










## ORIGINAL ARTICLE

# Interaction of two strongly divergent archaeellins stabilizes the structure of the *Halorubrum* archaeellum

Mikhail G. Pyatibratov<sup>1</sup>  | Alexey S. Syutkin<sup>1</sup>  | Tessa E. F. Quax<sup>2</sup>  |  
 Tatjana N. Melnik<sup>1</sup>  | R. Thane Papke<sup>3</sup> | Johann Peter Gogarten<sup>3</sup>  | Igor I. Kireev<sup>4</sup>  |  
 Alexey K. Surin<sup>1,5,6</sup>  | Sergei N. Beznosov<sup>1</sup>  | Anna V. Galeva<sup>1</sup>  | Oleg V. Fedorov<sup>1†</sup>

<sup>1</sup>Institute of Protein Research, Russian Academy of Sciences, Pushchino, Moscow Region, Russia

<sup>2</sup>Archaeal Virus-Host Interactions, Faculty of Biology, University of Freiburg, Freiburg, Germany

<sup>3</sup>Department of Molecular and Cell Biology, University of Connecticut, Storrs, CT, USA

<sup>4</sup>A.N. Belozersky Institute of Physico-chemical Biology, M.V. Lomonosov Moscow State University, Moscow, Russia

<sup>5</sup>Pushchino Branch, Shemyakin-Ovchinnikov Institute of Bioorganic Chemistry, Russian Academy of Sciences, Pushchino, Moscow Region, Russia

<sup>6</sup>State Research Center for Applied Microbiology & Biotechnology, Obolensk, Moscow Region, Russia

**Correspondence**

Mikhail G. Pyatibratov, Institute of Protein Research, RAS, Institutskaya st., 4, Pushchino, Moscow Region 142290, Russia. Email: bratov@vega.protres.ru

**Funding information**

Russian Foundation for Basic Research, Grant/Award Number: 19-04-00613 A; RFBR, Grant/Award Number: 19-04-01327 A; Emmy Noether grant from the Deutsche Forschungsgemeinschaft, Grant/Award Number: 411069969; MSU Development Program, Grant/Award Number: PNR 5.13; and United Pushchino Center "Structural and functional studies of proteins and RNA," Grant/Award Number: 584307.

**Abstract**

Halophilic archaea from the genus *Halorubrum* possess two extraordinarily diverged archaeellin genes, *flaB1* and *flaB2*. To clarify roles for each archaeellin, we compared two natural *Halorubrum lacusprofundi* strains: One of them contains both archaeellin genes, and the other has the *flaB2* gene only. Both strains synthesize functional archaeella; however, the strain, where both archaeellins are present, is more motile. In addition, we expressed these archaeellins in a *Haloferax volcanii* strain from which the endogenous archaeellin genes were deleted. Three *Hfx. volcanii* strains expressing *Hrr. lacusprofundi* archaeellins produced functional filaments consisting of only one (FlaB1 or FlaB2) or both (FlaB1/FlaB2) archaeellins. All three strains were motile, although there were profound differences in the efficiency of motility. Both native and recombinant FlaB1/FlaB2 filaments have greater thermal stability and resistance to low salinity stress than single-component filaments. Functional supercoiled *Hrr. lacusprofundi* archaeella can be composed of either single archaeellin: FlaB2 or FlaB1; however, the two divergent archaeellin subunits provide additional stabilization to the archaeellum structure and thus adaptation to a wider range of external conditions. Comparative genomic analysis suggests that the described combination of divergent archaeellins is not restricted to *Hrr. lacusprofundi*, but is occurring also in organisms from other haloarchaeal genera.

**KEYWORDS**

archaea, archaeellin, archaeellum, *Haloferax*, *Halorubrum*, motility

†Deceased 25 January 2020.

Mikhail G. Pyatibratov and Alexey S. Syutkin are co-first authors.

This paper is dedicated to our mentor and colleague, Oleg V. Fedorov.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2020 The Authors. *MicrobiologyOpen* published by John Wiley & Sons Ltd.

## 1 | INTRODUCTION

Archaeal flagella (archaella) are morphologically and functionally similar to bacterial flagella. However, the archaellum structure, assembly mechanism, and protein composition are fundamentally different from the flagellum and instead show similarity to type IV pili. Fundamental differences between archaeal and bacterial flagellar filaments are the presence of signal sequences in archaeal flagellins (archaellins) and the multiplicity of archaellin encoding genes in archaea (Jarrell & Albers, 2012). Recently, cryoelectron microscopy was used to determine the spatial structures of the archaella filaments of three archaea: two methanogens *Methanospirillum hungatei* (3.4 Å resolution) (Poweleit et al., 2016) and *Methanococcus maripaludis* (4 Å resolution) (Meshcheryakov et al., 2019) and a hyperthermophile *Pyrococcus furiosus* (4.2 Å resolution) (Daum et al., 2017). Furthermore, the crystal structure of the N-terminal truncation of archaellin FlaB1 *Methanocaldococcus jannaschii* has been determined at a resolution of 1.5 Å (Meshcheryakov et al., 2019). The structure of archaeal filaments differs significantly not only from bacterial flagella but also from bacterial type IV pili (Braun et al., 2016; Poweleit et al., 2016). The amino acid residues of archaellins responsible for intersubunit interactions were identified, as well as the protein regions forming the outside surface of the filament. The proposed models for the archaellar filament do not contain the long-pitch protofilaments found in bacterial flagellar filaments. To explain the archaella supercoiling, it was proposed to consider them as semiflexible filaments in a viscous medium (Coq, Du Roure, Marthelot, Bartolo, & Ferm, 2008; Tony, Lauga, & Hosoi, 2006; Wolgemuth, Powers, & Goldstein, 2000). For such structures, thrust can be generated by their rotation. Using molecular modeling, it was shown that conformational changes in the globular domain of the archaellin can lead to extension and compression, as well as bending of the filaments (Braun et al., 2016). However, the detailed mechanism of the archaellum supercoiling is not fully understood.

The proposed models of spatial archaellar filament structure do not explain the structural and functional role of multiple archaellins. Despite the presence of several archaellin genes in genomes of *M. hungatei*, *M. maripaludis*, and *P. furiosus*, protein products of only one of these genes were found incorporated in their filaments (Daum et al., 2017; Meshcheryakov et al., 2019; Poweleit et al., 2016). This raises the question about the importance of encoding multiple different archaellins. Interestingly, the presence of several copies of archaellin genes in archaeal genomes is very common.

Currently, almost 3,000 archaeal genomic sequences are deposited at the NCBI database (<https://www.ncbi.nlm.nih.gov/genome/browse/#!/prokaryotes/archaea>), including about 400 genomes of halophilic archaea. The majority of these archaeal genomes contain archaellin genes. The known genomes of crenarchaeota typically only possess a single archaellin gene. However, the large majority of euryarchaeal genomes contain multiple archaellin genes.

In haloarchaea, multiple archaellin genes appear to have originated from duplication events (Desmond, Brochier-Armanet, & Gribaldo, 2007). Duplicated genes can be located in the same or

different operons. For example, in three popular model haloarchaea, the organization of archaellin genes differs significantly. *Haloarcula hispanica* has three operons with 1 archaellin gene in each of them, two *Haloferax volcanii* archaellin genes are in one operon, and *Halobacterium salinarum* has two operons with 2 and 3 archaellin genes each. The similarity between the archaellin paralogs of the above species is still very high, indicating relatively recent duplication. Interestingly, several haloarchaea have multiple archaellin genes (most often, two) that are very divergent (such as *Halobiforma*, *Halopiger*, *Halorubrum*, *Natrialba*, and *Natronolimnobius* species). Even archaea belonging to the same genus can differ drastically from each other in the number and size of the archaellin genes. It has been suggested that the archaellum supercoiling, as for bacterial flagella, can be achieved through a combination of subfilaments of different lengths, constructed from different types of subunits (Tarasov, Pyatibratov, Beznosov, & Fedorov, 2004; Tarasov, Pyatibratov, Tang, Dyall-Smith, & Fedorov, 2000). For *Hbt. Salinarum*, it was demonstrated that both archaellins FlgA1 and FlgA2 are necessary for the formation of a functional supercoiled archaellum, and mutant strains with a single FlgA1 or FlgA2 archaellin had straight nonfunctional filaments. In the case of methanogenic archaea, the multiple archaellin genes were shown to encode major and minor structural components of the archaellum filament. A “hook”-like structure was observed in a number of methanogenic archaea. In *Methanococcus voltae* and *M. maripaludis*, the archaellum hook segment is built of the FlaB3 archaellin, while the FlaB1 and FlaB2 proteins are the main components of the filament (Bardy, Mori, Komoriya, Aizawa, & Jarrell, 2002; Chaban et al., 2007). Inactivation of either the *flaB1* or *flaB2* genes resulted in a loss of motility and cessation of archaellum synthesis (including the hook) (Chaban et al., 2007). Recently, it has been shown that FlaB1 is the predominant component of *Methanococcus maripaludis* filaments (Meshcheryakov et al., 2019). Inactivation of the *flaB3* gene does not lead to the cessation of filament synthesis and a noticeable change in motility on semisolid agar (Chaban et al., 2007). However, time-lapse microscopy showed impaired motility for this deletion strain (movement in a closed circle). The structures corresponding to the *Methanococcales* hooks were not found in the native archaella of other archaea. It is possible that the differentiation of one of the archaellins into a “hook protein” with a special structural role is a relatively late evolutionary event in the *Methanococcales* and is not typical for other archaea.

In *Hfx. Volcanii*, it was shown that the archaellum filament consists of one major (FlgA1) and one minor (FlgA2) component. However, the structural role of the minor component is unknown and it does not form a hook-like structure (Tripepi, Esquivel, Wirth, & Pohlschröder, 2013). Deletion of the *flgA2* gene leads to hypermotile cells by an unknown mechanism (Tripepi et al., 2013).

In contrast to all abovementioned examples, some halophilic archaea can form functional archaella with only one of the genomically encoded archaellin proteins (Pyatibratov et al., 2008; Syutkin, Pyatibratov, Beznosov, & Fedorov, 2012; Syutkin, Pyatibratov, Galzitskaya, Rodríguez-Valera, & Fedorov, 2014). These archaellins were thought to be partially redundant. However, in the case of

TABLE 1 Plasmids and strains

Plasmid/strain	Relevant properties	Reference
Plasmids		
pTA1228	Amp <sup>r</sup> , <i>pyrE2</i> and <i>hdrB</i> markers, inducible <i>ptna</i> promoter	Brendel et al. (2014)
pMT21	pTA963 (Amp <sup>r</sup> , <i>pyrE2</i> and <i>hdrB</i> markers, inducible <i>ptna</i> promoter) containing <i>flgA1flgA2His</i>	Tripepi et al. (2013)
pAS1	pTA1228 containing <i>flaB1B2</i> of <i>Hrr. saccharovororum</i>	This study
pAS2	pTA1228 containing <i>flaB1</i> of <i>Hrr. saccharovororum</i>	This study
pAS3	pTA1228 containing <i>flaB2</i> of <i>Hrr. saccharovororum</i>	This study
pAS4	pTA1228 containing <i>flaB2</i> of <i>Hrr. saccharovororum</i> (signal peptide was replaced to the FlaB1)	This study
pAS5	pTA1228 containing <i>flaB1B2</i> of <i>Hrr. lacusprofundi</i> DL18	This study
pAS6	pTA1228 containing <i>flaB1</i> of <i>Hrr. lacusprofundi</i> DL18	This study
pAS7	pTA1228 containing <i>flaB2</i> of <i>Hrr. lacusprofundi</i> DL18	This study
Strains		
DH5 $\alpha$	<i>E. coli</i> F <sup>-</sup> $\phi$ 80lacZDM15 ( <i>lacZYA-argF</i> ) U169 <i>recA1 endA1 hsdR17 (rK<sup>-</sup> mK<sup>+</sup>) phoA supE44 thi-1 gyrA96 relA1</i>	Invitrogen
MT2	<i>Hfx. volcanii</i> $\Delta$ <i>pyrE2</i> $\Delta$ <i>trp</i> $\Delta$ <i>flgA1</i> $\Delta$ <i>flgA2</i>	Tripepi et al. (2010)
MT45	MT2 containing pMT21	Tripepi et al. (2013)
AS1	MT2 containing pAS1	This study
AS2	MT2 containing pAS2	This study
AS3	MT2 containing pAS3	This study
AS4	MT2 containing pAS4	This study
AS5	MT2 containing pAS5	This study
AS6	MT2 containing pAS6	This study
AS7	MT2 containing pAS7	This study
ACAM 34	<i>Halorubrum lacusprofundi</i> (B-1753, ATCC 49,239)	Franzmann et al. (1988)
DL18	<i>Halorubrum lacusprofundi</i>	Erdmann et al. (2017)
ATCC 29,252	<i>Halorubrum saccharovororum</i> (B-1747)	Tomlinson and Hochstein (1976)

*Haloarcula marismortui* it was shown that these proteins function as ecoparalogs; that is, they are expressed under different environmental conditions and provide distinct stability advantages under varying salt concentrations (Syutkin et al., 2014, 2019).

In this work, we investigate the role of multiple archaellin genes of the *Halorubrum* genus. In contrast to the systems studied before, members of the *Halorubrum* group possess multiple archaellins with highly diverged protein sequences. Our preceding work has shown that functional supercoiled archaella filaments of *Hrr. lacusprofundi* ATCC49239 (ACAM 34) are formed from a protein encoded by a single archaellin gene (*flaB2*) (Syutkin et al., 2012). However, unlike *Hrr. lacusprofundi* ACAM 34, other *Halorubrum* species possess at least two archaellin genes (*flaB1* and *flaB2*) located in one operon. The

amino acid sequences of the FlaB1 and FlaB2 archaellins differ significantly from each other (<43% identical residues, the N-terminal region being more conserved). We use the *Hrr. lacusprofundi* archaella as a model to address the role of multiple highly divergent archaellin genes often found in genomes of haloarchaea.

Recently, *Hrr. lacusprofundi* strains (DL18 and R1S1) with two archaellin genes (which is more typical for *Halorubrum* species) were isolated (Tschitschko et al., 2018). Since the presence of a single archaellin gene in *Hrr. lacusprofundi* ACAM 34 is sufficient to form functional supercoiled archaella, the presence of the *flaB1* gene may seem redundant. FlaB1 could be responsible for: (a) formation of specific filaments that differ from FlaB2 in function (and, e.g., function as ecoparalogs) and (b) stabilization of the filament structure

together with FlaB, possibly as a result of constructive neutral evolution (Lukeš, Archibald, Keeling, Doolittle, & Gray, 2011).

In the present study, we characterized the archaella of the *Hrr. lacusprofundi* DL18 strain, containing two archaellin genes *flaB1* and *flaB2*, and compared them with archaella of the ACAM 34 strain whose FlaB2 is (except the signal sequence) completely identical to its counterpart from the DL18 strain. To clarify the role of FlaB1 and FlaB2 archaellins, we used the well-developed expression system of *Hfx. volcanii*. With this approach, we could show that either the FlaB1 or FlaB2 protein is sufficient to form functional archaella. However, the combination of the two proteins renders the archaellum filament structure much more stable which can help maintain motility in a wider range of conditions.

## 2 | MATERIALS AND METHODS

### 2.1 | Strains and growth conditions

The plasmids and strains used in this study are listed in Table 1. The strains *Halorubrum lacusprofundi* B-1753 (ACAM 34, ATCC 49,239) and *Halorubrum saccharovororum* B-1747 (ATCC 29,252) were from the All-Russian Collection of Microorganisms (VKM), Pushchino; strain *Halorubrum lacusprofundi* DL18 (Erdmann, Tschitschko, Zhong, Raftery, & Cavicchioli, 2017) was kindly provided by R. Cavicchioli; and strain *Haloferax volcanii* MT2 (Tripepi, Imam, & Pohlschröder, 2010) was kindly provided by M. Pohlschröder.

The *Hrr. lacusprofundi* cells were grown under moderate aeration at 37°C in a liquid medium containing 0.5% casamino acids, 0.5% yeast extract, 3.42 M (20%) NaCl, 27 mM KCl, 80 mM MgSO<sub>4</sub>, 12 mM sodium citrate, 6 mM sodium glutamate, and pH 7.2. Filter-sterilized aqueous solution of microelements (1.7 ml) containing 0.9 mM MnCl<sub>2</sub> · 7H<sub>2</sub>O and 17 mM FeCl<sub>3</sub> · 7H<sub>2</sub>O was added to 1 L of the medium after autoclaving. All *Hfx. volcanii* transformed strains were grown at 37°C in liquid or solid/semisolid agar medium containing 0.5% casamino acids, 2.91 M (17%) NaCl, 0.15 M MgSO<sub>4</sub>, 1 mM MnCl<sub>2</sub>, 50 mM KCl, 3 mM CaCl<sub>2</sub>, pH 7.2 (Mod-HV medium), and 1.2% (solid) or 0.24% (semisolid) agar. Heterologous archaella synthesis was induced by addition of tryptophan (concentration of 0.2–1 mg/ml).

In experiments testing motility comparing at various salinity, we used Mod-HV media with NaCl concentrations of 10, 15, 20, and 25% (1.71, 2.57, 3.42, and 4.28 M), while the concentrations of the remaining components did not change. The swarming diameters on semisolid agar plates were measured daily. Ten biological replicates were performed, and average diameter values and standard deviations were calculated. Data were analyzed using GraphPad Prism 8.0 (GraphPad Software Inc., San Diego, USA).

To isolate recombinant archaella from liquid medium (with the same composition, but 0.5% casamino acids were replaced by 1% yeast extract), a piece of agar from the edge of the spot was used for inoculation.

Life cell imaging was performed in the CA medium as described by Quax et al. (2018). In short, cells were grown until an OD of ~0.05

and imaged in a round DF 0.17 216mm microscopy dish (Bioprotechs) and observed at 20x magnification in the PH2 mode with a Zeiss Axio Observer 2.1 Microscope equipped with a heated XL-5 2000 Incubator running VisiView® software.

Time-lapse movies were recorded for 15 s, and the X, Y coordinates of cells were tracked with MetaMorph® software. From the X, Y coordinates, the average velocity of a given time frame was calculated for each cell using the Pythagoras theorem. Also the X, Y coordinates were used to calculate the number of turns larger than 90° that each cell made per second in the total time it was tracked. In case no 90° turn was observed within the time frame of the 15-s movie, the cell was automatically assigned the value >16. Velocity and turns per second values were averaged for technical replicates recorded in a single experiment on a single day. Each experiment was performed at least on three independent occasions. The percentage of motile cells was calculated by counting cells that displayed motility in a 15-s time-lapse movie and dividing this by the total number of visible cells in the frame of this movie. This was done for at least 10 movies (displaying ~50 cells each) for each strain for each experiment. The experiment was performed on at least three independent occasions.

### 2.2 | Preparation of DNA and polymerase chain reaction (PCR)

The plasmids for heterologous archaellin expression were assembled by the SLIC method (Li & Elledge, 2012) with modifications. These expression vectors included the inducible tryptophanase promoter (*tna*) to drive expression of these genes. The DNA fragments containing desired archaellin genes were amplified from *Hrr. lacusprofundi* or *Hrr. saccharovororum* genome with the primers described in Table 2. The PCR was performed with Q5 High-Fidelity DNA Polymerase (New England Biolabs); temperatures were as follows: 98°C for 30 s for initial denaturation, 25 cycles: 10 s at 98°C, 20 s at 65°C, 90 s at 72°C, and then 2 min at 72°C for final elongation. Resulting PCR products were purified from the reaction mixture and mixed with the pTA1228 vector preliminary linearized by NdeI treatment. Each assembly mix contained 100 fmol of both vector and insert, as well as 1.5 U of T4 DNA polymerase and NEBuffer 2.1 (New England Biolabs). The reaction was incubated for 4 min at 22°C and then stopped by the addition of 10 mM dCTP. The resulting mix was used for *E. coli* transformation. The colonies that appeared on the next day were analyzed by PCR with the pTA1228\_seqF and pTA1228\_seqR primers. The plasmids from positive colonies were isolated, and correct assembly of the plasmid was confirmed by sequencing with the primers used for colonies screening.

### 2.3 | Isolation of archaellar filaments

Archaellar filaments were prepared by precipitation with polyethylene glycol 6,000 (Gerl, Deutzmann, & Sumper, 1989). The protein preparations were dissolved in 10 mM Na-phosphate, pH 8.0

TABLE 2 Primers

Primer name	Sequence (5'-3')
pAS1_F	TCACATTTCGCGGACCTATTGCGCATATGTTTCGAAACAATACTCGACGAG
pAS1_R	CATGTGGTGGTGGTGGTGGTGCATATTAGAGCCGGACCGCTT
pAS2_R	CATGTGGTGGTGGTGGTGGTGCATATTACAGCCTCACCGAAGTCT
pAS3_F	TCACATTTCGCGGACCTATTGCGCATATGTTTCGