Allers, 2010; Strillinger et al., 2016; Haque et al., 2019, 2020; Karan et al., 2023). Furthermore, the lack of biofilm formation extends its utility to bioreactor applications, further underlining its versatility and potential for biotechnological endeavors (Strillinger et al., 2016; Haque et al., 2020).

Functional protein expression and purification are essential for exploring protein structure, function, and biotechnological applications (Du et al., 2022). Therefore, applying appropriate purification tags that enhance solubility and are cleavable in protein production is crucial. These tags help achieve high expression yields, target protein purity, save time, and reduce production costs (Gomari et al., 2020). However, it is important to note that the construction of fusion proteins can often affect the expression level, solubility, stability, and subsequent purification efficiency of recombinant proteins (Yadav et al., 2016; Köppl et al., 2022; Le et al., 2022). Even in well-studied expression systems like *E. coli*, there is no universal rule for determining the optimal expression conditions, purification tag types, and positions (Köppl et al., 2022). Regarding haloarchaeal proteins, expression levels are typically low, and the procedures for purifying these proteins can be complex and costly (Martínez-Espinosa, 2019). Systematic studies are missing to investigate the effects of purification tag positions and the influence of commonly used affinity and

Abbreviations: sfGFP, superfolder green fluorescent protein; His6 or his8 tag, hexa or octa histidine-tag, halophilic alcohol dehydrogenase (ADH); IMAC, immobilized metal affinity chromatography; SDS-PAGE, sodium dodecyl sulfate-polyacrylamide gel electrophoresis; SUMO, small ubiquitin-like modifier; TEV, tobacco etch virus; *Haloferax volcanii*, *H. volcanii*; *Escherichia coli*, *E. coli*.

solubility-enhancing tags on expression, solubility, and purification in haloarchaeal expression systems.

This study systematically investigates the impact of various purification tags on protein expression, solubility, and subsequent purification of *H. volcanii* expressed proteins. Consequently, fusion constructs using the superfolder green fluorescent protein (sfGFP) and a range of common purification tags (polyhistidine-tag or His-tag, Strep-tag®II, Twin-Strep-tag®, a FLAG-tag, and 3x-FLAG-tag), both at the N- and C-termini, with and without cleavage sites were designed (Skerra and Schmidt, 1999; Einhauer and Jungbauer, 2001; Kimple and Sondek, 2004; Pédelacq et al., 2006; Shahravan et al., 2008; Allers et al., 2010; Hermans et al., 2012; Schmidt et al., 2012, 2013; Young et al., 2012; Beznosov et al., 2013; Kimple et al., 2013; Zhao et al., 2013; Djender et al., 2014; Nguyen et al., 2014; Yeliseev et al., 2017; Mishra, 2020; Köppl et al., 2022). Next, the study also investigates the novel four amino acid short C-tag at the C-terminus of sfGFP to diversify the tag library further. Fluorescence measurements of sfGFP enable rapid and efficient expression, solubility, and purification estimations of these fusion proteins (Pédelacq et al., 2006). Furthermore, a comparison was conducted using 8xHis-tag and Strep-tag®II at both the N- and C-termini with the red fluorescent protein mCherry and a previously fully characterized halophilic alcohol dehydrogenase (ADH) from the deep-sea brine pool of the Red Sea (Shaner et al., 2004; Grötzinger et al., 2018; Akal et al., 2019). Additionally, the dual-affinity-tag approach by combining the 8xHis-tag and the Strep-tag®II at the N-terminus of sfGFP, mCherry, and ADH was explored. These investigations unveiled the substantial impact of different tags, tag lengths, and positions on protein expression and purification. The insights gained from this study hold significant promise in addressing the challenges associated with producing and purifying halophilic proteins in *H. volcanii*.

Materials and methods

Materials

All chemicals and solvents were purchased from Sigma-Aldrich (St. Louis, MO, United States). Water was desalted and purified using a milli-Q® (Merck, Darmstadt, Germany) system. This study used *Haloferax volcanii* H1895 and the vector pTA1992 (Haque et al., 2019).

Plasmid preparation

The purification-tag-fusion-sfGFP protein synthetic gene library (GenScript Biotech Corporation, HK), together with the mCherry and ADH constructs were codon-optimized using the java codon adaptation online tool Jcat for *Halobacterium* sp. (strain NRC-1/ATCC 700922/JCM 11081) (Grote et al., 2005) and cloned into the pTA1921 vector via NdeI and BamHI restriction enzymes (Supplementary Figure S1 and Supplementary Table S1).

Protein expression in *Haloferax volcanii*

The construct containing vectors was transformed into the *Haloferax volcanii* H1895 using the PEG/EDTA method (Dyall-Smith, 2008). *H. volcanii* and derivatives were cultured in the Hv-YPC medium at 45°C with shaking for 24–36 h, as previously described (Strillinger et al., 2016; Grötzinger et al., 2018; Akal et al., 2019; Karan et al., 2020; Alshehri et al., 2022). For solid media, 2% (w/v) agar was added. Stock cultures were maintained in glycerol at −80°C. For short-term use, cultures were maintained on stock plates.

Estimation of expression level and solubility

The expression level of the sfGFP or mCherry constructs was evaluated by measuring the fluorescence (ex 485 nm/em 507 nm) or mCherry (ex 587 nm/em 610 nm) of 24 h grown cell culture/OD_{600} with three different colonies. The expression level of the ADH constructs was evaluated by determining the final protein concentration after purification using NanoDrop absorption at 280 nm. The results were transformed into relative expression, with the highest fluorescence (for sfGFP N-ter Strep-tag, mCherry C-ter 8xHis-tag, and ADH N-ter 8xHis-tag) set to 100%.

The solubility of sfGFP and mCherry was evaluated by comparing the fluorescence after sonication, before and after centrifugation. For ADH, which lacks fluorescence, solubility was evaluated using SDS-PAGE and Western blotting to compare the soluble fraction against the pellet.

Cell harvesting and lysis

One liter of cell culture was grown until an OD_{600} of 1.0–1.5. Cells were harvested by centrifugation (4,000 × g, 4°C, 45 min) in a Legend XFR Centrifuge (Thermo Scientific, Waltham, United States) and disrupted in binding buffer (50 mM sodium phosphate buffer, pH 7.4 containing 200 mM or 2 M NaCl) containing Pierce™ Protease Inhibitor EDTA-free tablet (Thermo Scientific) using a sonicator (Model Q500, QSONICA, Newtown, CT, United States) with a 1.9 cm probe. Cell debris was removed by centrifugation (24,000 × g, 4°C, 45 min) in a Multifuge X1R Centrifuge (Thermo Scientific).

Purification of constructs

All constructs were purified using the ÄKTAprime plus chromatography system (GE Healthcare Life Sciences, Piscataway, NJ, United States) according to the purification tag in the binding buffer (50 mM sodium phosphate, pH 7.4, 200 mM or 2 M NaCl).

Purification of His-tag fused constructs

The polyHis and SUMO fused constructs were purified on a 1 mL HiTrap Ni^{2+} chelating column (GE Healthcare Life Sciences, Piscataway, NJ, United States). The supernatant was loaded at a 1.0 mL/min flow rate onto the column pre-equilibrated with a binding buffer (50 mM sodium phosphate, pH 7.4, 200 mM, or 2 M NaCl) containing 20 mM imidazole. After washing the column with binding buffer, the protein was eluted in one step with His-elution buffer (binding buffer containing 250 mM imidazole).

Purification of Strep-tag®II fused constructs

The Strep-tag®II and Twin-Strep-tag® tag fused constructs were purified on a 1 mL Strep-Tactin®XT 4Flow (GenScript Biotech Corporation, HK). The supernatant was loaded at a 1.0 mL/min flow rate onto the column pre-equilibrated with binding buffer (50 mM sodium phosphate, pH 7.4, 200 mM, or 2 M NaCl). After washing the column with binding buffer, the protein was eluted in one step with a strep-elution buffer (binding buffer containing 50 mM biotin).

Purification of FLAG-tag fused constructs

The FLAG-tag and 3X FLAG-tag fused constructs were purified on a 1 mL Anti-DYKDDDDK G1 Affinity Resin column (GenScript Biotech Corporation, HK). The supernatant was loaded at a 1.0 mL/min flow rate onto the column pre-equilibrated with binding buffer (50 mM sodium phosphate, pH 7.4, 200 mM, or 2 M NaCl). After washing the column with binding buffer, the protein was eluted in a one-step FLAG-elution buffer (100 mM Tris/HCl buffer, pH 12.0, 200 mM, or 2 M NaCl). To purify the FLAG/3xFLAG-tag fused constructs, MonoRab™ Anti-DYKDDDDK Magnetic Beads (Cat No. L00835, GenScript) were tested as well. The cell lysate was added to the pre-washed beads with bead binding buffer (50 mM sodium phosphate, pH 7.4, 200 mM or 2 M NaCl) and incubated on a shaker at 4°C for 2 h. After magnetic separation, the supernatant was removed, and the magnetic beads were washed with the bead binding buffer three times. Finally, the DYKDDDDK fused construct bound to the beads was eluted by bead elution buffer (100 mM Tris/HCl buffer, pH 12.0).

Purification of C-tag fused constructs

The C-tag fused constructs were purified on a 1 mL CaptureSelect™ C-tagXL column (Thermo Scientific, Waltham, United States). The supernatant was loaded at a 1.0 mL/min flow rate onto the column pre-equilibrated with binding buffer (50 mM sodium phosphate, pH 7.4, 200 mM, or 2 M NaCl). The protein was eluted in one step with a C-tag-elution buffer (50 mM sodium phosphate, pH 7.4, 2 M $MgCl_2$).

Supplementary steps for purification and analysis

The purified fractions were combined, further purified, and concentrated with Amicon® Ultra-4 Centrifugal Filter Units, 10 kDa. The protein was then dialyzed against dialysis buffer (20 mM sodium phosphate, pH 7.4, 200 mM, or 2 M NaCl). Protein concentration was determined using NanoDrop absorption at 280 nm.

Note: The purification procedures of each construct were not optimized (e.g., elution gradient, wash steps) to compare better the purity of sfGFP, respectively, mCherry and ADH, constructs for each purification tag.

Protein cleavage

For fusion constructs containing cleavable tags, the cleavage of the purification tags was performed overnight at 4°C in dialysis buffer (50 mM Tris-HCl, 1 mM DTT, pH 8.0) containing 200 mM NaCl or 2 M NaCl, with the respective SUMO-Protease, containing a His-tag, (ratio of 1:10 w/w) and TEV-Protease, containing a His-tag, (200 U for 1 mg of fusion protein), and then reapplied to the respective affinity column to remove the cleaved purification tag. The SUMO- and TEV-Protease were removed by the HiTrap Ni^{2+} chelating column (GE Healthcare Life Sciences, Piscataway, NJ, United States).

The FLAG tag contains an internal enterokinase (EK) cleavage site, which allows the removal of the affinity tag after purification. The eluted fractions were incubated with Enterokinase, containing a His-tag (Z03376-1, Genscript, 200 U for 1 mg of fusion protein) to remove the FLAG-tag. The Enterokinase was removed by the HiTrap Ni^{2+} chelating column (GE Healthcare Life Sciences, Piscataway, NJ, United States).

SDS-PAGE and western blotting analysis

Western blotting analysis was performed as previously described (DasSarma et al., 2013; Andar et al., 2017; Karan et al., 2020). The SDS-PAGE analysis was performed using Novex® Tris-glycine gels (4–20%, Invitrogen, Carlsbad, CA, United States). Proteins were then transferred to 0.45 µm nitrocellulose membranes (Millipore Corp., Boston, MA). Membranes were blocked for 15 min in Pierce Fast Blocking Buffer (Thermo Fisher Scientific) and then incubated in Anti-GFP Polyclonal Antibody (Thermo Fisher Scientific) at 1:2000 dilution, followed by three washing steps with TBST, then incubated with Goat anti-Rabbit IgG (H + L) Cross-Adsorbed Secondary Antibody, Alexa Fluor 488 (Invitrogen, Catalog # A-11008).

The purity of the proteins was evaluated using SDS-PAGE stained with Coomassie brilliant blue and quantified with ImageJ software[1] (Schneider et al., 2012). Purity was specified as a percentage of the total protein content, with the intensity of the target protein band compared to the total protein intensity on each lane of the gel.

Tryptic digest and LC-MS/MS analysis

The identification of corresponding peptides was performed by LC-MS/MS analysis. 10 µg of the sample was digested with trypsin using the FASP protocol (Wiśniewski et al., 2009). Peptides were measured using an LTQ-Orbitrap mass spectrometer (Thermo Fisher Scientific, Waltham, MA, United States) and analyzed by MASCOT v2.3 (Matrix Sciences Ltd., United Kingdom).

Results and discussion

Experimental design for systematic screening

In the study, *H. volcanii* strain H1895 as the host, in combination with the pTA1992 plasmid that features the exceptionally strong constitutive p.syn promoter (Haque et al., 2019, 2020) was employed. Next, a versatile purification-tag-fusion-protein library using the sfGFP protein as a reporter (Pédelacq et al., 2006), was designed and

1 https://imagej.nih.gov/ij/

constructed. Within this library, a range of purification tags, including 6xHis, 8xHis, Strep-tag®II, Twin-Strep-tag®, FLAG, and 3x-FLAG, strategically placed at both the N-terminus and C-terminus of sfGFP (Figure 1) were introduced.

His-tags are the most used and cost-effective affinity tags to purify recombinant proteins using an immobilized metal affinity chromatography (IMAC) column (Porath et al., 1975). Notably, protein activity is rarely affected by polyhistidine affinity tags because of their relatively small size and charge (Block et al., 2009). In the context of *H. volcanii* strains where naturally histidine-rich proteins PitA and Cdc48d have been replaced with orthologous genes containing fewer histidine residue, the 6xHis-tag has been widely adopted for recombinant protein purification (Grötzinger et al., 2018; Haque et al., 2019, 2020; Karan et al., 2020). The Small Ubiquitin-like Modifier (SUMO)-tag, which incorporated either a 6xHis-tag or an 8xHis-tag for expression and purification (Butt et al., 2005) was used for the library design. Strep-tag®II, an 8-amino acid peptide tag (WSHPQFEK), exhibits a strong affinity for Strep-Tactin® (or Strep-Tactin®XT), a specially engineered streptavidin under physiological buffer conditions (Skerra and Schmidt, 1999; Schmidt et al., 2013; Yeliseev et al., 2017). This interaction allows for the rapid one-step purification of nearly any recombinant protein at high purity and functionality. Moreover, the Strep-tag®II can be found as a repetitive motif, generating a Twin-Strep-tag®. The Twin-Strep-tag®, characterized by its relatively higher affinity for Strep-Tactin® matrices, offers enhanced purification capabilities for recombinant proteins.

The FLAG-tag is a small, highly soluble tag consisting of a hydrophilic octapeptide epitope (DYKDDDDK) with an enterokinase-cleavage site (DDDDK). This tag's hydrophilic nature positions it primarily on the surface of the recombinant protein, increasing accessibility to resins and minimizing potential adverse effects on the fusion protein (Hopp et al., 1988; Schmidt et al., 2012). Similarly, the 3xFLAG-tag is a robust epitope tag (DYKDHDG-DYKDHDI-DYKDDDDK) was fused with sfGFP at the N- and C-termini.

A new class of purification tags, the four-peptide (E-P-E-A) short C-tag, which exhibits a strong affinity for a camelid antibody fragment when expressed at the C-terminus of a protein was also explored (Djender et al., 2014).

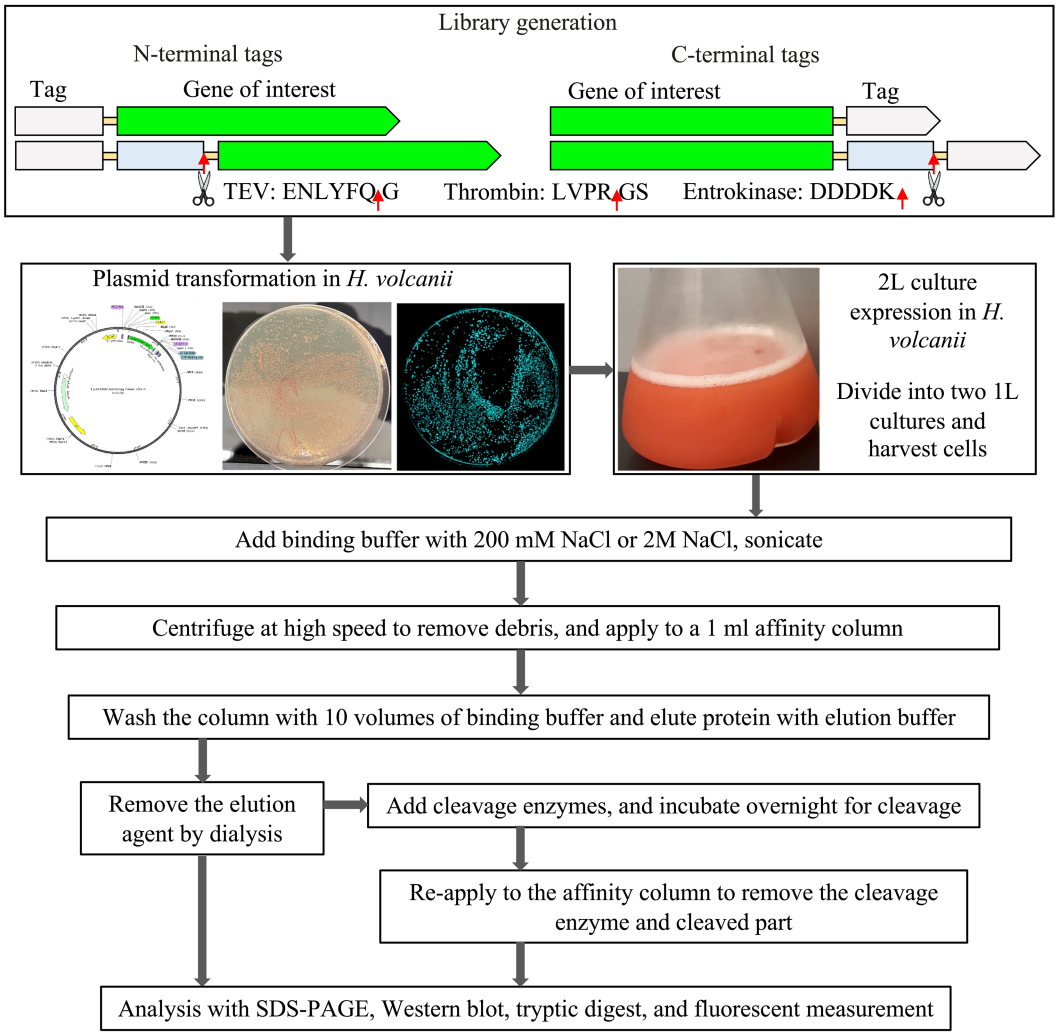


FIGURE 1
Schematic representation of constructs and purification process.

10.3389/fmicb.2024.1403623

TABLE 1 Relative expression, solubility, purity, and yield of sfGFP expressed with His-tag constructs.

Construct	Expression level (%)	Solubility (%)		Purity (%)		Yield (mg/L of culture)	
		Low salt	High salt	Low salt	High salt	Low salt	High salt
N-terminal purification His-tag/TEV cleavage site/ SUMO-tag							
1. 6xHis-sfGFP	51	97	94	97	75	2.4	2.8
2. 8xHis-sfGFP	95	98	93	90	85	4.9	3.5
5. 6xHis-TEV-sfGFP	49	100	92	90 (95)	85 (95)	2.5 (2.1)	2.3 (0.9)
6. 8xHis-TEV-sfGFP	75	100	96	90 (97)	85 (94)	2.8 (2.5)	1.7 (0.7)
9. 6xHis-SUMO-sfGFP	82	100	100	70 (85)	65 (80)	2.8 (2.4)	2.9 (2.1)
10. 8xHis-SUMO-sfGFP	87	100	99	65 (86)	75 (85)	3.3 (2.9)	3.6 (2.4)
C-terminal purification His-tag/TEV cleavage site							
15. sfGFP-6xHis	22	100	95	75	45	0.8	0.9
16. sfGFP-8xHis	23	100	97	70	40	0.8	1.1
20. sfGFP-TEV-6xHis	15	99	96	65 (96)	63 (90)	0.9 (0.6)	0.8 (0.4)
21. sfGFP-TEV-8xHis	17	100	97	80 (97)	85 (96)	0.8 (0.7)	0.7 (0.4)

The numbers in parentheses represent the values after cleavage and purification in the corresponding step.

We also scrutinized the effect of positioning the TEV cleavage site at the N-terminus and C-terminus of the sfGFP fusion constructs (Supplementary Table S1 and Supplementary Figure S1). A total of 26 distinct sfGFP constructs were studied, each evaluated for expression through fluorescence measurements (Supplementary Figure S2), further corroborated by Western blot analysis (Supplementary Figure S3). The different sfGFP constructs were compared for their relative expression level, solubility, and purification profile (Supplementary Figure S4). Tryptic digest analysis was additionally performed to confirm the identity of the fusion proteins (Supplementary Figure S5). The focus was primarily on purification at low salt (200 mM NaCl) and high salt concentrations (2 M NaCl). This comprehensive evaluation aimed to determine the most suitable purification tag for the native purification of extremophilic halophilic proteins, capitalizing on the suitability of *H. volcanii* as the expression host.

The initial systematic screening library was expanded to examine the applicability of the sfGFP results to the red fluorescent protein mCherry (Shaner et al., 2004) and a previously fully characterized halophilic alcohol dehydrogenase (ADH) from the deep-sea brine pool of the Red Sea (Grötzinger et al., 2018) (Supplementary Figure S7). These investigations revealed the significant impact of different tags, tag lengths, and positions on protein expression and purification. This resulted in a total of 37 screened constructs.

Expression and purification of sfGFP fusion proteins

The corresponding overview of the expression and purification of the sfGFP constructs is presented in the order of the different studied affinity purification tag systems.

Polyhistidine-tag (His-tag)

The experimental design explored the efficacy of both a 6xHis-tag and an 8xHis-tag at the N- and C-termini of sfGFP, both with and without a TEV cleavage site. Furthermore, the N-terminal solubility-enhancing and cleavable Small Ubiquitin-like Modifier (SUMO)-tag, which incorporated either a 6xHis-tag or an 8xHis-tag for expression and purification (Butt et al., 2005), was used. All the constructs underwent purification and cleavage, where applicable, under varying salt concentrations, ranging from low (200 mM NaCl) to high (2 M NaCl) in a 50 mM sodium phosphate buffer, pH 7.4. The expression level, solubility, purity, and final yield of the fusion sfGFP constructs are presented in Table 1 and Figure 1.

N-terminal vs. C-terminal and 6xHis- vs. 8xHis-tag

The expression of sfGFP is significantly influenced by the choice of tag and its location (Yilmaz and Arslanyolu, 2015). Notably, the highest level of sfGFP expression was achieved with the N-terminal 8xHis-tag (Table 1). Surprisingly, the introduction of the N-terminal 8xHis-SUMO-tag did not significantly enhance expression yield, and the addition of a cleavable TEV cleavage site had a noticeable impact on expression levels. Conversely, sfGFP expression with C-terminal His-tags demonstrated significantly lower levels, up to one-fifth, compared to the N-terminal 8xHis-tag. While the expression level of sfGFP varied regarding the position and length of the tag, it's noteworthy that these variations did not substantially impact solubility. All constructs exhibited consistently high solubility levels ranging from 92 to 100%, irrespective of salt concentration. The purity of the sfGFP fusion proteins displayed notable variation concerning His-tag length, position, and salt concentration in the purification buffer. The 8xHis-tag consistently outperformed all other His-tag constructs, resulting in high purity and yield in high and low salt buffers. This outcome aligns with expectations, as the increased number of histidine residues in the tag leads to more efficient binding on IMAC columns. Furthermore, it allows for more rigorous washes, enhancing the overall purity of the protein.

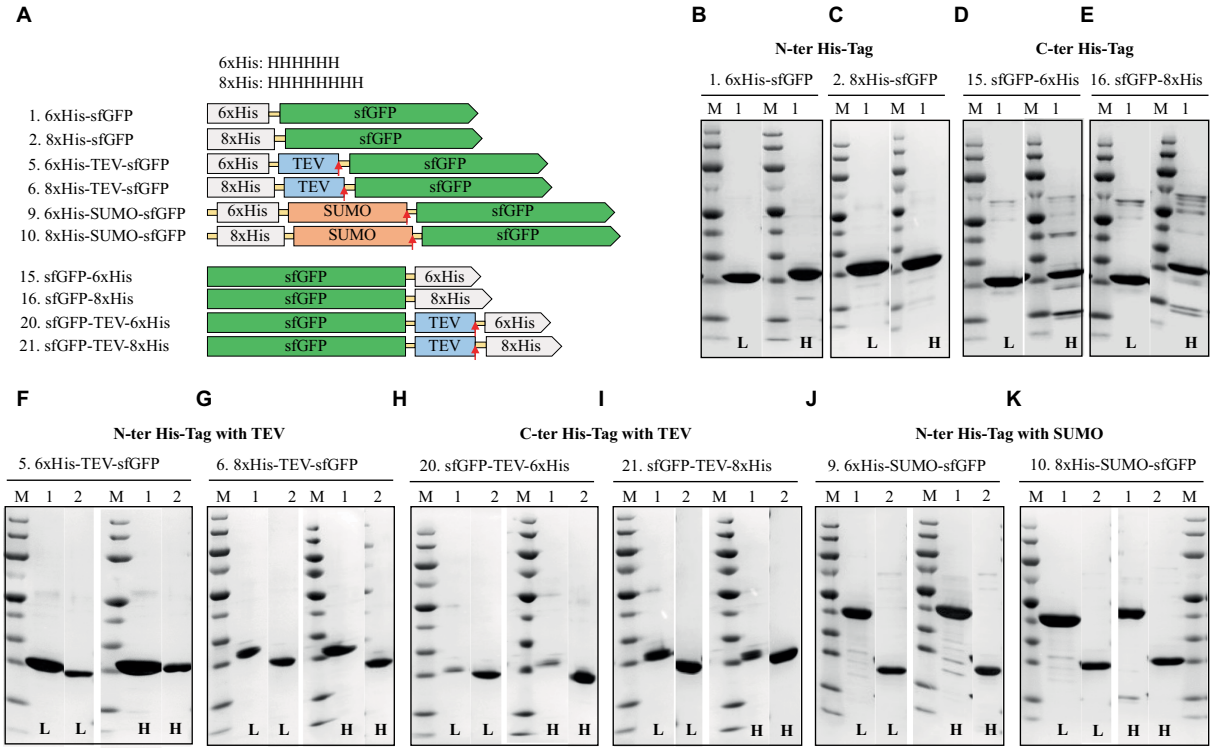


FIGURE 2
Schematic representation of His-tag constructs **(A)** and SDS-PAGE analysis of the purification of the sfGFP with N-terminal 6xHis **(B)**, 8xHis **(C)**, and C-terminal 6xHis **(D)** 8xHis **(E)**, N-terminal His-tag with cleavable tag, TEV, 6xHis, **(F)** 8xHis **(G)**, and C-terminal His-tag with TEV cleavage site, 6xHis **(H)**, 8xHis **(I)**, N-terminal His-tag with solubility and cleavable SUMO-tag with 6xHis **(J)**, or 8xHis **(K)**. Lane M: molecular weight marker; Lane 1: Elution from HisTrap affinity column using the binding buffer containing 250 mM imidazole; Lane 2: Flow-through from HisTrap affinity column loaded after TEV-/SUMO-Protease cleavage. L: Low salt (200 mM NaCl); H: High salt (2 M NaCl) in the binding, washing, and elution buffers.

Sumo-tag vs. TEV cleavage site and the impact of salt

A comparative assessment of the solubility-enhancing and cleavable SUMO-tag versus the TEV cleavage site yielded valuable insights into the performance of these constructs. The SUMO-tag exhibited robust performance, outperforming the cleavable TEV cleavage site. The SUMO-tag was pivotal in enhancing sfGFP expression and ensuring high purity, irrespective of the buffering composition. sfGFP constructs with the TEV cleavage site also displayed strong expression, achieving even higher purity levels. However, the TEV cleavage site did not cleave efficiently at high salt concentrations (Figure 2). This limitation resulted in approximately 40–50% final yields compared to the uncleaved constructs. Of note, the sfGFP construct featuring a C-terminal 6xHis-tag with a TEV cleavage site exhibited the lowest production level among all the tested His-tagged constructs.

Overall, the N-terminal 8xHis-tag proved the best performer, offering a simple, rapid, and highly efficient one-step purification method for sfGFP. This approach employed cost-effective materials and delivered approximately double the final yield compared to the 6xHis-tag and nearly five times the yield compared to the C-terminal 6xHis- or 8xHis-tag constructs (Table 1).

The 6xHis-tag remains the most favored purification tag for *H. volcanii* due to its compatibility with high-salt media (Haque et al., 2019). However, the results, which contradict earlier studies on N-terminal His-tagged GFP that indicated adverse effects on protein production, emphasize the ongoing challenge of selecting the optimal position and length for the His-tag within various expression systems (Nguyen et al., 2014; Phan et al., 2015; Tian et al., 2019; Le et al., 2022). This study underscores the importance of making informed choices regarding His-tag design to maximize expression levels and purification efficiency.

Strep-tag®II and twin-Strep-tag®

Next, the sfGFP constructs fused with Strep-tag®II and Twin-Strep-tag® were investigated.

N-terminal vs. C-terminal and Strep-tag®II vs. twin-Strep-tag®

The results of the sfGFP constructs fused with Strep-tag®II are summarized in Table 2 and Figure 3. Expression levels of sfGFP varied based on the position of the Strep-tag®II and Twin-Strep-tag®. Like the His-tag, the N-terminal Strep-tag®II yielded the highest sfGFP expression, surpassing the C-terminal Strep-tag®II or Twin-Strep-tag® by up to fivefold. Notably, while C-terminal Strep-tag®II has been utilized for protein production in *H. volcanii* and proven compatible with high salt concentrations,

TABLE 2 Relative expression, solubility, and purity of sfGFP expressed with Strep-tag®II and Twin-Strep-tag® constructs.

Construct	Expression level (%)	Solubility (%)		Purity (%)		Yield (mg/L of culture)	
		Low salt	High salt	Low salt	High salt	Low salt	High salt
N-terminal purification Strep-tag®II/TEV cleavage site/SUMO-tag							
3.Strep-sfGFP	100	100	85	95	75	5.3	4.3
4. Twin-Strep-sfGFP	32	100	75	84	85	2.2	1.9
7. Strep-TEV-sfGFP	27	94	88	90 (97)	87 (93)	1.9 (1.6)	2.0 (0.9)
8. Twin-Strep-TEV-sfGFP	22	100	90	79 (97)	85 (96)	1.5 (1.2)	1.3 (0.6)
11. Strep-SUMO-sfGFP	34	100	95	67 (96)	62 (95)	2.4 (1.5)	2.1 (1.2)
12. Twin-Strep-SUMO-sfGFP	27	97	100	65 (94)	71 (96)	2.3 (1.4)	2.0 (1.0)
C-terminal purification Strep-tag®II/TEV cleavage site							
17. sfGFP-Strep	21	91	100	90	93	1.6	1.5
18. sfGFP-Twin-Strep	17	97	94	75	70	1.7	1.9
22. sfGFP-TEV-Strep	16	100	97	70 (95)	65 (90)	1.6 (1.1)	1.7 (0.6)
23. sfGFP-TEV-Twin-Strep	15	100	100	52 (87)	64 (95)	1.9 (1.1)	1.3 (0.6)

The numbers in parentheses represent the values after cleavage and purification in the corresponding step.

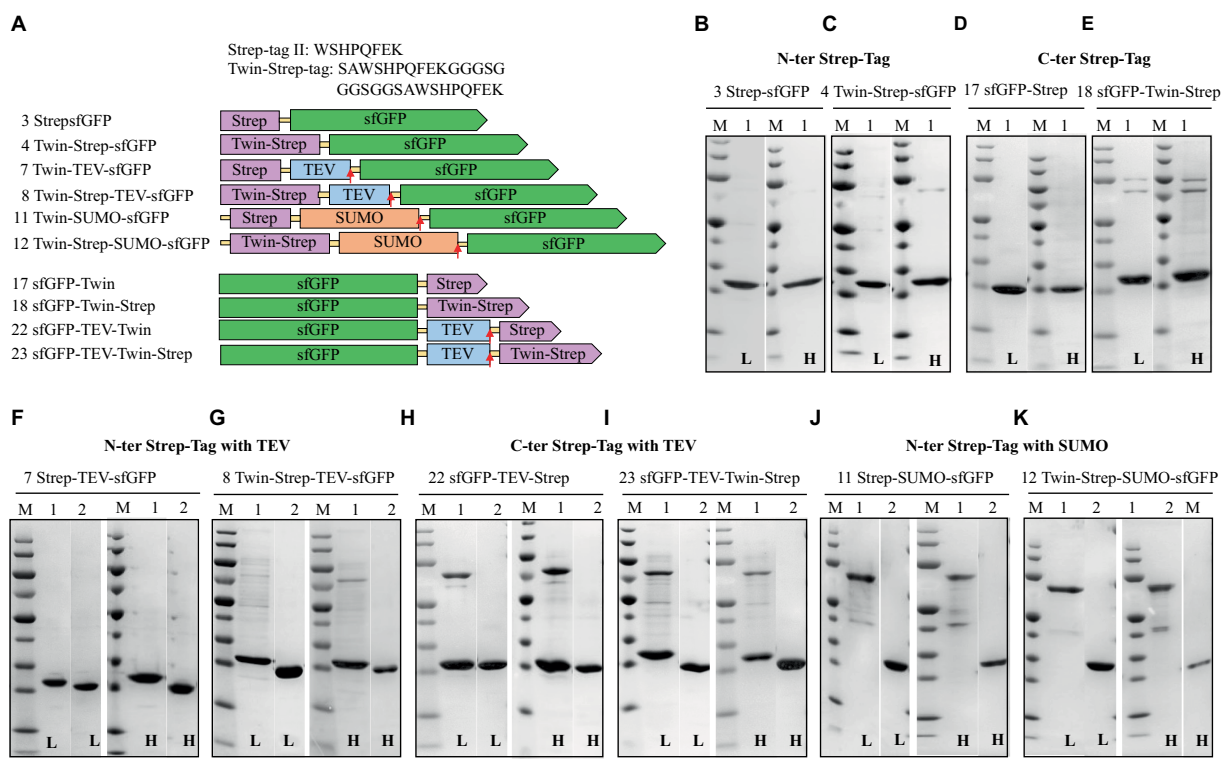


FIGURE 3
Schematic representation of Strep-tag (Strep-tag®II and Twin-Strep-tag®) constructs **(A)** and SDS-PAGE analysis of the purification of the sfGFP with N-Terminal Strep **(B)**, Twin-Strep-tag® **(C)**, and C-terminal Strep **(D)** Twin-Strep-tag®s **(E)**. N-terminal Strep with TEV cleavage site, Strep, **(F)** Twin-Strep-tag® **(G)**, and C-terminal Strep-tag®II with TEV cleavage site, Strep **(H)**, Twin-Strep-tag®s **(I)**, N-terminal Strep-tag®II with solubility and cleavable SUMO-Strep-tag **(J)**, and Twin-Strep-tag® **(K)**. Lane M: molecular weight marker; Lane 1: Elution from Strep-Tactin affinity column using the binding buffer containing 50 mM biotin; Lane 2: Flow-through from HisTrap and Strep-Tactin affinity column loaded after TEV-/SUMO-Protease cleavage. L: Low salt (200 mM NaCl); H: High salt (2 M NaCl) in the binding, washing, and elution buffers.

the study reveals that N-terminal Strep-tag®II results in approximately five times higher expression, suggesting that C-terminal usage in *H. volcanii* should be reconsidered (Braun et al., 2019). Intriguingly, Twin-Strep-tag®, which generally exhibits a higher affinity for Strep-Tactin® and greater tolerance to salts and detergents in buffers than Strep-tag®II (Yeliseev et al., 2017),

displayed an unexpected reduction in sfGFP expression to about one-third when compared to N-terminal Strep-tag®II. The introduction of Strep-tags (Strep-tag®II or Twin-Strep-tag®) appeared to have no significant effect on the solubility of the fusion proteins, regardless of salt conditions. Furthermore, these tags enabled high-purity purification processes (Table 2 and Figure 2).

Sumo vs. TEV cleavage site and the impact of salt

Next, experiments to assess the removal of Strep-tags using protease cleavage sites (TEV and SUMO) positioned between the sfGFP protein and the Strep-tags were conducted. Subsequently, a purification step was employed to separate the sfGFP protein from the cleaved Strep-tags. Consequently, the findings indicate that both Strep-tag®II and Twin-Strep-tag®, when fused with SUMO or TEV cleavage sites, could be efficiently removed through protease cleavage (SUMO- or TEV-Protease).

Among all the tested tags fused to the N-terminus of sfGFP, Strep-tag®II consistently yielded the highest levels of sfGFP expression.

Flag-tag

Next, the expression and solubility of sfGFP fused with the FLAG-tag and 3xFLAG-tag, in combination with the SUMO-tag or TEV cleavage site (Einhauer and Jungbauer, 2001; Schmidt et al., 2012; Beznosov et al., 2013) were assessed.

N-terminal vs. C-terminal FLAG-tag and 3xFLAG-tag

Comparable to His and Strep-tagged sfGFP findings, N-terminal FLAG, and 3xFLAG-tags exhibited higher expression levels (50 and 75%, respectively) than C-terminal tags (Table 3). Interestingly, N-terminal 3xFLAG-fused sfGFP demonstrated expression levels 1.5 to 3 times superior to N-terminal FLAG-tag, C-terminal FLAG, or 3xFLAG-tagged sfGFP. Additionally, all constructs displayed high solubility, exceeding 90%, regardless of whether the FLAG or 3xFLAG was positioned at the N-terminus or C-terminus of the sfGFP protein.

Cleaving the FLAG-tag/3xFLAG-tag and the impact of salt

The Enterokinase efficiently removed the FLAG-tag even at higher salt concentrations, which led to a relatively lower yield (Figure 4). All FLAG-tagged expressions resulted in highly pure sfGFP with no visible contaminating proteins on SDS-PAGE (Figure 4). To purify FLAG and 3xFLAG-tagged sfGFP from sonicated *H. volcanii* cell culture, the anti-DYKDDDDK magnetic beads (GenScript) were also tested. Remarkably, we were able to streamline the purification process by eliminating centrifugation or filtration steps and still achieved high purification (data not shown).

While FLAG-tags have previously been employed to purify recombinant proteins in various expression systems, including bacteria, baculovirus, mammalian cells, and yeast, there has been limited information regarding their use in halophilic systems (Einhauer and Jungbauer, 2001; Mishra, 2020). FLAG peptide has been tagged with modified archaellin genes (flaA1, flaA2, and flaB2) in haloarchaea *Halobacterium salinarum* cells and utilized for the identification of proteins within the archaellum (Beznosov et al., 2013). However, no reports existed on FLAG-tagged protein expression and purification in the halophilic context until this study. These findings suggest that the FLAG and 3xFLAG-tags suit protein expression and purification in halophilic systems. The FLAG-tag carries several advantages over His and Strep tags. Like the 6xHis-, 8xHis-, and Strep-tag®II, the FLAG-tag comprises a small hydrophilic peptide (8 amino acids), unlikely to impact protein folding or function. Furthermore, it can be easily removed with Enterokinase. However, it is worth noting that Enterokinase, while versatile, is sensitive to high salt concentrations compared to other proteases like the SUMO-Protease, limiting its application in the extremophilic environment (Shahravan et al., 2008).

C-tag

Next, the C-tag constructs were examined, and the results detailing the expression level, solubility, purity, and final yield of the fusion sfGFP constructs are outlined in Table 4 and Figure 5. The C-tag-binding resin facilitated the purification of pure proteins in both low-salt and high-salt buffers. Notably, under the 2 M NaCl purification buffer, the protein achieved high purity at 98%, albeit with a lower yield. However, for the 200 mM NaCl purification buffer, the purity of sfGFP could be significantly improved by

TABLE 3 Relative expression, solubility, and purity of sfGFP expressed with FLAG-tag and 3xFLAG-tag.

Construct	Expression level (%)	Solubility (%)		Purity (%)		Yield (mg/L of culture)	
		Low salt	High salt	Low salt	High salt	Low salt	High salt
N-terminal purification FLAG-tag							
13. FLAG-sfGFP	50	100	92	55 (95)	92 (97)	3.6 (2.2)	2.3 (1.6)
14. 3xFLAG-sfGFP	75	100	97	90 (97)	95 (98)	3.3 (2.9)	2.9 (1.9)
C-terminal purification FLAG-tag							
25. sfGFP-FLAG	26	100	92	40	65	2.6	2.1
26. sfGFP-3xFLAG	23	97	100	55	85	2.1	1.4

The numbers in parentheses represent the values after cleavage and purification in the corresponding step.

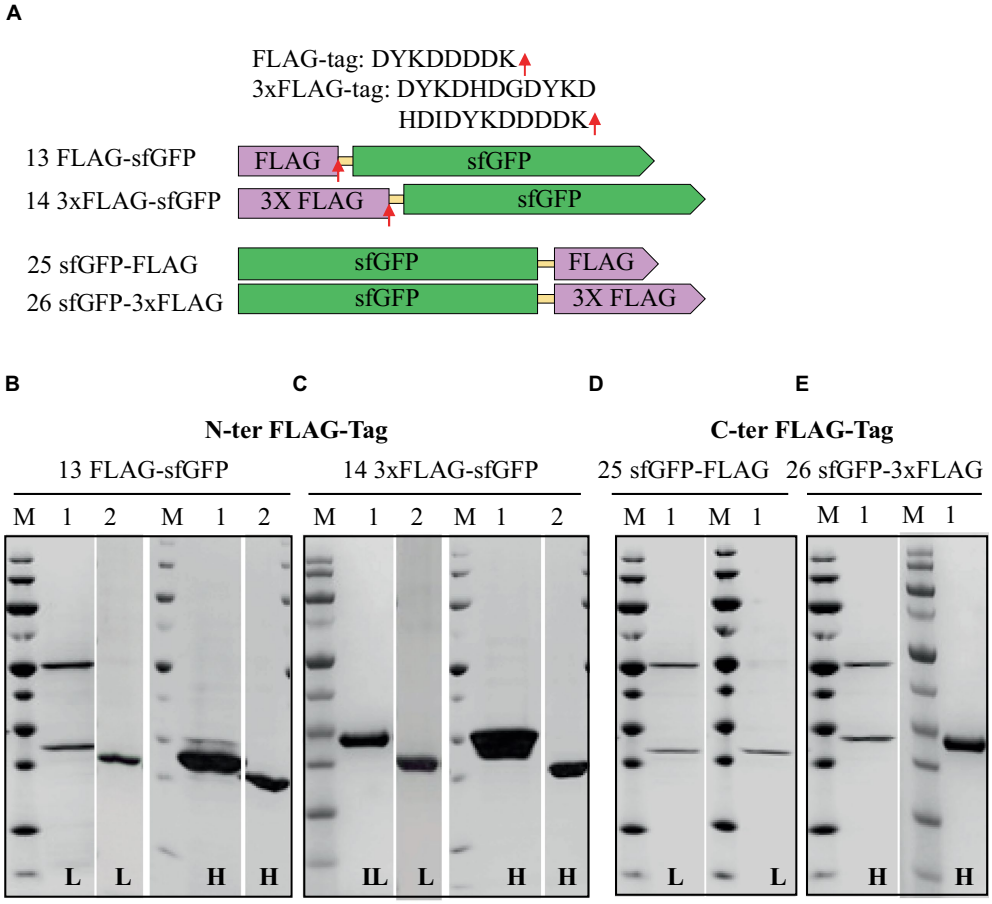


FIGURE 4
Schematic representation of FLAG-tag constructs **(A)** and SDS-PAGE analysis of the purification of the sfGFP with N-Terminal FLAG **(B)**, 3xFLAG **(C)**, and C-terminal FLAG **(D)** 3xFLAG with TEV cleavage site **(E)**. Lane M: molecular weight marker; Lane 1: Elution from anti-FLAG affinity column using 100 mM Tris/HCl buffer, pH 12.0; Lane 2: Flow-through from HisTrap affinity column loaded after TEV-Protease cleavage. L: Low salt (200 mM NaCl); H: High salt (2 M NaCl) in the binding, washing, and elution buffers.

TABLE 4 **Relative expression, solubility, and purity of sfGFP expressed with C-tag.**

Construct	Expression level (%)	Solubility (%)		Purity (%)		Yield (mg/L of culture)	
		Low salt	High salt	Low salt	High salt	Low salt	High salt
19. sfGFP-C-tag	20	100	92	92	98	0.9	0.6
24. sfGFP-TEV-C-tag	21	97	100	70 (85)	85 (95)	1.2 (0.9)	0.8 (0.4)

The numbers in parentheses represent the values after cleavage and purification in the corresponding step.

column washing with 1 M NaCl (Supplementary Figure S6). Moreover, including a TEV cleavage site aided in removing the C-tag, resulting in higher purity, although this came at the expense of yield, particularly in high-salt conditions due to the reduced activity of TEV-Protease in high-salt environments (Parks et al., 1995; Waugh, 2011).

Among all the affinity tags tested, the smallest among them, the C-tag comprising just four amino acids, proved to be a rapid and efficient tool for purification. The C-tag offers several advantages over more established tags such as His, Strep, and FLAG. It is the smallest affinity purification tag with high binding affinity and selectivity. Furthermore, its limited impact on protein folding and functionality due to its small size makes it an attractive choice for applications in protein purification (Hermans et al., 2012; Djender et al., 2014; Jin et al., 2017).

The systematic screening of the different purification tags and tag positions with sfGFP revealed a significant enhancement in protein production, purity, and yield in *H. volcanii* when an 8xHis-tag or Strep-tag®II was placed at the N-terminus of sfGFP. This enhancement is potentially linked to favorable folding kinetics of newly synthesized proteins and increased accessibility of these tags during affinity purification, which may facilitate higher yield and purity. To explore whether this trend was unique to the expression system or exclusive to sfGFP, the investigation was extended to include the red fluorescent protein mCherry and a previously fully characterized halophilic alcohol dehydrogenase (ADH) from the

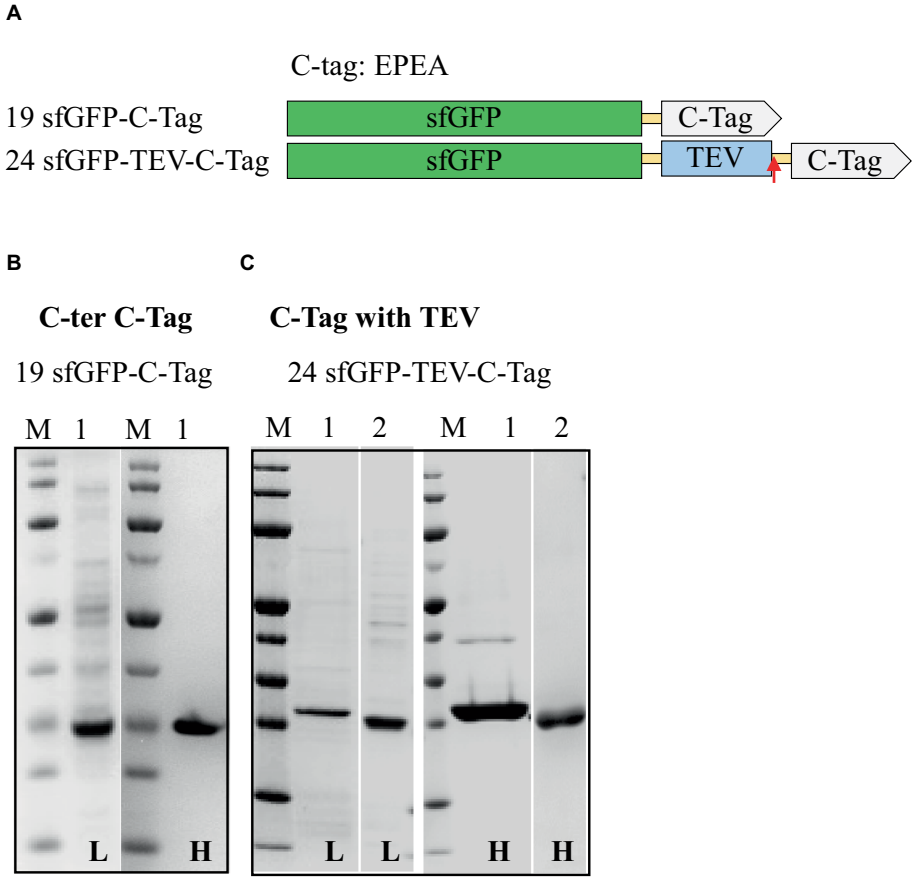


FIGURE 5
Schematic representation of C-tag constructs **(A)** and SDS-PAGE analysis of the purification of the sfGFP with C-tag **(B)** and C-tag with TEV cleavage site **(C)**. Lane M: molecular weight marker; Lane 1: Elution from C-tag XL affinity column using 2 M $MgCl_2$; Lane 2: Flow-through from HisTrap affinity column loaded after TEV-Protease cleavage. L: Low salt (200 mM NaCl); H: High salt (2 M NaCl) in the binding and washing buffers.

deep-sea brine pool of the Red Sea (Shaner et al., 2004; Grötzinger et al., 2018).

Transferring the expression and purification results from sfGFP to mCherry and ADH

The 8xHis-tag and Strep-tag®II were both positioned at the N- and C-termini to investigate the impact of tag position on mCherry and the ADH (Supplementary Table S2 and Supplementary Figure S1). Surprisingly, we discovered that mCherry demonstrated improved expression levels when the 8xHis-tag, or Strep-tag®II, was located at the C-terminus of the protein. Conversely, ADH exhibited more efficient expression when the 8xHis-tag or Strep-tag®II was situated at the N-terminus of the respective proteins (Table 5 and Figure 6). Although these proteins shared similar relative solubility levels, they exhibited divergent expression patterns, emphasizing the nuanced nature of tag effects in different protein contexts.

These findings underscore the critical role of purification tags, their positions, and lengths in determining fusion protein expression, purification, and yields. Moreover, this observation emphasizes the necessity of exploring various fusion strategies to optimize expression and purification efficiencies. We considered applying a dual-affinity-tag approach to create a more general purification tag for *H. volcanii*.

Impact of combining 8xHis-tag and twin-Strep-tag® at the N-terminus on the expression and purification of sfGFP, mCherry, and ADH: a dual-affinity-tag approach

Assessing different tag combinations, positions, and lengths for a novel protein can be resource-intensive and time-consuming (Bernier et al., 2018). To address this challenge, we applied a dual-affinity-tag approach by combining 8xHis-tag and Twin-Strep-tag® tags at the N-terminus of sfGFP, mCherry, and ADH (Supplementary Table S3 and Supplementary Figure S1). In the purification process first a HiTrap Ni^{2+} chelating column was used to exploit the 8xHis-tag, followed by a Strep-Tactin®XT 4Flow column to utilize the Twin-Strep-tag®, achieving high-purity protein. Surprisingly, this dual-affinity-tag approach consistently outperformed all previously tested tag configurations regarding expression level and purification yields (Table 6 and Figure 7). Moreover, the dual-affinity-tag seems to be a

TABLE 5 Relative expression, solubility, and purity of mCherry and ADH expressed with N- and C-terminal 8xHis-tag and Strep-tag®II.

Construct	Expression level[a,b] (%)	Solubility[c] (%)		Purity[d] (%)		Yield[e] (mg/L of culture)	
		Low salt	High salt	Low salt	High salt	Low salt	High salt
mCherry with 8xHis-tag							
27. 8xHis-mCherry	65	100	92	48	52	2.6	2.9
28. mCherry-8xHis	100	100	97	66	74	4.2	4.6
mCherry with Strep-tag®II							
29. Strep-mCherry	27	100	92	64	69	0.8	0.9
30. mCherry-Strep	96	97	100	68	75	2.4	2.8
ADH with 8xHis-tag							
31. 8xHis-ADH	100	100	92	98	99	35	34
32. ADH-8xHis	69	100	97	96	96	24	26
ADH with Strep-tag®II							
33. Strep-ADH	57	100	92	99	99	20	19
34. ADH-Strep	51	97	100	98	97	18	19

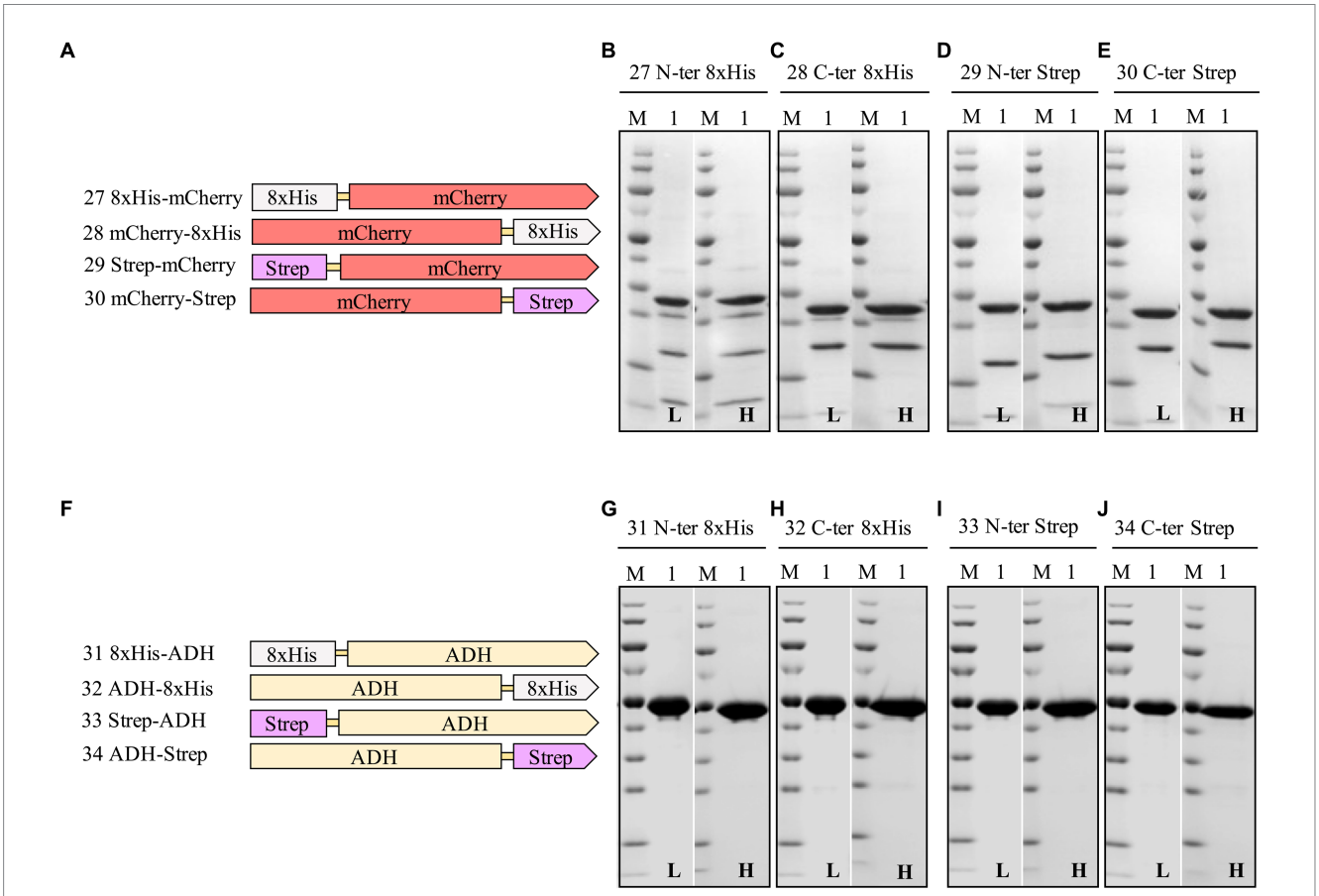


FIGURE 6
Schematic representation of N- and C-terminal 8xHis-tag and Strep-tag constructs of mCherry **(A)**, and ADH **(F)**, and SDS-PAGE analysis of the purification of the mCherry **(B–E)**, ADH **(G–J)** Lane M: molecular weight marker; Lane 1: Elution from affinity column. L: Low salt (200 mM NaCl); H: High salt (2 M NaCl) in the binding and washing buffers.

universal solution for protein expression and purification in *H. volcanii* since all tested proteins were notably expressed in higher amounts and yielded higher overall protein content.

The dual-affinity-tag approach is a method that streamlines the purification of recombinant proteins, yielding homogeneous preparations of the proteins of interest in *H. volcanii*. This is achieved

TABLE 6 Relative expression, solubility, and purity of sfGFP, mCherry, and ADH expressed with N-terminal dual-affinity-tag consisting of 8xHis-tag and Twin-Strep-tag®.

Construct	Expression level[a] (%)	Solubility[b] (%)		Purity[c] (%)		Yield[d] (mg/L of culture)	
		Low salt	High salt	Low salt	High salt	Low salt	High salt
35. 8xHis-Twin-Strep-sfGFP	135	100	97	99	99	7.1	6.9
36. 8xHis-Twin-Strep-mCherry	138	100	97	57	68	5.9	5.7
37. 8xHis-Twin-Strep-ADH	109	100	98	89	86	38	42

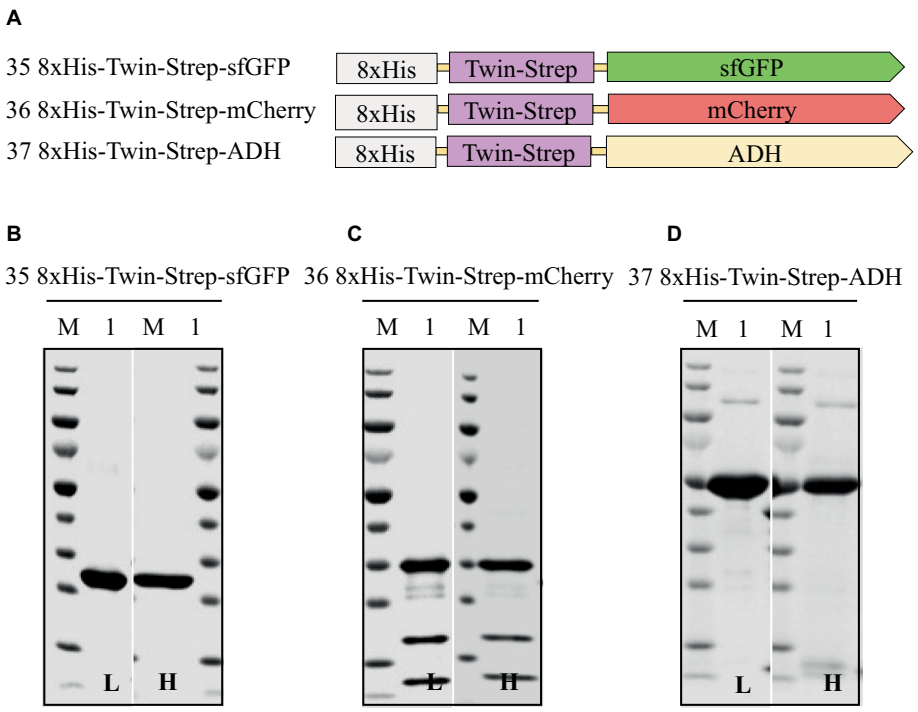


FIGURE 7
Schematic representation of N-terminal dual-affinity-tag consisting of 8xHis-tag and Twin-Strep-tag® constructs of sfGFP, mCherry, and ADH **(A)** and SDS-PAGE analysis of the purification of the sfGFP **(B)**, mCherry **(C)**, ADH **(D)**. Lane M: molecular weight marker; Lane 1: Elution from affinity column. L: Low salt (200 mM NaCl); H: High salt (2 M NaCl) in the binding and washing buffers.

by enabling two consecutive affinity chromatography steps, efficiently removing contaminating proteins.

Conclusion

In conclusion, the study demonstrates the efficient production and purification of sfGFP proteins utilizing various affinity tags within the *Haloferax volcanii* expression system. The choice of tag and its position significantly impacts the levels of protein production and purification, with the dual-affinity-tag configuration of the 8xHis-tag and Strep-tag®II at the N-terminus generally providing optimal results. This approach has proven effective not just for sfGFP but also for mCherry and ADH, highlighting the importance of tag configuration over mere expression levels.

The options for cleaving the purification tag at high salt concentrations are limited due to the salt sensitivity of commonly tested proteases, including enterokinase and Tobacco Etch Virus (TEV) protease. Other frequently used proteases, such as human rhinovirus 3C protease (HRV 3C) and thrombin, also exhibit salt sensitivity (Jenny et al., 2003; Dušeková et al., 2022). However, salt-insensitive alternatives like SUMO protease and Factor Xa exist (Jenny et al., 2003). Additionally, emerging alternatives, such as the Plum Pox Virus (PPV) NIa protease, demonstrate reduced salt sensitivity, presenting viable options for such conditions (Zheng et al., 2008).

By expanding the library with mCherry and a halophilic alcohol dehydrogenase, we observed that optimal tag positioning varies depending on the protein, with C-terminal tags more effective for mCherry and N-terminal tags better suited for ADH. This underscores the nuanced approach needed in tag placement and protease selection to accommodate different protein characteristics.

As the investigation advance, incorporating proteins with inherently poor solubility will be crucial to further test and validate the adaptability of the suggested tagging strategies. This future work

will enhance the general understanding of the expression system's capabilities, providing a more comprehensive foundation for biotechnological applications in extremophilic organisms.

In summary, the study not only reinforces the significance of strategic tag design and placement but also the critical role of compatible protease selection in the successful production and purification of proteins in extremophilic systems like *Haloferax volcanii*. By focusing on these elements, researchers can optimize protein yield and purity, thereby improving the overall efficiency of their expression and purification protocols.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding authors.

Author contributions

RK: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. DR: Conceptualization, Formal analysis, Investigation, Methodology, Supervision, Validation, Writing – review & editing. TA: Resources, Validation, Writing – review & editing. MR: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. The research reported in this publication was supported by funding from the King Abdullah University of Science and Technology.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2024.1403623/full#supplementary-material

References

Akal, A. L., Karan, R., Hohl, A., Alam, I., Vogler, M., Grötzinger, S. W., et al. (2019). A polyextremophilic alcohol dehydrogenase from the Atlantis II deep Red Sea brine pool. *FEBS Open Bio* 9, 194–205. doi: 10.1002/2211-5463.12557

Allers, T. (2010). Overexpression and purification of halophilic proteins in Haloferax volcanii. *Bioeng Bugs* 1, 290–292. doi: 10.4161/bbug.1.4.11794

Allers, T., Barak, S., Liddell, S., Wardell, K., and Mevarech, M. (2010). Improved strains and plasmid vectors for conditional overexpression of His-tagged proteins in Haloferax volcanii. *Appl. Environ. Microbiol.* 76, 1759–1769. doi: 10.1128/AEM.02670-09

Alshehri, S., Karan, R., Ghalayini, S., Kahin, K., Khan, Z., Renn, D., et al. (2022). Air-loaded gas vesicle nanoparticles promote cell growth in three-dimensional bioprinted tissue constructs. *Int J Bioprint* 8:489. doi: 10.18063/ijb.v8i3.489

Andar, A. U., Karan, R., Pecher, W. T., DasSarma, P., Hedrich, W. D., Stinchcomb, A. L., et al. (2017). Microneedle-assisted skin permeation by nontoxic Bioengineerable gas vesicle nanoparticles. *Mol. Pharm.* 14, 953–958. doi: 10.1021/acs.molpharmaceut.6b00859

Baeshen, M. N., Al-Hejin, A. M., Bora, R. S., Ahmed, M. M. M., Ramadan, H. A. I., Saini, K. S., et al. (2015). Production of biopharmaceuticals in E. coli: current scenario and future perspectives. *J. Microbiol. Biotechnol.* 25, 953–962. doi: 10.4014/jmb.1412.12079

Bernier, S. C., Cantin, L., and Salesse, C. (2018). Systematic analysis of the expression, solubility and purification of a passenger protein in fusion with different tags. *Protein Expr. Purif.* 152, 92–106. doi: 10.1016/j.pep.2018.07.007

Beznosov, S., Pyatibratov, M., Veluri, P., Mitra, S., and Fedorov, O. (2013). A way to identify archaellins in Halobacterium salinarum archaella by FLAG-tagging. *Open Life Sci* 8, 828–834. doi: 10.2478/s11535-013-0202-0

Block, H., Maertens, B., Spriestersbach, A., Brinker, N., Kubicek, J., Fabis, R., et al. (2009). Immobilized-metal affinity chromatography (IMAC): a review. *Methods Enzymol.* 463, 439–473. doi: 10.1016/S0076-6879(09)63027-5

Braun, F., Thomalla, L., van der Does, C., Quax, T. E. F., Allers, T., Kaever, V., et al. (2019). Cyclic nucleotides in archaea: cyclic di-AMP in the archaeon Haloferax volcanii and its putative role. *Microbiology* 8:e00829. doi: 10.1002/mbo3.829

Butt, T. R., Edavettal, S. C., Hall, J. P., and Mattern, M. R. (2005). SUMO fusion technology for difficult-to-express proteins. *Protein Expr. Purif.* 43, 1–9. doi: 10.1016/j.pep.2005.03.016

DasSarma, S., Karan, R., DasSarma, P., Barnes, S., Ekulona, F., and Smith, B. (2013). An improved genetic system for bioengineering buoyant gas vesicle nanoparticles from haloarchaea. *BMC Biotechnol.* 13:112. doi: 10.1186/1472-6750-13-112

Djender, S., Beugnet, A., Schneider, A., and De Marco, A. (2014). The biotechnological applications of recombinant single-domain antibodies are optimized by the C-terminal fusion to the EPEA sequence (C tag). *Antibodies* 3, 182–191. doi: 10.3390/antib3020182

Du, M., Hou, Z., Liu, L., Xuan, Y., Chen, X., Fan, L., et al. (2022). Progress, applications, challenges and prospects of protein purification technology. *Front. Bioeng. Biotechnol.* 10:28691. doi: 10.3389/fbioe.2022.1028691

Dušeková, E., Berta, M., Sedláková, D., Řeha, D., Dzurillová, V., Shaposhnikova, A., et al. (2022). Specific anion effect on properties of HRV 3C protease. *Biophys. Chem.* 287:106825. doi: 10.1016/j.bpc.2022.106825

Dyall-Smith, M. (2008). *The Halohandbook: Protocols for haloarchaeal genetics*. Haloarchaeal Genetics Laboratory: Melbourne, VI, Australia, 14.

Einhauer, A., and Jungbauer, A. (2001). The FLAG™ peptide, a versatile fusion tag for the purification of recombinant proteins. *J. Biochem. Biophys. Methods* 49, 455–465. doi: 10.1016/S0165-022X(01)00213-5

Elleuche, S., Schröder, C., Sahm, K., and Antranikian, G. (2014). Extremozymes—biocatalysts with unique properties from extremophilic microorganisms. *Curr. Opin. Biotechnol.* 29, 116–123. doi: 10.1016/j.copbio.2014.04.003

Gomari, M. M., Saraygord-Afshari, N., Farsimadan, M., Rostami, N., Aghamiri, S., and Farajollahi, M. M. (2020). Opportunities and challenges of the tag-assisted protein purification techniques: applications in the pharmaceutical industry. *Biotechnol. Adv.* 45:107653. doi: 10.1016/j.biotechadv.2020.107653

Grote, A., Hiller, K., Scheer, M., Munch, R., Nortemann, B., Hempel, D. C., et al. (2005). JCat: a novel tool to adapt codon usage of a target gene to its potential expression host. *Nucleic Acids Res.* 33, W526–W531. doi: 10.1093/nar/gki376

Grötzinger, S. W., Karan, R., Strillinger, E., Bader, S., Frank, A., Al Rowaihi, I. S., et al. (2018). Identification and experimental characterization of an extremophilic brine pool alcohol dehydrogenase from single amplified genomes. *ACS Chem. Biol.* 13, 161–170. doi: 10.1021/acschembio.7b00792

Haque, R. U., Paradisi, F., and Allers, T. (2019). Haloferax volcanii as immobilised whole cell biocatalyst: new applications for halophilic systems. *Appl. Microbiol. Biotechnol.* 103, 3807–3817. doi: 10.1007/s00253-019-09725-y

Haque, R. U., Paradisi, F., and Allers, T. (2020). Haloferax volcanii for biotechnology applications: challenges, current state and perspectives. *Appl. Microbiol. Biotechnol.* 104, 1371–1382. doi: 10.1007/s00253-019-10314-2

Hermans, P., Clasen, R., and Detmers, F. (2012). Reinventing affinity tags: innovative technology designed for routine purification of C-terminal EPEA-tagged recombinant proteins. *Gen. Eng. Biotechnol. News* 32, 48–49. doi: 10.1089/gen.32.17.20

Hopp, T. P., Prickett, K. S., Price, V. L., Libby, R. T., March, C. J., Pat Cerretti, D., et al. (1988). A short polypeptide marker sequence useful for recombinant protein identification and purification. *Bio/Technology* 6, 1204–1210. doi: 10.1038/nbt1088-1204

Ibrahim, T. N. B. T., Abas, A. B., and Razak, N. F. A. (2023). Escherichia coli cell factory for synthesis of biomolecules. *Biomanuf. Sustain. Prod. Biomol.* 7, 141–163. doi: 10.1007/978-981-19-7911-8_7

Jenny, R. J., Mann, K. G., and Lundblad, R. L. (2003). A critical review of the methods for cleavage of fusion proteins with thrombin and factor Xa. *Protein Expr. Purif.* 31, 1–11. doi: 10.1016/S1046-5928(03)00168-2

Jin, J., Hjerrild, K. A., Silk, S. E., Brown, R. E., Labbé, G. M., Marshall, J. M., et al. (2017). Accelerating the clinical development of protein-based vaccines for malaria by efficient purification using a four amino acid C-terminal C-tag. *Int. J. Parasitol.* 47, 435–446. doi: 10.1016/j.ijpara.2016.12.001

Karan, R., Mathew, S., Muhammad, R., Bautista, D. B., Vogler, M., Eppinger, J., et al. (2020). Understanding high-salt and cold adaptation of a polyextremophilic enzyme. *Microorganisms* 8:1594. doi: 10.3390/microorganisms8101594

Karan, R., Renn, D., Nozue, S., Zhao, L., Habuchi, S., Allers, T., et al. (2023). Bioengineering of air-filled protein nanoparticles by genetic and chemical functionalization. *J. Nanobiotechnol.* 21:108. doi: 10.1186/s12951-023-01866-7

Kimple, M. E., Brill, A. L., and Pasker, R. L. (2013). Overview of affinity tags for protein purification. *Curr. Protoc. Protein Sci.* 73, 9.9.1–9.9.23. doi: 10.1002/0471140864.ps0909s73

Kimple, M. E., and Sondek, J. (2004). Overview of affinity tags for protein purification. *Curr. Protoc. Protein Sci.* 36, 9.9.1–9.9.19. doi: 10.1002/0471140864.ps0909s36

Köppl, C., Lingg, N., Fischer, A., Kröß, C., Loibl, J., Buchinger, W., et al. (2022). Fusion tag design influences soluble recombinant protein production in Escherichia coli. *Int. J. Mol. Sci.* 23:7678. doi: 10.3390/ijms23147678

Kruglikov, A., Wei, Y., and Xia, X. (2022). Proteins from thermophilic Thermus thermophilus often do not fold correctly in a mesophilic expression system such as Escherichia coli. *ACS Omega* 7, 37797–37806. doi: 10.1021/acsomega.2c04786

Le, N. T. P., Phan, T. T. P., Phan, H. T. T., Truong, T. T. T., Schumann, W., and Nguyen, H. D. (2022). Influence of N-terminal His-tags on the production of recombinant proteins in the cytoplasm of Bacillus subtilis. *Biotechnol. Rep.* 35:e00754. doi: 10.1016/j.btre.2022.e00754

Martínez-Espinosa, R. M. (2019). Heterologous and homologous expression of proteins from haloarchaea: denitrification as case of study. *Int. J. Mol. Sci.* 21:82. doi: 10.3390/ijms21010082

Mishra, V. (2020). Affinity tags for protein purification. *Curr. Protein Pept. Sci.* 21, 821–830. doi: 10.2174/1389203721666200606220109

Nguyen, N. H., Phan, T. T. P., Tran, T. L., and Nguyen, H. D. (2014). Investigating the expression of GFP fused with HIS-TAG at N-or Cterminus using plasmid pHT253 and pHT254 in Bacillus subtilis. *VNUHCM J. Sci. Technol. Dev.* 17, 5–11. doi: 10.32508/stdj.v17i4.1550

Parks, T. D., Howard, E. D., Wolpert, T. J., Arp, D. J., and Dougherty, W. G. (1995). Expression and purification of a recombinant tobacco etch virus NIa proteinase: biochemical analyses of the full-length and a naturally occurring truncated proteinase form. *Virology* 210, 194–201. doi: 10.1006/viro.1995.1331

Pédelacq, J.-D., Cabantous, S., Tran, T., Terwilliger, T. C., and Waldo, G. S. (2006). Engineering and characterization of a superfolder green fluorescent protein. *Nat. Biotechnol.* 24, 79–88. doi: 10.1038/nbt1172

Pérez-Arnaiz, P., Dattani, A., Smith, V., and Allers, T. (2020). Haloferax volcanii—a model archaeon for studying DNA replication and repair. *Open Biol.* 10:293. doi: 10.1098/rsob.200293

Phan, T. T. P., Tran, L. T., Schumann, W., and Nguyen, H. D. (2015). Development of P grac 100-based expression vectors allowing high protein production levels in Bacillus subtilis and relatively low basal expression in Escherichia coli. *Microb. Cell Factories* 14, 1–9. doi: 10.1186/s12934-015-0255-z

Porath, J., Carlsson, J., Olsson, I., and Belfrage, G. (1975). Metal chelate affinity chromatography, a new approach to protein fractionation. *Nature* 258, 598–599. doi: 10.1038/258598a0

Puetz, J., and Wurm, F. M. (2019). Recombinant proteins for industrial versus pharmaceutical purposes: a review of process and pricing. *PRO* 7:476. doi: 10.3390/pr7080476

Schmidt, T. G. M., Batz, L., Bonet, L., Carl, U., Holzapfel, G., Kiem, K., et al. (2013). Development of the twin-Strep-tag® and its application for purification of recombinant proteins from cell culture supernatants. *Protein Expr. Purif.* 92, 54–61. doi: 10.1016/j.pep.2013.08.021

Schmidt, P. M., Sparrow, L. G., Attwood, R. M., Xiao, X., Adams, T. E., and McKimm-Breschkin, J. L. (2012). Taking down the FLAG! How insect cell expression challenges an established tag-system. *PLoS One* 7:e37779. doi: 10.1371/journal.pone.0037779

Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012). NIH image to ImageJ: 25 years of image analysis. *Nat. Methods* 9, 671–675. doi: 10.1038/nmeth.2089

Schultz, J., dos Santos, A., Patel, N., and Rosado, A. S. (2023). Life on the edge: bioprospecting extremophiles for astrobiology. *J Indian Inst Sci* 103, 721–737. doi: 10.1007/s41745-023-00382-9

Shahravan, S. H., Qu, X., Chan, I. S., and Shin, J. A. (2008). Enhancing the specificity of the enterokinase cleavage reaction to promote efficient cleavage of a fusion tag. *Protein Expr. Purif.* 59, 314–319. doi: 10.1016/j.pep.2008.02.015

Shaner, N. C., Campbell, R. E., Steinbach, P. A., Giepmans, B. N. G., Palmer, A. E., and Tsien, R. Y. (2004). Improved monomeric red, orange and yellow fluorescent proteins derived from Discosoma sp. red fluorescent protein. *Nat. Biotechnol.* 22, 1567–1572. doi: 10.1038/nbt1037

Skerra, A., and Schmidt, T. G. M. (1999). Applications of a peptide ligand for streptavidin: the Strep-tag. *Biomol. Eng.* 16, 79–86. doi: 10.1016/S1050-3862(99)00033-9

Strillinger, E., Grötzinger, S. W., Allers, T., Eppinger, J., and Weuster-Botz, D. (2016). Production of halophilic proteins using Haloferax volcanii H1895 in a stirred-tank bioreactor. *Appl. Microbiol. Biotechnol.* 100, 1183–1195. doi: 10.1007/s00253-015-7007-1

Sysoev, M., Grötzinger, S. W., Renn, D., Eppinger, J., Rueping, M., and Karan, R. (2021). Bioprospecting of novel extremozymes from prokaryotes—the advent of culture-independent methods. *Front. Microbiol.* 12:630013. doi: 10.3389/fmicb.2021.630013

Tian, R., Liu, Y., Chen, J., Li, J., Liu, L., Du, G., et al. (2019). Synthetic N-terminal coding sequences for fine-tuning gene expression and metabolic engineering in Bacillus subtilis. *Metab. Eng.* 55, 131–141. doi: 10.1016/j.ymben.2019.07.001

Wang, Y., Qian, J., Shi, T., Wang, Y., Ding, Q., and Ye, C. (2024). Application of extremophile cell factories in industrial biotechnology. *Enzym. Microb. Technol.* 175:110407. doi: 10.1016/j.enzmictec.2024.110407

Waugh, D. S. (2011). An overview of enzymatic reagents for the removal of affinity tags. *Protein Expr. Purif.* 80, 283–293. doi: 10.1016/j.pep.2011.08.005

Wiśniewski, J. R., Zougman, A., Nagaraj, N., and Mann, M. (2009). Universal sample preparation method for proteome analysis. *Nat. Methods* 6, 359–362. doi: 10.1038/nmeth.1322

Yadav, D. K., Yadav, N., Yadav, S., Haque, S., and Tuteja, N. (2016). An insight into fusion technology aiding efficient recombinant protein production for functional proteomics. *Arch. Biochem. Biophys.* 612, 57–77. doi: 10.1016/j.abb.2016.10.012

Yeliseev, A., Zoubak, L., and Schmidt, T. G. M. (2017). Application of Strep-Tactin XT for affinity purification of twin-Strep-tagged CB(2), a G protein-coupled cannabinoid receptor. *Protein Expr. Purif.* 131, 109–118. doi: 10.1016/j.pep.2016.11.006

Yilmaz, G., and Arslanyolu, M. (2015). Efficient expression of codon-adapted affinity tagged super folder green fluorescent protein for synchronous protein localization and affinity purification studies in Tetrahymena thermophila. *BMC Biotechnol.* 15:22. doi: 10.1186/s12896-015-0137-9

Young, C. L., Britton, Z. T., and Robinson, A. S. (2012). Recombinant protein expression and purification: a comprehensive review of affinity tags and microbial applications. *Biotechnol. J.* 7, 620–634. doi: 10.1002/biot.201100155

Zhao, X., Li, G., and Liang, S. (2013). Several affinity tags commonly used in chromatographic purification. *J. Anal. Methods Chem.* 2013, 1–8. doi: 10.1155/2013/581093

Zheng, N., Pérez, J. J., Zhang, Z., Domínguez, E., Garcia, J. A., and Xie, Q. (2008). Specific and efficient cleavage of fusion proteins by recombinant plum pox virus NIa protease. *Protein Expr. Purif.* 57, 153–162. doi: 10.1016/j.pep.2007.10.008

frontiers | Frontiers in Microbiology

TYPE Original Research
PUBLISHED 14 August 2024
DOI 10.3389/fmicb.2024.1422844

OPEN ACCESS

EDITED BY
Maria Ines Gimenez,
National University of Mar del Plata, Argentina

REVIEWED BY
Rosa María Martínez-Espinosa,
University of Alicante, Spain
Ansgar Poetsch,
Ruhr University Bochum, Germany

*CORRESPONDENCE
Anita Marchfelder
✉ anita.marchfelder@uni-ulm.de
Christof Lenz
✉ christof.lenz@med.uni-goettingen.de

†These authors have contributed equally to this work and share first authorship

‡These authors share last authorship

RECEIVED 24 April 2024
ACCEPTED 31 July 2024
PUBLISHED 14 August 2024

CITATION
Sailer A-L, Jevtic Z, Stoll B, Wörtz J, Sharma K, Urlaub H, Dyall-Smith M, Pfeiffer F, Marchfelder A and Lenz C (2024) Iron starvation results in up-regulation of a probable *Haloferax volcanii* siderophore transporter.
Front. Microbiol. 15:1422844.
doi: 10.3389/fmicb.2024.1422844

Iron starvation results in up-regulation of a probable *Haloferax volcanii* siderophore transporter

Anna-Lena Sailer[1†], Zivojin Jevtic[2,3†], Britta Stoll[1], Julia Wörtz[1], Kundan Sharma[4], Henning Urlaub[4,5], Mike Dyall-Smith[6,7], Friedhelm Pfeiffer[1,6], Anita Marchfelder[1*‡] and Christof Lenz[4,5*‡]

[1]Biology II, Ulm University, Ulm, Germany, [2]Department of Biomedicine, University Children's Hospital, University of Basel, Basel, Switzerland, [3]Department of Biomedicine, University of Basel, Basel, Switzerland, [4]Bioanalytical Mass Spectrometry Group, Max Planck Institute for Multidisciplinary Sciences, Göttingen, Germany, [5]Bioanalytics Group, Department of Clinical Chemistry, University Medical Center Göttingen, Göttingen, Germany, [6]Computational Systems Biochemistry, Max Planck Institute for Biochemistry, Martinsried, Germany, [7]Veterinary Biosciences, Faculty of Science, Melbourne Veterinary School, University of Melbourne, Parkville, VIC, Australia

The response of the haloarchaeal model organism *Haloferax volcanii* to iron starvation was analyzed at the proteome level by data-independent acquisition mass spectrometry. Cells grown in minimal medium with normal iron levels were compared to those grown under low iron conditions, with samples being separated into membrane and cytoplasmic fractions in order to focus on import/export processes which are frequently associated with metal homeostasis. Iron starvation not only caused a severe retardation of growth but also altered the levels of many proteins. Using a comprehensive annotated spectral library and data-independent acquisition mass spectrometry (DIA-MS), we found that iron starvation resulted in significant changes to both the membrane and the soluble proteomes of *Hfx. volcanii*. The most affected protein is the RND family permease HVO_A0467, which is 44-fold enriched in cells grown under iron starvation. The gene HVO_A0467 can be deleted suggesting that it is not essential under standard conditions. Compared to wild type cells the deletion strain shows only slight changes in growth and cell morphologies show no differences. Molecular docking predictions indicated that HVO_A0467 may be an exporter of the siderophore schizokinen for which a potential biosynthesis cluster is encoded in the *Hfx. volcanii* genome. Together, these findings confirm the importance of iron for archaeal cells and suggest HVO_0467 as a siderophore exporter.

KEYWORDS

***Haloferax volcanii*, proteome, iron starvation, data-independent acquisition mass spectrometry, DIA-MS, import/export processes, metal homeostasis**

1 Introduction

Despite being one of the most frequent elements on earth, the bioavailability of iron is often limited in natural environments due to the low solubility of ferric iron (Fe^{3+}) which is the predominant oxidation state in oxic environments. Iron is an essential element for most microbial organisms due to its key role in, e.g., iron–sulfur clusters, iron centers, and heme groups. Consequently, organisms have developed various iron-acquisition pathways, among others high-affinity iron uptake systems which often use highly specific iron-binding

compounds, i.e., siderophores (Miethke and Marahiel, 2007). Understanding the response of organisms toward iron limitation conditions (hereafter referred to as iron starvation) provides insight into how they have adapted to cope with stresses in their natural environment.

A lot of attention has been devoted to the response mechanisms of bacteria to iron starvation, including changes in the biosynthesis, secretion and uptake of siderophores to sequester iron. Surprisingly little is known, however, about the iron starvation response of archaea, e.g., halophilic archaea (here referred to as haloarchaea). These commonly live in poorly oxygenated environments at extremely high salt concentrations. Several studies have dealt with iron biology of haloarchaea, mainly using the model species *Haloferax volcanii* and *Halobacterium salinarum*. The effect of iron limitation on the respiratory chain of *Hbt. salinarum* has been analyzed (Hubmacher et al., 2003). Iron uptake has been measured in *Hbt. salinarum* strain JW5 and was attributed to an energy-dependent iron uptake process (Hubmacher et al., 2007). However, the responsible gene has not yet been identified. In *Halobacterium*, the SirR/DtxR family transcription regulators Idr1 and Idr2 were characterized and shown to be involved in iron homeostasis (Schmid et al., 2011).

In *Hfx. volcanii*, the genes *iucA* and *iucC* were identified by the antiSMASH server as predicting the synthesis of the siderophore schizokinen (Blin et al., 2023). A detailed bioinformatic reconstruction (Niessen and Soppa, 2020) identified a six-gene cluster which codes for a siderophore biosynthetic pathway. The products of four genes (*iucABDC*) were identified by bioinformatics as converting the precursor diaminopropane to schizokinen and the other two genes (*dat*, *bdb*) as specifying enyzmes that generate diaminopropane from aspartate-semialdehyde. The transcription of the cluster is induced under iron limitation conditions and deletion of the *iucABDC* genes resulted in a growth defect at very low concentrations of Fe^{3+} but not of Fe^{2+} (Niessen and Soppa, 2020). This is consistent with schizokinen having an extremely high affinity for Fe^{3+} (Mullis et al., 1971; Storey et al., 2006; Chuljerm et al., 2019, 2020). More recently, three diphtheria toxin repressor proteins Idr, SirR, and TroR were identified as potential transcriptional regulators of iron homeostasis (Martinez Pastor et al., 2024).

Hfx. volcanii is a model organism for halophilic archaea, the genome is sequenced (Hartman et al., 2010) and its biology is under intense study (Haque et al., 2020), leading to increased functional annotation and understanding. One key method to study the response of an organism to different environmental conditions is quantitative proteomics. Several proteomic analyses of *Hfx. volcanii* have been reported, such as the response to oxidative stress induced by hypochlorite (McMillan et al., 2018), changes in salt concentration and temperature (Jevtic et al., 2019), and the transition from exponential to stationary phase (Cerletti et al., 2018). Multiple proteome data sets available for *Hfx. volcanii* have been analyzed in a community effort and combined in ArcPP, the archaeal proteome project (Schulze et al., 2020). None of these global data sets, however, included *Hfx. volcanii* cultivated under conditions of iron starvation. We therefore set out to perform a comprehensive proteome study under these conditions using label-free data-independent acquisition mass spectrometry (DIA-MS). DIA-MS is a well-established method that provides high coverage of identified and quantified proteins without disturbing the experimental system. In addition, we selected to study the impact of iron starvation on membrane-associated and soluble proteins separately following fractionation, since many known iron scavenging and import processes in other species rely on dedicated transmembrane import systems. We therefore chose a fractionated approach to obtain higher coverage of membrane-associated proteins, and closely interrogated the data both for general stress response mechanisms and, more specifically, for evidence of iron uptake mechanisms. Furthermore, we investigated the proteins most affected under iron starvation in more detail.

2 Materials and methods

2.1 Strains and growth conditions

Hfx. volcanii strain H119 was used as wild type strain for all experiments, as reported in a previous publication (Jevtic et al., 2019). In short, it is derived from wild-type strain DS2 by removal of plasmid pHV2 and deletion of three genes that can be used as selection markers (Δ*pyrE2*, Δ*trpA*, Δ*leuB*). H119 was grown aerobically with shaking (200 rpm) to exponential phase (0.55 OD_{650}) at 45°C in 2.2 M salt in minimal medium Hv-min (Allers et al., 2004) containing normal or low iron (normal: 0.0083 mM $FeSO_4$*$7H_2O$; low iron: $FeSO_4$*$7H_2O$ was omitted; no efforts to further minimize residual iron were taken). Two biological replicates were prepared for each sample. In an additional control experiment, cells were grown in complex medium Hv-YPC (Allers et al., 2004) at three biological replicates per sample. Refer to Supplementary Table S3 for an overview of the samples.

2.2 Protein isolation and in-solution digestion

Cells were harvested by centrifugation, resuspended in 18% saltwater (2.46 M NaCl, 88 mM $MgCl_2$, 85 mM $MgSO_4$, 56 mM KCl, 12 mM Tris–HCl, pH 7.5), sonicated and subsequently incubated with sodium taurodeoxycholate to a final concentration of 0.006% as reported previously (Jevtic et al., 2019). Membranes were collected by centrifugation (100,000 × g for 1 h at 4°C), and membrane pellets and supernatants were processed separately. Membranes were resuspended by sonication in 5 mL 18% saltwater. Both samples were incubated with DNAse I, exonuclease III, and RNase A to digest nucleic acids. Aliquots of 0.5 mL were frozen in liquid nitrogen and stored at −80°C. Aliquots were thawed and proteins were precipitated by adding 4.5 mL 100% cold acetone and incubation at −20°C overnight. Proteins were pelleted by centrifugation at 6,000 × g for 60 min at 4°C. Pellets were washed four to five times with 80% acetone and in a final washing step with 5 mL 80% ethanol. Pellets were air-dried and frozen in liquid nitrogen until further use.

Forty micrograms of precipitated protein per biological replicate were solubilized using 3-[(2-methyl-2-undecyl-1,3-dioxolan-4-yl) methoxy]-1-propanesulfonate cleavable surfactant (Rapigest, Waters). Following reduction and alkylation with dithiothreitol and iodoacetamide, proteins were digested using sequencing grade porcine trypsin (Promega) at an enzyme-to-substrate ratio of 1:50 (w:w). Following acidic cleavage of the surfactant, the resulting fatty acids were pelleted by centrifugation and removed. Supernatant peptide mixtures were dried in a SpeedVac concentrator and stored at −20°C prior to analysis.

2.3 Mass spectrometry acquisition

The mass spectrometric analysis was performed as described previously (Jevtic et al., 2019). For generation of an annotated MS/MS spectral library, peptide aliquots of three biological replicates of a *Hfx. volcanii* culture grown in complex medium were pooled to a total amount of 80 µg, and separated into eight fractions using a reversed phase spin column (Pierce High pH Reversed-Phase Peptide Fractionation Kit, Thermo Fisher Scientific). These fractions, as well as all biological replicates of membrane and soluble samples, were spiked with a synthetic peptide standard (iRT Standard, Biognosys) for retention time alignment.

Samples were analyzed on a nanoflow chromatography system (Eksigent nanoLC425) hyphenated to a hybrid triple quadrupole-TOF mass spectrometer (TripleTOF 5,600+) equipped with a Nanospray III ion source (Ionspray Voltage 2,400 V, Interface Heater Temperature 150°C, Sheath Gas Setting 12) and controlled by Analyst TF 1.7.1 software build 1,163 (all Sciex). In brief, peptides were dissolved in loading buffer (2% acetonitrile, 0.1% formic acid v/v in water) to a concentration of 0.3 µg/µL. For each analysis, 1.5 µg of digested protein was enriched on a precolumn (0.18 mm ID × 20 mm, Symmetry C18, 5 µm, Waters) and separated on an analytical RP-C18 column (0.075 mm ID × 250 mm, HSS T3, 1.8 µm, Waters) using a 90 min linear gradient of 5–35% acetonitrile/0.1% formic acid (v:v) at 300 nL/min.

Qualitative LC/MS/MS analysis was performed to construct an annotated MS/MS spectral library to improve detection rates. A top25 data-dependent acquisition (DDA) method was used, with an MS survey scan of *m/z* 350–1,250 accumulated for 350 ms at a resolution of 30,000 full-width at half-maximum (FWHM). MS/MS scans of *m/z* 180–1,600 were accumulated for 100 ms at a resolution of 17,500 FWHM and a precursor isolation width of 0.7 FWHM, resulting in a total cycle time of 2.9 s. Precursors above a threshold MS intensity of 125 with charge states 2+, 3+ or 4+ were selected for MS/MS; the dynamic exclusion time was set to 30 s. MS/MS activation was achieved by CID using nitrogen as a collision gas and the manufacturer's default rolling collision energy settings. Fractionated rich media samples were analyzed in technical duplicate, and membrane and soluble biological samples in technical triplicate by repeated injections.

For quantitative analysis, LC/MS/MS data were acquired by data-independent acquisition (DIA) using 65 variable size windows (Zhang et al., 2015) across the 400–1,050 *m/z* range. Fragments were produced using rolling collision energy settings for charge state 2+, and fragments acquired over an *m/z* range of 350–1,400 for 40 ms per segment. In combination with a 100 ms survey scan, the overall cycle time was 2.75 s. Each biological replicate was analyzed by triplicate injection.

2.4 Mass spectrometry data processing

Mass spectrometry data were processed using Spectronaut v.14.2 software (Biognosys) (Bruderer et al., 2015). A hybrid spectral library was generated using the Pulsar search engine by searching all DDA and DIA data against an in-house *Hfx. volcanii* protein sequence database (version 06/2019) containing 4,107 protein entries and 79 additional noncoding ORFs[1] augmented with 52 common lab contaminants. Protein identification was performed using default settings (enzyme specificity Trypsin/P, 2 missed cleavages; fixed Iodoacetamide (C), variable Oxidation (M), Acetylation (N-term); dynamic mass and retention time recalibration) to an estimated False Discovery Rate (FDR) of 1%, resulting in a spectral library containing 46,172 precursors mapping to 27,039 primary peptide sequences.

For protein quantitation, the spectral library was used to extract fragment ion chromatograms from the DIA data (dynamic mass and retention time alignment, up to 6 fragments per peptide, up to 10 peptides per protein, 1% FDR) resulting in quantitative values for 29,333 precursors mapping to 20,374 primary peptide sequences and 2,262 protein groups (2,254 *Hfx. volcanii*, 8 laboratory background). Quantitative data were normalized using quantile normalization, resulting in median coefficients of variation within each biological state below 20.1 and 30.0% on the peptide and protein levels, respectively.

2.5 Data analysis and functional annotation

Protein quantitation results were analyzed in Perseus v1.5.6.0 (Max Planck Institute of Biochemistry) (Tyanova and Cox, 2018). After grouping of individual data files for biological states and logarithmic transformation, missing values were imputed from a low-level normal distribution (width 0.3, downshift 1.8) and abundance changes analyzed for statistical significance by pairwise Student's *t*-testing with an FDR-based multiple testing correction ($p < 0.05$, S0 = 1). Differentially abundant protein populations were analyzed for enrichment of arCOGs (Makarova et al., 2007, 2015) using FunRich v3.1.3 software (Pathan et al., 2015), and for enrichment of KEGG classes (Kanehisa and Goto, 2000) using DAVID Functional Annotation Tool v6.8 (Huang da et al., 2009).

2.6 Generation and analysis of deletion strain ΔHVO_A0467

For the deletion of HVO_A0467, plasmid pTA131-up.do (HVO_A0467) (Supplementary Table S1) was generated by cloning a fragment of 467 bp upstream (including the first base pair of HVO_A0467 since this is also part of the upstream gene HVO_A0466), and a fragment of 426 bp downstream of HVO_A0467 into the pTA131 vector. Subsequently strain H119 was transformed with this plasmid and plated on plates lacking uracil to mediate integration of the plasmid into the genome. Pop-in candidates were validated by PCR and streaked out on agar plates containing 5-fluoroorotic acid to force pop-out of the plasmid backbone. Pop-out candidates were investigated further by Southern blot analysis using 10 µg of genomic DNA digested with SalI and separated on a 0.8% agarose gel. DNA fragments were transferred to a nylon membrane Hybond-N+ (GE Healthcare) by capillary blotting. For verification of the knockout, two probes were used that bind either to the gene or next to the deleted region. A 362 bp fragment of the region upstream of HVO_A0467 was amplified using primers HVO_A0467-up-fwd and HVO_A0467-up-rev. A 255 bp fragment of the gene itself was amplified using oligonucleotides HVO_A0467-intern-fwd and HVO_A0467-intern-rev to serve as a gene-specific probe. The PCR fragments were labeled with [α-^{32}P]-dCTP using the DECAprime II DNA labeling kit (Thermo Fisher Scientific). The labeled products were used as hybridization probes. The resulting deletion strain was termed

1 https://doi.org/10.5281/zenodo.3565631

HV118 (Supplementary Figure S1). The deletion strain was compared to wild type cells using light microscopy and growth analysis. For phase-contrast-microscopy, cells were grown in YPC or adjusted media preparations (see "Strains and Growth conditions") until they reached the desired OD_{650} indicating the culture to be in the exponential or stationary growth phase. The OD_{650} for exponential and stationary phase differed depending on the medium used: exponential phase: YPC: OD_{650}: 0.4–0.6; Hv-min: OD_{650}: 0.2–0.4; Hv-min Low iron: OD_{650}: 0.1–0.3; stationary phase: YPC: $OD_{650} > 0.9$; Hv-min: $OD_{650} > 0.5$; Hv-min Low iron: $OD_{650} > 0.3$. A 2 µL sample of culture was placed on an agarose pad prepared by dropping ~50 µL of 1% agarose containing 18% BSW (Dyall-Smith, 2008) onto a glass slide at room temperature, and a clean 22 × 50 mm number 1.5 glass coverslip placed on top. Images were acquired using a 100 × / 1.4 NA oil immersion objective and phase contrast optics using a Leica DM6 B Microscope (Leica Biosystems).

For growth experiments strains H119 (wild type) and ΔHVO_A0467 were precultured in YPC to $OD_{650} = 0.4–0.6$, washed once with the adjusted medium (see above "Strains and growth conditions") and then diluted to $OD_{650} = 0.05$ and transferred to microtiter plates in biological triplicates and 5 technical replicates. These were subsequently cultured in a heated plate reader (Epoch 2 NS Microplate Spectrophotometer, Agilent Technologies) (aerobically, orbital shaking at 425 rpm, 45°C) while OD_{650} was measured every 30 min. Outer wells were filled with salt water as evaporation barriers (Jaschinski et al., 2014).

2.7 SDS-page

Cells were grown in Hv-min with or without the addition of $FeSO_4$ to an OD_{650} of 0.4–0.6 before they were harvested at 10,000 × g at 4°C for 30 min. Cell pellets were washed in enriched PBS and then lysed by resuspending in lysis buffer and ultrasonication. After centrifugation (30 min, 18,000 × g, 4°C) the protein concentration of the supernatant fraction was determined using a Roti-Nanoquant assay according to the manufacturer's protocol. 5 µg were loaded onto a 10% polyacrylamide gel. Subsequently the gel was stained with Coomassie Brilliant blue.

2.8 Ligand docking

The AlphaFold predicted 3D structure of *Hfx. volcanii* protein HVO_A0467 was downloaded as a pdb file (AF-D4GRD2-F1-model_v3.pdb, accessed 24/5/2023) from UniProtKB entry D4GRD2.[2] The protein was inspected using PyMOL (v.2.4.0, Schrödinger) and checked to ensure there were no extraneous molecules. The schizokinen structure was downloaded from the Zinc15 database[3] in SMILES format. The CB-Dock2 server (cadd.labshare.cn/cb-dock2) was used for protein-ligand blind docking (Liu et al., 2022). The SMILES structure of schizokinen was input into CB-Dock2 using the 'draw ligand' option. CB-Dock2 automates the preparation of ligand and protein for docking, determines potential binding pockets of the protein, and uses AutoDock Vina for ligand docking (Eberhardt et al., 2021). The docked ligand was examined further using Pymol and PLIP (Adasme et al., 2021).

2 www.uniprot.org/uniprotkb/D4GRD2/entry

3 zinc15.docking.org/substances/ZINC000014692536

2.9 Analysis of HVO_A0467 and HVO_A0466 homologs by BLASTp and phylogenetic tree reconstructions

Twenty one species with a completely sequenced genome were selected that represent 18 genera from 3 orders of *Halobacteriaceae* (Supplementary Table S2). First, RND permeases were selected based on their annotation from 11 species which are under continuous annotation review (category A) (Pfeiffer and Oesterhelt, 2015). BLASTp analysis confirmed that this collection is exhaustive. Similarly, all HVO_A0466 homologs were selected (annotated as RND-T-associated protein). BLASTp analysis identified two additional solo homologs (annotated as uncharacterized). From the genomes of category B, all proteins having the term RND in the protein name were also selected. The complete set of HVO_A0467 homologs was subjected to BLASTp analysis against all represented species in order to identify additional homologs. BLASTp matches which are attributed to extensive coiled-coil regions were ignored. Similarly, the complete set of already identified HVO_A0466 homologs was subjected to BLASTp analysis against all represented species in order to identify additional homologs.

For phylogenetic tree reconstructions, the webservice at https://mafft.cbrc.jp/alignment/server/ was used. Sequences in fasta format were uploaded, aligned with the MAFFT aligner (default settings) and trees inferred using the Neighbor-Joining algorithm (JTT model, 100 bootstrap sampling), and viewed using the archaeopteryx.js tool (Kuraku et al., 2013).

2.10 Additional bioinformatic analysis

Protter software[4] (Omasits et al., 2014) was used to illustrate the transmembrane topology of HVO_A0467. For HVO_A0467, transmembrane domains were loaded from UniProtKB entry D4GRD2. The localization of the N-terminus was set manually to cytoplasmic, so that the large extramembrane domains are periplasmic. This is consistent with other RND permeases with available 3D structure (PDB:6AJJ; PDB:5KHS). This is also consistent with the topology prediction of TMHMM[5] (Krogh et al., 2001) and DeepTMHMM[6] (Hallgren et al., 2022), but opposite to the topology prediction by Phobius[7] (Kall et al., 2007).

3 Results

3.1 Biological characterization of *Haloferax volcanii* under iron starvation conditions

Growth of *Hfx. volcanii* strain H119 was analyzed in minimal medium under conditions of iron starvation and at normal iron levels, and for comparison in complex medium (Figure 1A). While growth in minimal medium was reduced compared to complex

4 https://wlab.ethz.ch/protter/start/

5 https://services.healthtech.dtu.dk/services/TMHMM-2.0/

6 https://dtu.biolib.com/DeepTMHMM

7 https://phobius.sbc.su.se/

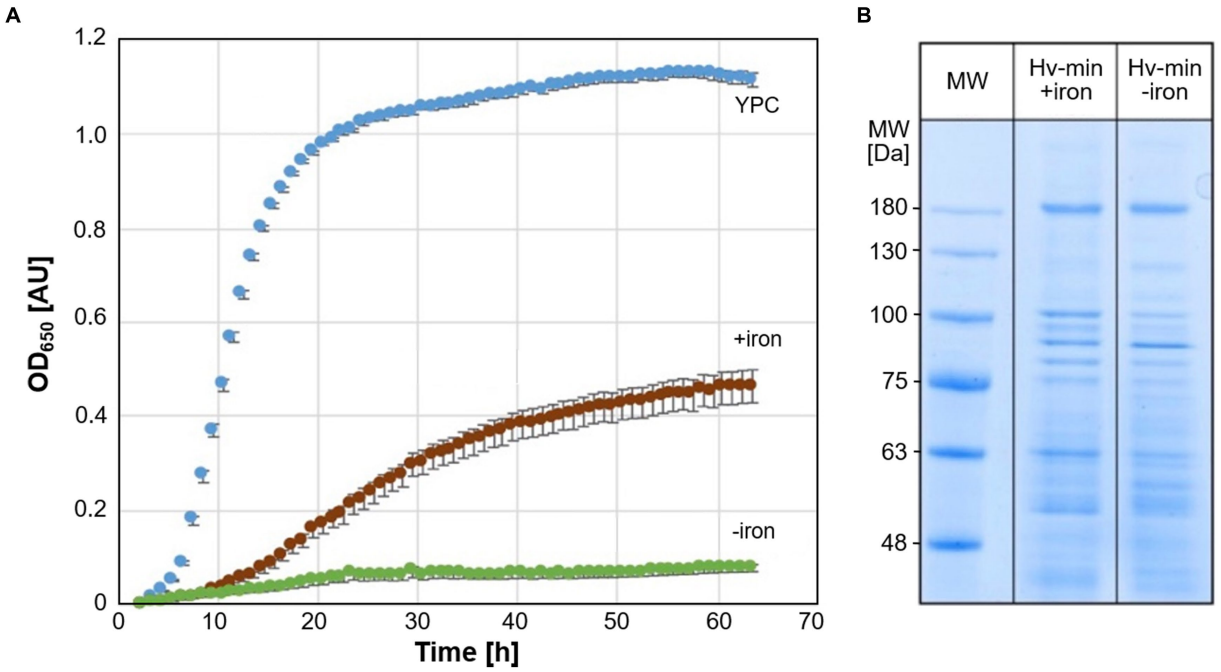


FIGURE 1
Growth of *Hfx. volcanii* and SDS-PAGE analysis of cell proteins. **(A)** Growth of *Hfx. volcanii* H119 in minimal medium (Hv-min) with addition of $FeSO_4$ (brown, +iron) or without added iron (green, −iron), and for comparison in complex medium (blue, YPC) at 45°C. $n = 3$ biological replicates, with 5 technical replicates each. y-axis: optical density at 650 nm, x-axis: time of growth in hours. **(B)** SDS-PAGE analysis of *Hfx. volcanii* H119 grown in minimal medium with (+iron) or without (−iron) addition of $FeSO_4$. MW: protein apparent molecular weight marker.

medium, the growth of cells under iron starvation was severely attenuated. SDS-PAGE of cell extracts obtained from cultures grown with and without addition of $FeSO_4$ exhibited only moderate differences in protein patterns (Figure 1B).

Cells are generally smaller when grown in minimal medium instead of complex medium, but iron starvation did not cause any discernible differences in cell morphology (Figure 2). Cells were observed in late exponential and stationary phase; analysis at early exponential phase was not performed due to severe growth attenuation in the absence of iron.

3.2 Large-scale proteome analysis of *Haloferax volcanii* under iron starvation conditions

Hfx. volcanii cells grown in minimal medium with or without added iron were separated into soluble and membrane fractions, and their respective proteomes analyzed. As a first step, comprehensive protein identification was performed using both data-dependent acquisition (DDA) and data-independent acquisition (DIA) data of the fractionated samples, as well as of a reference culture grown in complex medium. We added the latter to maximize proteome coverage, as some proteins may not have been highly expressed upon growth in minimal media, and thus would not have been represented in the annotated MS/MS spectral library used for quantitative data extraction from the DIA data.

After mapping to UniprotKB accessions, this identified a total of 2,784 non-redundant *Haloferax* proteins (i.e., protein groups) evidenced by 27,040 proteotypic peptide sequences at a False Discovery Rate (FDR) of 1%. This corresponds to a coverage of 68.3% of the theoretical proteome, and compares favorably with the numbers reported by the community Archaeal Proteome Project (ArcPP) (Schulze et al., 2020). This data set includes the first mass spectrometric evidence for 12 protein products from the *Hfx. volcanii* protein sequence database which are not contained in the ArcPP data set. The identified peptide and protein identifications were transcribed into an annotated MS/MS spectral library.

Next, we performed DIA-MS-based quantitative profiling of the membrane and soluble proteomes of *Hfx. volcanii* under iron starvation as well as under normal iron condition. Using the above spectral library, we obtained quantitative data for a total of 2,254 *Haloferax* proteins (1,786 ± 147 per injection), spanning 5 orders of magnitude linear abundance range as estimated by their iBAQ values (Schwanhausser et al., 2011) and corresponding to 55.4% of the theoretical proteome (Supplementary Table S3). To assess any methodological bias in the detected proteins, a comparison of the functional classes of this set to those of the entire proteome was conducted (see Supplementary Text S1; Supplementary Figure S1). No significant bias was detected. We also tested for the enrichment of membrane-associated proteins in the membrane fraction, and found a systematic enrichment of, e.g., ABC-type transport system proteins corresponding to the *dppx*/*tsgX*/*znuX* genes as *bona fide* membrane or membrane-associated proteins in the pellet fraction to validate our experimental approach (Figure 3).

3.3 Differential protein expression under iron starvation conditions

We examined abundance changes in the membrane and soluble fractions caused by iron starvation for statistical significance by pairwise

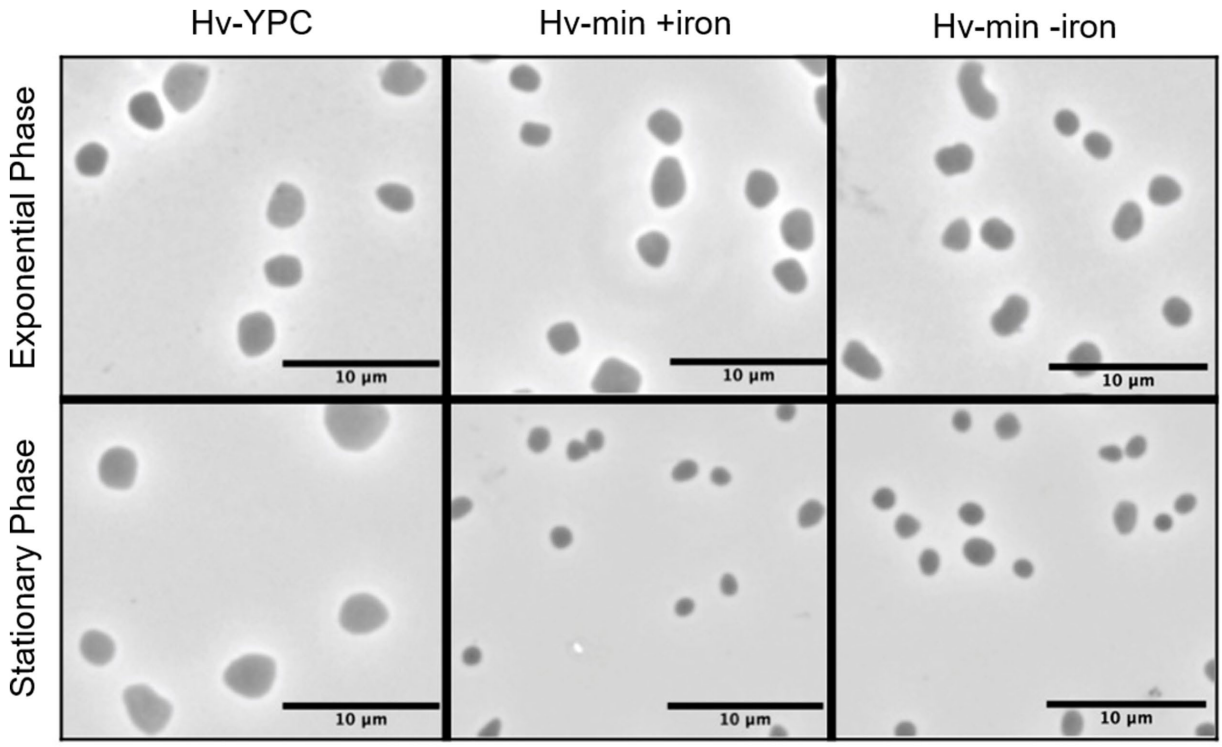


FIGURE 2
Microscopic analysis of cell morphology in the presence and absence of $FeSO_4$. Phase contrast images of *Hfx. volcanii* H119 cells grown in complex medium (YPC, left panel), in minimal medium (Hv-min) with (+iron, middle panel) or without added $FeSO_4$ (–iron, right panel), to exponential (upper panels) or stationary phase (lower panels). The size bar represents 10 µm.

testing (Supplementary Table S4; Figure 4), and observed a total of 1,429 differentially expressed proteins: 1,138 exclusively in the soluble fraction, 136 exclusively in the pelleted membranes, and 155 in both fractions. The large number of differently expressed proteins indicates a significant fine-tuning of the *Hfx. volcanii* proteome under iron starvation conditions.

To find commonalities in the complex data, we subjected the populations of differentially abundant proteins to functional enrichment analysis, both with regard to their associated arCOG classes and to their KEGG metabolic circuits (see Supplementary Text S2). ABC transporters are among the most prominent class of proteins which were found differentially regulated. Interestingly, we do not see major changes in proteins related to stress response.

The most highly regulated proteins in our data set in either the membrane or soluble fraction were examined in detail. In total, 118 proteins were highly regulated (≥4-fold abundance change), with 79 proteins being highly up-regulated and 39 proteins being highly down-regulated (Supplementary Table S5). Table 1 lists the 19 proteins with at least 10-fold abundance changes.

HVO_A0467 is the most highly induced protein, having an outstanding up-regulation (44-fold) specifically in the membrane fraction. It is an integral membrane protein, distantly related to the superfamily of RND family permeases, and further evidence provided below (next section) indicate it is a likely siderophore exporter. The second-most highly up-regulated protein is the SMC-like protein Sph3, which was recently shown to be required for rod formation (Schiller et al., 2024). However, cell morphology analysis was not performed in early exponential phase (where cells would form rods in the presence of Sph3) because of the limited growth at low iron conditions, with OD_{650} typically well below 0.1. Among the remaining 19 proteins showing at least 10-fold up-regulation is an ABC transporter permease subunit (HVO_B0197, fhuG2). The *fhuCGD* ABC transporters have an iron-loaded siderophore as a probable substrate according to regulon-based substrate assignment, and all *fhu* genes/operons are preceded by a *ferR* regulon signal (Leyn and Rodionov, 2015) (see Supplementary Text S5). HVO_B0197 is one of two paralogous permeases and shows 12.7-fold up-regulation. There are 11 paralogous substrate-binding proteins of the iron-loaded siderophore ABC transporter (*fhuD*), 5 of which are up-regulated more than 4-fold (HVO_1464, 8.0-fold; HVO_A0557, 7.0-fold, HVO_B0047, 5.2-fold, HVO_B0144, 4.9-fold and HVO_B0198, 4.5-fold). None of the remaining, highly up-regulated proteins with known function appear to show a direct relation to iron transport and metabolism. Many of the regulated proteins lack a definite function assignment or are totally uncharacterized, i.e., are annotated as conserved hypothetical protein.

The most highly down-regulated protein (HVO_A0584, 13.8-fold) has no obvious connection to iron starvation. Down-regulation of a general transcription factor (HVO_1676, *tfb2*, 12.5-fold) may indicate multi-layer gene regulation, as general transcription factors have been implicated in gene regulation in haloarchaea (Coker and DasSarma, 2007; Lu et al., 2008). Other transcription factors are also down-regulated (HVO_2268, *tfb4*, 5.9-fold, HVO_1052, *tfb1*, 4.7-fold). The down-regulation of HVO_B0041 (*iucC*, 10.4-fold) is counter-intuitive for a siderophore biosynthesis protein. Another protein from the same operon, HVO_B0044 (*iucA*) is oppositely regulated (3.8-fold up-regulation). More details about the siderophore biosynthesis operon are given in Supplementary Text S6. Finally, one subunit of a respiratory complex I homolog (Nuo

10.3389/fmicb.2024.1422844

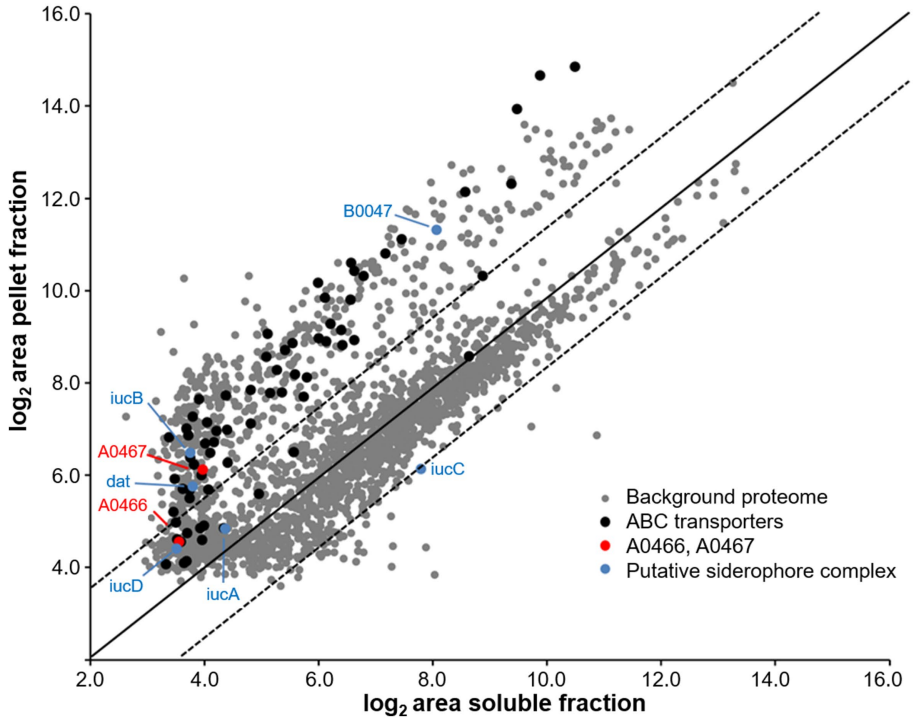


FIGURE 3
Enrichment of membrane-associated proteins in the pellet fraction. Observed log_2 protein areas in the soluble and pellet fractions. Black, ABC type transport proteins (dppX/tsgX/znuX genes); red, proteins HVO_A0466 and HVO_A0467 investigated in this study; blue, putative siderophore complex proteins (Niessen and Soppa, 2020). Black line indicates equal abundance, dashed lines a 3-fold enrichment or depletion in the membrane fraction.

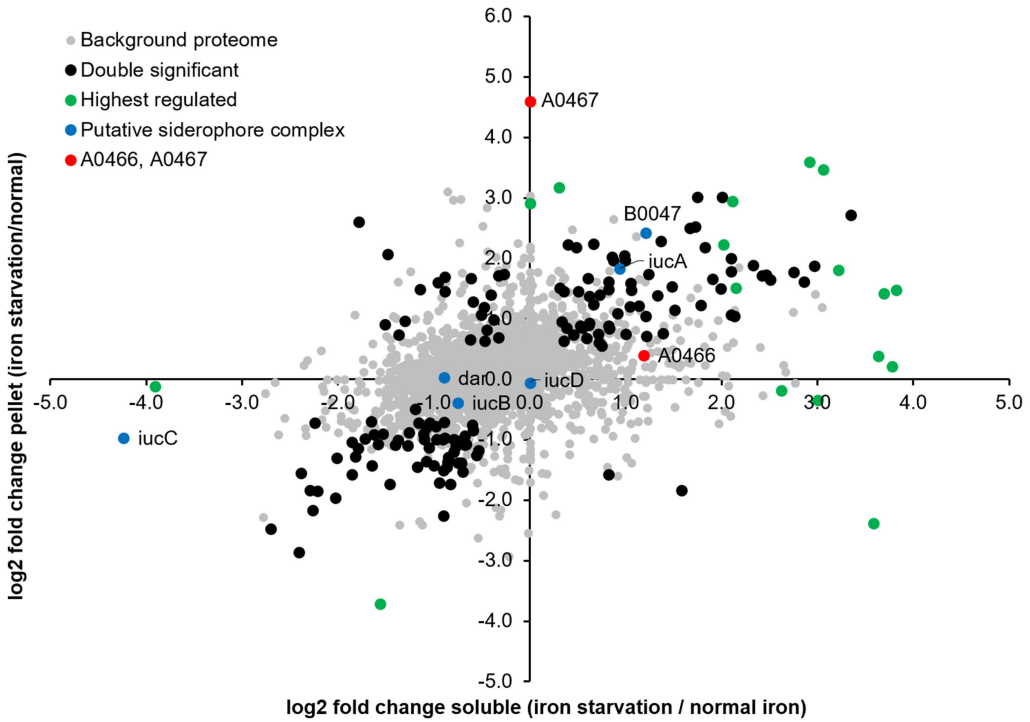


FIGURE 4
Differential protein expression under iron starvation condition. log_2 fold changes observed in the soluble and membrane fractions relative to normal iron conditions. Black, proteins significantly differentially expressed in both soluble and pellet fractions; red, proteins HVO_A0466 and HVO_A0467 investigated in this study; green, other highest regulated proteins; blue, putative siderophore complex proteins (Niessen and Soppa, 2020).

TABLE 1 Proteins that are most affected by iron starvation.

Gene	Fold regulation	Protein name
Proteins more than 10 fold up-regulated under iron starvation condition		
HVO_A0467	+43.7	RND superfamily permease
HVO_2175	+17.7	Smc-like protein Sph3
HVO_1017	+15.6	Ketosamine kinase domain protein
HVO_1958	+14.5	Pyruvoyl-dependent arginine decarboxylase
HVO_0834	+14.4	DUF35 family protein
HVO_1004	+14.1	Conserved hypothetical protein
HVO_2009	+14.0	Conserved hypothetical protein
HVO_A0090	+13.4	Gentisate 1,2-dioxygenase
HVO_A0549	+13.0	NosL family protein
HVO_0655	+12.7	GalE family epimerase/dehydratase
HVO_B0197	+12.7	ABC-type transport system permease protein (probable substrate siderophore)
HVO_2354	+12.5	Conserved hypothetical protein
HVO_1422	+12.3	XerC/D-like integrase
HVO_2795	+12.0	GNAT family acetyltransferase
HVO_1859	+11.6	UPF0104 family protein
HVO_2359	+11.6	Conserved hypothetical protein
Proteins more than 10 fold down-regulated under iron starvation		
HVO_A0584	−13.8	Inosine-5′-monophosphate dehydrogenase
HVO_1676	−12.5	Transcription initiation factor TFB
HVO_B0041	−10.4	Siderophore biosynthesis protein IucC

Data for the soluble and membrane fraction are combined, and the higher of the two regulation factors is given for proteins which were found significantly regulated in both samples. Upper set, proteins more than 10 fold up-regulated under iron starvation condition. Lower set, proteins more than 10 fold down-regulated under iron starvation.

complex) is strongly down-regulated (NuoN, 9.9-fold). Of the 11 subunits of the Nuo complex, 9 were found to be down-regulated. Also, additional subunits of the respiratory chain were found to be down-regulated (Supplementary Text S4).

Besides looking at the highest differentially abundant protein, we also examined the expression of known iron-containing proteins. Indeed we consistently found slightly (~2fold), but statistically significantly reduced expression of a number of cytochrome complex proteins (e.g., HVO_0038 *cyc1*, HVO_0841 *petD*, HVO_0842 *petB*), ferredoxins (HVO_2995 *fdx*, HVO_1831 *ferA4*, HVO_2729 *ferB2*) and ferredoxin-related proteins (HVO_0869 *gltB*, HVO_0887 *korB*, HVO_0888 *korA*, HVO_1304 *porB*, HVO_1305 *porA*) as well as other putative iron–sulfur proteins (HVO_A0083, HVO_1695) in the cytosol under iron starvation conditions. This reduced expression of key components of the redox chain reflects the effects of iron starvation on cellular metabolism, and highlights the importance of elucidating potential siderophore systems in more detail.

3.4 HVO_A0467, a potential siderophore exporter, is the most highly regulated protein

As pointed out above, HVO_A0467 is the most highly regulated protein, being up-regulated 44-fold under iron starvation conditions (Figure 4; Table 1). It is an 827 amino acid integral membrane protein with 12 predicted transmembrane domains (Figure 5) but was undetected in many previous proteome studies, such as the 12 MS datasets used for the original Archaeal Proteome Project (arcPP) (Schulze et al., 2020). A later dataset (Schulze et al., 2021) detected two peptides, however each only with a single peptide-to-spectrum match. This is in line with its observation at very low levels in our data under normal iron concentrations. In our proteomic data set it is reliably detected with 3 peptides (^{387}LVCDNVRTR^{395}, ^{578}IYDELFASSASGR^{590}, ^{632}ATSTGDVVVR^{641}).

To gain further insight into the function of this protein, a bioinformatic approach was used to detect relationships with distant homologs. HVO_A0467 is an RND superfamily transport protein, based on BLASTp comparison to the transporter classification database (TCDB) (Saier et al., 2021). The top hits are transporters (permeases) of the RND (Resistance/Nodulation/Cell division) superfamily. There are only extremely distant homologs in the annotated SwissProt section of UniProtKB, with the best hit (26% sequence identity) being uncharacterized. The N-terminal and C-terminal halves of RND transporters share sequence similarities, each half consisting of an MMPL (Mycobacterial Membrane Protein Large) domain (InterPro: IPR004869; Domenech et al., 2005). *Mycobacterium* codes for 12 RND transporters, one of which was shown to be involved in siderophore export (Wells et al., 2013).

Protein-ligand docking was used to investigate whether HVO_A0467 might function as an exporter of the *Hfx. volcanii* siderophore schizokinen. The identity of this siderophore was proposed earlier (Niessen and Soppa, 2020) and is consistent with functional assignment by the antiSMASH server (Blin et al., 2023). Schizokinen was docked *in silico* to the 3D structure of HVO_A0467 predicted by AlphaFold using the CB-Dock2 server (Liu et al., 2022). Five potential binding pockets were detected on the surface of HVO_A0467, with the largest (C1, 3,612 Å^3) being at the base of the transmembrane domain (TMD), a region that includes a central channel as well as residues predicted to be projecting into the cytoplasm. The C1 pocket produced the highest free binding energy for schizokinen by far (Vina score − 7.5 kcal/mol), with the other pockets located in the head domains showing lower binding energies of −5.5 kcal/mol or less. The free binding energy of −7.5 kcal/mol for the C1 pocket is similar to free binding energies of −7.7 to −8.2 kcal/mol, which have been obtained with other ligand-transporter interactions (Kumar et al., 2017; Zhang et al., 2019; Ple et al., 2022). Figure 6 shows schizokinen bound in the central channel of the C1 pocket, near the base of the TMD. It is predicted to have binding interactions with 10 sidechains.

Schizokinen is a hydroxamate-type siderophore, and is hydrophilic (Khan et al., 2018), which makes it different from the well-studied lipophilic substrates that have been used to study RND transporters, such as the mycolic acid transporter MmpL3 (Klenotic et al., 2021). Lipophilic molecules enter the cell membrane and are then taken up by RND transporters via channels that open into the membrane. Schizokinen is likely to bind initially to the cytoplasmic base of HVO_A0467, although it remains unknown if transport can be performed by a single protein, or if it needs to assemble into a homodimer or trimer.

10.3389/fmicb.2024.1422844

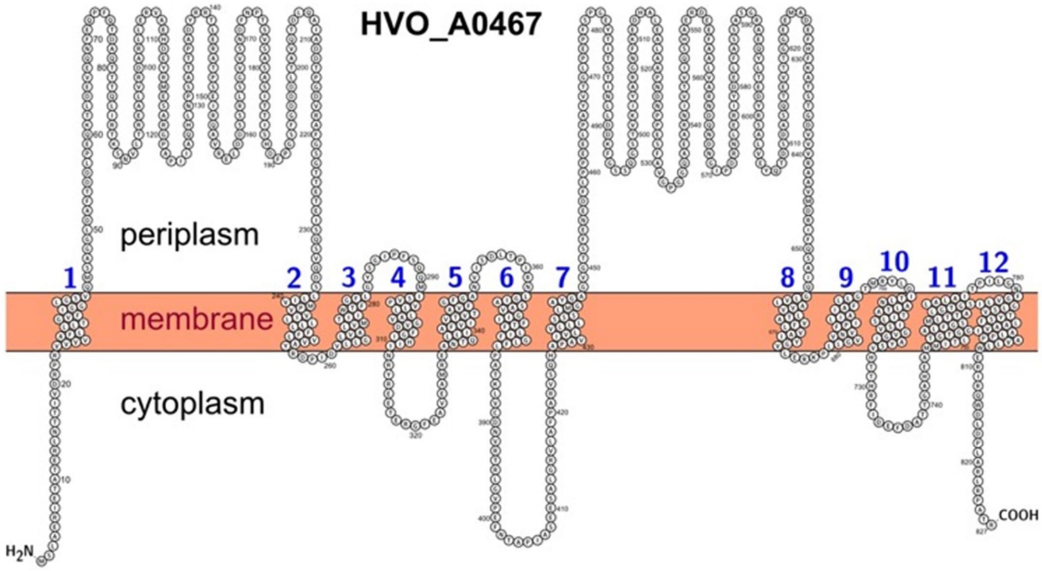


FIGURE 5
Predicted membrane topology of HVO_0467 as depicted using the PROTTER webserver. The transmembrane domains are based on UniProt features and are numbered from 1 to 12. The N-terminus was set to intracellular as predicted by DeepTMHMM and TMHMM, contrary to a Phobius prediction.

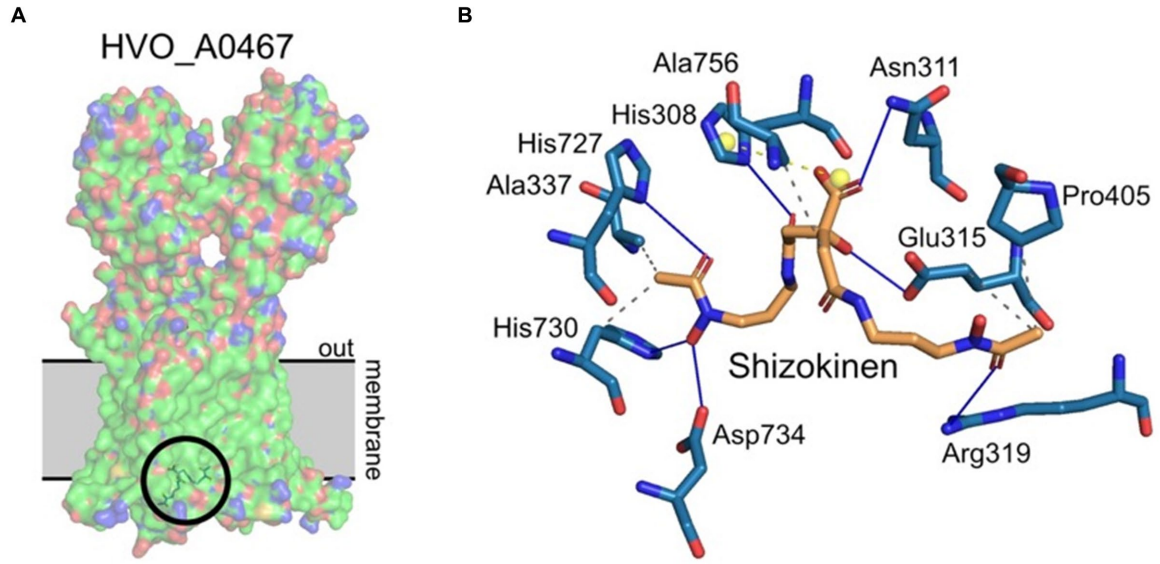


FIGURE 6
Protein/ligand docking of schizokinen to HVO_A0467. **(A)** Lowest energy binding pose of schizokinen (circled) to HVO_A0467, as predicted by CB-Dock2. The position of the archaeal membrane (gray) was predicted using PPM3.0. The ligand lies at the base of the internal channel of HVO_A0467 formed by the 12 transmembrane domains of the protein. Surface coloring by electrostatic potential, with the hydrophobic membrane spanning trunk region largely in green. **(B)** Predicted interactions of schizokinen with the residues of HVO_A0467. Image obtained using PLIP (Adasme et al., 2021). Blue lines, hydrogen bonds; yellow balls, charge centers; gray dashed lines, hydrophobic interactions.

3.5 HVO_A0467 is not essential under standard conditions

Deletion of the HVO_A0467 gene produced a viable strain of *Hfx. volcanii* (ΔHVO_A0467), showing that this gene is not essential under standard growth conditions. No pronounced differences could be detected between cell shapes from the deletion strain and its parent under different growth conditions (data not shown). Interestingly, the deletion mutant grows slightly better in minimal medium and low iron (Figure 7).

3.6 HVO_A0467 forms an operon with HVO_0466

The genomic context of HVO_A0467 was found to be intriguing. HVO_A0467 forms an operon with HVO_A0466 (Figure 8), which

encodes an uncharacterized protein that was shown to be glycosylated (Schulze et al., 2021) and is probably secreted. The genes have a one-base overlap, with the last base of the HVO_A0466 stop codon being the first base of the HVO_A0467 start codon.

An initial analysis detected two paralogs of HVO_A0466 in *Hfx. volcanii*, each of them also encoded adjacent to an RND permease. Therefore, we systematically analyzed a selection of 21 genomes from 18 genera (Figure 9; Supplementary Table S2 "Species Homologs"). With very few exceptions, HVO_A0466 homologs are encoded adjacent to an RND permease, and many RND permeases are encoded adjacent to a HVO_A0466 homolog, suggesting a shared molecular function of HVO_A0466 with HVO_A0467.

4 Discussion

Hfx. volcanii is a well-studied archaeal microorganism which has shown remarkable adaptability to changes in environmental conditions such as hypoxia, salinity or temperature stress. No information has been available, however, about its response to iron scarcity, a key environmental factor limiting microbial growth with a broad range of response mechanisms reported in bacteria. Bioinformatic analysis indicated the presence of a gene cluster for biosynthesis of the siderophore schizokinen in *Hfx. volcanii*, but solid evidence and information on the corresponding siderophore export and ferri-siderophore reimport systems has been critically lacking.

In order to deepen our understanding of the response of *Hfx. volcanii* to iron scarcity, we present the results of a differential proteome analysis conducted under conditions of iron scarcity compared to normal iron concentration. The chosen conditions lead to severe growth attenuation and thus are suitable addressing this biological phenomenon. We conducted fractionation of the samples to put increased focus on the detection of membrane proteins. Surprisingly, iron scarcity did not produce a stringent proteomics response similar to the one observed under salinity and temperature stress (Jevtic et al., 2019). More specifically, even the *iucABCD* gene cluster predicted to be responsible for schizokinen production did not show a uniform pattern of regulation, with *iucC* being strongly down-regulated.

The protein showing the highest factor of up-regulation under iron scarcity was HVO_A0467, an RND superfamily integral membrane protein with 12 predicted transmembrane proteins. Intriguingly, computational docking of schizokinen to a predicted structure of HVO_A0467 identified a potential binding pocket close to the base of the internal channel of HVO_A0467 formed by the 12 transmembrane domains of the protein, making the protein a strong candidate for export of schizokinen. Deleting HVO_A0467, however, did confer neither a growth advantage nor a disadvantage under iron scarcity, indicating that there may be alternative mechanisms of sequestering iron in *Hfx. volcanii*. Additional experimentation will be necessary to finally confirm both the production of schizokinen and its export by HVO_A0467.

The RND permease HVO_A0467 shows a strong genetic association with the upstream gene HVO_A0466 that encodes a secreted glycoprotein (Supplementary Table S6; Figure 9). This is also worth investigating, as they may interact functionally within the pseudoperiplasmic space. AlphaFold predicts HVO_A0466 to be a long, fiber-like protein consisting of four regularly

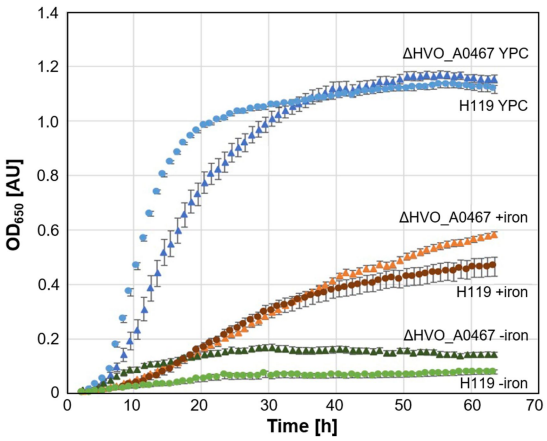


FIGURE 7
Growth of parent strain H119 and deletion strain ΔHVO_A0467 in different media. Growth of *Hfx. volcanii* H119 (circles) and ΔHVO_A0467 (triangles) in complex medium (blue, YPC), minimal medium (Hv-min) with (brown, +iron) or without added $FeSO_4$ (green, –iron). $n = 3$ biological replicates, with 5 technical replicates each. y-axis: optical density at 650 nm, x-axis: time of growth in hours.

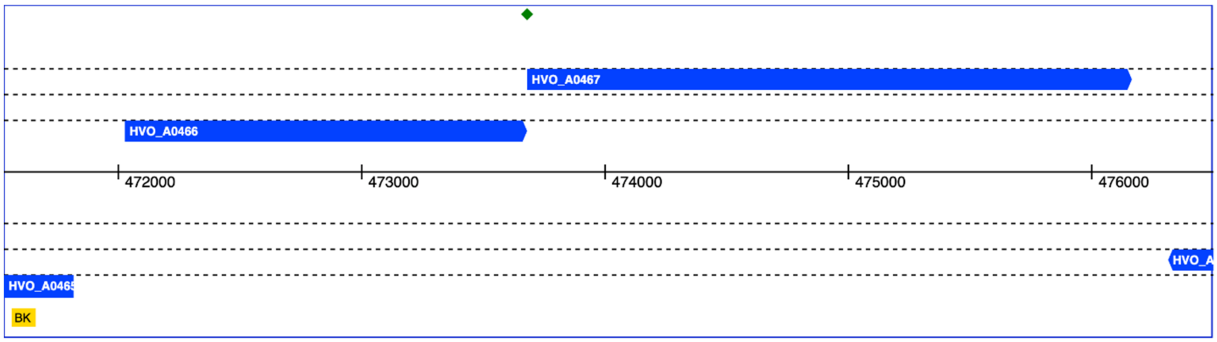


FIGURE 8
Genomic location of HVO_A0466 and HVO_A0467. Both genes overlap by one nucleotide (green diamond). The genomic coordinates of pHV4 are shown in the middle. Data obtained from HaloLex (Pfeiffer et al., 2008).

10.3389/fmicb.2024.1422844

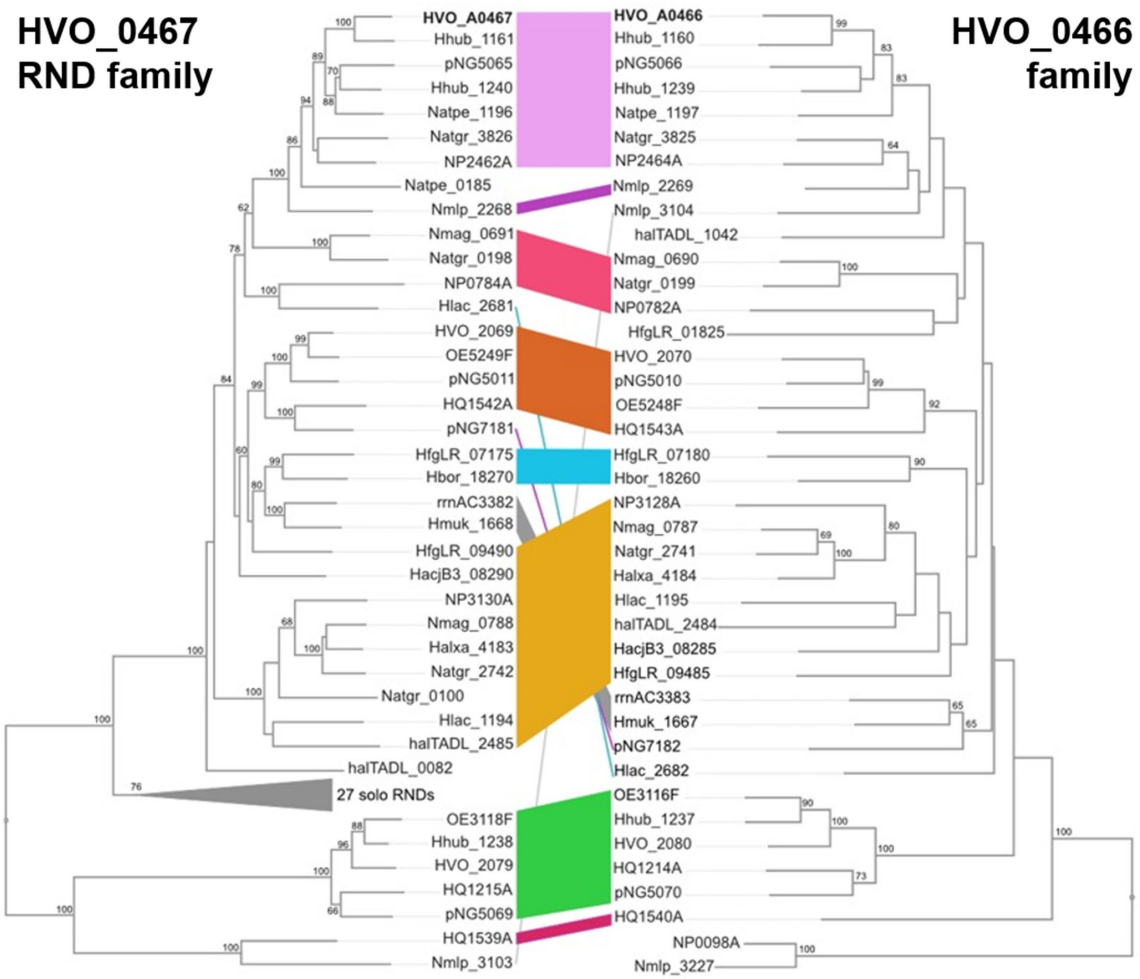


FIGURE 9
Comparison of inferred phylogenies of HVO_A0467 (RND) family proteins (left) and HVO_A0466 homologs (right). Colored stripes indicate corresponding proteins from the same genome, with many such pairs forming related groupings. For example, HVO_A0467 and HVO_A0466 (top) are bolded, and belong to a cluster of related, corresponding protein pairs (pink stripe). NJ algorithm with bootstrap values (100 replicates) of >60 shown at branch points.

spaced, closely folded domains (Ig-folds) and attached to the membrane via a C-terminal transmembrane domain (Supplementary Figure S4). The finding that such an association between RND proteins and HVO_A0466-like proteins is widespread makes it likely there is a functional interaction between RND and its cognate HVO_A0466-like protein.

A recent study of the *Hfx. volcanii* transcriptional regulatory network (TRN) identified three diphtheria toxin repressor (DtxR) proteins as potential transcriptional regulators of iron homeostasis, Idr (HVO_0538), SirR (HVO_0819) and TroR (HVO_0863) (Martinez Pastor et al., 2024). We were able to detect SirR and TroR in our proteomic data set, however neither of the two proteins showed significantly differential abundance upon iron starvation in either the soluble or the membrane fraction. This discrepancy between the transcriptome and proteome points to more complex mechanisms employed by *Hfx. volcanii* to maintain iron homeostasis under conditions of scarcity.

Data availability statement

The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium via the PRIDE (Perez-Riverol et al., 2019) partner repository with the dataset identifier PXD029036.

Author contributions

A-LS: Data curation, Formal analysis, Investigation, Methodology, Visualization, Writing – original draft, Writing – review & editing. ZJ: Data curation, Formal analysis, Investigation, Methodology, Writing – review & editing. BS: Formal analysis, Investigation, Methodology, Writing – review & editing. JW: Formal analysis, Investigation, Methodology, Writing – review & editing. KS: Formal analysis, Investigation, Writing – review & editing. HU: Conceptualization, Funding acquisition, Resources, Supervision, Validation, Writing

– review & editing. MD-S: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing – original draft, Writing – review & editing. FP: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing – original draft, Writing – review & editing. AM: Conceptualization, Data curation, Formal analysis, Funding acquisition, Methodology, Project administration, Resources, Supervision, Validation, Writing – original draft, Writing – review & editing. CL: Conceptualization, Data curation, Formal analysis, Methodology, Project administration, Resources, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This research was funded by the German Research Council (Deutsche Forschungsgemeinschaft, DFG) in the framework of the priority program SPP 2002, "Small Proteins in Prokaryotes, an Unexplored World" to AM (MA 1538/24-1).

Acknowledgments

ZJ and CL would like to thank Monika Raabe at the Max Planck Institute for Multidisciplinary Sciences, Göttingen and Lisa Neuenroth at the University Medical Center Göttingen for expert technical assistance. AM would like to thank Manuela Weishaupt for expert technical assistance.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2024.1422844/full#supplementary-material

References

Adasme, M. F., Linnemann, K. L., Bolz, S. N., Kaiser, F., Salentin, S., Haupt, V. J., et al. (2021). PLIP 2021: expanding the scope of the protein-ligand interaction profiler to DNA and RNA. *Nucleic Acids Res.* 49, W530–W534. doi: 10.1093/nar/gkab294

Allers, T., Ngo, H. P., Mevarech, M., and Lloyd, R. G. (2004). Development of additional selectable markers for the halophilic archaeon *Haloferax volcanii* based on the leuB and trpA genes. *Appl. Environ. Microbiol.* 70, 943–953. doi: 10.1128/AEM.70.2.943-953.2004

Blin, K., Shaw, S., Augustijn, H. E., Reitz, Z. L., Biermann, F., Alanjary, M., et al. (2023). antiSMASH 7.0: new and improved predictions for detection, regulation, chemical structures and visualisation. *Nucleic Acids Res.* 51, W46–W50. doi: 10.1093/nar/gkad344

Bruderer, R., Bernhardt, O. M., Gandhi, T., Miladinovic, S. M., Cheng, L. Y., Messner, S., et al. (2015). Extending the limits of quantitative proteome profiling with data-independent acquisition and application to acetaminophen-treated three-dimensional liver microtissues. *Mol. Cell. Proteomics* 14, 1400–1410. doi: 10.1074/mcp.M114.044305

Cerletti, M., Giménez, M. I., Tröetschel, C., D' Alessandro, C., Poetsch, A., de Castro, R. E., et al. (2018). Proteomic study of the exponential-stationary growth phase transition in the haloarchaea *Natrialba magadii* and *Haloferax volcanii*. *Proteomics* 18:e1800116. doi: 10.1002/pmic.201800116

Chuljerm, H., Chen, Y. L., Srichairatanakool, S., Hider, R. C., and Cilibrizzi, A. (2019). Synthesis and iron coordination properties of schizokinen and its imide derivative. *Dalton Trans.* 48, 17395–17401. doi: 10.1039/c9dt02731a

Chuljerm, H., Deeudom, M., Fucharoen, S., Mazzacuva, F., Hider, R. C., Srichairatanakool, S., et al. (2020). Characterization of two siderophores produced by *Bacillus megaterium*: a preliminary investigation into their potential as therapeutic agents. *Biochim. Biophys. Acta Gen. Subj.* 1864:129670. doi: 10.1016/j.bbagen.2020.129670

Coker, J. A., and DasSarma, S. (2007). Genetic and transcriptomic analysis of transcription factor genes in the model halophilic archaeon: coordinate action of TbpD and TfbA. *BMC Genet.* 8:61. doi: 10.1186/1471-2156-8-61

Domenech, P., Reed, M. B., and Barry, C. E. III. (2005). Contribution of the *Mycobacterium tuberculosis* MmpL protein family to virulence and drug resistance. *Infect. Immun.* 73, 3492–3501. doi: 10.1128/IAI.73.6.3492-3501.2005

Dyall-Smith, M. L. (2008). The Halohandbook [Online]. Available at: https://haloarchaea.com/wp-content/uploads/2018/10/Halohandbook_2009_v7.3mds.pdf (Accessed 2012).

Eberhardt, J., Santos-Martins, D., Tillack, A. F., and Forli, S. (2021). AutoDock Vina 1.2.0: new docking methods, expanded force field, and python bindings. *J. Chem. Inf. Model.* 61, 3891–3898. doi: 10.1021/acs.jcim.1c00203

Hallgren, J., Tsirigos, K. D., Pedersen, M. D., Almagro Armenteros, J. J., Marcatili, P., Nielsen, H., et al. (2022). DeepTMHMM predicts alpha and beta transmembrane proteins using deep neural networks. *bioRxiv*. doi: 10.1101/2022.04.08.487609

Haque, R. U., Paradisi, F., and Allers, T. (2020). *Haloferax volcanii* for biotechnology applications: challenges, current state and perspectives. *Appl. Microbiol. Biotechnol.* 104, 1371–1382. doi: 10.1007/s00253-019-10314-2

Hartman, A. L., Norais, C., Badger, J. H., Delmas, S., Haldenby, S., Madupu, R., et al. (2010). The complete genome sequence of *Haloferax volcanii* DS2, a model archaeon. *PLoS One* 5:e9605. doi: 10.1371/journal.pone.0009605

Huang da, W., Sherman, B. T., and Lempicki, R. A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* 4, 44–57. doi: 10.1038/nprot.2008.211

Hubmacher, D., Matzanke, B. F., and Anemuller, S. (2003). Effects of iron limitation on the respiratory chain and the membrane cytochrome pattern of the Euryarchaeon *Halobacterium salinarum*. *Biol. Chem.* 384, 1565–1573. doi: 10.1515/BC.2003.173

Hubmacher, D., Matzanke, B. F., and Anemuller, S. (2007). Iron-uptake in the Euryarchaeon *Halobacterium salinarum*. *Biometals* 20, 539–547. doi: 10.1007/s10534-006-9064-5

Jaschinski, K., Babski, J., Lehr, M., Burmester, A., Benz, J., Heyer, R., et al. (2014). Generation and phenotyping of a collection of sRNA gene deletion mutants of the Haloarchaeon *Haloferax volcanii*. *PLoS One* 9:e90763. doi: 10.1371/journal.pone.0090763

Jevtic, Z., Stoll, B., Pfeiffer, F., Sharma, K., Urlaub, H., Marchfelder, A., et al. (2019). The response of *Haloferax volcanii* to salt and temperature stress: a proteome study by label-free mass spectrometry. *Proteomics* 19:e1800491. doi: 10.1002/pmic.201800491

Kall, L., Krogh, A., and Sonnhammer, E. L. (2007). Advantages of combined transmembrane topology and signal peptide prediction--the Phobius web server. *Nucleic Acids Res.* 35, W429–W432. doi: 10.1093/nar/gkm256

Kanehisa, M., and Goto, S. (2000). KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 28, 27–30. doi: 10.1093/nar/28.1.27

Khan, A., Singh, P., and Srivastava, A. (2018). Synthesis, nature and utility of universal iron chelator - siderophore: a review. *Microbiol. Res.* 212-213, 103–111. doi: 10.1016/j.micres.2017.10.012

Klenotic, P. A., Moseng, M. A., Morgan, C. E., and Yu, E. W. (2021). Structural and functional diversity of resistance-nodulation-cell division transporters. *Chem. Rev.* 121, 5378–5416. doi: 10.1021/acs.chemrev.0c00621

Krogh, A., Larsson, B., von Heijne, G., and Sonnhammer, E. L. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* 305, 567–580. doi: 10.1006/jmbi.2000.4315

Kumar, N., Su, C. C., Chou, T. H., Radhakrishnan, A., Delmar, J. A., Rajashankar, K. R., et al. (2017). Crystal structures of the *Burkholderia multivorans* hopanoid transporter HpnN. *Proc. Natl. Acad. Sci. USA* 114, 6557–6562. doi: 10.1073/pnas.1619660114

Kuraku, S., Zmasek, C. M., Nishimura, O., and Katoh, K. (2013). aLeaves facilitates on-demand exploration of metazoan gene family trees on MAFFT sequence alignment server with enhanced interactivity. *Nucleic Acids Res.* 41, W22–W28. doi: 10.1093/nar/gkt389

Leyn, S. A., and Rodionov, D. A. (2015). Comparative genomics of DtxR family regulons for metal homeostasis in Archaea. *J. Bacteriol.* 197, 451–458. doi: 10.1128/JB.02386-14

Liu, Y., Yang, X., Gan, J., Chen, S., Xiao, Z. X., and Cao, Y. (2022). CB-Dock2: improved protein-ligand blind docking by integrating cavity detection, docking and homologous template fitting. *Nucleic Acids Res.* 50, W159–W164. doi: 10.1093/nar/gkac394

Lu, Q., Han, J., Zhou, L., Coker, J. A., DasSarma, P., DasSarma, S., et al. (2008). Dissection of the regulatory mechanism of a heat-shock-responsive promoter in Haloarchaea: a new paradigm for general trancription factor directed archaeal gene regulation. *Nucleic Acids Res.* 36, 3031–3042. doi: 10.1093/nar/gkn152

Makarova, K. S., Sorokin, A. V., Novichkov, P. S., Wolf, Y. I., and Koonin, E. V. (2007). Clusters of orthologous genes for 41 archaeal genomes and implications for evolutionary genomics of archaea. *Biol. Direct* 2:33. doi: 10.1186/1745-6150-2-33

Makarova, K. S., Wolf, Y. I., and Koonin, E. V. (2015). Archaeal clusters of orthologous genes (arCOGs): an update and application for analysis of shared features between *Thermococcales*, *Methanococcales*, and *Methanobacteriales*. *Life (Basel)* 5, 818–840. doi: 10.3390/life5010818

Martinez Pastor, M., Sakrikar, S., Hwang, S., Hackley, R. K., Soborowski, A. L., Maupin Furlow, J. A., et al. (2024). TroR ist he primary regulator of the iron homeostasis transcription network in the halophilic archaeon *Haloferax volcanii*. *Nucleic Acids Res.* 52, 125–140. doi: 10.1093/nar/gkad997

McMillan, L. J., Hwang, S., Farah, R. E., Koh, J., Chen, S., and Maupin-Furlow, J. A. (2018). Multiplex quantitative SILAC for analysis of archaeal proteomes: a case study of oxidative stress responses. *Environ. Microbiol.* 20, 385–401. doi: 10.1111/1462-2920.14014

Miethke, M., and Marahiel, M. A. (2007). Siderophore-based iron acquisition and pathogen control. *Microbiol. Mol. Biol. Rev.* 71, 413–451. doi: 10.1128/MMBR.00012-07

Mullis, K. B., Pollack, J. R., and Neilands, J. B. (1971). Structure of schizokinen, an iron-transport compound from *Bacillus megaterium*. *Biochemistry* 10, 4894–4898. doi: 10.1021/bi00802a010

Niessen, N., and Soppa, J. (2020). Regulated iron siderophore production of the halophilic archaeon *Haloferax volcanii*. *Biomol. Ther.* 10:1072. doi: 10.3390/biom10071072

Omasits, U., Ahrens, C. H., Muller, S., and Wollscheid, B. (2014). Protter: interactive protein feature visualization and integration with experimental proteomic data. *Bioinformatics* 30, 884–886. doi: 10.1093/bioinformatics/btt607

Pathan, M., Keerthikumar, S., Ang, C. S., Gangoda, L., Quek, C. Y., Williamson, N. A., et al. (2015). FunRich: an open access standalone functional enrichment and interaction network analysis tool. *Proteomics* 15, 2597–2601. doi: 10.1002/pmic.201400515

Perez-Riverol, Y., Csordas, A., Bai, J., Bernal-Llinares, M., Hewapathirana, S., Kundu, D. J., et al. (2019). The PRIDE database and related tools and resources in 2019: improving support for quantification data. *Nucleic Acids Res.* 47, D442–D450. doi: 10.1093/nar/gky1106

Pfeiffer, F., Broicher, A., Gillich, T., Klee, K., Mejia, J., Rampp, M., et al. (2008). Genome information management and integrated data analysis with HaloLex. *Arch. Microbiol.* 190, 281–299. doi: 10.1007/s00203-008-0389-z

Pfeiffer, F., and Oesterhelt, D. (2015). A manual curation strategy to improve genome annotation: application to a set of haloarchael genomes. *Life (Basel)* 5:1427–1444. doi: 10.3390/life5021427

Ple, C., Tam, H. K., Vieira Da Cruz, A., Compagne, N., Jimenez-Castellanos, J. C., Muller, R. T., et al. (2022). Pyridylpiperazine-based allosteric inhibitors of RND-type multidrug efflux pumps. *Nat. Commun.* 13:115. doi: 10.1038/s41467-021-27726-2

Saier, M. H., Reddy, V. S., Moreno-Hagelsieb, G., Hendargo, K. J., Zhang, Y., Iddamsetty, V., et al. (2021). The transporter classification database (TCDB): 2021 update. *Nucleic Acids Res.* 49, D461–D467. doi: 10.1093/nar/gkaa1004

Schiller, H., Hong, Y., Kouassi, J., Rados, T., Kwak, J., DiLucido, A., et al. (2024). Identification of structural and regulatory cell-shape determinants in Haloferax volcanii. *Nat Commun*. 15:1414. doi: 10.1038/s41467-024-45196-0

Schmid, A. K., Pan, M., Sharma, K., and Baliga, N. S. (2011). Two transcription factors are necessary for iron homeostasis in a salt-dwelling archaeon. *Nucleic Acids Res.* 39, 2519–2533. doi: 10.1093/nar/gkq1211

Schulze, S., Adams, Z., Cerletti, M., De Castro, R., Ferreira-Cerca, S., Fufezan, C., et al. (2020). The archaeal proteome project advances knowledge about archaeal cell biology through comprehensive proteomics. *Nat. Commun.* 11:3145. doi: 10.1038/s41467-020-16784-7

Schulze, S., Pfeiffer, F., Garcia, B. A., and Pohlschroder, M. (2021). Comprehensive glycoproteomics shines new light on the complexity and extent of glycosylation in archaea. *PLoS Biol.* 19:e3001277. doi: 10.1371/journal.pbio.3001277

Schwanhausser, B., Busse, D., Li, N., Dittmar, G., Schuchhardt, J., Wolf, J., et al. (2011). Global quantification of mammalian gene expression control. *Nature* 473, 337–342. doi: 10.1038/nature10098

Storey, E. P., Boghozian, R., Little, J. L., Lowman, D. W., and Chakraborty, R. (2006). Characterization of 'Schizokinen'; a dihydroxamate-type siderophore produced by *Rhizobium leguminosarum* IARI 917. *Biometals* 19, 637–649. doi: 10.1007/s10534-006-9001-7

Tyanova, S., and Cox, J. (2018). Perseus: a bioinformatics platform for integrative analysis of proteomics data in Cancer research. *Methods Mol. Biol.* 1711, 133–148. doi: 10.1007/978-1-4939-7493-1_7

Wells, R. M., Jones, C. M., Xi, Z., Speer, A., Danilchanka, O., Doornbos, K. S., et al. (2013). Discovery of a siderophore export system essential for virulence of *Mycobacterium tuberculosis*. *PLoS Pathog.* 9:e1003120. doi: 10.1371/journal.ppat.1003120

Zhang, B., Li, J., Yang, X., Wu, L., Zhang, J., Yang, Y., et al. (2019). Crystal structures of membrane transporter MmpL3, an anti-TB drug target. *Cell* 176:e613, 636–648.e13. doi: 10.1016/j.cell.2019.01.003

Zhang, Y., Bilbao, A., Bruderer, T., Luban, J., Strambio-De-Castillia, C., Lisacek, F., et al. (2015). The Use of Variable Q1 Isolation Windows Improves Selectivity in LC-SWATH-MS Acquisition. *J Proteome Res.* 14, 4359–4371. doi: 10.1021/acs.jproteome.5b00543

frontiers | Frontiers in Microbiology

TYPE Original Research
PUBLISHED 12 November 2024
DOI 10.3389/fmicb.2024.1474570

OPEN ACCESS

EDITED BY
Maria Ines Gimenez,
National University of Mar del Plata, Argentina

REVIEWED BY
Tushar Kant Beuria,
Institute of Life Sciences (ILS), India
Yaodong Chen,
Northwest University, China

*CORRESPONDENCE
Sonja-Verena Albers
✉ sonja.albers@biologie.uni-freiburg.de

†PRESENT ADDRESSES
Megha Patro,
Structural and Computational Biology Unit,
European Molecular Biology Laboratory,
Heidelberg, Germany
Solenne Ithurbide,
Departement de Microbiologie, Infectiologie
et Immunologie, Université de Montréal,
Montréal, QC, Canada

RECEIVED 01 August 2024
ACCEPTED 02 October 2024
PUBLISHED 12 November 2024

CITATION
Patro M, Grünberger F, Sivabalasarma S, Gfrerer S, Rodriguez-Franco M, Nußbaum P, Grohmann D, Ithurbide S and Albers S-V (2024) MinD2 modulates cell shape and motility in the archaeon *Haloferax volcanii*. *Front. Microbiol.* 15:1474570.
doi: 10.3389/fmicb.2024.1474570

MinD2 modulates cell shape and motility in the archaeon *Haloferax volcanii*

Megha Patro[1,2†], Felix Grünberger[3], Shamphavi Sivabalasarma[1,2], Sabrina Gfrerer[1], Marta Rodriguez-Franco[4], Phillip Nußbaum[1], Dina Grohmann[3], Solenne Ithurbide[1†] and Sonja-Verena Albers[1,5*]

[1]Molecular Biology of Archaea, Institute of Biology, Faculty of Biology, University of Freiburg, Freiburg, Germany, [2]Spemann Graduate School of Biology and Medicine, University of Freiburg, Freiburg, Germany, [3]Institute of Biochemistry Genetics and Microbiology, Institute of Microbiology and Archaea Centre, Single-Molecule Biochemistry Lab & Biochemistry Centre Regensburg, University of Regensburg, Regensburg, Germany, [4]Cell Biology, Institute of Biology, Faculty of Biology, University of Freiburg, Freiburg, Germany, [5]Signalling Research Centres BIOSS and CIBSS, University of Freiburg, Freiburg, Germany

In bacteria and archaea, proteins of the ParA/MinD family of ATPases regulate the spatiotemporal organization of various cellular cargoes, including cell division proteins, motility structures, chemotaxis systems, and chromosomes. In bacteria, such as *Escherichia coli*, MinD proteins are crucial for the correct placement of the Z-ring at mid-cell during cell division. However, previous studies have shown that none of the 4 MinD homologs present in the archaeon *Haloferax volcanii* have a role in cell division, suggesting that these proteins regulate different cellular processes in haloarchaea. Here, we show that while deletion of MinD2 in *H. volcanii* (Δ*minD2*) does not affect cell growth or division, it impacts cell shape and motility by mispositioning the chemotaxis arrays and archaellum motors. Finally, we explore the links between MinD2 and MinD4, which has been previously shown to modulate the localization of chemosensory arrays and archaella in *H. volcanii*, finding that the two MinD homologues have synergistic effects in regulating the positioning of the motility machinery. Collectively, our findings identify MinD2 as an important link between cell shape and motility in *H. volcanii* and further our understanding of the mechanisms by which multiple MinD proteins regulate cellular functions in haloarchaea.

KEYWORDS

Haloferax volcanii, archaea, cell shape, shape transition, light and fluorescence microscopy, protein localisation, motility

Introduction

A proper spatial distribution of cellular components is essential for the optimal functioning of cells. In bacteria and archaea, the ParA/MinD family of ATPases is crucial for the spatiotemporal organization of various cellular cargoes. For example, ParA proteins are involved in plasmid partitioning and chromosome segregation (Baxter and Funnell, 2014; Jalal and Le, 2020), while MinD is known for its role in regulating the placement of the bacterial divisome (Lutkenhaus, 2007). However, ParA/MinD proteins are not restricted to these functions and have been shown to modulate the positioning of several other cellular components, including flagella (Pulianmackal et al., 2023; Pradhan et al., 2024), chemotaxis systems (Ringgaard et al., 2011), and the conjugation machinery (Atmakuri et al., 2007).

Although several ParA/MinD homologs are encoded in archaeal genomes, their distribution and functions are still being elucidated. Notably, while ParA can be found in almost all archaea, only Euryarchaeota encode for MinD homologs (Nußbaum et al., 2020). Furthermore, although a few structural analyses have been carried on archaeal MinD homologs (*Pyrococcus horikoshii*, *Pyrococcs furiosus* and *Archaeoglobus fulgidus*) (Jeoung et al., 2009; Szklarczyk et al., 2017), the functional roles of MinD have so far only been studied in *H. volcanii* (Nußbaum et al., 2020). *H. volcanii* encodes for 4 MinD homologs: *minD1* (HVO_0225), *minD2* (HVO_0595), *minD3* (HVO_1634) and *minD4* (HVO_0322). Notably, in contrast to the critical role of MinDs in regulating cell division in bacteria, deletion of all 4 MinD homologs had no role in cell division or growth in *H. volcanii* (Nußbaum et al., 2020). However, deletion of MinD4 reduced archaeal swimming motility due to the mispositioning of chemotaxis arrays and archaellum motors, suggesting a significant role of MinD4 in governing these processes (Nußbaum et al., 2020). However, the role of other MinD homologues remains unclear.

In this study, we characterized the functions of the MinD2 protein of *H. volcanii*. Since MinD2 was observed to not directly affect FtsZ localization (and thus cell division) or cell growth, it suggests a role in other cellular pathways (Nußbaum et al., 2020). Using genetic mutants, we showed it has a crucial role in determining cell shape, influencing the transition from rod-shaped to plate-shaped cells. Furthermore, using fluorescently tagged MinD2 variants, we demonstrated that the protein has a diffused localization pattern, which suggests a potential regulatory mechanism for its cellular functions, possibly involving interactions with partner proteins. Additionally, we demonstrated that MinD2 synergizes with MinD4 to modulate chemosensory array localization, archaella assembly and motility, further illustrating the role of MinD2 homologues in spatial organization. Overall, our findings contribute to a deeper understanding of the function of MinD2 in *H. volcanii*, highlighting its multifaceted role in coordinating cellular morphology and motility.

Results

MinD2 impacts cell shape

To elucidate the function of MinD2 in *H. volcanii*, we started by characterizing the growth of a mutant strain lacking MinD2 (Δ*minD2*). Growth rate measurements showed that the Δ*minD2* strain exhibited a growth pattern similar to that of the wild-type (WT) strain H26 (Supplementary Figure 1), suggesting that MinD2 does not directly influence growth in *H. volcanii* as previously shown by Nußbaum et al. (2020).

Then, we assessed the impact of MinD2 on morphology, using phase contrast microscopy to compare the cell shape and size of WT vs. Δ*minD2* cells during different growth stages. As previously described, we found that WT *H. volcanii* (H26) undergoes growth-dependent alterations in its cell morphology. The time point of this shape transition has been observed to differ greatly based on medium components and is also affected by the presence of plasmids (Li et al., 2019; de Silva et al., 2021; Patro et al., 2023). During the early log phase, the liquid culture is predominantly composed of rod-shaped cells (R), which gradually transition into an intermediate state (I) and then into plate-shaped cells (P) as the optical density (OD_{600}) of the culture increases. The categorization of cell shape was conducted as previously described in Patro et al. (2023). Indeed, H26 (WT) cells predominantly exhibited rod-shaped morphology until an OD_{600} of 0.03, with a minority showing I or P shapes ($R = 41\%$; $I = 26\%$; $P = 33\%$) (Figure 1 and Supplementary Figure 2a). Starting from an OD_{600} of 0.06, a noticeable transition towards P cells became evident, with plates representing the majority ($P = 55\%$). This trend persisted as the culture progressed, with the frequency of P cells reaching 81% at an OD_{600} of 0.2. By contrast, the majority of Δ*minD2* cells exhibited a plate-shaped morphology even in the early log phase (OD_{600} 0.01), with few I and R cells ($R = 17\%$; $I = 33\%$; $P = 50\%$) (Figure 1 and Supplementary Figure 2a). The frequency of P cells continued to increase at higher culture densities, with more than 80% of the cell population adopting a plate-shaped morphology at OD_{600} above 0.06 (Figure 1 and Supplementary Figure 1).

Additionally, a comparison of the cell area in both strains displayed a similar trend, with Δ*minD2* cells having significantly smaller areas than WT cells at an OD_{600} of 0.01 (Supplementary Figure 2b). Nevertheless, the areas of both WT and Δ*minD2* cells were comparable at higher densities ($OD_{600 \geq}$ 0.03).

Collectively, these analyses demonstrate that while MinD2 does not affect cell growth, it significantly influences cell shape, with the deletion of *minD2* resulting in the loss of the ability to maintain a rod-shaped morphology, particularly in the early growth phase.

Previously, we found that the presence of a plasmid played a transient role in maintaining rod-shaped morphology in *H. volcanii* (Patro et al., 2023). To test this, we looked at the effects of deleting the MinD2 gene on cell shape and size, with or without the empty plasmid pTA1392. In the WT (H26) cells, when the plasmid was present, the cells changed from rods to plates between OD_{600} 0.06 and 0.1. Without the plasmid, this switch happened earlier, between OD_{600} 0.03 and 0.06 (Supplementary Figure 1b; Patro et al., 2023). Similarly, in MinD2 mutants, more cells were rod-shaped at OD_{600} 0.01 when pTA1392 was present. Without the plasmid, these mutant cells were mostly plate-shaped at OD_{600} 0.01. In MinD2 mutants with pTA1392, cells shifted to an intermediate shape (39%) at OD_{600} 0.03 and mostly plate-shaped (40%) at OD_{600} 0.06. This shape change happened one generation earlier in the MinD2 mutants with pTA1392 compared to WT cells with the plasmid. These results highlight MinD2's role in modulating cell shape, especially in becoming plate-shaped. The plasmid increases the number of rod-shaped cells in the MinD2 mutant, promoting rod development. Furthermore, while cells that contained plasmids maintained rod-shaped morphology longer throughout the different growth phases, the absence of *minD2* still reduced the ability to maintain a rod-shaped morphology throughout all phases (Supplementary Figures 1b, 2c,d).

MinD2 deletion impacts the assembly of archaella

In *H. volcanii,* cell shape is linked to swimming motility. During the early logarithmic phase, rod-shaped *H. volcanii* cells exhibit motility, powered by the assembly of polar bundles of archaella (Li et al., 2019). However, as cells transition into plate shape during stationary phase, the archaellum filaments are lost, resulting in motility cessation. Given the observed impact of MinD2 deletion on cell shape, we investigated whether the Δ*minD2* also showed changes

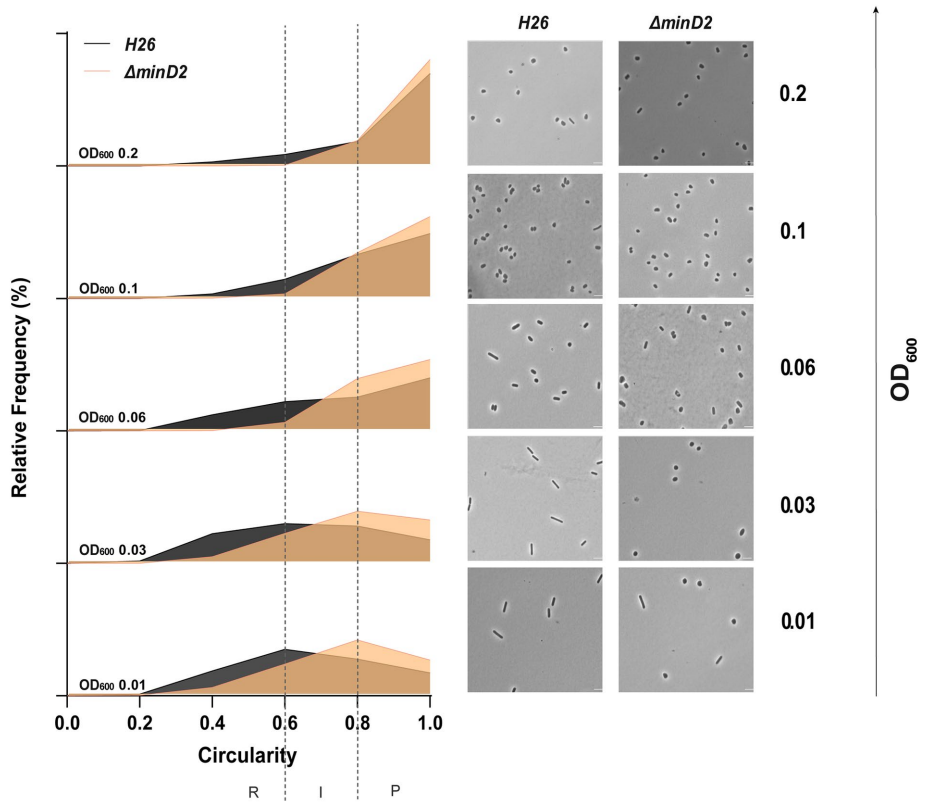


FIGURE 1
Cell shape analysis in Δ*minD2* strain and wild-type H26 cells throughout the growth curve. Left: Relative frequency distribution of cell circularity comparing H26 (grey) and Δ*minD2* (orange) analyzed from micrographs. Vertical dashed line represents the different cell type R, rods; I, intermediates and P, Plates. Sum of the graph height per OD_{600} equals 100% and *Y*-axis indicates the percentage of cell population per cell type. Right: Phase contrast micrographs showing H26 and Δ*minD2* at different growth stages from OD_{600} 0.01 to 0.2 (bottom to top). Scale bar: 4 μm. $n_{H26} > 1745$, $n_{\Delta minD2} > 1{,}535$. Three independent experiments with biological triplicates were carried out for both the strains.

in archaella assembly and motility. For this, we used transmission electron microscopy (TEM) to characterize WT and Δ*minD2* cells collected at an early log phase (OD_{600} 0.03). However, *H. volcanii* cells express several pili throughout different growth phases (Esquivel et al., 2016), which are difficult to differentiate in diameter and size from archaella. Therefore, we further deleted *pilB3*, the pilus assembly ATPase, in both the H26 and Δ*minD2* strain. These Δ*pilB3* mutants lack pili, which facilitates visualization of archaella in these strains.

The results revealed that the rod-shaped H26Δ*pilB3* cells, both with and without plasmid pTA1392, displayed archaellation in 80 and 83.7% of the cells, respectively (Figures 2a,c). By contrast, the Δ*minD2*Δ*pilB3* strain, characterized by a discoid-shaped phenotype, only the few rod-shaped cells observed showed archaellation (~20% of the cells). In the presence of pTA1392, where the percentage of rod-shaped cells is slightly increased (24%; Supplementary Figures 2a,b), the percentage of archaellated cells is also observed to increase to ~32% (Figures 2b,c).

Given the observed effects of MinD2 deletion on archaellation levels, we investigated whether this impacted the motility of the Δ*minD2* strain, by growing the cells on soft agar plates. First, we found that Δ*minD2* strain grown in nutrient-rich YPC media showed reduced motility compared to H26 (data not shown). However, due to the composition of the medium, cell growth was rapid, making it difficult to differentiate growth from swimming halos in soft agar plates. Therefore, we switched to a different medium (CA), where strains grow slower. In order to support growth on CA medium, which lacks uracyl, strains were transformed with a plasmid (pTA1392) containing the *pyrE2* locus (which enables uracyl biosynthesis). Using this strategy, we observed that Δ*minD2* + pTA1392 cells exhibited a swimming defect, with their motility being only ~25% of what was observed for H26 + pTA1392 cells (Figure 2d). We were able to complement the Δ*minD2* motility defect by expression of MinD2 on plasmid pTA1392 (pSVA6011, Figure 3). Additionally, the observed reduction in motility in Δ*minD2* was consistent with the lower number of mutant cells assembling archaella.

The hallmark of ParA/MinD superfamily proteins is the Walker A (WA) and Walker B (WB) motif (Leipe et al., 2002; Hu and Lutkenhaus, 2003). Mutation of both the motif in MinD4 had previously displayed motility phenotype and disrupted MinD4's localization at poles (Nußbaum et al., 2020). We wanted to investigate the role of the WA and WB motifs of MinD2 by generating plasmids with WA (K16A) mutant and WB (D117A). Expressing these in Δthe minD2 strain, we observed that mutation of WA did not have an effect on the swimming motility, and the strain had 100% activity as visualized in the wild type (H26). Whereas the WB mutant showed a defect in motility (~30%) similar to Δthe *minD2* deletion mutant strain (Supplementary Figure 4).

Thus, these results confirm that MinD2 not only modulates cells shape but also impacts cell motility, and that Walker B motif of MinD2 plays a role in the motility phenotype.

10.3389/fmicb.2024.1474570

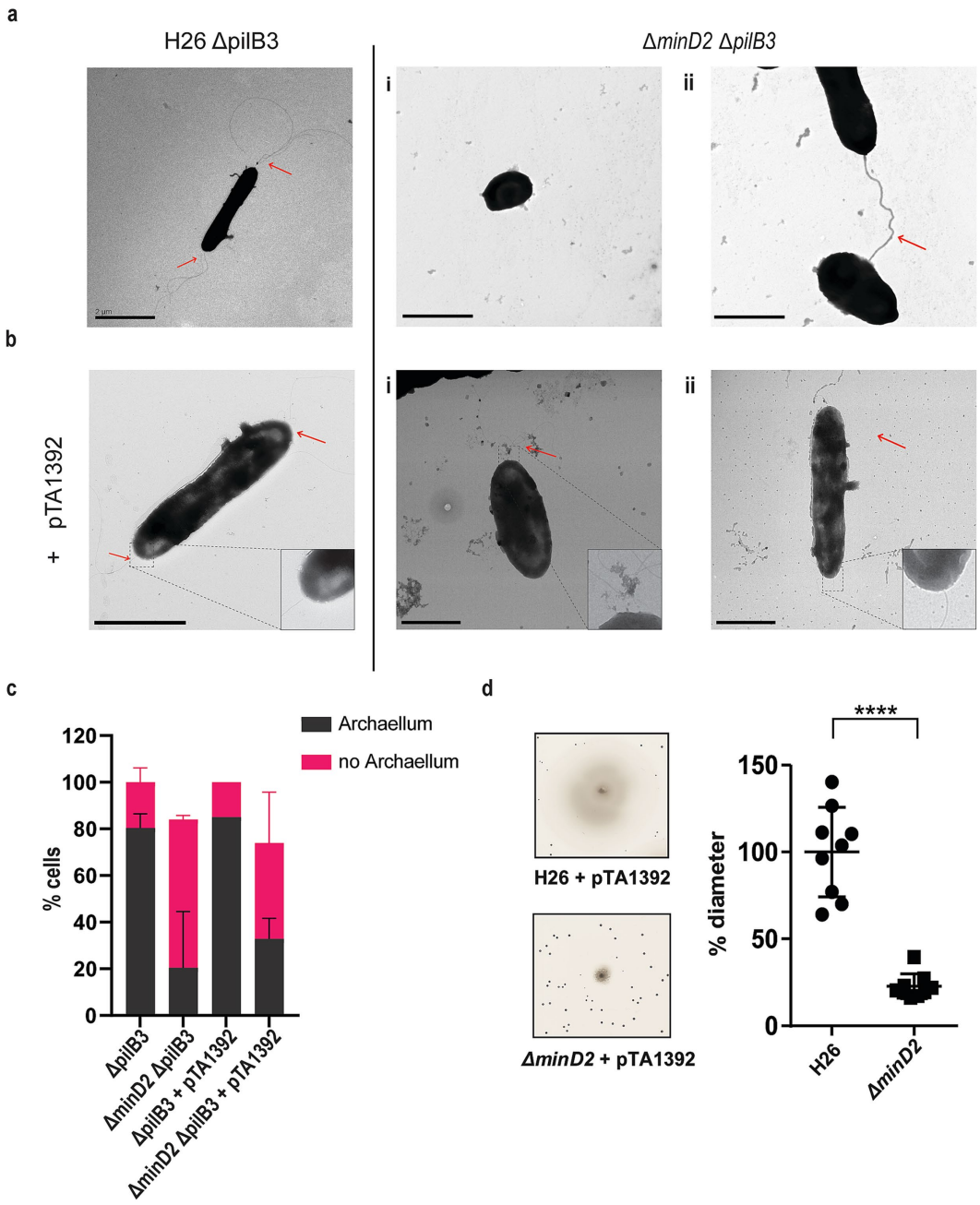


FIGURE 2
The deletion of *minD2* affects cell motility and archaellation. (a) Transmission electron microscopy of H26 *ΔpilB3* cells showing archaella and *ΔpilB3ΔminD2* (i) plate-shaped cell without archaella and (ii) rod shaped cell with archaella. (b) H26 *ΔpilB3* + pTA1392 cells showing archaella and *ΔpilB3ΔminD2* + pTA1392 (i) cell in transition rod to round (ii) rod shaped cell showing archaella. All cells were visualized at an early exponential phase (OD_{600}: 0.03). Scale bar: 2 µm. (c) Distribution of cells with or without archaella for H26 ΔpilB3 and ΔminD2 in the presence and absence of plasmid pTA1392 from TEM analyzed cells. $n_{\Delta pilB3} = 51$, $n_{\Delta minD2} = 54$, $n_{\Delta pilB3+pTA1392} = 51$ $n_{\Delta pilB3+pTA1392} = 40$ and $n_{\Delta minD2+pTA1392} = 59$. (d) Semi-solid agar-based motility assay to visualize the swimming ability of H26 and Δ*minD2*. Left panel: representative inserts; right panel: average diameter of the motility rings. Graph represents values from 3 technical replicates from 3 different biological replicates. **** $p < 0.0001$. Red arrow indicates the archaellum.

Fluorescently tagged MinD2 shows diffused localization

The observed impact of MinD2 on motility is reminiscent of the role of MinD4, which we previously showed to be a MinD homologue that oscillates along the cell axis in *H. volcanii* and stimulates the formation of chemosensory arrays and archaella at the cell poles, thereby regulating motility (Nußbaum et al., 2020). Therefore, we proceeded to examine the localization pattern of MinD2, by generating fluorescent fusion proteins. For this, the MinD2 protein was tagged at the N-or C-terminus with various tags and linkers, which were then expressed in the Δ*minD2* deletion mutant. To test whether the fluorescent tag had an impact on the function of MinD2, we performed complementation experiments

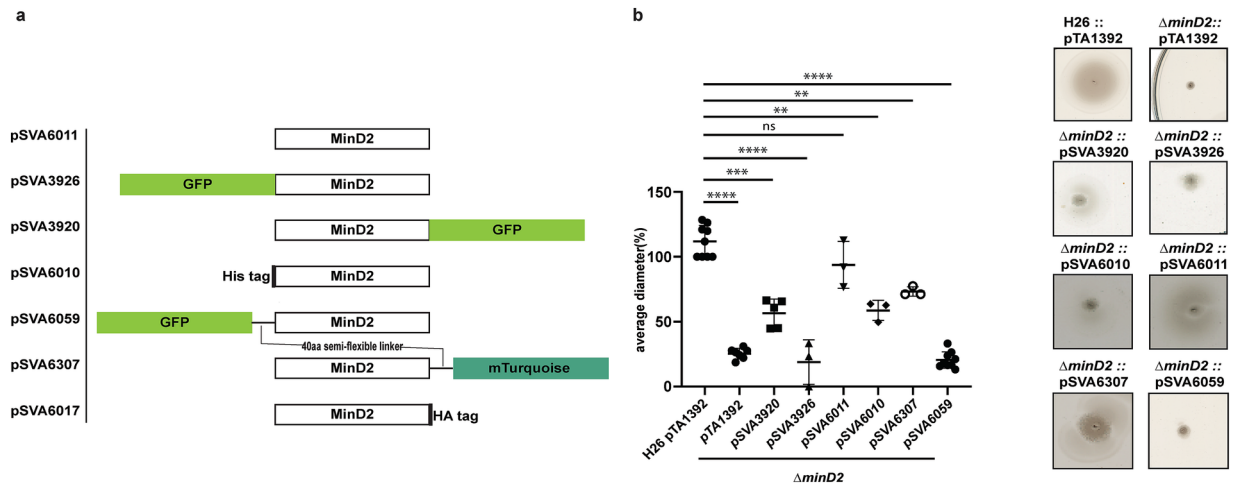


FIGURE 3
MinD2 plasmids and complementation assay. (a) Schematic overview of MinD2 plasmids with different tags. (b) Average diameter of motility rings measured relative to the wild type, from different *minD2* strains harboring different tagged MinD2 plasmids. Values are from 3 independent experiments including 3 biological replicates each. (*p*-values: **** <0.0001, *** 0.0002, ** 0.002, * 0.0332).

and assessed their motility using motility assays (Figure 3b). Notably, when a tag-less MinD2 variant was used (pSVA6011), the swimming phenotype could be restored to ~91% of the activity observed in WT cells. Expression of an N-terminal His-tag MinD2 variant (pSVA6010) restored ~62% of the motility activity, whereas the strain expressing an N-terminally tagged GFP variant (pSVA3926) had a motility defect similar to the Δ*minD2* strain (Figure 3b) and showed diffuse localization in the cells (Supplementary Figure 3). Expression of a variant in which MinD2 was fused to an N-terminal GFP tag with a semi-flexible linker (Ithurbide et al., 2024) (pSVA6059) could also not restore motility. Tagging MinD2 on the C-terminus was, in general, more effective at restoring motility. Using only an HA tag at the C-terminus complemented the swimming back to 100% (Supplementary Figure 4). And, the strain expressing a C-terminally GFP tagged MinD2 variant (pSVA3920) retained ~60% of the motility displayed by WT cells, and the introduction of a semi-flexible linker between MinD2 and a C-terminal mTurquoise tag (pSVA6307) led to a restored swimming efficiency up to ~80% of that observed in H26 cells (Figure 3 and Supplementary Table 4).

We then proceeded to characterize the localization of the most functional C-term fluorescently tagged MinD2 variants. For this, samples were collected at similar OD_{600}s as that of the cell shape experiments. Visualization of both the C-terminally GFP tagged and C-terminally mTurquoise tagged MinD2 variants showed a diffused localization throughout the cells (Supplementary Figure 3b) and from early log phase through mid-log phase. Based on our observations, MinD2-linker-mturquoise did not show a change in localization during the transition from rod to plate phases and does not show a different localization in rods versus in plates. Collectively, these data show that MinD2 does not localize to the cell poles, suggesting that its impact on the polar organization of the motility machinery likely requires the interaction of MinD2 with other proteins. Furthermore, the inability of N-terminally tagged MinD2 to complement the swimming phenotype suggests that the N-terminus of MinD2 is important for its activity and interaction with protein partners *in vivo*.

Possible interaction partners of MinD2

HVO_0596

HVO_0596 is a protein of unknown function encoded downstream of the *minD2* gene, likely within the same operon (Babski et al., 2016; Figure 4a). Synteny report on MinD2 and HVO_0596 shows the two gene to be conserved in Haloarchaea (Supplementary Table 5). Furthermore, an Alphafold 3 (Abramson et al., 2024) prediction indicated a possible interaction between MinD2 and HVO_0596, based on 6 hydrogen bonds between 5 residues of HVO_0596 at the C-terminus and the C-terminus of MinD2 with an iPTM score = 0.83 and pTM = 0.63 (Figure 4b). To explore a potential role of HVO_0596 in MinD2 functionality, we generated a deletion mutant for HVO_0596 and a double deletion mutant *Δhvo_0596ΔminD2*. The *Δhvo_0596* deletion mutant showed no discernible impairment of cell growth (Figure 4c) and to the contrary of *minD2* deletion, *the deletion of HVO_0596 has no effect on* cell shape development and transition during growth (Figure 4d).

Additionally, no soft-agar motility defect could be observed for the single mutant *Δhvo_0596 and* the double deletion mutant *Δhvo_0596ΔminD2* exhibited the same motility defect as the *ΔminD2* deletion mutant (Figure 4e). These results suggest that HVO_0596 does not directly affect motility.

To elucidate the localization pattern of HVO_0596, we generated a fluorescently tagged version of HVO_0596 at its N-terminus with mNeongreen (pSVA6051). Localization experiments showed diffuse fluorescence in both deletion mutants *Δhvo_0596, ΔminD2* and *Δhvo_0596ΔminD2* (Figure 4f).

CetZ5 and CetZ6

Haloferax volcanii has 6 paralogues of CetZs of which CetZ1 and CetZ2 have been studied for their role as cytoskeletal proteins and role in motility (Brown et al., 2024). An accompanying study by Brown et al. (2024, unpublished) suggests a possible interaction between MinD2 and CetZ1. Specifically, MinD2 was found to influence the cellular positioning of CetZ1, impacting its polar localization. While

10.3389/fmicb.2024.1474570

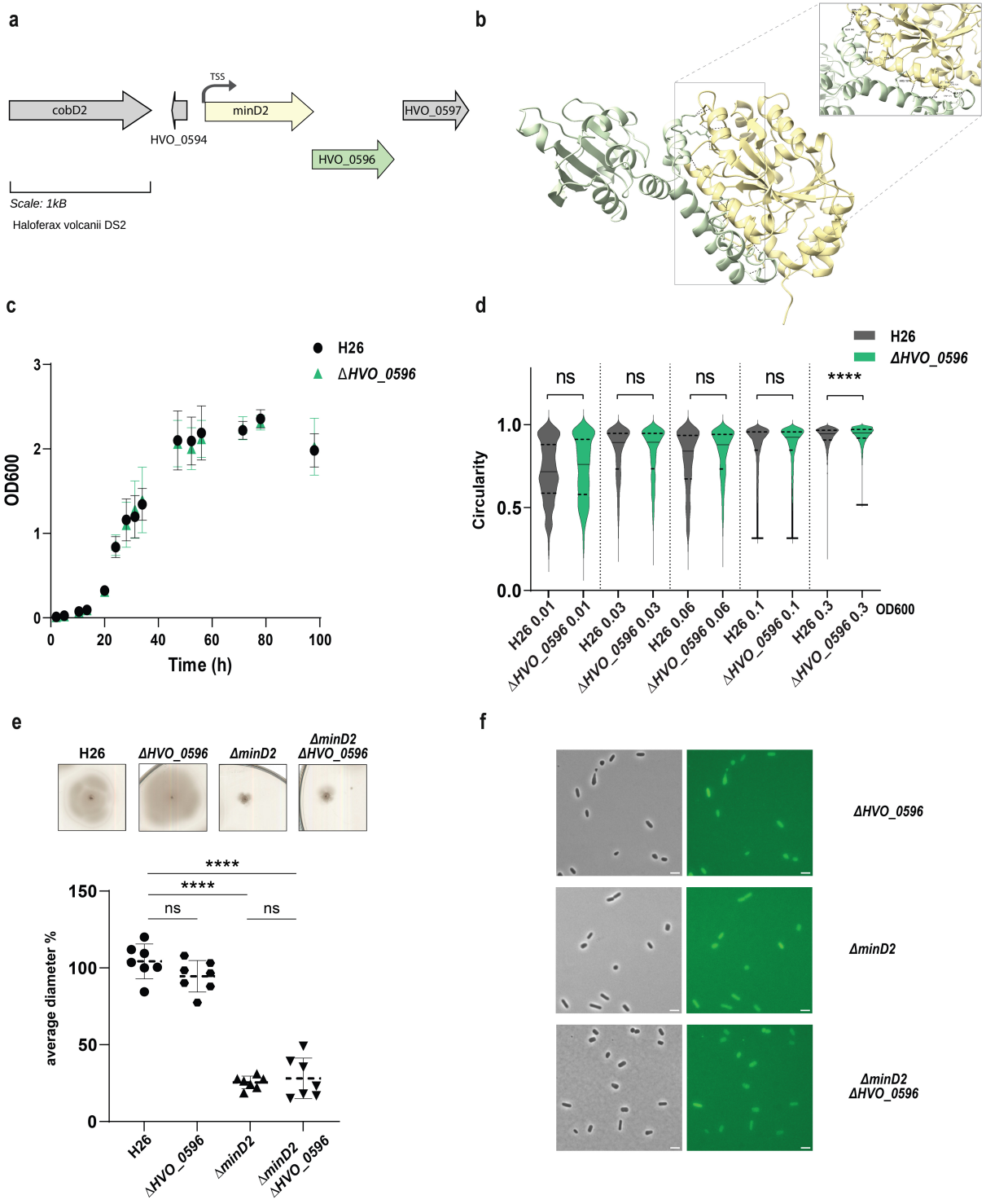


FIGURE 4
HVO_0596 might interacts with MinD2 but its deletion mutant has no phenotypes. (a) Schematic representation of the gene neighborhood showing MinD2 (HVO_0595) (yellow) and HVO_0596 (green) showing the TSS (black arrow) present 28 bp upstream of start. (b) Alphafold 3 prediction revealing interaction between MinD2 and HVO_0596. Black dashed lines indicates the intermolecular H-bond at the C-terminus of both proteins and red dashed line indicates the interaction found with relaxed angle criteria in ChimeraX. (c) Growth curve of H26 (black) and Δ*HVO_0596* (green). (d) Distribution of cell circularity (%) at different ODs; $n > 1100$. (e) Quantification (bottom panel) of motility diameter for by the different mutants and WT (inserts: Top Panel). Calculations were made using 3 independent experiments including >2 biological replicates each. Black line indicates mean, lower, and upper lines the standard deviation. (f) Representative inserts of mNeonGreen-HVO_0596 localization in H26(WT) and deletion mutants Δ*hvo_0596,* Δ*minD2* and Δ*HVO_0596minD2*. Scale bar: 4 μm.

the function of the other four CetZs is uncharacterized, CetZ5 was hypothesized to be a cytoskeletal protein by Schiller et al. (2024). It is therefore possible that MinD2 interacts with these CetZ proteins in the cell. Therefore, we wanted to address possible interactions with CetZ5 and CetZ6.

CetZ5 and CetZ6, initially characterized as FtsZ7 and FtsZ8, belong to the tubulin/FtsZ family, and previous studies have shown that deleting these proteins has no effect on cell division (Duggin et al., 2015). To characterize the role of these proteins with respect to MinD2, we generated single deletion mutants (*ΔcetZ5* and *ΔcetZ6*) and double deletion mutants with *minD2* (*ΔcetZ5ΔminD2* and *ΔcetZ6ΔminD2*). We found no motility defects for either *ΔcetZ5* or *ΔcetZ6* (Figures 5a,c) as observed previously (Duggin et al., 2015). The double deletion mutants (*ΔcetZ5ΔminD2* and *ΔcetZ6ΔminD2*) showed reduced motility similar to that of the single *ΔminD2* mutant, indicating that the impact in swimming ability is due to deletion of MinD2 rather than these CetZ proteins.

To gain insight into the cellular positioning of the CetZ proteins, we created N-terminally GFP tagged versions of CetZ5 (pSVA6040) and CetZ6 (pSVA6042). In both cases, the localization of the proteins was diffused across the cells (Figures 5c,d). To check if cell shape has an effect on CetZ localization, we further visualized the distribution of the tagged proteins at different ODs. However, both GFP-CetZ5 and GFP-CetZ6 displayed diffused fluorescence throughout all the growth phases analyzed (Figures 5c,d).

Collectively, our experiments indicate no direct interaction of HVO_0596, CetZ5, and CetZ6 with MinD2.

MinD2 regulates the positioning of the motility and chemotactic machineries

Given that the ΔminD2 mutant exhibited a motility defect and displayed reduced assembly of archaella, we hypothesized that MinD2 may have a function similar to that of MinD4 with respect to the cellular positioning of the motility and chemotactic machineries. Therefore, we investigated the localization of the motility machinery in ΔminD2. For this, we used strains expressing ArlD-GFP, a fluorescently tagged version of a well-established marker protein, which is part of the cytoplasmic archaellum motor complex, used as an indicator to identify cells with archaella (Li et al., 2019). Upon expression of ArlD-GFP in the Δ*arlD* strain at OD_{600} 0.01, we observed that most cells (77%) had fluorescent foci at the cell poles, with only a few cells displaying diffused fluorescence (23%) (Figures 6a,b), similar to results previously observed in *H. volcanii* (Li et al., 2019). However, when ArlD-GFP was expressed in *ΔminD2ΔarlD* cells, fluorescent foci at the poles were only detected in 25% of the cells (Figures 6a,b). This result agrees with the low abundance of archaella observed in TEM, where these structures were present in only ~32% of *ΔminD2ΔpilB3* (+ pTA1392) cells (Figure 2).

As described above, in previous studies we established a correlation between MinD4 and the positioning of the archaellum machinery. Therefore, we decided to explore the links between MinD2 and MinD4 in the regulation of this process, by comparing the localization of the motility machinery in cells lacking either one of the MinD homologues (Δ*minD2*Δ*arlD* or ΔminD4Δ*arlD*) or both homologues (Δ*minD2*Δ*minD4*Δ*arlD*). In the Δ*minD4*Δ*arlD* mutant, we observed that the number of cells with ArlD polar foci was ~20% at an OD_{600} of 0.01 (Nußbaum et al., 2020). In the *ΔminD2ΔminD4* mutant, the formation of polar foci decreased, being present in only ~7% of cells at an OD_{600} of 0.01, and continuing to decrease with increasing OD_{600}. Additionally, it was previously observed that all cells in the stationary phase form foci corresponding to the archaellum motor complex. We observed comparable results in the Δ*arlD* and *ΔminD2ΔarlD* mutants, with all cells having foci upon reaching the stationary phase at OD_{600} 0.6 (Figure 6b). However, in the *ΔminD2ΔminD4ΔarlD* mutants, we observed a consistent reduction in foci formation even at high OD_{600}, which was not observed in the individual mutants (Figure 6b). Which suggests a synergistic effect or partially redundant role of MinD2 and MinD4 in the archaellum polar assembly.

In *H. volcanii*, chemotaxis involves the assembly of chemosensory arrays, which are preferentially localized at the cell poles during the early log phase and become diffused as the cells enter stationary phase (Li et al., 2019). Therefore, we investigated whether MinD2 also influences the localization of chemosensory arrays, using the chemotaxis protein CheW as a marker for these clusters (Li et al., 2019). For this, we expressed GFP-CheW in cells lacking CheW (Δ*cheW*) and cells lacking both CheW and MinD2 (*ΔminD2ΔcheW*). In Δ*cheW* cells, expression of GFP-CheW led to polar foci in 78% of cells and diffused localization in 22% of cells (Li et al., 2019 and Figure 6c). By contrast, in *ΔminD2ΔcheW* cells, expression of GFP-CheW led to only 5% of the cells displaying polar chemosensory foci, while the remaining 95% of *ΔminD2ΔcheW* cells showed diffused GFP-CheW localization, suggesting the absence of chemosensory arrays localization (Figure 6c). Further analysis showed that the number of cells with CheW foci formation decreased substantially in Δ*cheW* as the OD increased, while the localization of CheW reduces 5–0% in *ΔminD2ΔcheW* cells (Figure 6d).

Since previous studies have shown that MinD4 regulates the positioning of chemosensory arrays, we further investigated the links between MinD2 and MinD4 in modulating CheW localization. For this, we created the triple deletion mutant Δ*minD2*Δ*minD4*Δ*cheW*, which has then been used to study the localization of GFP-CheW. Our results showed a further reduction in CheW foci formation in *ΔminD2ΔminD4ΔcheW* cells compared to *ΔminD2ΔcheW* cells, with only 2% of cells showing polar GFP-CheW foci at OD_{600} of 0.01. In the transition of OD_{600} from 0.01 to 0.03, we observed GFP-CheW foci formation to be completely absent, with no cells displaying CheW foci (Figure 6d).

Collectively, these data show that MinD2 impacts the localization of both the motility and chemosensory machineries, in addition to the effect that MinD4 has on the localization of these cellular components.

RNA-seq analysis reveals limited transcriptional changes in archaellum and chemotaxis gene upon *minD2* deletion

Given the pronounced effects of *minD2* deletion on cell shape and motility, we explored whether these phenotypic changes can be explained by transcriptional alterations in the relevant genes. Therefore, RNA-seq analysis was performed on WT and Δ*minD2* strains at early-log (OD_{600}: 0.02), mid-log (OD_{600}: 0.3) and stationary phase (OD_{600}: 2) to identify any differentially expressed genes (log_2-fold changes: Supplementary Table 6, transcript per million values:

10.3389/fmicb.2024.1474570

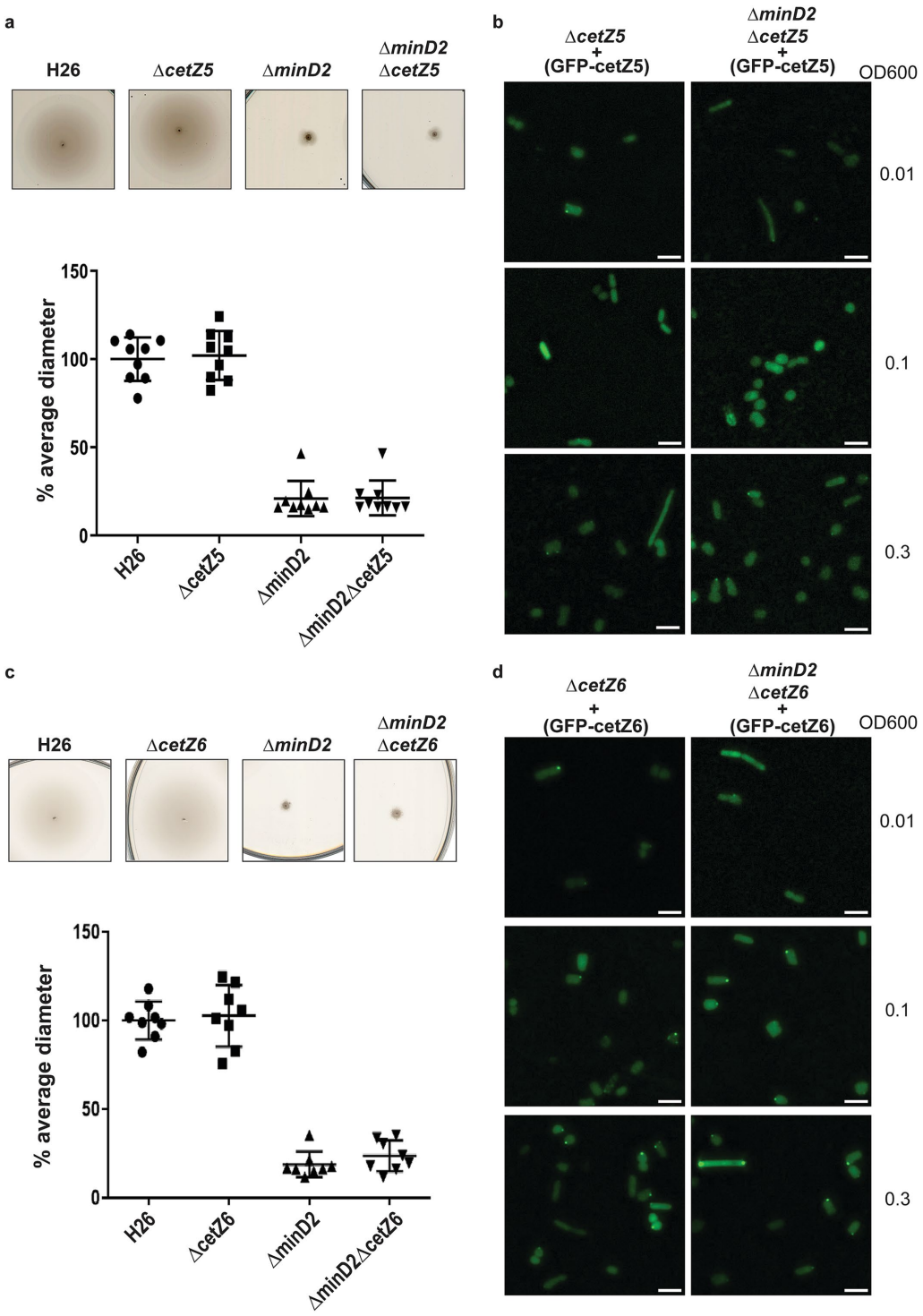


FIGURE 5
CetZ5 and CetZ6 do not have a motility phenotype and display diffused fluorescence. (a) Semi-solid agar-based motility assay for H26, Δ*CetZ5* and double mutant, Δ*minD2CetZ5*. (b) Localization of GFP-CetZ5 in Δ*CetZ5* and Δ*minD2CetZ5*. (c) Semi-solid agar-based motility assay for H26, Δ*CetZ6* and double mutant, Δ*minD2CetZ6*. (d) Localization of GFP-CetZ6 in Δ*CetZ6* and Δ*minD2CetZ6*. Scale bar: 4 μm. Calculations were made using 3 independent experiments including 3 biological replicates each.

Supplementary Table 7). Despite the clear morphological and motility defects, no significant transcriptional changes in the core genes related to archaellum assembly or chemotaxis were consistent over all three growth stages (Figure 7a). Instead, some of the genes showed stage-dependent differences, including downregulation of CetZ2 (log_2-fold change: −2.3) and upregulation of PssA (log_2-fold change: 1.1) during early growth phase, and upregulation of CheW (log_2-fold change: 0.9) during stationary phase. In general, many significant transcriptional changes were observed in the early (down: 21.9%; up: 20.9%) and stationary phase (down: 12.1%; up: 19.8%),

10.3389/fmicb.2024.1474570

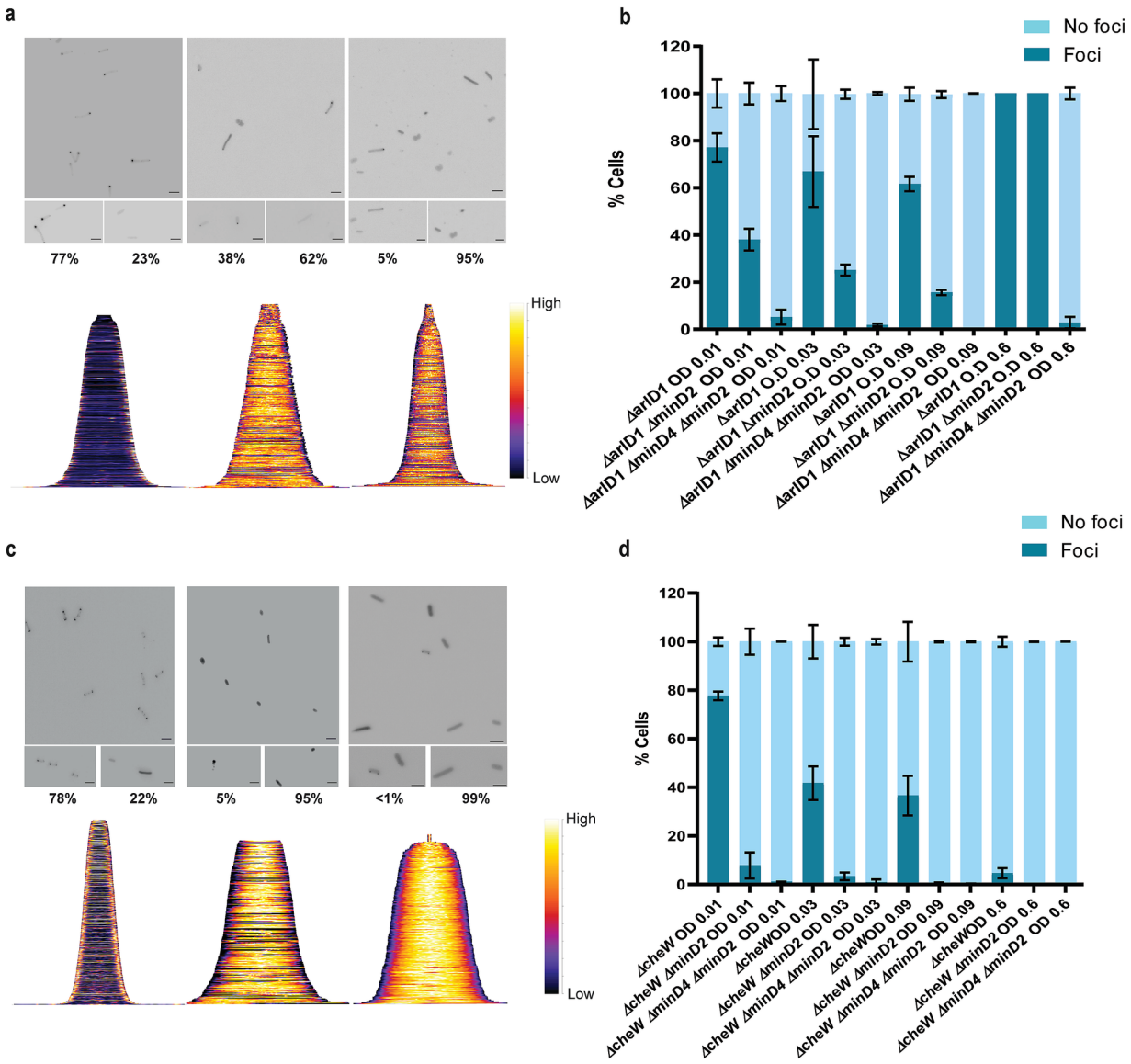


FIGURE 6
MinD2 affects the archaellum and chemotaxis machinery. (a) Fluorescent image of ArlD-GFP in Δ*arlD*, Δ*arlDminD2* and Δ*arlDminD2minD4* strain. (b) Graphical analysis of the microscopic images to represent the % of cells with or without ArlD-GFP polar foci. (c) Fluorescent image of GFP-CheW in Δ*cheW*, Δ*cheWminD2*, and Δ*cheWminD2minD4* strain. (d) Graphical analysis of the microscopic images to represent the % of cells with or without GFP-CheW polar foci. Scale bar: 4 μm. (a,c) Lower panel—demographic analysis of the foci distribution showing spatial distribution of the proteins (yellow) arranged in an ascending order of cell length. Cells are arranged in ascending order (b,d). Calculations were made using 3 independent experiments including 3 biological replicates each.

but not at mid-exponential (down: 0.4%; up: 0.4%) (Figure 7b). From the 33 differentially regulated genes at OD_{600} of 0.2 (mid), the loss of polar localization of the motility machinery could not be explained (Figure 7c). To detect any general trends, a functional enrichment analysis using clusters of orthologous groups was performed (Figure 7d and Supplementary Table 8). This analysis revealed the absence of consistently regulated and functionally related genes upon *minD2* deletion. Instead, a counter-regulation of metabolism groups C (energy production and conversion), G (carbohydrate transport and metabolism), E (amino acid transport and metabolism), and I (lipid transport and metabolism) could be observed, with genes that are downregulated during early growth phase but are later upregulated.

Together, these RNA-seq findings complement the previous localization studies and suggest that MinD2 regulates these processes at the level of protein localization rather than gene expression. This underscores the importance of MinD2 in maintaining cell morphology and motility through mechanisms beyond transcriptional control.

Discussion

Haloferax volcanii encodes four homologs of the MinD protein, which unlike their counterparts in bacterial cells, are not involved in cell division and do not regulate the localization of the divisome (Nußbaum et al., 2020). Previous studies have started to elucidate the

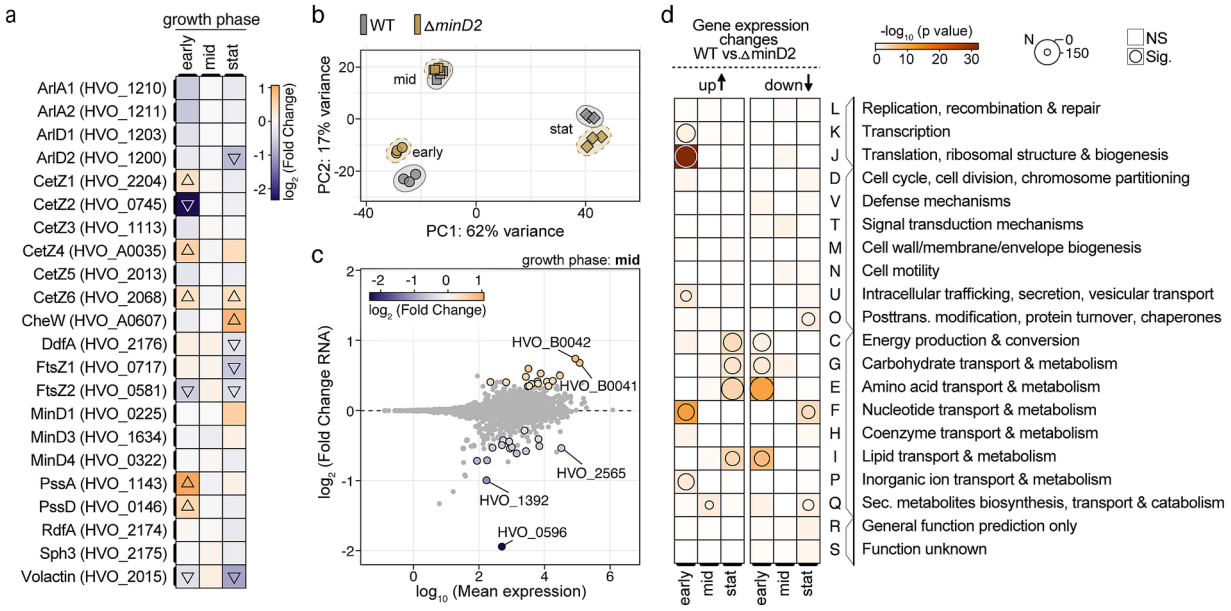


FIGURE 7
Deletion of minD2 leads to growth-phase-dependent transcriptomic changes with limited influence on archaellum and chemotaxis genes. (a) Heatmap showing the $\log_2$-fold changes in expression levels of selected cell division, cell cycle-related, chemotaxis, archaellation and shape-determining genes across three growth phases: early (OD_{600}: 0.02), mid (OD_{600}: 0.3), and stationary (stat, OD_{600}: 2). The color gradient represents downregulation (blue) and upregulation (orange), with triangles indicating significant changes in the ΔminD2 strain relative to wild type (WT). (b) Principal component analysis (PCA) based on transcriptomic data for WT (gray) and ΔminD2 (brown) strains across growth phases indicated by different shaped (early: circle, mid: rectangle, and stat: triangle). (c) MA plot ($\log_2$-fold change vs. $\log_{10}$ mean expression) for the mid-exponential phase, highlighting differentially expressed genes in the ΔminD2 strain compared to WT. Significantly regulated genes (adjusted *p*-value <0.05) are shown with a black outline and are colored. (d) Archaeal clusters of orthologous groups (arCOG) enrichment analysis of gene expression changes (WT vs. ΔminD2) across growth phases. Overrepresentation of groups is shown for upregulated (left) and downregulated (right) genes in the ΔminD2 strain. The size of the circles represents the number of differentially expressed genes, and the color intensity corresponds to the significance of enrichment.

functions of MinD proteins in *H. volcanii,* particularly MinD4 (HVO_0322), which governs the precise positioning of both the archaellum and chemotaxis machineries, which are indispensable for enabling directional and purposeful motility (Nußbaum et al., 2020) in rod cells. Here, we further extend the characterization of the functions of MinD proteins in *H. volcanii* cells, focusing on MinD2. Our analyses underscored MinD2 as a regulatory protein in cell shape morphology (Figure 1) and motility (Figure 2), by enabling cells to retain a rod shape, particularly in the early growth phase. Furthermore, we demonstrate that MinD2 shows diffuse localization across the cell (Supplementary Figure 3) and likely interacts with a variety of partner proteins to mediate its effects (Figures 4, 5). By analyzing the localization of a variety of proteins involved in the formation of the archaellum and chemotaxis complexes, we also reveal that MinD2 modulates the placement of the motility and chemotaxis machineries (Figure 6). Interestingly, RNA-seq analysis indicated no transcriptional changes in genes related to morphology or motility, despite the phenotype observed (Figure 7). Finally, we investigate the links between MinD2 and MinD4, showing that the two MinD homologues have synergistic roles in linking cell shape and motility in *H. volcanii* (Figure 6).

Studies on *H. volcanii* provide insights into the regulation of cell shape in response to environmental cues. The cells undergo remarkable transformations in cell shape during different growth phases and conditions. These changes, from rod-shaped to flat, polygonal pleomorphic disks (plate shaped), have been a subject of interest due to their potential roles in adaptation and survival strategies (Halim et al., 2017; Li et al., 2019; de Silva et al., 2021). However, our understanding of the molecular mechanisms controlling archaeal cell shape determination is still developing. Previous studies identified proteins such as CetZ1, LonB, ArtA, PssA, and PssD as important regulators of this process (Duggin et al., 2015; Halim et al., 2017; Ferrari et al., 2020; Brown and Duggin, 2023). More recently, cell-shape mutants lacking the ability to form plates have been studied, including DdfA (disk determining factor), which is likely involved in the signaling pathways that determine cell shape. Additionally, studies on RdfA (rod-determining factor) and Sph3 (SMC-like protein) indicate that these proteins are also involved in the signal cascade that potentially regulates cell shape (Schiller et al., 2024). However, how these and other cell-shape determinants interact with each other, and how their activity is regulated by environmental conditions, remains unclear.

Our findings identify MinD2 as another protein that regulates cell shape in *H. volcanii* cells (Figure 1 and Supplementary Figure 2). One of the most distinctive features of the *ΔminD2* mutant is its preponderance to form plate-shaped cells, including in the early log phase, highlighting the importance of MinD2 in maintaining the rod-shaped morphology characteristic of *H. volcanii* cells. Furthermore, while the presence of a plasmid delayed the rod-to-plate cell shape transition, particularly in early log phase, the majority of *ΔminD2*+pTA1392 cells still displayed plate shape as the optical density of the cultures increased (Figure 1b). Therefore, although the presence of plasmid can partially prevent the loss of rod shape in the

Δ*minD2* mutant early on, the absence of MinD seems to dominate the phenotype, resulting in the majority of cells being plates.

While our findings demonstrate that MinD2 regulates cell shape, the mechanism by which MinD2 operates remains unclear. Notably, we find that MinD2 shows a diffuse localization pattern across the cell, rather than localizing to specific foci, such as the cell pole (Figure 3). Furthermore, our experiments with various variants of MinD2 suggest that its N-terminus is important for activity, potentially through interactions with protein partners (Figure 3). Indeed, we identified several potential interacting partners of MinD2, including HVO_0596, CetZ5, and CetZ6. However, mutants lacking HVO_0596, a protein transcribed from the same operon as MinD2, or either of the CetZs, showed no discernible phenotypes with regards to morphology (Figures 4, 5). Furthermore, our experiments with single vs. double mutants show that the observed impacts of motility in these mutants is primarily attributable to MinD2 (Figures 4, 5).

While future studies are needed to further elucidate how MinD2 regulates morphology and motility in *H. volcanii*, our findings provide some insights into the links between these two cellular processes. For example, we found that discoid *ΔminD2* cells have a significant decrease in the number of archaella, which results in decreased motility (Figure 2). In addition, MinD2 deletion also impacted the positioning of the chemotaxis machinery (Figure 6). One possibility is that these changes in the localization of archaella and chemosensory arrays result from the observed changes in morphology in *ΔminD2* cells, particularly their inability to retain rod shape. This possibility is supported by studies in other microbial species linking cell shape and the special organization of motility machinery. For example, in *E. coli,* chemosensory arrays preferably localize to the curved membrane (Strahl et al., 2015; Draper and Liphardt, 2017). Functional enrichment analysis further supported the notion that MinD2's effects are not driven by consistent changes in gene expression (Figure 7), indicating that MinD2 participates in the spatial organization of the archaeal cell pole. For example, some proteins, like bacterial ParA/MinD homologs (such as FlhG and ParC) rely on polar landmark proteins for localization (Lutkenhaus, 2012). In *H. volcanii*, the MinD4 homolog oscillates along the cell axis and is hypothesized to function similarly to polar landmark proteins, thereby influencing the proper positioning of both archaella and chemosensory arrays (Nußbaum et al., 2020). Notably, here we show that *ΔminD2ΔminD4* mutants have stronger defects in the positioning of the motility and chemosensory machinery (Figure 6) than *ΔminD4* single mutants (Nußbaum et al., 2020; Figure 6), suggesting that MinD2 and MinD4 have non-redundant roles in these processes.

Based on our findings and previous studies, we propose a model for how MinD2 controls cell shape and motility in *H. volcanii* (Figure 8). In this model, while MinD4 is important for promoting the adequate positioning of polar proteins, some of these proteins can still find the 'probable' pole even in the absence of MinD4. This is in agreement with previous studies, which showed that while a reduced frequency of *ΔminD4* mutant cells express archaella (20%) and chemosensory array (40%), some cells can still assemble these structures (Nußbaum et al., 2020). In the case of *ΔminD2* mutants, we postulate that the loss of rod shape resulting from MinD2 absence influences the positioning of these machineries in most cells, although some can still assemble archaella and chemosensory arrays guided by polar cap proteins. In the absence of both MinD2 and MinD4, cells are always plate shaped and polar MinD4 patches are not established. This combination may render proteins targeted to the poles unable to detect 'probable' poles and/or impair their interaction with different polar cap proteins. Under this scenario, the lack of both MinD2 and MinD4 leads to a synergistic effect, resulting in an almost complete absence of cells with archaellum (5%) or chemosensory (1%) machineries.

Morphological integrity and the determination of cell shape are pivotal aspects of microbial physiology. While these mechanisms are fairly well understood in bacteria, less is known about their archaeal counterparts. This study further advances our understanding of cell morphology determination in archaea by highlighting the pivotal role of the archaeal MinD2 homologue in modulating *H. volcanii* morphology. Notably, while bacterial Min proteins predominantly influence cell division, our observations suggest that MinD2 in archaea may have evolved or diversified to have more influence on cell morphology. Additionally, our findings support a model in which MinD2 contributes to cells retaining a rod-shape morphology, which enables polar cap proteins to recognize the cell poles, thereby supporting the assembly and adequate positioning of the motility and chemotaxis machinery. Furthermore, our findings add support to previous studies on MinD4 suggesting that archaeal MinD homologues have non-redundant roles in influencing cell shape and the positioning of motility and chemotaxis machinery. These findings contribute to a deeper understanding of the intricate regulatory network governing cellular processes in archaea, and set the stage for future studies aimed at uncovering the detailed molecular mechanisms by which MinD proteins modulate haloarchaeal physiology.

Materials and methods

All chemicals have been purchased from Roth or Sigma unless stated otherwise.

Strain and growth condition

Escherichia coli strains were cultured in LB (Luria Broth)-medium or grown on LB agar plates, with the necessary antibiotics (100 μg/mL ampicillin, 30 μg/mL chloramphenicol, 25 μg/mL kanamycin) and grown at 37° C. Liquid cultures were constantly shaken at 150 rpm.

Haloferax volcanii H26 cells were grown in YPC medium [0.5% (w/v) yeast extract (Difco), 0.1% (w/v) peptone (Oxoid), and 0.1%(w/v) casamino acids (Difco)] dissolved in 18% buffered Salt Water (SW) (144 g/L NaCl, 18 g/L MgCl2 * 6 H2O, 21 g/L MgSO4 * 7 H2O, 4.2 g/L KCl, 12 mM Tris/HCl, pH 7.5), supplemented with 3 mM $CaCl_2$, adjusted to a pH of 7.2 with KOH for transformations. For experiments, CAB medium was used [i.e. CA medium (0.5% (w/v) casamino acid) dissolved in 18% SW, supplemented with 3 mM $CaCl_2$, and 0.8 μg/mL of thiamine, and 0.1 μg/mL of biotin, adjusted to a pH of 7.2 with KOH and supplemented with trace elements solutions (Duggin et al., 2015)] was used. For each experiment, a single colony was inoculated into 5 mL medium and diluted to a larger volume on the subsequent day. This dilution was crucial to ensure an appropriate cell density for subsequent experiments. By adjusting the OD_{600} to the desired value, it was possible to obtain a consistent starting point on the day of the experiment.

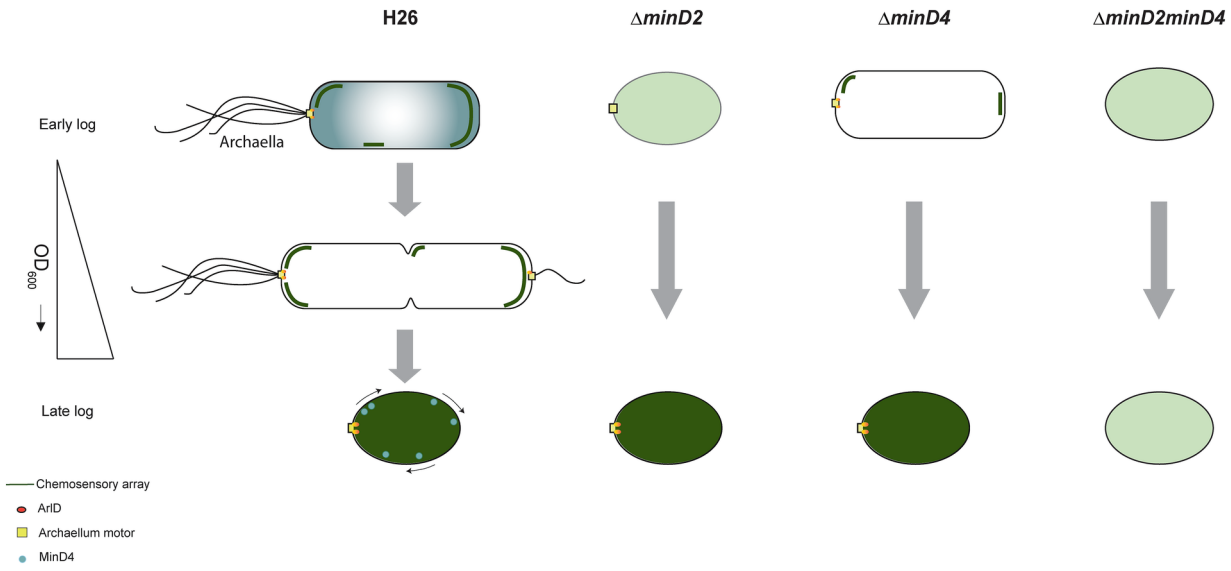


FIGURE 8
Proposed model for MinD2 function in *Haloferax volcanii*. *H. volcanii* cells show growth-dependent cell shape morphology transitioning from rods to plate-shaped cells. We show that MinD2 has a strong effect on shape, with Δ*minD2* mutant cells predominantly adopting a plate shape. We propose that the shape phenotype impacts the proper positioning of the archaellum (archaellum motor: yellow square and ArlD: orange ovals) and chemotactic machinery (dark green lines). According to this model, in Δ*minD2* cells, the localization of both machineries is diffused (light green) with a very few cells showing the ability to localize the chemosensory arrays in early log phase. As the growth stage progresses, only chemosensory arrays remain diffused (dark green) and the archaellum motor is able to localize at a pole. Previous studies on MinD4 (Nußbaum et al., 2020; Nußbaum et al., 2020), shows the Δ*minD4* mutant cells has an effect on the positioning on archaellum and chemosensory arrays. Together, we postulate that deleting both MinD homologues (Δ*minD2minD4*) has a synergistic effect, resulting in these cells being unable to localize both archaellum and chemosensory arrays to the pole at any stage of the growth phase.

For strains with an auxotrophic mutation grown in CA/CAB medium, the medium was supplemented with 50 µg/mL uracil for *ΔpyrE2*. Alternatively, the strains were transformed with a plasmid carrying the respective gene for viable growth. For growth curve, cells were grown in 15 mL culture volume and measured with cell growth quantifier (CGQ) (Aquila biolabs GmbH) at 45°C and shaking at 120 rpm.

Genetic modification of *Haloferax volcanii*

Transformation in H26 based on uracil selection via the Polyethylene glycol 600 (PEG600) along with gene deletion and expression studies were conducted as described previously (Allers et al., 2004). For transformation into *H. volcanii*, non-methylated plasmids were extracted from *E. coli dam*$^-$/*dcm*$^-$ (C2925I, NEB). Mutant strains generated and used are described in Supplementary Table 1. Plasmids created for knockout mutants are described in Supplementary Table 2. Primers to create knockout plasmids were based on pTA131, are described in Supplementary Table 3.

Growth curve

Glycerol stocks were streaked on solid agar medium substituted with uracil and incubated at 45°C for 5 days. A single colony from plates was used to inoculate 5 mL of media on day 1. Strains without plasmid were grown in CAB with 50 µg/mL uracil and strains with plasmid pTA1392 (containing *pfdx-pyrE2*) were grown in CAB medium without supplements. To generate a growth curve, the obtained culture was inoculated on day 2 at a starting OD_{600} of 0.05. The cell density was measured with cell growth quantifier (CGQ) (Aquila biolabs GmbH) at 45°C and shaking at 120 rpm and measurement were taken every 300 s.

Spot survival assay

To assess the growth and viability of the H26 and *ΔminD2* (transformed with pTA1392) strains, cells were cultured in 5 mL CAB medium at 45°C to an OD_{600} of 0.2. On the next day, cell cultures were set to a theoretical OD_{600} of 0.2. A serial dilution was prepared until a dilution of 10^5 and 5 µL of each dilution per strain was spotted on CA-plates. The plates were incubated for 2 days at 45°C in sealed plastic bags.

Microscopy

The cell shape was analyzed by imaging the cells with an inverted phase contrast light microscope (Zeiss Axio Observer Z.1). The cells were grown in 5 mL of the respective medium and diluted in 20 mL media volume the next day in order to achieve an OD_{600} of 0.01 the day after. For each culture, 5 µL sample was collected from different growth phases and dropped at the center of an agarose pad (0.3% (w/v) agarose dissolved in 18% SW). On drying, the pad was covered with a cover slip and imaged. The images were acquired at 100x magnification using the oil immersion phase contrast (PH3) channel. All sampled were analyzed in triplicated. Fluorescence microscopy images were acquired on Zeiss Axio Observer Z.1 (ex: 450–490 nm

em: 500–550 nm filter from Chroma®), equipped with a heated XL-52000 Incubator running VisiVIEW® software for MinD4 and CheW.

Image analysis

The phase contrast images from the microscopy were analyzed using Fiji (Schindelin et al., 2012) combined with MicrobeJ plugin (Ducret et al., 2016). For the analysis, cells that formed aggregates or were fragmented were discarded from the calculation. The circularity of the cells was automatically calculated. The diameter of each analyzed cell was thereon calculated and grouped into 6 bins in the range interval of 0.1 to 1. The parameters used for circularity were as previously defined in Patro et al. (2023).

Transmission electron microscopy

Cells were harvested at 2000 g for 15 min. The resulting pellet was resuspended to a theoretical O.D of 10. Five µL of cells of the cell suspension was applied to a glow discharged carbon coated copper grid (Plano GmbH, Wetzlar Germany) and incubated for 10 s. The excess liquid was blotted away. The cells were then stained with 2% uranyl acetate (w/v). Cells were imaged using Zeiss Leo 912 Omega (tungsten) operated at 80 kV and images were taken using Dual speed 2K on Axis charge-coupled device (COD) camera (TRS, Sharp-Eye).

Motility assay

Semi-solid agar plates were prepared using 0.33% agar in CA medium supplemented with 1 mM tryptophan. Cultures grown at OD_{600} 0.3 were inoculated into the plates using stab techniques, and the plates were then incubated at 45°C for 4 days. To compare the motility of different strains, all strains were spotted on the same plate. For each strain, a minimum of 3 technical replicates and 3 biological replicates were conducted. After 4 days, the diameter of the motility ring was assessed.

RNA sequencing for differential gene expression analysis

RNA extraction

RNA was extracted from the *minD2* deletion strain during lag phase (OD_{600} 0.02), exponential phase (OD_{600} 0.3), and stationary phase (OD_{600} 2), and WT H26 was used as control. Cultures were scaled to 1 L for lag phase, 20 mL for exponential phase, and 3 mL for stationary phase, with 5 replicates per strain and growth phase. Once the cells reached the desired optical density, they were harvested and resuspended to a theoretical OD_{600} of 5. Approx. 2 mL of the resuspended cells were transferred to a new tube and centrifuged. RNA was isolated using the RNeasy® Plus Mini Kit from QIAGEN. The cell pellets were resuspended in 600 µL of RLT-plus buffer, and RNA extraction was performed according to the Qit. The RNA was then eluted in 31 µL of RNase-free water (Roth®), rapidly frozen in liquid nitrogen, and stored at −80°C.

Library preparation and sequencing

RNA quality was assessed using a Bioanalyzer, with only samples exhibiting RNA integrity number (RIN) values of 8.5 or higher being used. Prior to library preparation, RNA samples underwent Turbo DNase treatment according to the manufacturer's instructions (Ambion, 1 unit) to remove any residual DNA. For each experimental condition, four independent biological replicates were prepared. To ensure the removal of ribosomal RNA, 2 µg of input RNA was treated with a Ribopool designed specifically for *Haloferax volcanii* (siTOOLs) according to the manufacturer's instructions.

Library preparation and RNA-sequencing were carried out as described in the Illumina "Stranded mRNA Prep Ligation" Reference Guide, the Illumina NextSeq 2000 Sequencing System Guide (Illumina, Inc., San Diego, CA, USA), and the KAPA Library Quantification Kit - Illumina/ABI Prism (Roche Sequencing Solutions, Inc., Pleasanton, CA, USA).

In brief, omitting the initial mRNA purification step with oligo(dT) magnetic beads, approximately 5 ng of rRNA depleted archaeal RNA was fragmented to an average insert size of 200–400 bases using divalent cations under elevated temperature (94°C for 8 min). Next, the cleaved RNA fragments were reverse transcribed into first strand complementary DNA (cDNA) using reverse transcriptase and random hexamer primers. Thereby Actinomycin D was added to allow RNA-dependent synthesis and to improve strand specificity by preventing spurious DNA-dependent synthesis. Blunt-ended second strand cDNA was synthesized using DNA Polymerase I, RNase H and dUTP nucleotides. The incorporation of dUTP, in place of dTTP, quenches the second strand during the later PCR amplification, because the polymerase does not incorporate past this nucleotide. The resulting cDNA fragments were adenylated at the 3′ ends and the pre-index anchors were ligated. Finally, DNA libraries were created using a 15 cycles PCR to selectively amplify the anchor-ligated DNA fragments and to add the unique dual indexing (i7 and I5) adapters. The bead purified libraries were quantified using the KAPA Library Quantification Kit. Equimolar amounts of each library were sequenced on an Illumina NextSeq 2000 instrument controlled by the NextSeq 2000 Control Software (NCS) v1.5.0.42699, using one 50 cycles P3 Flow Cell with the dual index, single-read (SR) run parameters. Image analysis and base calling were done by the Real Time Analysis Software (RTA) v3.10.30. The resulting .cbcl files were converted into .fastq files with the bcl2fastq v2.20 software. Library preparation and RNA-sequencing were performed at the Genomics Core Facility "KFB–Center of Excellence for Fluorescent Bioanalytics" (University of Regensburg, Regensburg, Germany[1]).

Differential gene expression analysis

Raw sequencing reads in FASTQ format were initially processed for quality control and trimming using fastp (v. 0.23.2) (Chen et al., 2018). This step removed low-quality bases and adapter sequences with parameters set to `--cut_front --cut_tail-q 30` to ensure high-quality data. The filtered reads were then aligned to the *Haloferax volcanii* DS2 reference genome using Bowtie2 (v. 2.5.0) with default settings (Langmead and Salzberg, 2012). The resulting sequence alignment files in SAM format were converted to BAM using samtools (v. 1.17) for efficient

1 www.kfb-regensburg.de

handling and analysis (Li et al., 2009). To identify differentially expressed genes, we followed the guidelines provided in the Bioconductor vignette for the DESeq2 package (v. 1.42.1) (Love et al., 2014). Gene counts were derived from RNA-seq data using featureCounts (part of RSubread package v. 2.16.1) with a custom GTF file (Liao et al., 2019). This file was created by filtering the *H. volcanii* DS2 GFF annotation file, obtained from the NCBI, to include only features with the type `gene` or `gene.` Principal component analysis (PCA) was performed on variance-stabilizing transformed data to assess overall data structure and identify outliers. Specifically, outlier replicates were removed based on visual inspection, including wild type early phase replicate 4, wild type mid phase replicate 1, wild type stationary phase replicate 1, Δ*minD2* early phase replicate 4 and Δ*minD2* stationary phase replicate 1. Differential expression analysis was conducted by comparing the wild type and Δ*minD2* strains across different growth conditions to identify growth-dependent changes in gene expression.

Functional enrichment analysis based on arCOG classification

To elucidate the functional characteristics of differentially expressed genes, we performed a functional enrichment analysis using the Archaeal Clusters of Orthologous Genes (arCOG) classification, as described previously (Makarova et al., 2015; Grünberger et al., 2023). Briefly, arCOGs for *H. volcanii* were retrieved from Makarova et al. (2015) and gene set enrichment analysis performed with the goseq package (v. 1.54.0) in R (Young et al., 2010). For each growth condition, a background file was generated from all genes that could be detected. Next, *p*-values for overrepresentation of arCOG terms in the differentially expressed genes were calculated separately for up-and downregulated genes based on RNA-seq data. Significance was determined using a threshold of 0.05 to identify terms that were significantly enriched.

Structural analyses

The Alphafold models for MinD2 and HVO_0596 were generated with Alphafold 3 Google colab (Abramson et al., 2024) and further analyzed by ChimeraX (Meng et al., 2023).

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author. RNA sequencing data are available at the European Nucleotide Archive (ENA, https://www.ebi.ac.uk/ena) under project accession number PRJEB79934.

Author contributions

MP: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. FG: Formal analysis, Investigation, Visualization, Writing – review & editing. SS: Formal analysis, Investigation, Writing – review & editing. SG: Formal analysis, Investigation, Writing – review & editing. MR-F: Formal analysis, Investigation, Writing – review & editing. PN: Formal analysis, Investigation, Writing – review & editing. DG: Supervision, Resources, Writing – review & editing. SI: Conceptualization, Formal analysis, Investigation, Supervision, Writing – review & editing. S-VA: Conceptualization, Funding acquisition, Methodology, Project administration, Resources, Supervision, Validation, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. MP was supported by the German Research Foundation on grant number AL1206/4-3. SG was supported by the German Science Foundation (DFG) project number 505545313 (AL1206/14-1) to S-VA. PN and SI were supported by the VW Foundation by a Momentum grant to S-VA (AZ 94993). SS was supported by the German Research Foundation under project number 403222702-SFB 1381 to S-VA. The TEM (Hitachi HT7800) was funded by the DFG project number 42689454 and is operated by the faculty of biology at the University of Freiburg as a partner unit within the Microscopy and Image Analysis Platform (MIAP) and Life Imaging Centre (LIC), Freiburg.

Acknowledgments

We thank Katharina Vogel for her technical assistance. Molecular graphics and analyses performed with UCSF ChimeraX, developed by the Resource for Biocomputing, Visualization, and Informatics at the University of California, San Francisco, with support from National Institutes of Health R01-GM129325 and the Office of Cyber Infrastructure and Computational Biology, National Institute of Allergy and Infectious Diseases. We acknowledge support by the Open Access Publication Fund of the University of Freiburg.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2024.1474570/full#supplementary-material

References

Abramson, J., Adler, J., Dunger, J., Evans, R., Green, T., Pritzel, A., et al. (2024). Accurate structure prediction of biomolecular interactions with AlphaFold 3. *Nature* 630, 493–500. doi: 10.1038/s41586-024-07487-w

Allers, T., Ngo, H. P., Mevarech, M., and Lloyd, R. G. (2004). Development of additional selectable markers for the halophilic archaeon *Haloferax volcanii* based on the leuB and trpA genes. *Appl. Environ. Microbiol.* 70, 943–953. doi: 10.1128/AEM.70.2.943-953.2004

Atmakuri, K., Cascales, E., Burton, O. T., Banta, L. M., and Christie, P. J. (2007). Agrobacterium ParA/MinD-like VirC1 spatially coordinates early conjugative DNA transfer reactions. *EMBO J.* 26, 2540–2551. doi: 10.1038/sj.emboj.7601696

Babski, J., Haas, K. A., Näther-Schindler, D., Pfeiffer, F., Förstner, K. U., Hammelmann, M., et al. (2016). Genome-wide identification of transcriptional start sites in the haloarchaeon *Haloferax volcanii* based on differential RNA-Seq (dRNA-Seq). *BMC Genomics* 17:629. doi: 10.1186/s12864-016-2920-y

Baxter, J. C., and Funnell, B. E. (2014). Plasmid partition mechanisms. *Microbiol. Spectr.* 2. doi: 10.1128/microbiolspec.PLAS-0023-2014

Brown, H. J., and Duggin, I. G. (2023). Diversity and potential multifunctionality of archaeal CetZ tubulin-like cytoskeletal proteins. *Biomol. Ther.* 13:34. doi: 10.3390/biom13010134

Brown, H. J., Islam, M. I., Ruan, J., Baker, M. A. B., Ithurbide, S., and Duggin, V. O. R. C. I. D. P. I. G. (2024). CetZ1-dependent assembly and positioning of the motility machinery in haloarchaea. *bioRxiv*. doi: 10.1101/2024.05.02.592137

Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). Fastp: an ultra-fast all-in-one FASTQ preprocessor., in. *Bioinformatics* 34, i884–i890. doi: 10.1093/bioinformatics/bty560

de Silva, R. T., Abdul-Halim, M. F., Pittrich, D. A., Brown, H. J., Pohlschroder, M., and Duggin, I. G. (2021). Improved growth and morphological plasticity of *Haloferax volcanii*. *Microbiology* 167:001012. doi: 10.1099/MIC.0.001012

Draper, W., and Liphardt, J. (2017). Origins of chemoreceptor curvature sorting in *Escherichia coli*. *Nat. Commun.* 8:14838. doi: 10.1038/ncomms14838

Ducret, A., Quardokus, E. M., and Brun, Y. V. (2016). MicrobeJ, a tool for high throughput bacterial cell detection and quantitative analysis. *Nat. Microbiol.* 1:16077. doi: 10.1038/NMICROBIOL.2016.77

Duggin, I. G., Aylett, C. H. S., Walsh, J. C., Michie, K. A., Wang, Q., Turnbull, L., et al. (2015). CetZ tubulin-like proteins control archaeal cell shape. *Nature* 519, 362–365. doi: 10.1038/nature13983

Esquivel, R. N., Schulze, S., Xu, R., Hippler, M., and Pohlschroder, M. (2016). Identification of *Haloferax volcanii* pilin N-glycans with diverse roles in pilus biosynthesis, adhesion, and microcolony formation. *J. Biol. Chem.* 291, 10602–10614. doi: 10.1074/jbc.M115.693556

Ferrari, M. C., Cerletti, M., Paggi, R. A., Troetschel, C., Poetsch, A., and De Castro, R. E. (2020). The LonB protease modulates the degradation of CetZ1 involved in rod-shape determination in *Haloferax volcanii*. *J. Proteome* 211:103546. doi: 10.1016/j.jprot.2019.103546

Grünberger, F., Schmid, G., Ahmad, Z.El, Fenk, M., Vogl, K., Reichelt, R., et al. (2023). *MBio*:14. doi: 10.1128/mbio.02174-23, Uncovering the temporal dynamics and regulatory networks of thermal stress response in a hyperthermophile using transcriptomics and proteomics, 14

Halim, M. F. A., Stoltzfus, J. D., Schulze, S., Hippler, M., and Pohlschroder, M. (2017). ArtAdependent processing of a tat substrate containing a conserved tripartite structure that is not localized at the C terminus. *J. Bacteriol.* 199:e00802-16. doi: 10.1128/JB.00802-16

Hu, Z., and Lutkenhaus, J. (2003). A conserved sequence at the C-terminus of MinD is required for binding to the membrane and targeting MinC to the septum. *Mol. Microbiol.* 47, 345–355. doi: 10.1046/j.1365-2958.2003.03321.x

Ithurbide, S., de Silva, R. T., Brown, H. J., Shinde, V., and Duggin, I. G. (2024). A vector system for single and tandem expression of cloned genes and multi-colour fluorescent tagging in *Haloferax volcanii*. *Microbiology* 170:170. doi: 10.1099/mic.0.001461

Jalal, A. S. B., and Le, T. B. K. (2020). Bacterial chromosome segregation by the ParABS system. *Open Biol.* 10:200097. doi: 10.1098/rsob.200097

Jeoung, J. H., Giese, T., Grünwald, M., and Dobbek, H. (2009). CooC1 from *Carboxydothermus hydrogenoformans* is a nickel-binding ATPase. *Biochemistry* 48, 11505–11513. doi: 10.1021/bi901443z

Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with bowtie 2. *Nat. Methods* 9, 357–359. doi: 10.1038/nmeth.1923

Leipe, D. D., Wolf, Y. I., Koonin, E. V., and Aravind, L. (2002). Classification and evolution of P-loop GTPases and related ATPases. *J. Mol. Biol.* 317, 41–72. doi: 10.1006/jmbi.2001.5378

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi: 10.1093/bioinformatics/btp352

Li, Z., Kinosita, Y., Rodriguez-Franco, M., Nußbaum, P., Braun, F., Delpech, F., et al. (2019). Positioning of the motility machinery in halophilic archaea. *mBio* 10:e00377-19. doi: 10.1128/mBio.00377-19

Liao, Y., Smyth, G. K., and Shi, W. (2019). The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. *Nucleic Acids Res.* 47:e47. doi: 10.1093/nar/gkz114

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15:550. doi: 10.1186/s13059-014-0550-8

Lutkenhaus, J. (2007). Assembly dynamics of the bacterial minCDE system and spatial regulation of the z ring. *Annu. Rev. Biochem.* 76, 539–562. doi: 10.1146/annurev.biochem.75.103004.142652

Lutkenhaus, J. (2012). The ParA/MinD family puts things in their place. *Trends Microbiol.* 20, 411–418. doi: 10.1016/j.tim.2012.05.002

Makarova, K. S., Wolf, Y. I., and Koonin, E. V. (2015). Archaeal clusters of orthologous genes (arCOGs): an update and application for analysis of shared features between thermococcales, methanococcales, and methanobacteriales. *Life* 5, 818–840. doi: 10.3390/life5010818

Meng, E. C., Goddard, T. D., Pettersen, E. F., Couch, G. S., Pearson, Z. J., Morris, J. H., et al. (2023). UCSF ChimeraX: tools for structure building and analysis. *Protein Sci.* 32:e4792. doi: 10.1002/pro.4792

Nußbaum, P., Ithurbide, S., Walsh, J. C., Patro, M., Delpech, F., Rodriguez-Franco, M., et al. (2020). An oscillating MinD protein determines the cellular positioning of the motility machinery in Archaea. *Curr. Biol.* 30, 4956–4972.e4. doi: 10.1016/j.cub.2020.09.073

Patro, M., Duggin, I. G., Albers, S.-V., and Ithurbude, S. (2023). Influence of plasmids, selection markers and auxotrophic mutations on *Haloferax volcanii* cell shape plasticity. *Front. Microbiol.* 14:1270665. doi: 10.3389/fmicb.2023.1270665

Pradhan, P., Taviti, A. C., and Beuria, T. K. (2024). The bacterial division protein MinDE has an independent function in flagellation. *J. Biol. Chem.* 300:107117. doi: 10.1016/j.jbc.2024.107117

Pulianmackal, L. T., Limcaoco, J. M. I., Ravi, K., Yang, S., Zhang, J., Tran, M. K., et al. (2023). Multiple ParA/MinD ATPases coordinate the positioning of disparate cargos in a bacterial cell. *Nat. Commun.* 14:3255. doi: 10.1038/s41467-023-39019-x

Ringgaard, S., Schirner, K., Davis, B. M., and Waldor, M. K. (2011). A family of ParA-like ATPases promotes cell pole maturation by facilitating polar localization of chemotaxis proteins. *Genes Dev.* 25, 1544–1555. doi: 10.1101/gad.2061811

Schiller, H., Kouassi, J., Hong, Y., Rados, T., Kwak, J., DiLucido, A., et al. (2024). Identification and characterization of structural and regulatory cell-shape determinants in *Haloferax volcanii*. *Nat Commun*. 15:1414. doi: 10.1038/s41467-024-45196-0

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., et al. (2012). Fiji: an open-source platform for biological-image analysis. *Nat. Methods* 9, 676–682. doi: 10.1038/nmeth.2019

Strahl, H., Ronneau, S., González, B. S., Klutsch, D., Schaffner-Barbero, C., and Hamoen, L. W. (2015). Transmembrane protein sorting driven by membrane curvature. *Nat. Commun.* 6:8728. doi: 10.1038/ncomms9728

Szklarczyk, D., Morris, J. H., Cook, H., Kuhn, M., Wyder, S., Simonovic, M., et al. (2017). The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res.* 45, D362–D368. doi: 10.1093/nar/gkw937

Young, M. D., Wakefield, M. J., Smyth, G. K., and Oshlack, A. (2010). Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biol.* 11:R14. doi: 10.1186/gb-2010-11-2-r14

frontiers | Frontiers in Microbiology

TYPE Original Research
PUBLISHED 22 November 2024
DOI 10.3389/fmicb.2024.1474697

OPEN ACCESS

EDITED BY
Damien Paul Devos,
Spanish National Research Council (CSIC), Spain

REVIEWED BY
Arthur Charles-Orszag,
University of California, San Francisco, United States
Burak Avcı,
Aarhus University, Denmark
Alex Bisson,
Brandeis University, United States

*CORRESPONDENCE
Iain G. Duggin
✉ iain.duggin@uts.edu.au

RECEIVED 02 August 2024
ACCEPTED 28 October 2024
PUBLISHED 22 November 2024

CITATION
Brown HJ and Duggin IG (2024) MinD proteins regulate CetZ1 localization in *Haloferax volcanii*. *Front. Microbiol.* 15:1474697. doi: 10.3389/fmicb.2024.1474697

MinD proteins regulate CetZ1 localization in *Haloferax volcanii*

Hannah J. Brown and Iain G. Duggin*

Australian Institute for Microbiology and Infection, University of Technology Sydney, Ultimo, NSW, Australia

CetZ proteins are archaea-specific homologs of the cytoskeletal proteins FtsZ and tubulin. In the pleomorphic archaeon *Haloferax volcanii*, CetZ1 contributes to the development of rod shape and motility, and has been implicated in the proper assembly and positioning of the archaellum and chemotaxis motility proteins. CetZ1 shows complex subcellular localization, including irregular midcell structures and filaments along the long axis of developing rods and patches at the cell poles of the motile rod cell type. The polar localizations of archaellum and chemotaxis proteins are also influenced by MinD4, the only previously characterized archaeal member of the MinD family of ATPases, which are better known for their roles in the positioning of the division ring in bacteria. Using *minD* mutant strains and CetZ1 subcellular localization studies, we show here that a second *minD* homolog, *minD2*, has a strong influence on motility and the localization of CetZ1. Knockout of the *minD2* gene altered the distribution of a fluorescent CetZ1-mTq2 fusion protein in a broad midcell zone and along the edges of rod cells, and inhibited the localization of CetZ1-mTq2 at the cell poles. MinD4 had a similar but weaker influence on motility and CetZ1-mTq2 localization. The *minD2/4* mutant strains formed rod cell shapes like the wildtype at an early log stage of growth. Our results are consistent with distinct roles for CetZ1 in rod shape formation and at the poles of mature rods, that are positioned through the action of the MinD proteins and contribute to the development of swimming motility in multiple ways. They represent the first report of MinD proteins controlling the positioning of tubulin superfamily proteins in archaea.

KEYWORDS

cytoskeleton, motility, protein localization, tubulin superfamily, halophile, archaea

Introduction

Spatial and temporal regulation of protein subcellular localization is critical for the proper function of all cells. In bacteria, there are several different classes of proteins from the ParA/MinD superfamily (Leipe et al., 2002), which coordinate positioning of flagella (Schuhmacher et al., 2015a,b), chemosensory arrays (Ringgaard et al., 2011; Roberts et al., 2012), chromosomes (Jalal and Le, 2020), plasmids (Baxter and Funnell, 2015), cell division cytoskeletal proteins (Lutkenhaus, 2007; Hu and Lutkenhaus, 2001), and various other molecules or structures (Lutkenhaus, 2012; Perez-Cheeks et al., 2012; Laloux and Jacobs-Wagner, 2014; Thompson et al., 2006; Savage et al., 2010). A key characteristic of ParA/MinD superfamily proteins is their ATPase activity (Loose et al., 2011; Howard and Kruse, 2005) and dimerization (Hu and Lutkenhaus, 2001; Dunham et al., 2009) involving amino acid residues within a deviant Walker A motif (Leipe et al., 2002; Lutkenhaus, 2012). In MinD proteins, a subset of proteins within the ParA/MinD superfamily, this ATP binding and hydrolysis drives intracellular Min system dynamics that can generate concentration gradients of Min proteins

(Baxter and Funnell, 2015; Loose et al., 2011; Howard and Kruse, 2005). In bacteria, these gradients are critical for the correct positioning of FtsZ, a key cell division protein and tubulin superfamily homolog (Bi et al., 1991; Erickson, 1995; Nogales et al., 1998), at mid-cell to help ensure symmetric division of the mother cell into two daughter cells (Hu and Lutkenhaus, 2001; Raskin and De Boer, 1999; De Boer et al., 1989).

The Min system of *Escherichia coli* is the prototypical example of Min system dynamics and function in bacteria. MinD oscillates between poles of the rod-shaped cells, controlled by its regulator MinE, and carries MinC an FtsZ antagonist. Together, these proteins generate dynamic concentration gradients of MinCD which is on average highest at cell poles and lowest at mid-cell, preventing Z-ring formation anywhere other than at mid-cell (Lutkenhaus, 2007). MinE accelerates the ATPase activity of MinD (Loose et al., 2011) and drives its release from the membrane which it is bound to via a membrane targeting sequence (Lackner et al., 2003; Park et al., 2011; Szeto et al., 2002). Together these proteins constitute a reaction-diffusion system driven by the hydrolysis of ATP. Disruption of MinD ATPase activity blocks oscillation and function completely (Zhou et al., 2005; Lutkenhaus and Sundaramoorthy, 2003).

ParA/MinD family homologs have been identified across the archaeal domain, but MinD homologs are particularly abundant in *Euryarchaea* (Nußbaum et al., 2020). In the model halophilic archaeon, *Haloferax volcanii*, there are four MinD paralogs. The functions of MinD1-3 (HVO_0225, HVO_0595, and HVO_1634, respectively) have not been described and no clear protein partners equivalent to bacterial MinE have been identified. MinD4 (HVO_0322) has been characterized and is known to oscillate between cell poles, as well as concentrating with a cap-like appearance at the poles (Nußbaum et al., 2020). Surprisingly, the oscillation was not dependent on the ATPase Walker A or Walker B motifs, but the cap-like structures were, suggesting that the oscillation is driven by an alternative energy source (Nußbaum et al., 2020). Furthermore, although *H. volcanii* possess FtsZ homologs, none of the MinD paralogs were found to have a role in the mid-cell positioning of FtsZ; simultaneous deletion of all four paralogs had no detected effect on FtsZ positioning. Instead, deletion of *minD4* caused a motility defect and reduced the polar positioning of key protein constituents of the archaellum and chemotaxis arrays, ArlD and CheW, respectively (Nußbaum et al., 2020). Furthermore, disruption of MinD4 predicted ATPase activity through mutations to the Walker A or Walker B motifs, or truncation of its extended C-terminal tail all result in motility defects. Potential roles of the other MinD paralogs in motility of *H. volcanii* have not been reported.

Although the archaellum and chemosensory arrays are considered the core motility machinery in *H. volcanii*, there are many other contributing factors, and a range of proteins with a variety of primary functions have been shown to influence motility (Tripepi et al., 2010, 2012; Abdul Halim et al., 2013, 2020; Esquivel and Pohlschroder, 2014; Cerletti et al., 2014; Chatterjee et al., 2024; Hackley et al., 2024; Duggin et al., 2015; Schiller et al., 2024). CetZ cytoskeletal proteins, a group of tubulin superfamily proteins specific to archaea and abundant in Haloarchaea (Duggin et al., 2015; Aylett and Duggin, 2017; Brown and Duggin, 2023), have been implicated in the control of cell shape and motility (Duggin et al., 2015; de Silva et al., 2024). The most well-studied and highly conserved CetZ homolog, CetZ1, is necessary for rod-shape development and swimming motility (Duggin et al., 2015; de Silva et al., 2024). Our recent results have suggested that CetZ1 is an important factor for the proper assembly and positioning of motility proteins ArlD and CheW (Brown et al., 2024). It is not yet clear whether CetZ1 contributes to motility solely via its essential role in rod shape development (Duggin et al., 2015), or whether it has other functions that promote motility. It also is not yet clear how MinD4 and CetZ1 might coordinate their activities to regulate positioning and assembly of motility machinery.

Currently, no connection between Min proteins and FtsZ or other tubulin superfamily homologs, such as the CetZs, has been reported in archaea. Here, we investigate the potential for MinD homologs of *H. volcanii* to control positioning of CetZ1 and identify MinD2 as a strong regulator of CetZ1 localization in rod cells. This is consistent with a model in which a hierarchy or network of proteins is involved in promoting the development of motility structures at the cell poles.

Methods

Growth and culturing

Haloferax volcanii was routinely grown in Hv-Cab (de Silva et al., 2021) medium supplemented with uracil (50 μg/ml) to fulfill auxotrophic requirements if necessary, or L-Tryptophan (1 mM) to induce expression of genes under the control of the p.*tna* promotor. Liquid cultures were grown at 42°C with shaking (200 rpm).

Strain and plasmid construction

Strains, plasmids and oligonucleotides used in this study are listed in Supplementary Tables S1–S3, respectively. *H. volcanii* H26 (Δ*pyrE2* ΔpHV2) was used as the wildtype parent strain for construction of mutant strains. All strains investigated in experiments contained a pHV2-based vector (pTA962 vector or derivatives) to complement the *pyrE2* auxotrophic marker and better match the original *H. volcanii* wild-type genotype; previous work has shown that the base strains with Δ*pyrE2*, Δ*hdrB* or ΔpHV2 show poorer expression of the cell morphology and motility phenotypes, which are sensitive to nutrient and culture conditions (Duggin et al., 2015; de Silva et al., 2021; Patro et al., 2023).

To create chromosomal point mutations in *minD2*, generating strains ID810 (*minD2WA**) and ID811 (*minD2WB**), H26 was transformed with vectors pTA131_minD2WA or pTA131_minD2WB, respectively, using previously described transformation methods (Cline et al., 1989) and the homologous recombination-based "pop-in pop-out" method (Allers et al., 2004) was used to replace the native *minD2* gene with Walker A (K16A, AAG to GCC) or Walker B (D117A, GAT to GCC) mutant alleles. Primers MinD2_USflank_extF (forward) and MinD2_DSflank_extR (reverse), designed to bind externally to the

regions of homologous recombination, were then used to amplify the *minD2* locus. The resultant PCR products were purified, and then sequenced by Sanger sequencing (Australian Genome Research Facility) to identify transformants carrying the expected mutations. All strains using *minD2WA** (ID810) as the parent strain were also verified by checking for complete conversion of alleles and the absence of potential suppressor mutations (SNPs, indels) using Illumina whole-genome sequencing with ~50X coverage.

Plasmids for the expression of *minD2*, *minD2WA**, and *minD2WB** were generated by amplifying these open reading frames from *H. volcanii* H26 (ID621), ID810, and ID811, respectively. Forward and reverse primers for this reaction incorporated NdeI and BamHI restriction sites at 5′ and 3′ ends, respectively. ORFs were then ligated between the same restriction sites in pTA962 (Allers et al., 2010), generating pHJB63-65 which were sequence verified.

pHJB72 for the dual expression of *cetZ1-G-mTq2* and *minD2* was generated by amplifying the *minD2* open reading frame from pHJB63, incorporating an XbaI site at the 5′ end. The amplicon was then ligated between NheI and NotI of pHVID135. Plasmids and oligonucleotides used in this study are listed in Supplementary Tables S2, S3, respectively.

Soft-agar motility assays

Soft-agar motility assays were carried out as described in Brown et al. (2024), using Hv-Cab with 0.25% (w/v) agar and 1 mM L-Tryptophan. Motility was quantified by averaging perpendicular measurements of the motility halo diameter after 3 days of incubation of the plates inside a plastic bag at 42°C.

Light microscopy

To prepare samples for microscopy, single colonies were first used to inoculate 3 ml Hv-Cab medium, each representing a single culture (biological) replicate. These cultures were allowed to grow to OD_{600} ~0.6. Then, 3 µl of this culture was taken to inoculate 10 ml Hv-Cab supplemented with 1 mM L-Tryptophan and grown for 18 h (OD_{600} ~0.1). A 2 µl droplet of cells was spotted onto a 1.5% agarose pad containing 18% BSW as described in Duggin et al. (2015). Samples were then covered with a clean #1.5 glass coverslip.

Imaging for phase-contrast and fluorescence microscopy was carried out using a V3 DeltaVision Elite (GE Healthcare) with an Olympus UPLSAPO 100X NA 1.4 objective and pco.edge 5.5 sCMOS camera. A filter set for mTurquoise2/CFP (excitation = 400–453, emission = 463–487) was used for imaging of CetZ1-mTq2 expressing strains. 3D Structured-illumination microscopy (3D-SIM) was conducted as described previously (de Silva et al., 2024) using a DeltaVision OMX SR microscope (GE Healthcare) and a DAPI filter for excitation (405 nm) and an AF488 filter for emission (504–552 nm), suitable for imaging of mTurquoise2. Raw images were reconstructed using SoftWorx (Applied Precision, GE Healthcare), and 3D renders were generated (Schmid et al., 2019) using Bitplane IMARIS (v9.6.0).

Image analysis

Image analysis was conducted using the MicrobeJ (v5.13I) (Ducret et al., 2016) plugin for FIJI (v2.14.0/1.54f) (Schindelin et al., 2012). Cell shape characteristics were analyzed using phase-contrast images and default settings. Fluorescent foci detection was carried out after background subtraction with a ball size of 50 pixels, and the *point* detection method was used for analysis of CetZ1-mTq2 localization (Tolerance = 200).

Fluorescence intensity profiles along the long axis of each cell were generated using the *XStatProfile* function in MicrobeJ (the long axis was normalized between 0 and 1). The mean and standard deviation of the raw fluorescence intensity over all cells were then determined (Stat = mean ± stdv, *Y*-axis = MEDIAL.intensity.ch2, number of bins = 50). The intensity data were also normalized (0–100%) and the median fluorescence intensity along the long axis determined for each culture replicate (Stat = median, *Y*-axis = MEDIAL.intensity.ch2, number of bins = 50), and then the mean and standard error of all culture replicates were plotted. These data were also displayed as the difference in fluorescence intensity between *minD* deletion or Walker mutant backgrounds and the wildtype background.

Heatmaps of foci localization were generated using pooled data from all culture replicates and the *XYCellDensity* function (Rendering = Density) with default settings.

Results

Deletion of *minD2* or *minD4* results in motility defects with normal rod cell shape

We began by assessing the effect of *H. volcanii minD* gene deletions on motility, focussing on MinD2 and MinD4, which are expressed significantly in standard culture conditions (Babski et al., 2016; Cerletti et al., 2018; Laass et al., 2019). MinD4 has previously been studied in this context (Nußbaum et al., 2020). The *minD* deletion strains grew at rates like the wildtype in starter cultures, and motility was assessed using soft-agar motility assays in which cells develop into a motile rod form (Duggin et al., 2015) and then swim out through the gel from a central inoculation point, forming a halo.

Deletion of *minD2* resulted in a motility defect (~25% of the halo diameter compared to wildtype), which was significantly stronger than that caused by deletion of *minD4* (~50% of wildtype motility; Figures 1A, B). Deletion of both *minD2* and *minD4* resulted in a stronger motility defect (~16% of wildtype motility). The deletion of all four *minD* homologs (Δ*minD1/2/3/4*) was indistinguishable to that of the Δ*minD2/4*.

H. volcanii is pleomorphic and cultures can show a range of cell shapes including irregular disk, or plate-shaped cells, and rods with varying widths and degrees of elongation (Duggin et al., 2015;

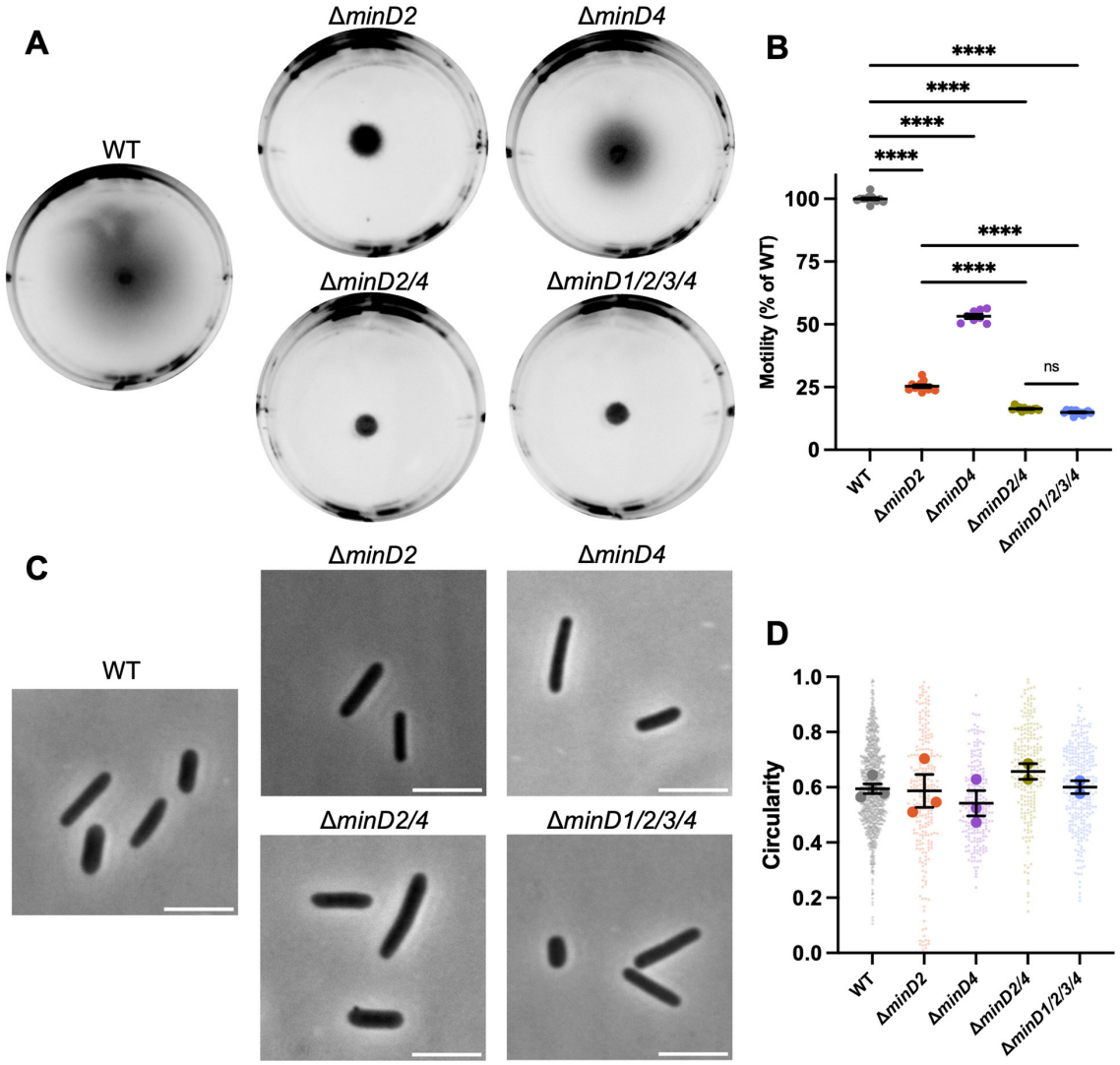


FIGURE 1
MinD2 and MinD4 have roles in motility without influencing shape. **(A)** H26 wildtype and *minD* deletion mutant backgrounds, carrying an empty pTA962 vector, were used to inoculate the center of a soft agar plate. **(B)** The diameter of their motility halos was measured across six independent liquid starter culture replicates (represented by individual points), each determined from the mean of at least two replicate agar plates. Mean and standard error is shown. **(C)** Phase-contrast microscopy of cells grown in Hv-Cab liquid medium with 1 mM L-tryptophan, sampled at OD_{600} ~0.1. Scale bars = 5 μm. **(D)** Cell circularity distributions, quantified from phase-contrast images as a proxy for cell elongation. Large data points represent mean circularity of one culture replicate, small data points represent individual cells from all culture replicates for the indicated strain. Mean and standard error of culture replicate means are shown (bars). The number of individual cells measured from pooled culture replicates were as follows: WT + pTA962, n = 996; Δ*minD2* + pTA962, n = 247; Δ*minD4* + pTA962, n = 217; Δ*minD2/4* + pTA962, n = 223; Δ*minD1/2/3/4* + pTA962, n = 279. In **(B, D)**, one-way ANOVA was used as a statistical test, and only significant and/or relevant comparisons are shown. ****p < 0.00005, ns, not significant.

de Silva et al., 2021). Furthermore, a motile-rod cell type with regular cell widths develops under certain conditions to promote swimming (Duggin et al., 2015; Li et al., 2019). To determine whether the two MinD proteins are required for rod formation, which might contribute to their effects on motility, we sampled wild-type and Δ*minD2/4* cells from the early-log phase of liquid batch cultures (OD ~0.1), where motile rods are prevalent (de Silva et al., 2021; Li et al., 2019). By quantifying cell circularity as an indicator of the extent of cell elongation, we observed no significant cell shape differences at this stage of growth in any of the strains (Figures 1C, D).

MinD2 and MinD4 influence the cellular positioning of CetZ1

MinD4 has been shown to have a role in the polar positioning of the archaellum and chemotaxis arrays through the use of tagged marker proteins ArlD and CheW, respectively (Nußbaum et al., 2020). We recently found that CetZ1 also influences the proper assembly and positioning of ArlD and CheW (Brown et al., 2024). It is unknown whether MinD4 and CetZ1 cooperate or work independently to promote motility. To begin investigating this, we sought to determine whether

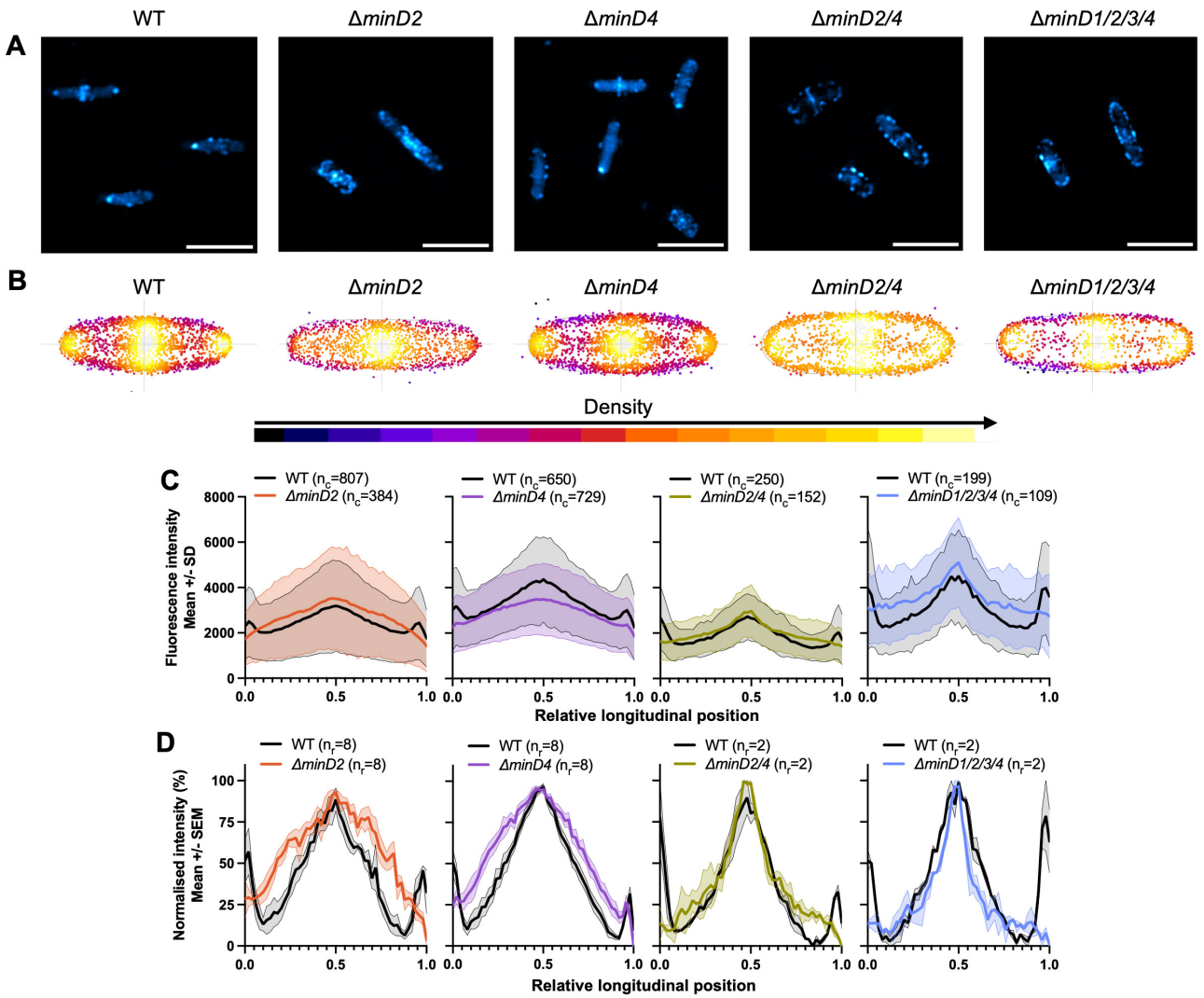


FIGURE 2
MinD2 controls polar positioning of CetZ1. H26 wildtype and *minD* deletion strains containing pHVID135 (for expression of CetZ1-mTq2) were grown in Hv-Cab with 1 mM L-Tryptophan and sampled at an approximate OD_{600} of 0.1. **(A)** Epifluorescence images of CetZ1-mTq2 localization in the indicated *minD* knockout strains (scale bar 5 μm). **(B)** Heatmaps representing the position and number (heatmap density) of detected CetZ1-mTq2 fluorescence peak intensities (foci) along the long (vertical) and lateral (horizontal) axes of all cells combined. **(C)** The median fluorescence intensity of CetZ1-mTq2 along the normalized long axis of each cell was used to generate a plot of the mean (bold line) and standard deviation (shading) of the raw fluorescence intensity data from all cells (n_c; combined from at least two replicate cultures). Cultures of the wildtype strain were repeated in parallel alongside each mutant independently. **(D)** The same intensity data were also normalized (0–100%) for each culture replicate and plotted as the mean (bold line) and standard error (shading) of all culture replicates (n_r) to enable comparison of protein distributions amongst the various strains.

the absence of MinD2 or MinD4 affects the localization of CetZ1.

To do this, we observed the localization of a CetZ1-mTurquoise2 (mTq2) fusion protein (Ithurbide et al., 2024) in the wildtype and various *minD* deletion strain backgrounds. Fluorescence microscopy of CetZ1-mTq2 in wildtype cells imaged in early-log growth (OD_{600} ~0.1; Figure 2A), showed the expected patchy localization with relatively high fluorescence at cell poles and the mid-cell region (de Silva et al., 2024; Brown et al., 2024; Ithurbide et al., 2024). To characterize these patterns, we analyzed the whole population of cells, in the wildtype and *minD* mutant backgrounds, by generating heatmaps that represent the position and frequency of fluorescent foci within a normalized cell shape (reflecting the averaged aspect ratio; Figure 2B). We also generated plots of the total and normalized mean fluorescence intensity of CetZ1-mTq2 along the long-axis of the cells (Figures 2C, D). In the Δ*minD2* background, CetZ1-mTq2 localization remained patchy and concentrated in a broad zone near midcell, which was noticeably broader compared to the wildtype, and the high-intensity polar localisations of CetZ1-mTq2 were not observed (e.g., Figures 2A, C, D). These patterns were also reflected in the number of fluorescent foci detected along the cell lengths (Figure 2B); the wildtype background displayed clear peaks of fluorescence around midcell and at the cell poles, but in the Δ*minD2* background the midcell foci were more spread out along the cell length, and showed almost complete loss of the polar clusters of foci.

To confirm whether the changes to CetZ1-mTq2 localization patterns were resultant from the loss of *minD2* specifically, we observed CetZ1-mTq2 localization when *minD2* deletion was complemented by expression of *minD2* from a plasmid containing the open reading frames for *cetZ1-mTq2* and *minD2* in tandem, under the control of the inducible p.*tna* promotor (Supplementary Figure S1). Dual expression of these genes in the Δ*minD2* background using 1 mM L-Tryptophan did not clearly restore the CetZ1-mTq2 localization defect (Supplementary Figures S1a, b). However, when 2 mM L-Tryptophan was used, the localization of CetZ1-mTq2 was partially restored (Supplementary Figures S1c–g), confirming that *minD2* regulates the localization of CetZ1. The requirement for the additional tryptophan to observe partial complementation is likely due to a sensitivity of the phenotype to the expression level of *minD2*, which is likely to be influenced by its second position in the tandem expression plasmid (Ithurbide et al., 2024).

A weaker defect of CetZ1-mTq2 polar localization was also observed in the Δ*minD4* background (compare the polar intensities between mutant and wildtype in Figure 2D). Additional control experiments demonstrated that the Δ*minD* strains producing CetZ1-mTq2 had no significant differences in shape compared to the wildtype control (OD ~0.1), and we also observed an equivalent motility defect with or without the CetZ1-mTq2 fusion (Supplementary Figures S2a, b). The results in Figure 2 also revealed that the zone of CetZ1-mTq2 localization around the midcell region was broadened in the absence of either MinD2 or MinD4 (Figures 2B, D). Interestingly, such broadening was not apparent in strains in which both *minD2* and *minD4* were deleted, although the polar localization of CetZ1-mTq2 was still strongly inhibited (Figures 2B, D). This suggests that the midcell zone of CetZ1-mTq2 localization is impacted when only one of MinD2 or MinD4 is present and implies a genetic interaction or interference between *minD2* and *minD4* that influences midcell CetZ1.

Structural characteristics of CetZ1-mTq2 localization in individual cells lacking *minD2* or *minD4*

We further resolved structural characteristics of the CetZ1-mTq2 polar localizations and the effects of the *minD* knock-outs, using three-dimensional structured-illumination microscopy (3D-SIM). Multiple cells are shown as maximum intensity projections in Figures 3A–C to visualize the effect of the mutations on CetZ1 structures in individual cells and facilitate comparisons to the aggregate data shown in Figures 2B–D.

In wildtype early-log cells, CetZ1-mTq2 often appeared as a bright focus, doublet, multi-lobed, or cap-like structure at the cell poles, suggestive of a ring or disc-like structure at the tip of the cell pole (Figures 3A, D). Patches and filaments were also observed along the cell edges in the midcell zone. In the absence of MinD2, the foci of CetZ1-mTq2 appeared more randomly distributed around the cells, with no clear preference for the cell poles or the midcell zone (Figure 3B). In the absence of MinD4, this effect was weaker, and some of the cells still appeared to show the relatively bright polar tip structures of CetZ1-mTq2 (Figure 3C). The Δ*minD2* and Δ*minD4* backgrounds also showed noticeably more of the CetZ1-mTq2 patches or foci throughout the cell area compared to the wildtype (Figures 3A–C). By inspecting the slice series of images to resolve the 3D detail (Figure 3E, Supplementary Video 1), CetZ1-mTq2 foci appeared noticeably more prevalent at the top and bottom surfaces along the length of the *minD* mutant cells, compared to the wildtype, and potentially in the cytoplasm too (Supplementary Videos 2–4). These observations are consistent with the broadening of the midcell zone and reduced polar positioning of CetZ1-mTq2 in the *minD* mutants seen in the aggregated epifluorescence data (Figures 2B–D).

ATPase active-site residues of MinD2 are required for normal swimming motility and polar localization of CetZ1

MinD proteins typically rely on ATP hydrolysis to drive their dynamic localization patterns, such as oscillation between cell poles, generating concentration gradients which form the basis for their spatial and temporal control of downstream proteins (Lutkenhaus, 2012; Lutkenhaus and Sundaramoorthy, 2003; Nußbaum et al., 2020). When the Walker A and Walker B motifs of *H. volcanii* MinD4 were previously mutated to disrupt ATP binding and hydrolysis, swimming motility was reduced and polar localization of MinD4 was disrupted, although not its oscillation (Nußbaum et al., 2020).

Given that MinD2 had the strongest effects on swimming motility and CetZ1 localization, we chose to further investigate the importance of the ATPase active-site for MinD2 functionality in motility and CetZ1 localization. We generated chromosomal point-mutations in the endogenous *minD2* gene in the ATPase active site via the deviant Walker A (K16A) and Walker B (D117A) motifs, here forth referred to as *minD2WA** and *minD2WB**, respectively. These mutations are equivalent to those previously made in *minD4* (Nußbaum et al., 2020).

Both *minD2WA** and *minD2WB** strains (containing the empty vector pTA962) partially perturbed motility, resulting in ~78 and ~36% of the wildtype motility halo diameter, respectively (Figures 4A, C), compared to the complete deletion of *minD2* (~25%, Figures 1A, B). The motility defect of Δ*minD2* was complemented by expressing *minD2* from a plasmid (pTA962 backbone), confirming the important role of *minD2* in motility (Figures 4B, C).

When *minD2WA** was expressed from a plasmid in the Δ*minD2* background, this strain had a more pronounced motility defect (~35%) compared to the chromosomal mutant *minD2WA** (~78%). This might be related to different expression levels of *minD2WA** in the plasmid with p.*tna* compared to its native chromosomal site. Expression of *minD2WB** from pTA962 in Δ*minD2* produced equivalent motility to the complete *minD2* deletion, consistent with its much stronger effects on motility observed with the chromosomal *minD2WB** mutant (Figures 4B, C).

Phase-contrast microscopy and analysis of cell shape (circularity) confirmed that, like Δ*minD2*, the *minD2WA**

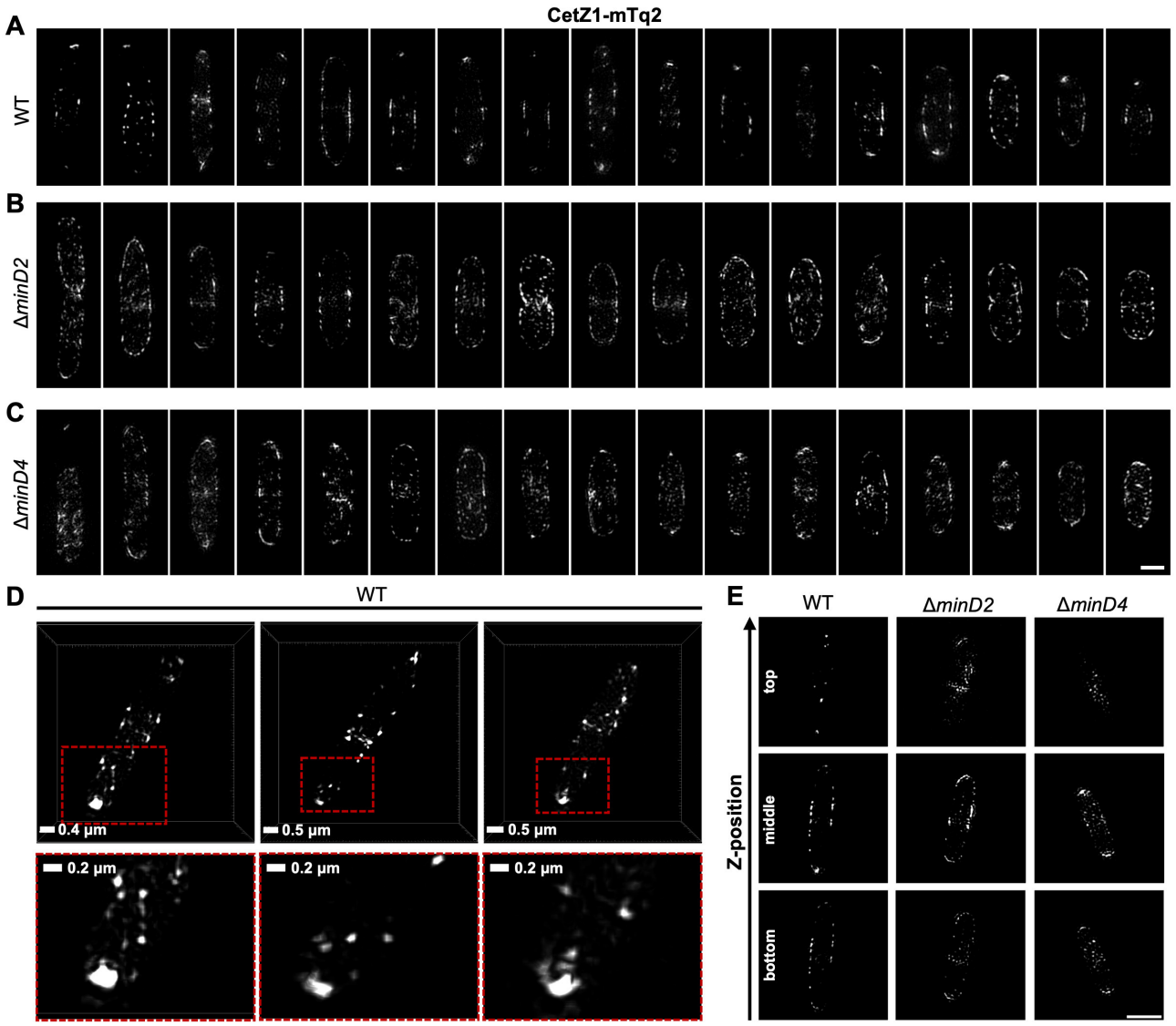


FIGURE 3
Subcellular structures of CetZ1-mTq2 in wildtype and *minD* mutant backgrounds, by 3D-SIM imaging. **(A–C)** Maximum intensity projections of CetZ1-mTq2 in H26 wildtype, Δ*minD2*, and Δ*minD4* backgrounds, viewed along the imaging *z*-axis (i.e., *XY*-views). Scale bar in panel **(C)** is 1 μm and applies to panels **(A–C)**. **(D)** *XY*-views of 3D rendered images, showing non-circular fluorescence foci of CetZ1-mTq2 at the cell poles in H26 wildtype background. Red boxes indicate zoomed area (bottom). **(E)** Examples of 3D-SIM slice sections of CetZ1-mTq2 in the indicated strains at approximate bottom, middle, and top *z*-positions of the cells (scale bar is 2 μm and applies to all images).

and *minD2WB** mutations did not significantly affect rod shape (at OD ~0.1), whether expressed chromosomally (Supplementary Figures S3a, c) or from the plasmid (Supplementary Figures S3b, d).

We then assessed the effect of the *minD2WA** and *minD2WB** chromosomal mutants on CetZ1-mTq2 localization. Control experiments showed that expression of CetZ1-mTq2 did not impact cell circularity or motility of either the *minD2WA** or *minD2WB** strains (Supplementary Figures S3c, e). Fluorescence microscopy (Figure 5A) showed that both Walker mutations resulted in a substantial loss of CetZ1-mTq2 fluorescence intensity (Figure 5C) and foci localization at cell poles and showed the midcell zone broadening (Figure 5B) like in the Δ*minD2* background (Figure 2).

Discussion

MinD proteins and related protein families are important for spatial and temporal positioning of a variety of molecules but have mainly been characterized in bacteria (Lutkenhaus, 2012). One of four MinD paralogs in the archaeal model organism *H. volcanii*, MinD4, has previously been described to affect motility via regulating the positioning of archaellum and chemotaxis arrays, via unknown mechanisms (Nußbaum et al., 2020). The tubulin-like cytoskeletal protein CetZ1 is another factor promoting the proper assembly and positioning of these proteins at the cell poles (Brown et al., 2024). CetZ1 is needed for rod cell shape formation (Duggin et al., 2015), which may contribute to motility in several ways, and shows complex subcellular localization along the cell body and then

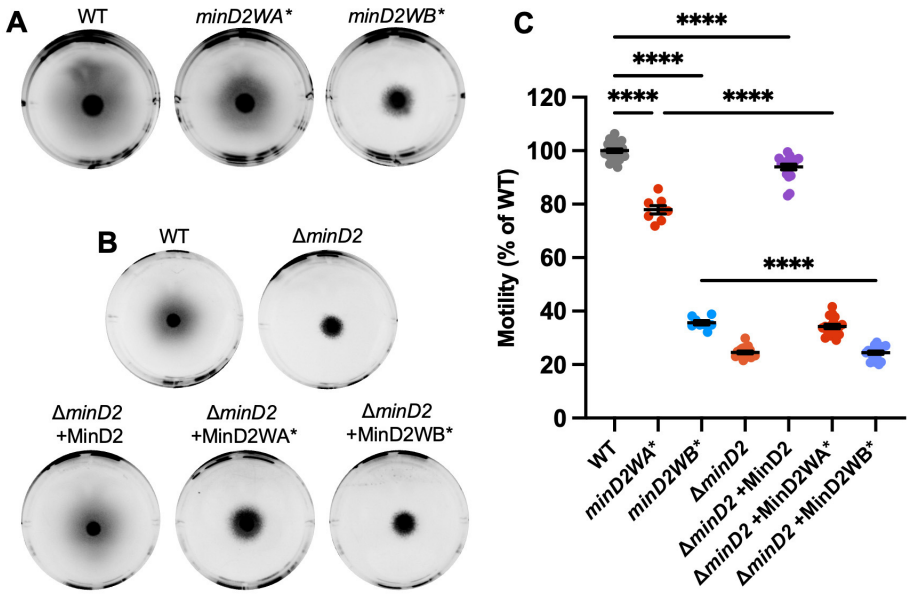


FIGURE 4
Chromosomal Walker mutants of *minD2* perturb motility. **(A)** H26 wildtype, *minD2WA**, and *minD2WB** backgrounds containing pTA962, **(B)** and Δ*minD2* containing pTA962 were compared to Δ*minD2* containing pHJB63-65 for expression of *minD2*, *minD2WA**, and *minD2WB**, respectively, and inoculated onto Hv-Cab soft agar (0.25 % w/v) supplemented with 1 mM L-Tryptophan. **(C)** Halo diameters were quantified across four starter culture replicates with two technical replicates (motility agar plates) each (indicated by individual points). Mean and standard error are shown. One-way ANOVA was used as a statistical test, and only significant and relevant comparisons are shown. ****$p < 0.00005$.

at the cell poles of motile rods, where it is hypothesized to form a scaffold-like structure, though its mechanism of action at these sites is unclear.

We found here that MinD2 is important for motility and for the subcellular positioning of CetZ1. The potential relationship between these two functions is yet to be elucidated. Deletion of *minD2* was found to have a more severe motility defect than deletion of *minD4*, and deletion of both genes resulted in an even greater motility defect, suggesting distinct roles or additive contributions of these two MinD paralogs in motility. MinD2 and to a lesser extent MinD4 also influenced the subcellular positioning of CetZ1 in the midcell region and at the poles, in a manner that was independent of cell shape. CetZ1 has been implicated in distinct roles in establishing rod-shape during motile cell differentiation, and then at the poles of the mature motile rods (Duggin et al., 2015; Brown et al., 2024). Our results suggest that the effects of MinD2 on CetZ1 localization do not have an essential role in maintaining rod shape, because Δ*minD2* cells could still form normal shaped rods at an early-mid log stage of culture.

We found that mutations in the Walker A and B motifs of the MinD2 ATPase both strongly influenced CetZ1 positioning along the cell length and poles, like the *minD2* deletion. However, only the Walker B mutant (D117A) strongly perturbed motility, whereas the Walker A (K16A) mutant (in its chromosomal context) resulted in only a mild motility defect. This suggests that MinD2 has an impact on motility that is largely independent of the K16 residue and CetZ1 positioning. The K16 residue is expected to be involved in nucleotide dependent dimerization (Lutkenhaus and Sundaramoorthy, 2003), so this function may not be critical to the overall motility function of MinD2, and instead plays a role in CetZ1 positioning for an as yet unresolved purpose. Notably, the equivalent MinD4 Walker A and B mutants reduced motility equivalently to *minD4* deletion (Nußbaum et al., 2020). The MinD2 and MinD4 proteins could therefore have multiple, differing functions that contribute to the proper development of motile *H. volcanii* cells.

The accompanying study by Patro et al. (2024) found that MinD2 is required for normal motility and the proper polar localization of the archaellum and chemotaxis proteins, ArlD and CheW, like previously seen with MinD4 (Nußbaum et al., 2020). They also reported a difference in cell shape during the development of rod cells in the Δ*minD2* compared to the wildtype. This likely represents an earlier phase of *H. volcanii* cultures, where cells are actively shapeshifting into motile rods (de Silva et al., 2021; Li et al., 2019), compared to the current study where rods had already formed, and we found that their cell shapes in the wildtype and Δ*minD2* strains were indistinguishable. Though we note that since *H. volcanii* morphogenesis is highly sensitive to exact media composition and culture conditions (de Silva et al., 2021; Patro et al., 2023), excessive comparisons of the exact magnitudes or timing of events from separate studies should be avoided. Nevertheless, if the rate or timing of rod formation is influenced by MinD2 (Patro et al., 2024), this may affect another developmental stage of motility, such as the placement of Arl/Che proteins at the poles. However, it is still unclear whether any effects on cell shape development directly contribute to the *minD2*-dependent reduction of motility.

Considering the main findings from this and other recent studies (Nußbaum et al., 2020; Patro et al., 2024), we can

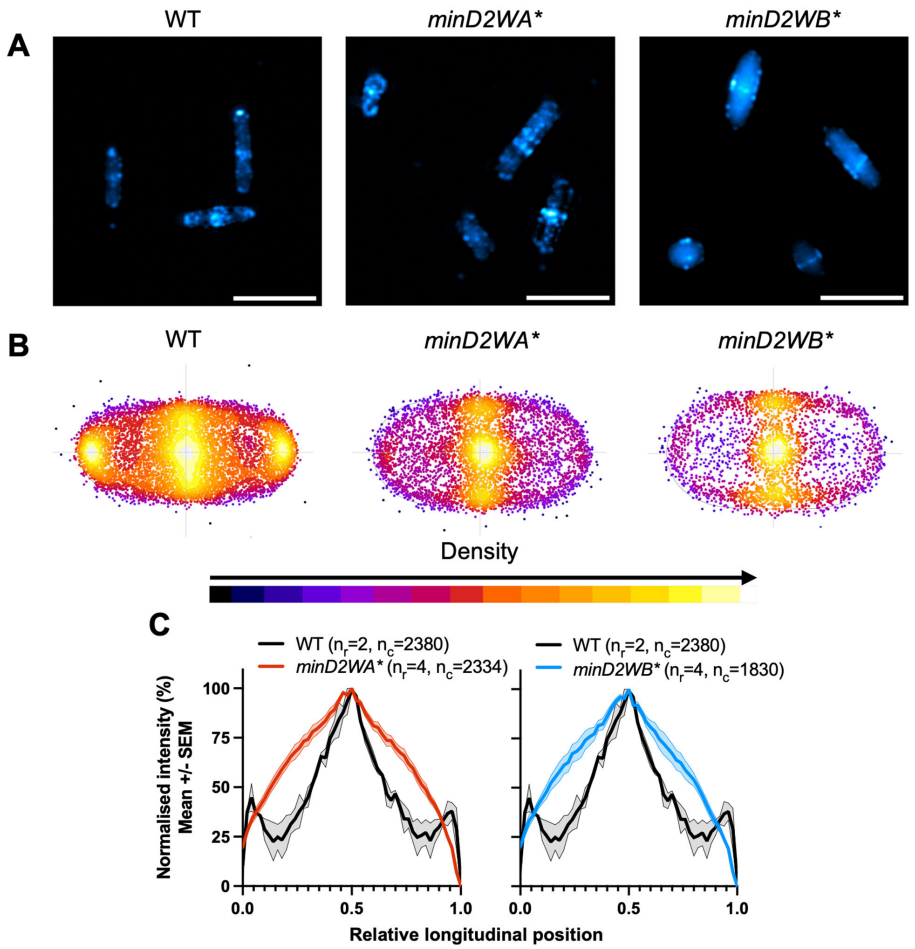


FIGURE 5
Chromosomal Walker mutants of *minD2* disrupt polar localization of CetZ1. H26 wildtype and *minD2* walker mutant strains containing pHVID135 (for expression of CetZ1-mTq2) were grown in Hv-Cab with 1 mM L-Tryptophan and sampled at OD_{600} ~0.1. **(A)** Cells were imaged using epifluorescence to observe CetZ1-mTq2 localization. Scale bar 5 μm. **(B)** Heat map representations of detected foci of CetZ1-mTq2 and their localisations. Heatmaps are colored by density, referring to subcellular localization along the relative length of lateral and longitudinal axes only and do not fluorescence intensity. **(C)** Normalized median fluorescence intensity of CetZ1-mTq2 along the long axis of the cell, represented as the mean (bold line) and standard deviation (shading) of biological replicates (as in Figure 2D). Two biological replicates were carried out for the wildtype background, and four for each of *minD2WA** and *minD2WB**, and data for H26 wildtype is included on both graphs for reference. The number of culture replicates (n_r) and number of cells (n_c) measured from pooled replicates are indicated in the key of panel **(C)**.

summarize several possibilities for the role of MinD proteins in motility and CetZ1 localization (Figure 6). In this model, MinD2 and MinD4 help direct the archaellum (Arl) and chemotaxis (Che) proteins to the cell poles directly or via positioning and assembly of polar CetZ1. We recently found that CetZ1 plays a role in the proper placement of the motility machinery (Brown et al., 2024). However, the present finding that the chromosomal *minD2WA** mutation blocked polar CetZ1 localization but only weakly affected motility suggests that MinD proteins might separately control both the position of CetZ1 and another aspect of motility, such as placement of the motility machinery at the cell poles, independently of MinD2 regulated CetZ1 positioning.

Previous work showed that MinD4 forms dynamic cell cap-like structures in an ATPase-dependent manner, and yet oscillates in the cells independently of ATPase function (Figure 6), suggesting that it interacts with another oscillator that drives the movement (Nußbaum et al., 2020). The localization patterns of MinD2 are currently unknown due to the absence of functional fluorescently tagged MinD2 (Patro et al., 2024). However, our observations that the complex patterns of CetZ1 localization are affected by MinD2 at multiple sites and at a cellular scale would suggest that similar cellular level MinD2 pattern formation might exist.

In summary, our findings have showed that multiple MinD homologs can have differing functions and regulate the subcellular positioning of the archaeal tubulin-like protein CetZ1 in the context of the development of swimming motility. They represent the first demonstration of MinD proteins controlling the position of tubulin superfamily proteins in archaea. The MinD-controlled positioning of CetZ1 in distinct subcellular locations is consistent with multiple functions of CetZ1, evoking parallels with the multiple functions of tubulins and other eukaryotic cytoskeletal proteins that are extensively regulated in time and space.

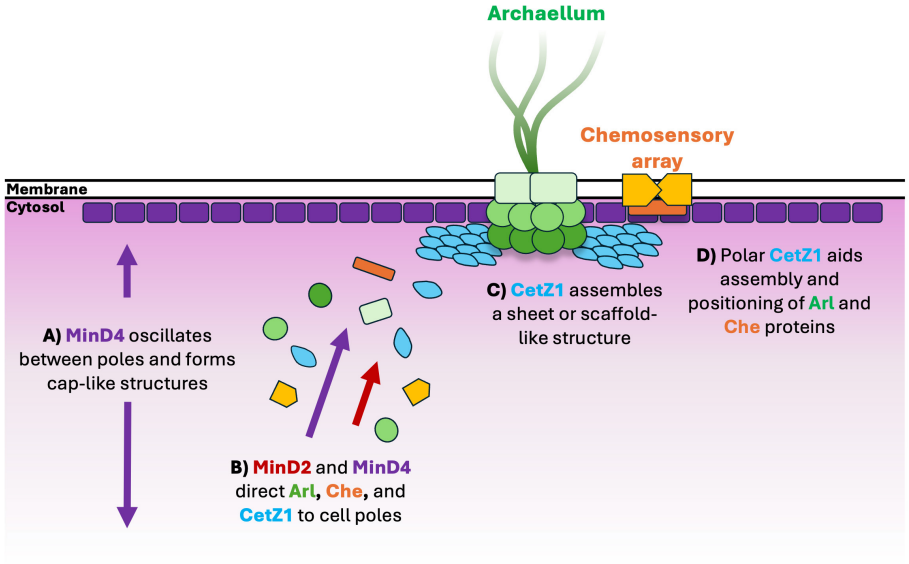


FIGURE 6
Schematic model for assembly of motility proteins at cell poles. **(A)** MinD4 oscillates between cell poles and forms cap-like localisations at cell poles in an ATPase dependent manner (Nußbaum et al., 2020). **(B)** MinD4 (Nußbaum et al., 2020) and MinD2 direct protein constituents of the archaellum ("Arl") and chemosensory arrays ("Che") such as ArlD and CheW, respectively, to cell poles (Patro et al., 2024). CetZ1 is also directed to cell poles, likely by both MinD paralogs, but more dominantly by MinD2. **(C)** At cell poles, CetZ1 is hypothesized to assemble as a sheet or disc-like structure, **(D)** where it may contribute to assembly and positioning of motility structures via an unknown mechanism (Brown et al., 2024).

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

HB: Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing – original draft, Writing – review & editing. ID: Conceptualization, Funding acquisition, Methodology, Project administration, Supervision, Validation, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. The work was supported by the Australian Research Council (DP160101076 and FT160100010 to ID).

Acknowledgments

The authors would like to acknowledge facilities support from L. Cole and A. Bottomley and the UTS Microbial Imaging Facility.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2024.1474697/full#supplementary-material

References

Abdul Halim, M. F., Pfeiffer, F., Zou, J., Frisch, A., Haft, D., Wu, S., et al. (2013). *Haloferax volcanii* archaeosortase is required for motility, mating, and C-terminal processing of the S-layer glycoprotein. *Mol. Microbiol.* 88, 1164–1175. doi: 10.1111/mmi.12248

Abdul Halim, M. F., Schulze, S., DiLucido, A., Pfeiffer, F., Bisson Filho, A. W., Pohlschroder, M., et al. (2020). Lipid anchoring of archaeosortase substrates and midcell growth in haloarchaea. *MBio* 11:e00349-20. doi: 10.1128/mBio.00349-20

Allers, T., Barak, S., Liddell, S., Wardell, K., and Mevarech, M. (2010). Improved strains and plasmid vectors for conditional overexpression of His-tagged proteins in *Haloferax volcanii*. *Appl. Environ. Microbiol.* 76, 1759–1769. doi: 10.1128/AEM.02670-09

Allers, T., Ngo, H.-P., Mevarech, M., and Lloyd, R. G. (2004). Development of additional selectable markers for the halophilic archaeon *Haloferax volcanii* based on the leuB and trpA genes. *Appl. Environ. Microbiol.* 70, 943–953. doi: 10.1128/AEM.70.2.943-953.2004

Aylett, C. H., and Duggin, I. G. (2017). "The tubulin superfamily in archaea," in *Prokaryotic Cytoskeletons*, eds. J. Löwe, and L. Amos (Cham: Springer), 393–417. doi: 10.1007/978-3-319-53047-5_14

Babski, J., Haas, K. A., Nather-Schindler, D., Pfeiffer, F., Forstner, K. U., Hammelmann, M., et al. (2016). Genome-wide identification of transcriptional start sites in the haloarchaeon *Haloferax volcanii* based on differential RNA-Seq (dRNA-Seq). *BMC Genomics* 17:629. doi: 10.1186/s12864-016-2920-y

Baxter, J. C., and Funnell, B. E. (2015). "Plasmid partition mechanisms," in *Plasmids: Biology and Impact in Biotechnology and Discovery*, eds. M. E. Tolmasky, and J. C. Alonso (Hoboken, NJ: Wiley), 133–155. doi: 10.1128/9781555818982.ch8

Bi, E., Dai, K., Subbarao, S., Beall, B., and Lutkenhaus, J. (1991). FtsZ and cell division. *Res. Microbiol.* 142, 249–252. doi: 10.1016/0923-2508(91)90037-B

Brown, H. J., and Duggin, I. G. (2023). Diversity and potential multifunctionality of archaeal CetZ Tubulin-like cytoskeletal proteins. *Biomolecules* 13:134. doi: 10.3390/biom13010134

Brown, H. J., Islam, M. I., Ruan, J., Baker, M. A., Ithurbide, S., Duggin, I. G., et al. (2024). CetZ1-dependent polar assembly of the motility machinery in haloarchaea. *bioRxiv* 2024.05.02.592137. doi: 10.1101/2024.05.02.592137

Cerletti, M., Gimenez, M. I., Da Troetschel, C., D' Alessandro, C., Poetsch, A., De Castro, R. E., et al. (2018). Proteomic study of the exponential-stationary growth phase transition in the Haloarchaea *Natrialba magadii* and *Haloferax volcanii*. *Proteomics* 18:e1800116. doi: 10.1002/pmic.201800116

Cerletti, M., Martínez, M. J., Giménez, M. I., Sastre, D. E., Paggi, R. A., and De Castro, R. E. (2014). The LonB protease controls membrane lipids composition and is essential for viability in the extremophilic haloarchaeon *Haloferax volcanii*. *Environ. Microbiol.* 16, 1779–1792. doi: 10.1111/1462-2920.12385

Chatterjee, P., Garcia, M., Cote, J., Yun, K., Legerme, G., Habib, R., et al. (2024). Involvement of ArlI, ArlJ, and CirA in archaeal type-IV pilin-mediated motility regulation. *bioRxiv* 2024.03.04.583388. doi: 10.1101/2024.03.04.583388

Cline, S. W., Lam, W. L., Charlebois, R. L., Schalkwyk, L. C., and Doolittle, W. F. (1989). Transformation methods for halophilic archaebacteria. *Can. J. Microbiol.* 35, 148–152. doi: 10.1139/m89-022

De Boer, P. A., Crossley, R. E., and Rothfield, L. I. (1989). A division inhibitor and a topological specificity factor coded for by the minicell locus determine proper placement of the division septum in *E. coli*. *Cell* 56, 641–649. doi: 10.1016/0092-8674(89)90586-2

de Silva, R. T., Abdul-Halim, M. F., Pittrich, D. A., Brown, H. J., Pohlschroder, M., Duggin, I. G., et al. (2021). *I*mproved growth and morphological plasticity of *Haloferax volcanii*. *Microbiology* 167:001012. doi: 10.1099/mic.0.001012

de Silva, R. T., Shinde, V., Brown, H. J., Liao, Y., and Duggin, I. G. (2024). Dynamic self-association of archaeal tubulin-like protein CetZ1 drives *Haloferax volcanii* morphogenesis. *bioRxiv* 2024.04.08.588506. doi: 10.1101/2024.04.08.588506

Ducret, A., Quardokus, E. M., and Brun, Y. V. (2016). MicrobeJ, a tool for high throughput bacterial cell detection and quantitative analysis. *Nat. Microbiol.* 1, 1–7. doi: 10.1038/nmicrobiol.2016.77

Duggin, I. G., Aylett, C. H., Walsh, J. C., Michie, K. A., Wang, Q., Turnbull, L., et al. (2015). CetZ tubulin-like proteins control archaeal cell shape. *Nature* 519, 362. doi: 10.1038/nature13983

Dunham, T. D., Xu, W., Funnell, B. E., and Schumacher, M. A. (2009). Structural basis for ADP-mediated transcriptional regulation by P1 and P7 ParAn. *EMBO J.* 28, 1792–1802. doi: 10.1038/emboj.2009.120

Erickson, H. P. (1995). FtsZ, a prokaryotic homolog of tubulin? *Cell* 80, 367–370. doi: 10.1016/0092-8674(95)90486-7

Esquivel, R. N., and Pohlschroder, M. (2014). A conserved type IV pilin signal peptide H-domain is critical for the post-translational regulation of flagella-dependent motility. *Mol. Microbiol.* 93, 494–504. doi: 10.1111/mmi.12673

Hackley, R. K., Hwang, S., Herb, J. T., Bhanap, P., Lam, K., Vreugdenhil, A., et al. (2024). TbsP and TrmB jointly regulate gapII to influence cell development phenotypes in the archaeon *Haloferax volcanii*. *Mol. Microbiol.* 121, 742–766. doi: 10.1111/mmi.15225

Howard, M., and Kruse, K. (2005). Cellular organization by self-organization: mechanisms and models for Min protein dynamics. *J. Cell Biol.* 168:533. doi: 10.1083/jcb.200411122

Hu, Z., and Lutkenhaus, J. (2001). Topological regulation of cell division in *E. coli*: spatiotemporal oscillation of MinD requires stimulation of its ATPase by MinE and phospholipid. *Mol. Cell* 7, 1337–1343. doi: 10.1016/S1097-2765(01)00273-8

Ithurbide, S., de Silva, R. T., Brown, H. J., Shinde, V., and Duggin, I. G. (2024). A vector system for single and tandem expression of cloned genes and multi-colour fluorescent tagging in *Haloferax volcanii*. *Microbiology* 170:001461. doi: 10.1099/mic.0.001461

Jalal, A. S., and Le, T. B. (2020). Bacterial chromosome segregation by the ParABS system. *Open Biol.* 10:200097. doi: 10.1098/rsob.200097

Laass, S., Monzon, V. A., Kliemt, J., Hammelmann, M., Pfeiffer, F., Forstner, K. U., et al. (2019). Characterization of the transcriptome of *Haloferax volcanii*, grown under four different conditions, with mixed RNA-Seq. *PLoS ONE* 14:e0215986. doi: 10.1371/journal.pone.0215986

Lackner, L. L., Raskin, D. M., and De Boer, P. A. (2003). ATP-dependent interactions between *Escherichia coli* Min proteins and the phospholipid membrane *in vitro*. *J. Bacteriol.* 185, 735–749. doi: 10.1128/JB.185.3.735-749.2003

Laloux, G., and Jacobs-Wagner, C. (2014). How do bacteria localize proteins to the cell pole? *J. Cell Sci.* 127, 11–19. doi: 10.1242/jcs.138628

Leipe, D. D., Wolf, Y. I., Koonin, E. V., and Aravind, L. (2002). Classification and evolution of P-loop GTPases and related ATPases. *J. Mol. Biol.* 317, 41–72. doi: 10.1006/jmbi.2001.5378

Li, Z., Kinosita, Y., Rodriguez-Franco, M., Nußbaum, P., Braun, F., Delpech, F., et al. (2019). Positioning of the motility machinery in halophilic archaea. *MBio* 10:e00377-19. doi: 10.1128/mBio.00377-19

Loose, M., Fischer-Friedrich, E., Herold, C., Kruse, K., and Schwille, P. (2011). Min protein patterns emerge from rapid rebinding and membrane interaction of MinE. *Nat. Struct. Mol. Biol.* 18, 577–583. doi: 10.1038/nsmb.2037

Lutkenhaus, J. (2007). Assembly dynamics of the bacterial MinCDE system and spatial regulation of the Z ring. *Annu. Rev. Biochem.* 76, 539–562. doi: 10.1146/annurev.biochem.75.103004.142652

Lutkenhaus, J. (2012). The ParA/MinD family puts things in their place. *Trends Microbiol.* 20, 411–418. doi: 10.1016/j.tim.2012.05.002

Lutkenhaus, J., and Sundaramoorthy, M. (2003). MinD and role of the deviant Walker A motif, dimerization and membrane binding in oscillation. *Mol. Microbiol.* 48, 295–303. doi: 10.1046/j.1365-2958.2003.03427.x

Nogales, E., Downing, K. H., Amos, L. A., and Löwe, J. (1998). Tubulin and FtsZ form a distinct family of GTPases. *Nat. Struct. Mol. Biol.* 5:451. doi: 10.1038/nsb0698-451

Nußbaum, P., Ithurbide, S., Walsh, J. C., Patro, M., Delpech, F., Rodriguez-Franco, M., et al. (2020). An oscillating MinD protein determines the cellular positioning of the motility machinery in archaea. *Cur. Biol.* 30, 4956–4972. e4. doi: 10.1016/j.cub.2020.09.073

Park, K.-T., Wu, W., Battaile, K. P., Lovell, S., Holyoak, T., Lutkenhaus, J., et al. (2011). The Min oscillator uses MinD-dependent conformational changes in MinE to spatially regulate cytokinesis. *Cell* 146, 396–407. doi: 10.1016/j.cell.2011.06.042

Patro, M., Duggin, I. G., Albers, S. V., and Ithurbide, S. (2023). Influence of plasmids, selection markers and auxotrophic mutations on *Haloferax volcanii* cell shape plasticity. *Front. Microbiol.* 14:1270665. doi: 10.3389/fmicb.2023.1270665

Patro, M., Sivabalasarma, S., Gfrerer, S., Rodriguez-Franco, M., Nußbaum, P., Ithurbide, S., et al. (2024). MinD2 modulates cell shape and motility in the archaeon *Haloferax volcanii*. *bioRxiv* 2024.08.01.606218. doi: 10.1101/2024.08.01.606218

Perez-Cheeks, B. A., Planet, P. J., Sarkar, I. N., Clock, S. A., Xu, Q., and Figurski, D. H. (2012). The product of tadZ, a new member of the parA/minD superfamily, localizes to a pole in *Aggregatibacter* actinomycetemcomitans. *Mol. Microbiol.* 83, 694–711. doi: 10.1111/j.1365-2958.2011.07955.x

Raskin, D. M., and De Boer, P. A. (1999). Rapid pole-to-pole oscillation of a protein required for directing division to the middle of *Escherichia coli*. *Proc. Nat. Acad. Sci.* 96, 4971–4976. doi: 10.1073/pnas.96.9.4971

Ringgaard, S., Schirner, K., Davis, B. M., and Waldor, M. K. (2011). A family of ParA-like ATPases promotes cell pole maturation by facilitating polar localization of chemotaxis proteins. *Genes Dev.* 25, 1544–1555. doi: 10.1101/gad.2061811

Roberts, M. A., Wadhams, G. H., Hadfield, K. A., Tickner, S., and Armitage, J. P. (2012). ParA-like protein uses nonspecific chromosomal DNA binding to partition protein complexes. *Proc. Natl. Acad. Sci.* 109, 6698–6703. doi: 10.1073/pnas.1114000109

Savage, D. F., Afonso, B., Chen, A. H., and Silver, P. A. (2010). Spatially ordered dynamics of the bacterial carbon fixation machinery. *Science* 327, 1258–1261. doi: 10.1126/science.1186090

Schiller, H., Hong, Y., Kouassi, J., Rados, T., Kwak, J., DiLucido, A., et al. (2024). Identification of structural and regulatory cell-shape determinants in *Haloferax volcanii*. *Nat. Commun.* 15:1414. doi: 10.1038/s41467-024-45196-0

Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., et al. (2012). Fiji: an open-source platform for biological-image analysis. *Nat. Methods* 9:676. doi: 10.1038/nmeth.2019

Schmid, B., Tripal, P., Fraaß, T., Kersten, C., Ruder, B., Grüneboom, A., et al. (2019). 3Dscript: animating 3D/4D microscopy data using a natural-language-based syntax. *Nat. Methods* 16, 278–280. doi: 10.1038/s41592-019-0359-1

Schuhmacher, J. S., Rossmann, F., Dempwolff, F., Knauer, C., Altegoer, F., Steinchen, W., et al. (2015b). MinD-like ATPase FlhG effects location and number of bacterial flagella during C-ring assembly. *Proc. Nat. Acad. Sci.* 112, 3092–3097. doi: 10.1073/pnas.1419388112

Schuhmacher, J. S., Thormann, K. M., and Bange, G. (2015a). How bacteria maintain location and number of flagella? *FEMS Microbiol. Rev.* 39, 812–822. doi: 10.1093/femsre/fuv034

Szeto, T. H., Rowland, S. L., Rothfield, L. I., and King, G. F. (2002). Membrane localization of MinD is mediated by a C-terminal motif that is conserved across eubacteria, archaea, and chloroplasts. *Proc. Nat. Acad. Sci.* 99, 15693–15698. doi: 10.1073/pnas.232590599

Thompson, S. R., Wadhams, G. H., and Armitage, J. P. (2006). The positioning of cytoplasmic protein clusters in bacteria. *Proc. Nat. Acad. Sci.* 103, 8209–8214. doi: 10.1073/pnas.0600919103

Tripepi, M., Imam, S., and Pohlschröder, M. (2010). *Haloferax volcanii* flagella are required for motility but are not involved in PibD-dependent surface adhesion. *J. Bacteriol.* 192, 3093–3102. doi: 10.1128/JB.00133-10

Tripepi, M., You, J., Temel, S., Önder, Ö., Brisson, D., Pohlschröder, M., et al. (2012). N-glycosylation of *Haloferax volcanii* flagellins requires known Agl proteins and is essential for biosynthesis of stable flagella. *J. Bacteriol.* 194, 4876–4887. doi: 10.1128/JB.00731-12

Zhou, H., Schulze, R., Cox, S., Saez, C., Hu, Z., Lutkenhaus, J., et al. (2005). Analysis of MinD mutations reveals residues required for MinE stimulation of the MinD ATPase and residues required for MinC interaction. *J. Bacteriol.* 187, 629–638. doi: 10.1128/JB.187.2.629-638.2005

frontiers | Frontiers in Microbiology

TYPE Original Research
PUBLISHED 28 March 2025
DOI 10.3389/fmicb.2025.1547649

OPEN ACCESS

EDITED BY
Zhe Lyu,
North Carolina State University, United States

REVIEWED BY
Rosa María Martínez-Espinosa,
University of Alicante, Spain
Arthur Charles-Orszag,
University of California, San Francisco, United States

*CORRESPONDENCE
María Inés Giménez
migimen@mdp.edu.ar

†These authors have contributed equally to this work

RECEIVED 18 December 2024
ACCEPTED 04 March 2025
PUBLISHED 28 March 2025

CITATION
Costa MI, Cerletti M, Paggi RA, Frecha SD, Zoratti V, Latorre LL, De Castro RE and Giménez MI (2025) Rhomboid proteases: key players at the cell surface within haloarchaea. *Front. Microbiol.* 16:1547649. doi: 10.3389/fmicb.2025.1547649

Rhomboid proteases: key players at the cell surface within haloarchaea

Mariana Inés Costa†, Micaela Cerletti†, Roberto Alejandro Paggi, Sofia Denise Frecha, Valeria Zoratti, Lucas Leonel Latorre, Rosana Esther De Castro and María Inés Giménez*

Instituto de Investigaciones Biológicas, Universidad Nacional de Mar del Plata (UNMdP)-Consejo Nacional de Investigaciones Científicas y Técnicas (CONICET), Mar del Plata, Argentina

Introduction: Rhomboid proteases are intramembrane serine proteases that play a key role in regulating membrane proteins across all domains of life. However, their function in archaea remains poorly understood. The model halophilic archaeon Haloferax volcanii encodes two rhomboid homologs, rho1 (HVO_1474) and *rho2* (HVO_0727). Previous studies indicated that the deletion of rho2 resulted in mild alterations in motility, adhesion, biofilm formation, and cell morphology, suggesting potential functional compensation by *rho1*.

Materials and methods: To investigate the role of these proteases, we generated single (Δ*rho1*) and double (Δ*rho1* Δ*rho2*) deletion mutants. Phenotypic characterization included viability assays, motility tests, adhesion and biofilm formation studies, as well as morphological analysis using microscopy. Functional overlap between rho1 and rho2 was evaluated through genetic complementation/overexpression experiments in which each gene was expressed in trans in the mutant backgrounds.

Results: Both Δ*rho1* and Δ*rho1* Δrho2 mutants were viable, indicating that these genes are not essential in H. volcanii. The Δ*rho1* mutant exhibited increased motility, enhanced biofilm formation, reduced adhesion to glass surfaces, and significant morphological alterations, particularly in trace element-deficient conditions. The double mutant (Δ*rho1* Δ*rho2*) showed increased adhesion to surfaces, mild motility reduction, and fewer morphological abnormalities compared to Δ*rho1*. Complementation assays revealed that both rho1 and *rho2* could restore motility in Δ*rho2* and adhesion in Δrho1. However, only rho1 was able to complement the morphological defects, suggesting a degree of functional divergence between these homologs.

Discussion: This work highlights the role of rhomboid proteases in regulating critical cell surface processes in *H. volcanii*, including biofilm formation, surface adhesion, and cell shape determination. The ability of rho1 and *rho2* to compensate for each other in certain functions while maintaining distinct roles underscores a complex regulatory interplay. Future research will focus on identifying natural substrates and elucidating the molecular mechanisms underlying rhomboid protease function in haloarchaea.

KEYWORDS

***Haloferax volcanii*, rhomboid protease, haloarchaea, intramembrane proteases, archaeal morphology**

Introduction

Over the past two decades there has been a growing interest in a distinctive group of proteases known as Intramembrane Proteases (IMPs), based on their remarkable ability to cleave peptide bonds within the hydrophobic environment of cellular membranes. These proteases are ubiquitously represented across all domains of life and serve diverse functions including protein maturation, activation of signaling molecules and protein degradation. Categorized according to their catalytic mechanisms, these proteolytic enzymes are classified into four families: serine, metallo, aspartyl, and a glutamyl protease (Kühnle et al., 2019).

The rhomboid protease (Rho) family includes intramembrane serine proteases and various pseudo proteases. Rho-like proteins are universally present throughout evolution, serving as regulators that influence the fate of membrane proteins (Kandel and Neal, 2020). Among Rho, those found in eukaryotes have been extensively studied for their crucial roles in regulating significant human pathologies such as male infertility (Radaelli et al., 2023), Parkinson's disease (Shi et al., 2011), Alzheimer's disease (Paschkowsky et al., 2016), type-2 diabetes (Walder et al., 2005; Civitarese et al., 2010) and cancer (Cheng et al., 2014; Song et al., 2015; Noy et al., 2016). Additionally, Rho from protozoan parasites causing amoebiasis, malaria, and toxoplasmosis play a role in promoting host cell invasion by cleaving surface proteins critical for the parasite entry into host cells (Brossier et al., 2005; O'Donnell et al., 2006; Baker et al., 2006; Baxt et al., 2008).

The biological role of Rho in prokaryotes physiology has been relatively underexplored. The *Escherichia coli* Rho homolog GlpG was the first to be investigated. GlpG can efficiently process eukaryotic Rho substrates showing protease activity (Reddy and Rainey, 2012). Furthermore, the crystal structure of the GlpG core domain has been resolved with a resolution of 2.1 Å (Wang et al., 2006). However, up to date, endogenous targets of GlpG have not been identified. The AarA Rho from the pathogenic bacterium *Providencia stuartii* specifically cleaves the N-terminal extension of TatA, a membrane-bound component of the twin-arginine protein translocation pathway. The processing of TatA activates the translocation mechanism, facilitating the export of an unidentified *quorum*-sensing signal (Stevenson et al., 2007). In Mycobacteria, null Rho mutants exhibit compromised biofilm formation and increased sensitivity to antibiotics (Kateete et al., 2012). *Bacillus subtilis* Rho YqgP cleaves the magnesium transporter MgtE. Additionally, YqgP interacts with the membrane-anchored metalloprotease FtsH which in turn completes the degradation of this polytopic membrane protein, resembling eukaryotic endoplasmic reticulum-associated protein degradation (Began et al., 2020). Rho of *Shigella sonnei* are involved in membrane protein quality control by specifically targeting "orphan" components of respiratory complexes. These components display protected metastable transmembrane domains when incorporated into a functional complex. Initial cleavage by Rho allows subsequent degradation of the orphan substrate (Liu et al., 2020). Quantitative shotgun mass spectrometry was used in *Corynebacterium glutamicum* to compare a Rho double mutant with the wild type (wt) strain. This study revealed differences in the abundance of proteins that participate in diverse cellular functions including a decrease in ribosomal subunits and RNA polymerase, variations in iron uptake proteins and changes in the abundance of lipid and mycolic acid biosynthetic enzymes, suggesting a functional link between Rho and cell envelope lipid biosynthesis. This connection was confirmed by shotgun lipidomics (Luenenschloss et al., 2022). Shotgun proteomics was also employed to compare a *Brucella abortus* Rho mutant with the parental strain. Proteins identified as potential substrates included denitrification enzymes and/or high oxygen affinity cytochrome c oxidase which are required for growth in low oxygen conditions (Marchesini et al., 2022).

Although Rho homologs are widely distributed in archaeal genomes, the biological relevance of this protease family in archaea remains poorly understood. The chromosome of the model haloarchaeon *H. volcanii* encodes two Rho homologs, designated *rho1* and *rho2*. A null mutation in *rho2* resulted in a strain (previously denoted as MIG1, and referred to as Δ*rho2* in this study) that exhibited reduced adhesion and motility, heightened sensitivity to novobiocin, impaired recovery from UV light irradiation and a truncation in glycosylation at N732 of the S-layer glycoprotein (Parente et al., 2014). Despite the significance and reproducibility of these phenotypes, they were relatively mild, suggesting potential functional complementation of the mutation by *rho1*. Comparative proteomics analysis between Δ*rho2* and the parental strain revealed that the absence of *Rho*2 had a widespread impact on the proteome, influencing the quantity and/or electrophoretic mobility of proteins associated with the observed phenotypes, as well as proteins involved in various processes such as metal homeostasis and cell division (Costa et al., 2018).

This study aimed to advance our understanding of the role of Rho in archaea. To achieve this, two novel deletion mutants were generated in *H. volcanii*: a single mutant targeting the *rho1* gene (Δ*rho1*) and a double mutant targeting both *rho1* and *rho2* (Δ*rho1* Δ*rho2*). Phenotypic characterization and functional cross-complementation assays revealed a partial functional overlap between the two proteases, along with a significant role for Rho in regulating cell surface physiology, impacting cell shape, motility, adhesion, and biofilm formation.

Materials and methods

Strains and growth conditions

Details regarding the strains, plasmids, and primers used in this study are shown in Supplementary Table S1. *H. volcanii* strains were cultured in CAB (de Silva et al., 2021), CA (CAB with no trace elements) or MGM media (Allers, 2009) at 42°C with agitation at 150 rpm. For the construction of growth curves, 3 mL starter cultures were inoculated from fresh isolated colonies and allowed to grow until the late exponential phase of growth (OD_{600} 0.7–1.2) was reached. Subsequently, 3 mL cultures were initiated using the starter cultures, at a final OD_{600} of 0.01 and allowed to grow under the aforementioned conditions. The OD_{600} was monitored at specified intervals using an Ultrospec10 cell density meter (Biochrom). The duplication time calculation and slope comparisons were performed using the GraphPad Prism version 8.0.2 for windows. For motility assays, *H. volcanii* strains were stab-inoculated in 0.25% (w/v) agar CAB plates and incubated at 42°C for 2–3 days. Motility was determined by measuring the diameter of the swimming ring with the ImageJ program.

Generation of plasmid constructions

The chromosomal regions encompassing the ORFs of either the *rho1* (HVO_1474) or *rho2* (HVO_0727) genes were PCR-amplified using Fwrho1NdeI + Rvrho1BamHI (for *rho1*) and HVO0727NdeIF + HVO0727BamHIr (for *rho2*) primers (Supplementary Table S1) with iProof DNA polymerase (BioRad) and *H. volcanii* H26 genomic DNA as the template. The resulting amplicons (1,651 bp for *rho1* and 891 bp for *rho2*) were individually cloned into the TOPO-Blunt vector (Invitrogen). The identities of the cloned fragments were confirmed by sequencing (Macrogen, Korea). The fragments were then digested with NdeI and EcoRI (*rho1*) or NdeI and BamHI (*rho2*) (Thermo) and subsequently subcloned into the pTA963 vector (Allers et al., 2010), which had been digested with the same restriction enzymes. The ligation products were introduced into *E. coli* DH10β via transformation, passed through *E. coli* GM33 to obtain non-methylated plasmid DNA, and finally transformed into the indicated *H. volcanii* strains using the standard polyethylene glycol (PEG) transformation method (Allers, 2009).

Construction of *H. volcanii rho1* in-frame knock-out mutants

The knock-out constructs were generated following a previously described protocol (Allers et al., 2004). Briefly, ~800 bp flanking each end of the *rho1* gene were PCR-amplified and consecutively cloned into the EcoRI/HindIII (upstream region) and the BamHI/XbaI (downstream region) sites of the haloarchaeal suicide vector pTA131. The resulting construct (pIG1) was initially amplified in *E. coli* DH10β and subsequently propagated in *E. coli* GM33 to obtain a non-methylated plasmid. This plasmid was then transformed into *H. volcanii* H26 or Δ*rho2* using the PEG method. Selection for a single homologous recombination event between one of the flanking regions on the knock-out construct and the chromosome (pop-in) was achieved by growth in CA medium lacking uracil (ura). Recombinants were subsequently cultured in liquid CA + ura (50 μg ml^{-1}) with two passages to fresh medium to facilitate a second recombination event leading to the excision of the plasmid from the chromosome (pop-out). Liquid cultures were then plated on CA with ura (10 μg ml^{-1}) and 5-FOA (50 μg ml^{-1}), followed by incubation at 42°C for 5–10 days. Colonies were screened by PCR using the external primers (Fwverifycorto and Rvverify2020) to confirm elimination of *rho1*.

Microscopy

Cells of *H. volcanii* H26 and mutant strains were inoculated from freshly isolated colonies in liquid medium until they reached the desired OD_{600}. Culture aliquots (5 μL) were placed on a coverslip, and a 0.5% (w/v) agarose CAB plug (3 mm thick) was added on top. The cells were analyzed using phase-contrast microscopy at 100× magnification (Eclipse Ti-S, NIKON). Cells were segmented using Omnipose (Cutler et al., 2022) with a custom model partially trained on our images. Segmented cells were analyzed using CellProfiler 4.2.5 (Carpenter et al., 2006) and the area (μm^2) and aspect ratio were determined. The aspect ratio, defined as the ratio of a cell's major to minor axis, increases with cell elongation. Statistical analysis was performed using the Kruskal-Wallis test to assess differences among multiple groups. Subsequently, Dunn's post-hoc test was employed to determine specific group differences. A p-value <0.001 was considered significant.

Cell adhesion

Surface adhesion was assessed using the method described by Tripepi et al. (2010). *H. volcanii* cultures (3 mL) at an OD_{600} of ~0.3 were incubated in 50 mL sterile screw-cap plastic tubes. Glass coverslips were inserted into each tube at a 90° angle and were incubated without shaking. After the specified time intervals coverslips were removed with forceps, submerged for 2 min in 2% (v/v) acetic acid and allowed to air dry. The dry coverslips were stained with 0.1% (w/v) crystal violet for 10 min, washed (three times) with distilled water, air-dried and then observed by light microscopy using a Nikon Eclipse Ti microscope. Cells were quantified using ImageJ.

Biofilm formation

H. volcanii biofilms were formed on pre-sterilized 96-well flat-bottom polystyrene microtiter plates (6 wells for each strain) following the method described elsewhere (Shukla, 2017). In brief, 200 μL of liquid CAB cultures (OD_{600} ~ 0.3) were dispensed into each well and 200 μL of distilled water was added to peripheral wells to minimize water loss. The microtiter plate was then incubated for 72 h at 42°C. Planktonic cells were removed by inverting the plate onto absorbent paper and biofilms were fixed with 99% (v/v) methanol. The plates were washed twice with CAB and air-dried. Subsequently, 200 μL of a 0.2% (w/v) crystal violet solution was added to all wells. After 5 min excess crystal violet was removed and plates were washed with distilled water (4 times) before air drying. Finally, the cell-bound crystal violet was dissolved in 33% (v/v) acetic acid. Biofilm growth was monitored in terms of Absorbance at 570 nm (A_{570}) using a microplate reader (BioTek ELx800).

Results

Rho homologs display functional complementation in *H. volcanii*

The phenotypes exhibited by the Δ*rho2* deletion mutant were consistent but mild, suggesting a degree of functional complementation by the other *H. volcanii* Rho homolog, *rho1* (Parente et al., 2014). To test this, we focused on the most prominent Δ*rho2* phenotype: motility on soft agar plates. Motility assays were performed using the H26 and Δ*rho2* strains transformed with the empty pTA963 vector, as well as the Δ*rho2* strain transformed with pTA963 harboring either the *rho1* or *rho2* gene. The rhomboid gene is predicted to be part of an operon with the *endV*, which encodes for a type-V endonuclease. Additionally, a construct containing the pTA963 vector with the complete *rho2*/*endV* operon under the control of the endogenous promoter (pMCF) was included. This construct was previously reported to complement the motility defect of the Δ*rho2* mutant (Parente et al., 2014). As shown in Figure 1, the three constructs

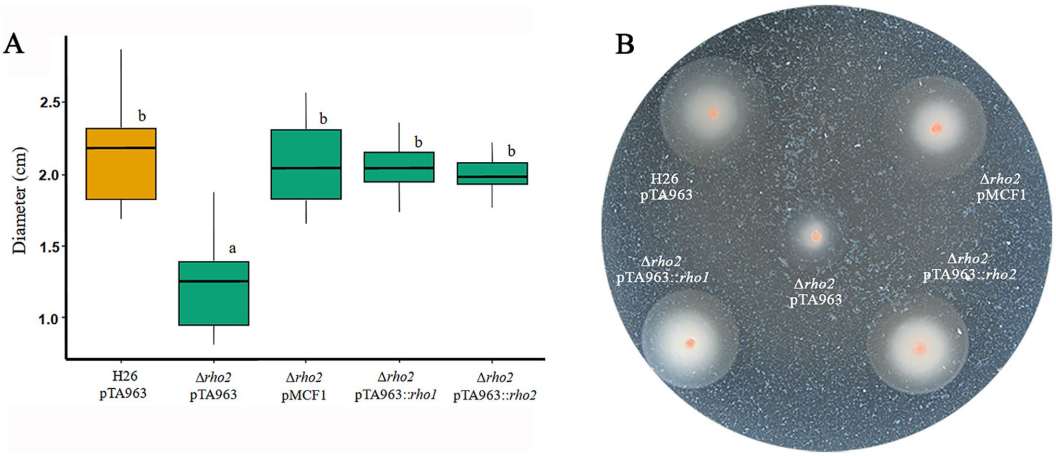


FIGURE 1
Complementation of the motility phenotype in the Δ*rho2* mutant with plasmids expressing *rho* homologs. *H. volcanii* H26 pTA963 and Δ*rho2* transformed with either pTA963, pTA963::*rho1*, pTA963::*rho2*, or pMCF1 (containing the complete *rho2*/*endV* operon, under the endogenous promoter) were cultured in CAB medium with 2 mM trp at 42°C overnight (ON) and subsequently adjusted to an OD_{600} = 0.5. For the motility assay, 1 μL of the adjusted culture was inoculated onto swimming agar plates with 2 mM trp and incubated for 48 h. Photographs were taken and the ring diameter was measured using ImageJ. All strains were inoculated on the same plate and three independent experiments with five replicates each were conducted. The results were analyzed using the R program. **(A)** Box plot of the distribution of the motility rings diameters for each strain. **(B)** Representative image of a soft agar motility plate.

successfully restored motility to wt-type (H26) levels. However, in the case of the *rho1* construct, it is important to note that since an endogenous copy of this gene is present in the Δ*rho2* mutant, phenotype complementation is achieved through over expression by the addition of an extra copy of *rho1* from the pTA963 vector.

It is worth noting that the diameter of motility halos in a given strain is influenced by both growth (cell division) and motility (Palma et al., 2022). To distinguish between these factors, we compared the specific growth rates of the plasmid-transformed strains. The Δ*rho2* + pTA963 and Δ*rho2* + pTA963::*rho1* strains showed no significant variation in growth compared to the H26 + pTA963 control. In contrast, the Δ*rho2* + pTA963::*rho2* strain displayed a significant growth defect ($p = 0.0007$) (Supplementary Figure S1). Despite this, both Rho genes successfully complemented the swimming defect of the Δ*rho2* + pTA963 mutant. This indicates that, although the pTA963::*rho2* strain exhibited impaired growth, *rho2* expression from the plasmid was still sufficient to counteract growth reduction and fully restore motility to wt-type levels.

Generation of single *rho1* and double *rho1 rho2* knock-out mutants

To identify the cellular processes regulated by Rho in *H. volcanii* we aimed to generate mutants devoid of the corresponding gene/s. Previously, we generated a single Δ*rho2* strain (Parente et al., 2014). In this work we constructed *H. volcanii* mutants lacking *rho1* in H26 (single *rho1* null mutant) or the Δ*rho2* strain (double null *rho1 rho2* mutant). Despite several initial unsuccessful attempts, we finally obtained the target strains by allowing one of the liquid medium passages (Materials and Methods section) to reach the stationary (S) growth phase (OD_{600} ~ 2). Gene deletion was confirmed via PCR using primers external to the knockout construct (Supplementary Figure S2) and the strains were denoted as "Δ*rho1*" and "Δ*rho1* Δ*rho2*." The successful generation of both mutant strains indicated that the presence of Rho is not required for *H. volcanii* viability at least under our standard laboratory conditions.

Characterization of the mutant strains

Growth in liquid and solid media

The growth of the mutant strains in liquid CAB medium was monitored by measuring OD_{600} over time (Figure 2A). While the growth curves of the Δ*rho2* and Δ*rho1* Δ*rho2* mutants did not exhibit differences compared to the parental H26 strain, the Δ*rho1* mutant showed a slight but statistically significant growth defect (p-value = 0.0276), with a duplication time of 4.27 h compared to 4.12 h for the H26 strain. This difference was evidenced by the variation in the slopes of the growth curves during the exponential (E) phase (Figure 2B). Additionally, growth was assessed under different salt conditions (4.8 M and 1.75 M NaCl) (Supplementary Figure S3). No significant differences were observed in the mutants' response to suboptimal salt concentrations (1.75 M NaCl). At high salt concentrations (4.8 M NaCl), the Rho mutants showed a slight improvement in growth. However, when considering all experiments together (three independent assays, each performed in triplicate) the data exhibited substantial variability, and thus the results were deemed inconclusive. Colony size, morphology, and appearance on solid medium were indistinguishable from those of the H26 strain (data not shown).

Cell morphology

To evaluate the impact of *rho* deletion on cell morphology, optical microscopy was used to compare *H. volcanii* H26 with Rho-deficient mutants (Δ*rho1*, Δ*rho2*, and Δ*rho1* Δ*rho2*) during early exponential (EE), E, late exponential (LE), and S growth phases in CA (Figure 3) and CAB media (Figure 4).

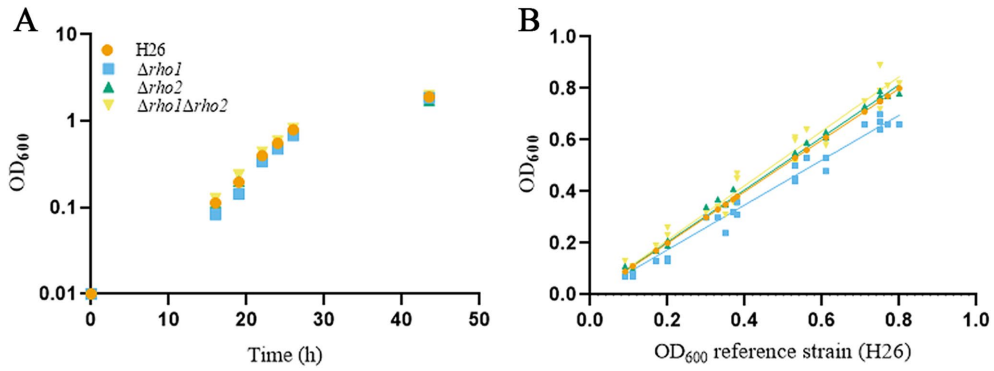


FIGURE 2
Evaluation of Rho deletion mutants growth in liquid medium. Cultures of the indicated strains were inoculated from ON starter cultures at an initial OD_{600} of 0.01 in 3 mL of CAB medium supplemented with ura (50 μg/mL). Growth was monitored by measuring OD_{600} at the indicated times and the values were used to generate the growth curve **(A)** Each culture was grown as three biological replicates, with experiments independently repeated at least three times. Growth rates were determined during the E phase by fitting linear regression models to the OD_{600} data and calculating the slopes, using GraphPad Prism software. For comparison, the H26, Δ*rho1,* Δ*rho2* and Δ*rho1* Δ*rho2* strains growth rate was plotted against the H26 growth curve, as a reference growth curve **(B)** Statistical significance was assessed using a two-tailed Student's *t*-test (*p*-value <0.05 considered significant).

The Δ*rho1* strain exhibited pronounced morphological abnormalities, including elongated rods, giant cells, and drop-like or reed-like shapes (Figures 3B, 4B). These aberrations were particularly evident in the absence of trace elements, aligning with previous reports that cell shape in CA medium is affected by such deficiencies (de Silva et al., 2021). Consistent with this, some abnormal cells were also observed in the parental H26 strain when grown in CA. However, aberrant morphologies were markedly more frequent and severe in the Δ*rho1* mutant. Interestingly, these morphological defects resolved as cultures entered the S phase, where all strains reverted to the characteristic disk-shaped morphology (Figure 3A). In CAB medium, which is supplemented with trace elements, the Δ*rho1* mutant displayed aberrant morphologies primarily during the EE and E phases (Figure 4A). These abnormalities were less pronounced than those observed in CA and disappeared earlier in the growth curve, further underscoring the influence of trace element availability on cellular morphology.

The Δ*rho2* strain did not exhibit substantial morphological changes compared to the parental H26 strain, aside from occasional drop-like cells observed during the EE phase. Notably, the double mutant Δ*rho1* Δ*rho2* did not display the severe morphological defects seen in the Δ*rho1* single mutant under either growth condition (Figures 3A, 4A), suggesting the presence of compensatory mechanisms in the absence of both genes.

Cells were segmented and the aspect ratio determined using CellProfiler 4.2.5 (refer to Materials and Methods for detailed methodology). Consistent with the previous observations, the Δ*rho1* strain exhibited a higher average aspect ratio (lengthened cells) across all growth phases, except in the S phase where cells of all the strains predominantly looked like disks (Figures 3A, 4A, right panel). These differences were confirmed to be statistically significant according to Kruskal-Wallis and Dunn tests. In addition, the cell area of the different mutants was calculated and compared to that of the parental strain (Supplementary Figure S4). It was evident that the Δ*rho1* cells exhibited larger areas compared to the other strains throughout the E phase in CA medium. When cell areas were measured in the presence of trace elements (CAB), *rho1* elimination led to a milder increase in cell area with a similar effect observed in the double Δ*rho1* Δ*rho2* mutant during the EE and E stages of growth.

Morphological abnormalities in the Δ*rho1* strain were attributable to the loss of *rho1*, as expression of the gene from the pTA963::*rho1* vector under 2 mM tryptophan (trp) induction restored normal morphology in CAB and CA media (Figure 5). In contrast, *rho2* expression using the same conditions failed to rescue the phenotype, indicating that functional redundancy between these proteases is limited.

Motility on soft agar plates

As mentioned earlier the Δ*rho2* strain exhibited reduced motility, a defect that could be restored by overexpressing *rho1* (Figure 1). To assess whether Rho1 also influences swimming in *H. volcanii*, we evaluated the motility of the mutant strains generated in this study using soft agar plates. As shown in Figure 6, Δ*rho1* displayed larger swimming halos than the parental strain, suggesting enhanced motility compared to H26. Notably, given that the Δ*rho1* strain demonstrates a mild reduction in growth (Figure 2), the observation of larger swimming halos suggests that motility is even further enhanced in this mutant. Interestingly, the double mutant Δ*rho1* Δ*rho2* exhibited a slight decrease in motility similar to that of the single mutant Δ*rho2*. When these results were compared to those shown in Figure 1, it was noted that, although the Δ*rho2* mutant formed significantly reduced swimming halos in both assays, the reduction was more pronounced when comparing Δ*rho2* pTA963 with H26 pTA963 than when comparing the strains without the plasmid. This difference could be attributed to variations in media composition (CAB + ura for H26 and Δ*rho2* vs. CAB for plasmid-transformed strains) and/or a differential effect of the presence of the pTA963 vector between strains.

Complementation assays were attempted, however, the Δ*rho1* pTA963 strain did not show significant differences in the diameter of halos compared to the control strain (H26 pTA963), which contrasted with the differences observed between the same strains lacking the plasmid (data not shown). This discrepancy may be attributed to variations in growth rates which were more pronounced in strains

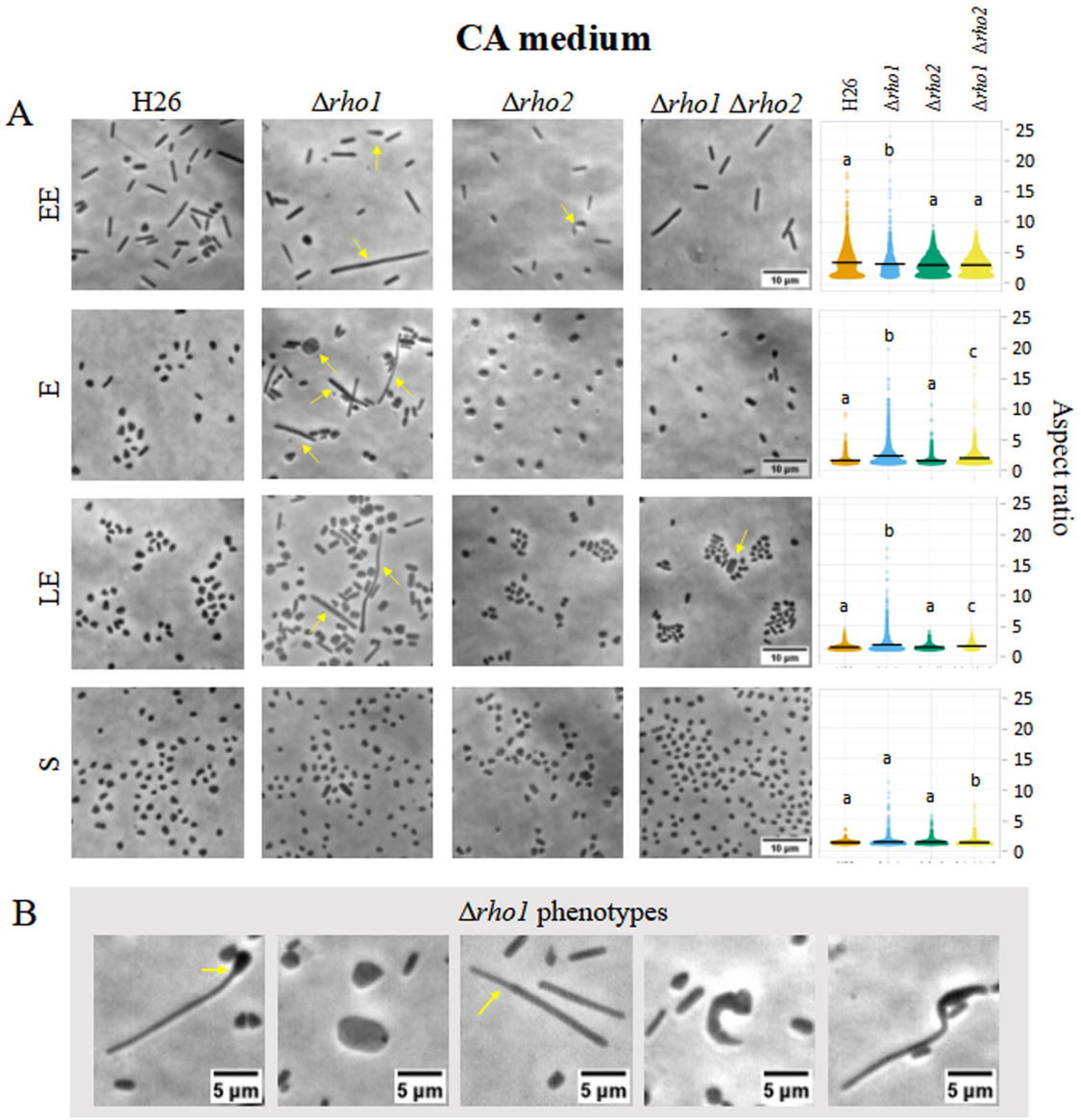


FIGURE 3
Cell morphology of *H. volcanii* Rho mutants in CA medium. *H. volcanii* single colonies of the specified strains were inoculated into 3 mL cultures of CA medium and samples were taken at EE (OD_{600} 0.1–0.3), E (OD_{600} 0.4–0.6), LE (OD_{600} 0.7–1), and S phases ($OD_{600} \geq 1.3$). **(A)** Cells were imaged in an inverted microscope under 0.5% (w/v) CAB-agarose pads. Right panel shows the distribution of the cell's aspect ratio for each strain and growth phase. At least 10 photograms from three independent cultures were analyzed for each strain and condition. **(B)** Different morphological abnormalities observed for the Δ*rho1* strain in CA medium.

harboring pTA963 (3 and 6% decrease in duplication time for Δ*rho1* and Δ*rho1* pTA963, respectively) and/or differential effects of the plasmid on the motility of H26 and Δ*rho1* strains. Consequently, functional complementation for this phenotype could not be investigated.

Cell adhesion and biofilm formation

We also assessed the capacity of the mutants to adhere to glass surfaces and to form biofilms. The Δ*rho1* strain evidenced a slightly reduced adhesion to glass coverslips similar to the Δ*rho2* strain. Surprisingly, the double mutant Δ*rho1* Δ*rho2* exhibited a significant increase in the number of adhered cells, suggesting a dysregulation of adhesion in the absence of both Rho homologs (Figures 7A,B). A copy of the *rho1* gene expressed from pTA963 vector could restore the H26 phenotype confirming that the reduced number of adhered cells was due to the absence of this gene (Figures 7C,D). Complementation of this phenotype was also accomplished by introducing an extra copy of the other Rho gene homolog (*rho2*) in the Trp-inducible expression vector pTA963, indicating functional overlap between these proteases (Figures 7C,D).

All the strains showed the ability to form immersed liquid biofilms (Supplementary Figure S5A) and exhibited honeycomb patterns when removing the Petri dish lid (Supplementary Figure S5B), as previously described for other *H. volcanii* strains (Schiller et al., 2020).

Early biofilm formation on plastic surfaces was assessed using the crystal violet method (Figure 8). The results showed that the single mutant Δ*rho1* exhibited enhanced capacity for biofilm formation

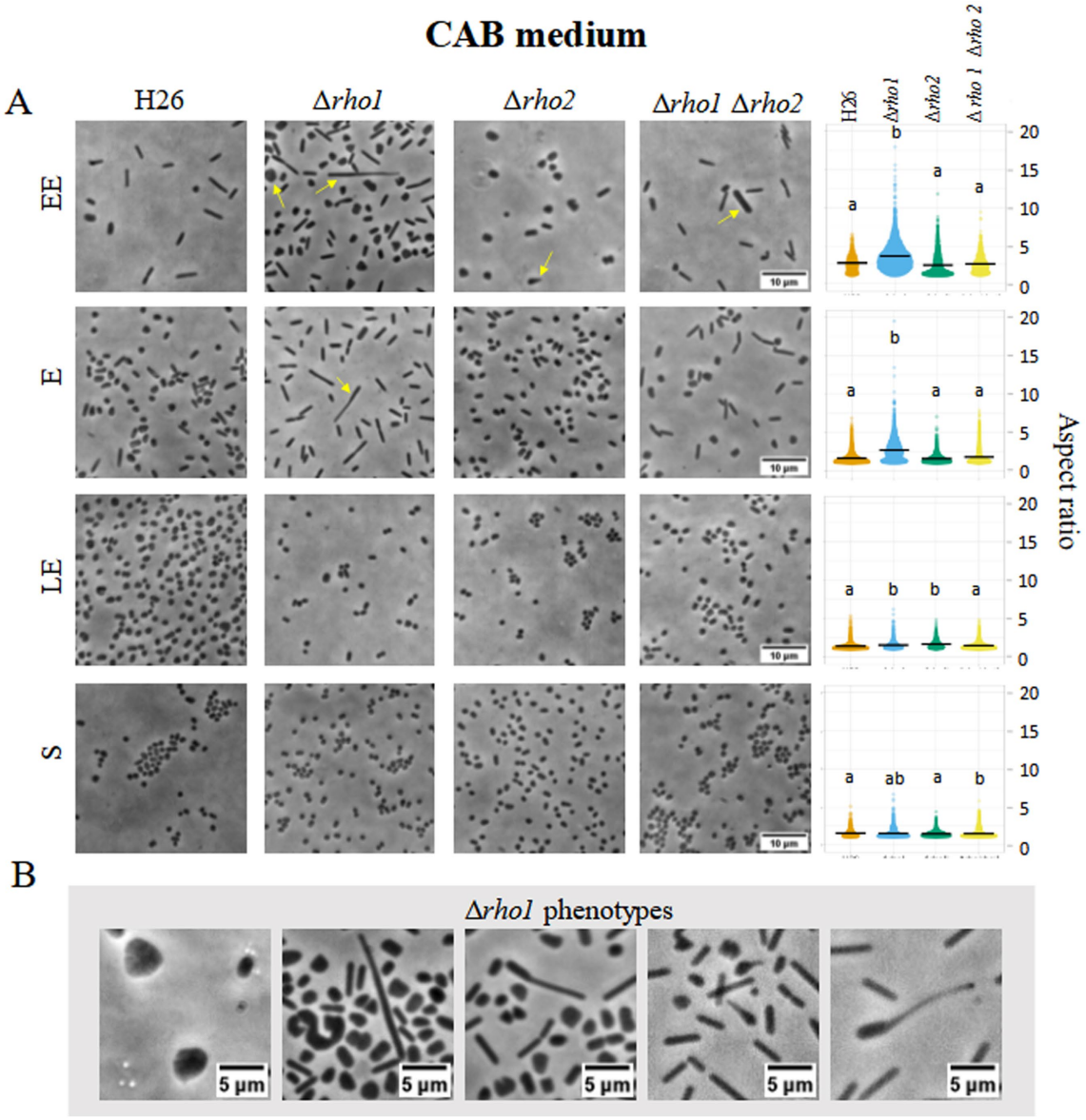


FIGURE 4
Cell morphology of *H. volcanii* Rho mutants in CAB medium. *H. volcanii* single colonies of the specified strains were inoculated into 3 mL cultures of CAB and samples were taken at the same stages of the growth curve depicted in Figure 3. **(A)** Cells were imaged in an inverted microscope under 0.5% (w/v) CAB-agarose pads. Right panel shows the distribution of cell aspect ratio for each strain and growth phase. At least 10 photograms from three independent cultures were analized for each strain and condition. **(B)** Morphological abnormalities observed for the Δ*rho1* strain in CAB medium.

compared to the parental strain and the other two mutants (Δ*rho2* and Δ*rho1* Δ*rho2*). Attempts to complement this phenotype by introducing the *rho1* gene in trans were unsuccessful, as the Δ*rho1* mutant carrying the pTA963 plasmid displayed a significant reduction in biofilm formation relative to the H26 pTA963 control (not shown). These findings, combined with the results from motility assays (see above), suggest that the presence of the pTA963 vector differentially influences the behavior of the H26 and Δ*rho1* strains.

Discussion

Rho are regulatory IMPs highly represented throughout archaeal, bacterial and eukaryotic genomes, however, the understanding of their biological role in archaea still remains limited. In this study, we used complementation assays, gene knockout, and mutant phenotypic characterization to gain insight into the relevance of the Rho family in haloarchaea.

The archaeal structure responsible for cell motility in *H. volcanii* is the archaellum. It consists of two proteins: Arch1, the primary archaellum component, and Arch2, which is less abundant and plays a regulatory role (Tripepi et al., 2013). Deletion of *arch1* renders *H. volcanii* cells non-motile (Tripepi et al., 2012). Interestingly, proteomic studies conducted by our group showed downregulation of *arch1 in the* Δ*rho2* mutant compared to the wt strain (Costa et al., 2018), consistent with the reduced motility phenotype observed previously (Parente et al., 2014). In addition, the protein levels of PibD, a peptidase involved in archaellins and pilins maturation and

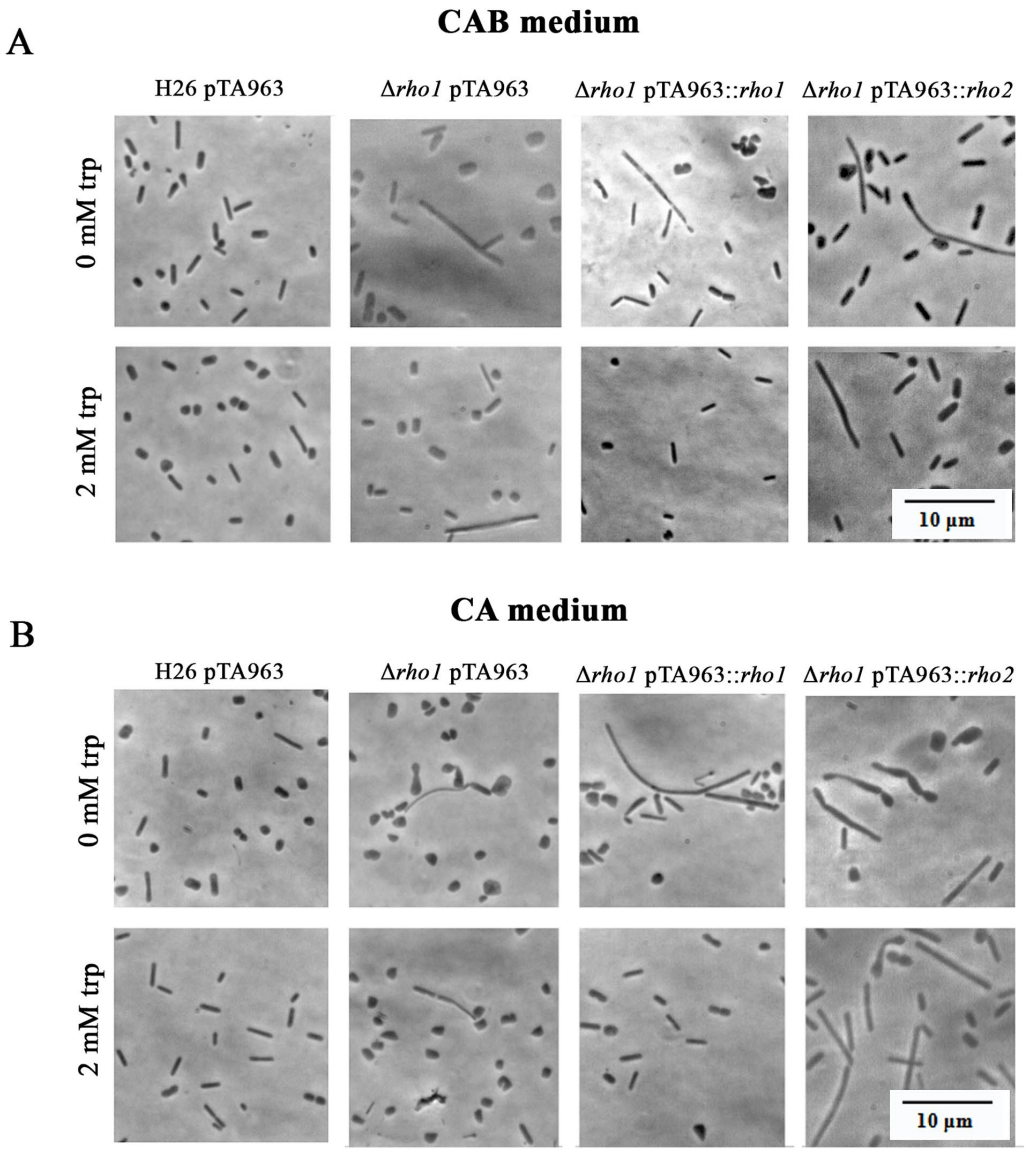


FIGURE 5
Complementation of the abnormal morphology phenotype in the Δ*rho1* strain. *H. volcanii* H26 and Δ*rho1* strains were transformed with either pTA963, pTA963::*rho1* or pTA963::*rho2* plasmids and cultured in CAB **(A)** or CA **(B)** medium until reaching the EE growth phase. Growth was conducted in the presence or absence of 2 mM trp to induce gene expression, as indicated on the left. Cell morphology was observed using optical microscopy as described in the Materials and Methods section. The photograms are representative of three independent cultures for each condition tested.

several glycosylases (involved in archaellin/pilin glycosylation) are also affected in the Δ*rho2* strain (Costa et al., 2018). In this study, we confirmed that the overexpression of a copy of *rho1* from a plasmid restores motility in the Δ*rho2* strain to the same extent as the *rho2* gene or a construct containing the complete operon *rho2/endV* (Figure 1), indicating functional complementation between Rho homologs. This observation suggests that despite differences in sequence, topology and functional domain organization (Parente et al., 2014), Rho1 and Rho2 probably participate in the regulation of common pathways. Consequently, to gain further insights into the function of Rho in haloarchaea, we generated two additional deletion mutants: a knockout in *rho1* (Δ*rho1*) and a double mutant affecting both *rho1* and *rho2* (Δ*rho1* Δ*rho2*).

The successful generation of both mutants implied that the genes encoding Rho are not essential for viability in *H. volcanii* under our standard laboratory conditions. It is noteworthy that the knockouts were only attainable when at least one of the three passages in liquid medium intended to alleviate the selection pressure (pop-out) reached at the late S growth phase. Our observations revealed altered cell morphology in the Δ*rho1* mutants during the E phase (Figures 3, 4) and these atypical cells may be outcompeted by those reverting to the wt genotype. Cells lacking *rho1* revert to a normal morphology during the S phase and therefore, mutant cells at this stage may not have a fitness disadvantage compared to the wt. This could facilitate the detection of mutant colonies on selective medium plates when at least one of the successive cultures reached the S phase. This hypothesis is supported by studies conducted in *M. smegmatis* Rho mutants, which showed normal growth in liquid medium but were outcompeted by the wt strain when they were co-cultured (Kateete et al., 2012). It is also possible that the mutants obtained in this work carried not only

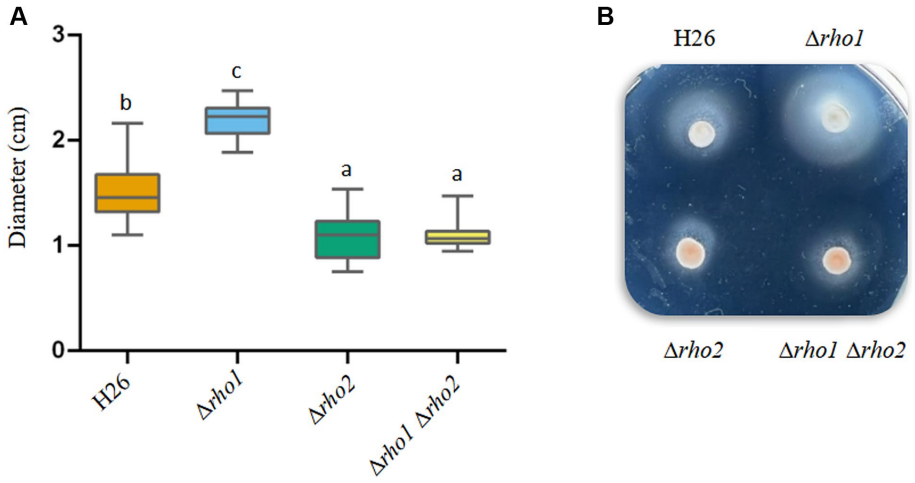


FIGURE 6
Effect of Rho gene elimination on cell motility in *H. volcanii*. Single colonies of the indicated strains were stab-inoculated on CAB soft agar plates containing 50 μg/mL ura and incubated at 42°C for 48–72 h. After incubation, photographs were taken, and the diameters of the growth halos were measured using ImageJ software. **(A)** Diameter values corresponding to three independent experiments, each with five replicates. Results were analyzed by LMMs (linear mixed models) with normal distribution to determine statistical significance between treatment and controls. A *post hoc* analysis was conducted using Tukey. All analyses were performed using R software version 3.6.1 (R Core Team, 2019). Statistically significant differences were determined at $p < 0.05$. **(B)** Representative photograph showing the motility halos generated by each strain.

the desired mutation but also secondary mutations that might act as second site suppressors. However, restoration of wt phenotypes (Figures 5, 7C,D) upon complementation with *rho1* in trans suggests that the phenotypes were specifically caused by the absence of *rho1*.

Archaeal cells exhibit a diverse array of distinctive morphologies including cocci (e.g., Borrell et al., 2023), rods (e.g., Strakova et al., 2023), squares (Walsby, 1980) and triangles (Takashina et al., 1990). *H. volcanii* displays motile rod-shaped cells in the EE/E growth phases which transition to flat, non-motile disks upon reaching LE and S phases (de Silva et al., 2021). These morphological changes likely represent an adaptation to variations in environmental conditions at each growth stage. However, the mechanisms underlying these changes in cell shape remain incompletely understood.

In this work, we showed that the Δ*rho1* mutant displays a spectrum of aberrant cell shapes during the E phase including giant cells, elongated filaments, irregular morphologies, drop-like forms as well as reed-like (tubular) shapes (Figures 3B, 4B). Some of these morphologies resemble those observed in mutants lacking tubulin-like homologs FtsZ1 and FtsZ2 as well as the divisome-interacting protein SepF (Lliao et al., 2021; Nussbaum et al., 2021) suggesting compromised cell division in the Δ*rho1* strain. Recently, two halofilins (HalA and HalB) proposed to contribute to the disk-to-rod transition in *H. volcanii* were described (Curtis et al., 2024). It is noteworthy that perturbations in cell shape observed in the Δ*rho1* strain evidenced some similarity to those of the *halB* null mutant. However, since the biological targets of Rho1 have yet to be identified, it is still not possible to determine at which point of *H. volcanii* cell morphology regulation this protease is involved and thus, explain the phenotypic similarities to the aforementioned mutants. Ongoing work will help to clarify this issue.

The irregular cells of the Δ*rho1* strain persisted during the E phase but reverted as cultures progressed to the LE (CAB medium with trace elements) or S phase (CA medium without trace elements) (see Figures 3A, 4A). These observations were consistent with our hypothesis aiming at the occurrence of a defect in cell division which may be overcome at the S phase when the cells are not actively duplicating. On the other hand, the differences observed between CA and CAB media are not surprising as the positive influence of trace elements on cell shape stability in *H. volcanii* has been previously documented by de Silva et al. (2021). Complementation of the abnormal cell shape phenotype by *rho1* expression from pTA963 vector was successful confirming that the variations in cell shape are a consequence of gene deletion. Cross complementation with an extra copy of the *rho2* homolog was unsuccessful, indicating that Rho2 is not involved in the generation of the cell shape phenotype (Figure 5).

We also investigated the effects of *rho2* elimination and the double gene knockout (Δ*rho1* Δ*rho2*) on cell morphology in CA and CAB media (Figures 3A, 4A). The Δ*rho2* mutant did not exhibit significant differences in morphology compared to the parental strain consistent with our previous findings in a nutrient-rich medium (Parente et al., 2014). Intriguingly, the double mutant showed a mild impact on cell morphology primarily reflected in slight differences in aspect ratio and area measurements compared to the parental strain, with significantly fewer occurrences of cells with abnormal phenotypes compared to the Δ*rho1* strain (Figures 3A, 4A). These observations suggest the existence of compensatory mechanisms regulating cell morphology when both Rho homologs are absent in *H. volcanii*. In addition to the findings presented in this study, previous research has reported the involvement of Rho in cell division and morphology determination in bacteria. Studies on *P. stuartii* and *B. subtilis* evidenced that Rho deletion mutants showed atypical filamentous cell shapes (Rather and Orosz, 1994; Mesak et al., 2004) indicating the relevance of Rho in prokaryotic cell shape determination.

The impact of Rho gene deletion on motility, adhesion, and early biofilm formation was also evaluated in this study. The Δ*rho1* mutant exhibited increased motility on soft agar plates (Figure 6), suggesting that Rho1 and Rho2 regulate motility in opposing ways, as Rho2 deficiency reduces swimming motility (Parente et al., 2014; Figure 1).

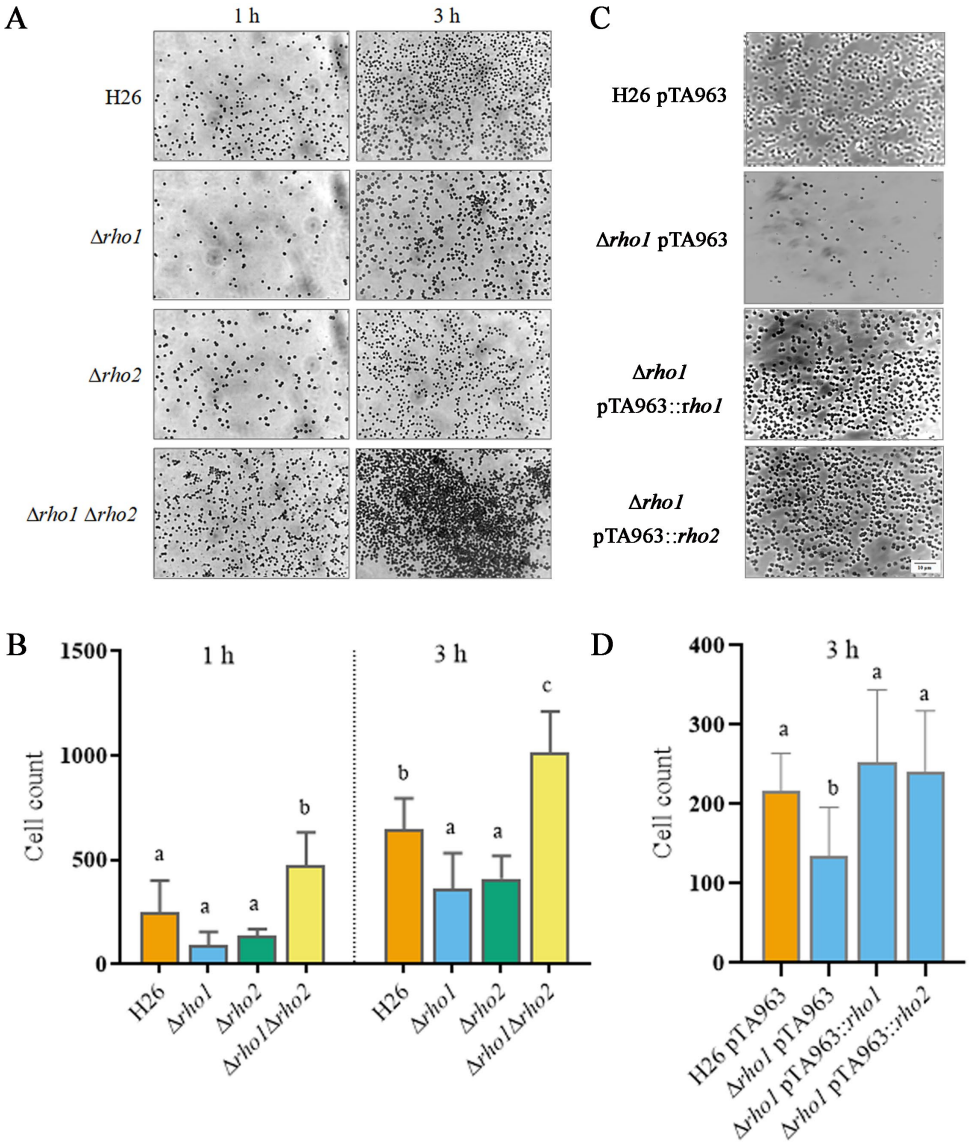


FIGURE 7
Evaluation of cell adhesion in Rho mutants. Glass coverslips were immersed in triplicate cultures of *H. volcanii* strains and incubated at 42°C for 1 or 3 h, as specified, followed by staining with crystal violet as described in the Materials and Methods section. **(A,C)** Optical microscopy images (1000×) showing cell adhesion to the glass surface for each strain. **(B,D)** Quantification of three independent experiments using Image J software. Data were analyzed using linear models (LM) with random effects, followed by Tukey's post hoc test. Statistical analyses were performed with R software version 3.6.1 (R Core Team, 2019). Differences were considered statistically significant at $p < 0.05$.

This finding contrasts with the ability of *rho1* overexpression to complement the Δ*rho2* motility defect (Figure 1). If the absence of *rho1* enhances motility (Figure 6), one would expect its overexpression to reduce swimming halos. The results obtained in this work suggest that both proteins are likely part of a regulatory network controlling motility. Since the double mutant exhibits a motility defect comparable to that of the Δ*rho2* mutant (Figure 6), it can be speculated that Rho1 functions downstream of Rho2 within this network. Additionally, the ability of either *rho1* or *rho2* to restore motility in both mutants suggests a degree of functional redundancy or a compensatory mechanism, which becomes evident when rhomboid genes are overexpressed. Future studies aimed at identifying potential substrates and/or interacting partners of Rho1 and Rho2 will help clarify this regulatory interplay.

H. volcanii can grow in biofilms which exhibit some differential traits, this includes changes in cellular morphology and an unusual form of social motility (Chimileski et al., 2014). The initial step of biofilm development involves adhesion to the abiotic surface and microcolony formation which depends on the presence and glycosylation of type-4 pilins (Esquivel et al., 2013, 2016). In this study, we investigated the role of Rho in cell adhesion and early-stage biofilm formation. We found that both single mutants Δ*rho1* and Δ*rho2* showed a mild effect on cell adhesion, evidenced by a reduced number of cells bound to glass during a 1–3 h incubation period (Figure 6; Costa et al., 2018). In previous work, we showed that Δ*rho2* was affected in pilin content (Costa et al., 2018) and had a defect in protein glycosylation (Parente et al., 2014), observations that possibly explain the reduced adhesion to glass

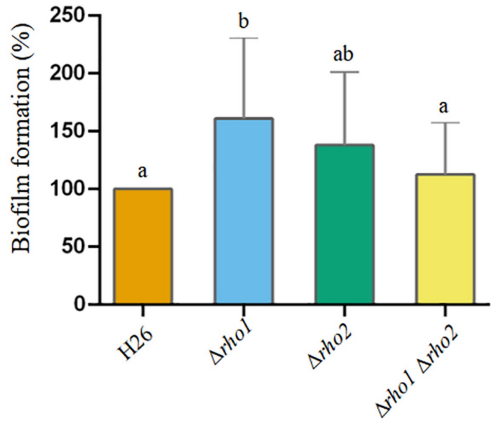


FIGURE 8
Effect of Rho gene deletion on biofilm formation in *H. volcanii*. EE phase cultures were transferred to 96-well polystyrene plates and incubated without agitation at 42°C for 72 h. Early-stage biofilms were stained with crystal violet according to the Materials and Methods section. Data was obtained from three independent experiments, each including six technical replicates. The amount of cell-bound colorant was quantified by measuring A_{574}, and values corresponding to the mutant strains were normalized to the average value of the parental strain (H26). Results were analyzed by LMMs (linear mixed models) with normal distribution to determine statistical significance between treatment and controls. A post hoc analysis was conducted using Tukey. All analyses were performed using R software version 3.6.1 (R Core Team, 2019). Statistically significant differences were determined at $p < 0.05$.

surfaces. Surprisingly, the double mutant showed a significant increase in the number of cells adhering to the glass coverslip (Figure 7). These results suggest that both proteases regulate the first step of biofilm biogenesis; it is likely that one Rho homolog is sufficient to modulate the level/processing of a regulatory factor that promotes cell adhesion and consequently, when both proteases are absent, this factor is dysregulated resulting in augmented number of bound cells. This is in agreement with the observation that both Rho genes are able to complement this phenotype when an extra copy is expressed from a plasmid (Figures 7C,D).

When the later stages of biofilm formation were assessed (72 h), this effect was no longer observed. The double mutant showed biofilm levels comparable to the parental strain, while the single Δ*rho1* mutant exhibited a slight increase in biofilm formation (Figure 8). Although these results may appear contradictory to those shown in Figure 7, it is important to note that these assays evaluate different stages of biofilm development. The adhesion assay focuses on the initial attachment phase, occurring within the first few hours of incubation with the coverslip, whereas the biofilm assay measures overall biofilm formation, including microcolony development and maturation, after 72 h of incubation. The transition from attachment to maturation is likely governed by distinct regulatory mechanisms, in which Rho proteins may play a role at different stages or in the transition itself. Further investigation will be needed to clarify the specific pathways involved. The involvement of Rho proteins in biofilm formation has been previously reported in *Mycobacterium smegmatis*. This Gram-positive bacterium encodes two Rho homologs, and in this organism, single deletion mutants exhibited reduced biofilm formation, while the double mutant displayed biofilm levels comparable to the wt-type strain (Kateete et al., 2012).

It is important to note that complementation (and cross-complementation with *rho2*) of the *rho1* motility and biofilm generation phenotypes could not be achieved. When the corresponding assays were attempted, strains harboring the empty pTA963 vector (used as experimental controls) did not exhibit the same behavior as the non-transformed strains. No differences were observed in motility on soft agar plates or in the formation of 72 h biofilms between H26 pTA963 and Δ*rho1* pTA963, in contrast to the variations observed between H26 and Δ*rho1* (Figures 6, 8). This may be explained by the presence of pTA963, as it has been reported that simply carrying a plasmid may affect cell metabolism, size, shape and growth of a given strain (Abdul-Halim et al., 2015; Patro et al., 2023). In this work, we observed that the Δ*rho1* pTA963 strain exhibited a more severe growth defect compared to H26 pTA963, in contrast to the smaller growth differences between Δ*rho1* and H26. These variations in growth may impact motility and/or biofilm formation. Additionally, the presence of pTA963 may also directly influence these processes. Consistent with this possibility, the Δ*rho2* mutant displayed a greater reduction in motility compared to control strains when transformed with the pTA963 vector (Figures 1, 6). Since this mutant does not exhibit growth differences relative to the wt strain (regardless of whether it carries the pTA963 vector, Figure 2), it is likely that the observed effect results from the plasmid differentially impacting the mutant and parental strain or from the presence/absence of ura (used as the plasmid selection marker) in the growth medium. These findings warrant further investigation.

Altogether, the data presented in this work provide evidence of the relevance of Rho in key processes occurring at the haloarchaeal cell surface including motility, cell adhesion, biofilm generation and cell shape determination. In addition, cross-complementation (homolog overexpression) assays indicated that even though Rho1 and Rho2 are involved in the regulation of related processes and to some extent display functional overlap, they also have distinctive roles, as for example regulation of cell shape. Ongoing research aims to identify and confirm natural targets of *H. volcanii* Rho1 and Rho2, which will contribute to clarify the specific role of these proteases in (halo) archaeal physiology and cell surface proteostasis.

Data availability statement

The original results are included in the article/Supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

MIC: Conceptualization, Formal analysis, Investigation, Methodology, Validation, Writing – original draft, Writing – review & editing. MC: Conceptualization, Formal analysis, Investigation,

Validation, Writing – original draft, Writing – review & editing. RP: Conceptualization, Investigation, Validation, Writing – review & editing. SF: Investigation, Validation, Writing – review & editing. VZ: Investigation, Writing – review & editing. LL: Investigation, Writing – review & editing. RC: Funding acquisition, Project administration, Writing – review & editing. MG: Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Supervision, Validation, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. This work was funded by ANPCyT grants PICT2018-2228 and PICT2021-GRC401 awarded to RC and MG, respectively, and UNMdP grant EXA1063/22 awarded to RC.

Acknowledgments

We want to thank the Bisson Lab, and in particular to John Mallon, for assistance with image processing and cell shape analysis. We also thank CONICET for awarding a postdoctoral fellowship to MIC and a doctoral fellowship to SF.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The authors declare that no Gen AI was used in the creation of this manuscript.

Publisher's note

Supplementary material

The Supplementary material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2025.1547649/full#supplementary-material

References

Abdul-Halim, M. F., Karch, K. R., Zhou, Y., Haft, D. H., Garcia, B. A., and Pohlschroder, M. (2015). 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. doi: 10.1128/JB.00849-15

Allers, T. (2009). The Halohandbook: Protocols for haloarchaeal genetics. Melbourne, VI, Australia: Haloarchaeal Genetics Laboratory.

Allers, T., Barak, S., Liddell, S., Wardell, K., and Mevarech, M. (2010). Improved strains and plasmid vectors for conditional overexpression of his-tagged proteins in *Haloferax volcanii*. *Appl. Environ. Microbiol.* 76, 1759–1769. doi: 10.1128/AEM.02670-09

Allers, T., Ngo, H. P., Mevarech, M., and Lloyd, R. G. (2004). Development of additional selectable markers for the halophilic archaeon *Haloferax volcanii* based on the leuB and trpA genes. *Appl. Environ. Microbiol.* 70, 943–953. doi: 10.1128/AEM.70.2.943-953.2004

Baker, R. P., Wijetilaka, R., and Urban, S. (2006). Two *plasmodium* rhomboid proteases preferentially cleave different adhesins implicated in all invasive stages of malaria. *PLoS Pathog.* 2:e113. doi: 10.1371/journal.ppat.0020113

Baxt, L. A., Baker, R. P., Singh, U., and Urban, S. (2008). An *Entamoeba histolytica* rhomboid protease with atypical specificity cleaves a surface lectin involved in phagocytosis and immune evasion. *Genes Dev.* 22, 1636–1646. doi: 10.1101/gad.1667708

Began, J., Cordier, B., Brezinova, J., Delisle, J., Hexnerova, R., et al. (2020). Rhomboid intramembrane protease YqgP licenses bacterial membrane protein quality control as adaptor of FtsH AAA protease. *EMBO J.* 39:e102935. doi: 10.15252/embj.2019102935

Borrell, G., Fadhhlaoui, K., Ben Hania, W., Gaci, N., Pehau-Arnaudet, G., Chaudhary, P. P., et al. (2023). *Methanomethylophilus alvi* gen. Nov., sp. nov., a novel Hydrogenotrophic methyl-reducing methanogenic Archaea of the order *Methanomassiliicoccales* isolated from the human gut and proposal of the novel family *Methanomethylophilaceae* fam. nov. *Microorganisms* 11:2794. doi: 10.3390/microorganisms11112794

Brossier, F., Jewett, T. J., Sibley, L. D., and Urban, S. (2005). A spatially localized rhomboid protease cleaves cell surface adhesins essential for invasion by *toxoplasma*. *Proc. Natl. Acad. Sci. USA* 102, 4146–4151. doi: 10.1073/pnas.0407918102

Carpenter, A. E., Jones, T. R., Lamprecht, M. R., Clarke, C., Kang, I. H., Friman, O., et al. (2006). CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biol.* 7:R100. doi: 10.1186/gb-2006-7-10-r100

Cheng, T. L., Lai, C. H., Jiang, S. J., Hung, J. H., Liu, S. K., Chang, B. I., et al. (2014). RHBDL2 is a critical membrane protease for anoikis resistance in human malignant epithelial cells. *ScientificWorldJournal* 2014:902987. doi: 10.1155/2014/902987

Chimileski, S., Franklin, M. J., and Papke, R. T. (2014). Biofilms formed by the archaeon *Haloferax volcanii* exhibit cellular differentiation and social motility, and facilitate horizontal gene transfer. *BMC Biol.* 12:65. doi: 10.1186/s12915-014-0065-5

Civitarese, A. E., Maclean, P. S., Carling, S., Kerr-Bayles, L., McMillan, R. P., Pierce, A., et al. (2010). Regulation of skeletal muscle oxidative capacity and insulin signaling by the mitochondrial rhomboid protease PARL. *Cell Metab.* 11, 412–426. doi: 10.1016/j.cmet.2010.04.004

Costa, M. I., Cerletti, M., Paggi, R. A., Trotschel, C., De Castro, R. E., Poetsch, A., et al. (2018). *Haloferax volcanii* proteome response to deletion of a rhomboid protease gene. *J. Proteome Res.* 17, 961–977. doi: 10.1021/acs.jproteome.7b00530

Curtis, Z., Escudeiro, P., Mallon, J., Leland, O., Rados, T., Dodge, A., et al. (2024). Halofilins as emerging bactofilin families of archaeal cell shape plasticity orchestrators. Available online at: doi: https://doi.org/10.1101/2024.01.22.576759.

Cutler, K. J., Stringer, C., Lo, T. W., Rappez, L., Stroustrup, N., Peterson, S. B., et al. (2022). Omnipose: a high-precision morphology-independent solution for bacterial cell segmentation. *Nat. Methods* 19, 1438–1448. doi: 10.1038/s41592-022-01639-4

de Silva, R. T., Abdul-Halim, M. F., Pittrich, D. A., Brown, H. J., Pohlschroder, M., and Duggin, I. G. (2021). Improved growth and morphological plasticity of *Haloferax volcanii*. *Microbiology* 167:1012. doi: 10.1099/mic.0.001012

Esquivel, R. N., Schulze, S., Xu, R., Hippler, M., and Pohlschroder, M. (2016). Identification of *Haloferax volcanii* pilin N-Glycans with diverse roles in pilus biosynthesis, adhesion, and microcolony formation. *J. Biol. Chem.* 291, 10602–10614. doi: 10.1074/jbc.M115.693556

Esquivel, R. N., Xu, R., and Pohlschroder, M. (2013). Novel archaeal adhesion pilins with a conserved N terminus. *J. Bacteriol.* 195, 3808–3818. doi: 10.1128/JB.00572-13

Kandel, R. R., and Neal, S. E. (2020). The role of rhomboid superfamily members in protein homeostasis: mechanistic insight and physiological implications. *Biochim. Biophys. Acta, Mol. Cell Res.* 1867:118793. doi: 10.1016/j.bbamcr.2020.118793

Kateete, D. P., Katabazi, F. A., Okeng, A., Okee, M., Musinguzi, C., Asimwe, B. B., et al. (2012). Rhomboids of mycobacteria: characterization using an aarA mutant of *Providencia stuartii* and gene deletion in *Mycobacterium smegmatis*. *PLoS One* 7:e45741. doi: 10.1371/journal.pone.0045741

Kühnle, N., Diederichs, K., and Breidenstein, A. (2019). Intramembrane proteases: Mechanisms and function in physiology and disease. *Biochimica et Biophysica Acta (BBA) - Biomembranes*, 1861:720–732. doi: 10.1016/j.bbamem.2018.11.008

Liu, G., Beaton, S. E., Grieve, A. G., Evans, R., Rogers, M., Strisovsky, K., et al. (2020). Bacterial rhomboid proteases mediate quality control of orphan membrane proteins. *EMBO J.* 39:e102922. doi: 10.15252/embj.2019102922

Lliao, Y., Ithurbide, S., Evenhuis, C., Lowe, J., and Duggin, I. G. (2021). Cell division in the archaeon *Haloferax volcanii* relies on two FtsZ proteins with distinct functions in division ring assembly and constriction. *Nat. Microbiol.* 6, 594–605. doi: 10.1038/s41564-021-00894-z

Luenenschloss, A., Ter Veld, F., Albaum, S. P., Neddermann, T. M., Wendisch, V. F., and Poetsch, A. (2022). Functional genomics uncovers pleiotropic role of rhomboids in *Corynebacterium glutamicum*. *Front. Microbiol.* 13:771968. doi: 10.3389/fmicb.2022.771968

Marchesini, M. I., Poetsch, A., Guidolin, L. S., and Comerci, D. J. (2022). *Brucella abortus* encodes an active rhomboid protease: proteome response after rhomboid gene deletion. *Microorganisms* 10:114. doi: 10.3390/microorganisms10010114

Mesak, L. R., Mesak, F. M., and Dahl, M. K. (2004). Expression of a novel gene, gluP, is essential for normal *Bacillus subtilis* cell division and contributes to glucose export. *BMC Microbiol.* 4:13. doi: 10.1186/1471-2180-4-13

Noy, P. J., Swain, R. K., Khan, K., Lodhia, P., and Bicknell, R. (2016). Sprouting angiogenesis is regulated by shedding of the C-type lectin family 14, member a (CLEC14A) ectodomain, catalyzed by rhomboid-like 2 protein (RHBDL2). *FASEB J.* 30, 2311–2323. doi: 10.1096/fj.201500122R

Nussbaum, P., Gerstner, M., Dingethal, M., Erb, C., and Albers, S. V. (2021). The archaeal protein SepF is essential for cell division in *Haloferax volcanii*. *Nat. Commun.* 12:3469. doi: 10.1038/s41467-021-23686-9

O'Donnell, R. A., Hackett, F., Howell, S. A., Treeck, M., Struck, N., Krnajski, Z., et al. (2006). Intramembrane proteolysis mediates shedding of a key adhesin during erythrocyte invasion by the malaria parasite. *J. Cell Biol.* 174, 1023–1033. doi: 10.1083/jcb.200604136

Palma, V., Gutiérrez, M. S., Vargas, O., Parthasarathy, R., and Navarrete, P. (2022). Methods to evaluate bacterial motility and its role in bacterial-host interactions. *Microorganisms* 10:563. doi: 10.3390/microorganisms10030563

Parente, J., Casabuono, A., Ferrari, M. C., Paggi, R. A., De Castro, R. E., Couto, A. S., et al. (2014). A rhomboid protease gene deletion affects a novel oligosaccharide N-linked to the S-layer glycoprotein of *Haloferax volcanii*. *J. Biol. Chem.* 289, 11304–11317. doi: 10.1074/jbc.M113.546531

Paschkowsky, S., Hamze, M., Oestereich, F., and Munter, L. M. (2016). Alternative processing of the amyloid precursor protein family by rhomboid protease RHBDL4. *J. Biol. Chem.* 291, 21903–21912. doi: 10.1074/jbc.M116.753582

Patro, M., Duggin, I. G., Albers, S.-V., and Ithurbide, S. (2023). Influence of plasmids, selection markers and auxotrophic mutations on *Haloferax volcanii* cell shape plasticity. *Front. Microbiol.* 14:1270665. doi: 10.3389/fmicb.2023.1270665

Radaelli, E., Assenmacher, C. A., Verrelle, J., Banerjee, E., Manero, F., Khiati, S., et al. (2023). Mitochondrial defects caused by PARL deficiency lead to arrested spermatogenesis and ferroptosis. *eLife* 12:4710. doi: 10.7554/eLife.84710

Rather, P. N., and Orosz, E. (1994). Characterization of aarA, a pleiotrophic negative regulator of the 2'-N-acetyltransferase in *Providencia stuartii*. *J. Bacteriol.* 176, 5140–5144. doi: 10.1128/jb.176.16.5140-5144.1994

Reddy, T., and Rainey, J. K. (2012). Multifaceted substrate capture scheme of a rhomboid protease. *J. Phys. Chem. B* 116, 8942–8954. doi: 10.1021/jp305077k

R Core Team (2019). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available at: https://www.R-project.org/

Schiller, H., Schulze, S., Mutan, Z., De Vaux, C., Runcie, C., Schwartz, J., et al. (2020). *Haloferax volcanii* immersed liquid biofilms develop independently of known biofilm machineries and exhibit rapid honeycomb pattern formation. *mSphere* 5:e00976-20. doi: 10.1128/mSphere.00976-20

Shi, G., Lee, J. R., Grimes, D. A., Racacho, L., Ye, D., Yang, H., et al. (2011). Functional alteration of PARL contributes to mitochondrial dysregulation in Parkinson's disease. *Hum. Mol. Genet.* 20, 1966–1974. doi: 10.1093/hmg/ddr077

Shukla, R. T. (2017). An improved crystal violet assay for biofilm quantification in 96-well Microtitre plate. Available online at https://doi.org/10.1101/100214.

Song, W., Liu, W., Zhao, H., Li, S., Guan, X., Ying, J., et al. (2015). Rhomboid domain containing 1 promotes colorectal cancer growth through activation of the EGFR signalling pathway. *Nat. Commun.* 6:8022. doi: 10.1038/ncomms9022

Stevenson, L. G., Strisovsky, K., Clemmer, K. M., Bhatt, S., Freeman, M., and Rather, P. N. (2007). Rhomboid protease AarA mediates quorum-sensing in *Providencia stuartii* by activating TatA of the twin-arginine translocase. *Proc. Natl. Acad. Sci. U. S. A.* 104, 1003–1008. doi: 10.1073/pnas.0608140104

Strakova, D., Galisteo, C., De La Haba, R. R., and Ventosa, A. (2023). Characterization of *Haloarcula terrestris* sp. nov. and reclassification of a *Haloarcula* species based on a taxogenomic approach. *Int. J. Syst. Evol. Microbiol.* 73:157. doi: 10.1099/ijsem.0.006157

Takashina, T. H., Otozai, K., Grant, W. D., and Horikoshi, K. (1990). *Haloarcula japonica* sp. nov., a new triangular halophilic Archaebacterium. *Syst. Appl. Microbiol.* 13, 177–181. doi: 10.1016/S0723-2020(11)80165-7

Tripepi, M., Esquivel, R. N., Wirth, R., and Pphlschroder, M. (2013). *Haloferax volcanii* cells lacking the flagellin FlgA2 are hypermotile. *Microbiology* 159, 2249–2258. doi: 10.1099/mic.0.069617-0

Tripepi, M., Imam, S., and Pohlschroder, M. (2010). *Haloferax volcanii* flagella are required for motility but are not involved in PibD-dependent surface adhesion. *J. Bacteriol.* 192, 3093–3102. doi: 10.1128/JB.00133-10

Tripepi, M., You, J., Temel, S., Onder, O., Brisson, D., and Pohlschroder, M. (2012). N-glycosylation of *Haloferax volcanii* flagellins requires known Agl proteins and is essential for biosynthesis of stable flagella. *J. Bacteriol.* 194, 4876–4887. doi: 10.1128/JB.00731-12

Walder, K., Kerr-Bayles, L., Civitarese, A., Jowett, J., Curran, J., Elliot, K., et al. (2005). The mitochondrial rhomboid protease PSARL is a new candidate gene for type 2 diabetes. *Diabetologia* 48, 459–468. doi: 10.1007/s00125-005-1675-9

Walsby, A. E. (1980). A square bacterium. *Nature* 283, 69–71. doi: 10.1038/283069a0

Wang, Y., Zhang, Y., and Ha, Y. (2006). Crystal structure of a rhomboid family intramembrane protease. *Nature* 444, 179–180. doi: 10.1038/nature05255

Frontiers in Microbiology

Explores the habitable world and the potential of microbial life

The largest and most cited microbiology journal which advances our understanding of the role microbes play in addressing global challenges such as healthcare, food security, and climate change.

Discover the latest Research Topics

See more →

Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact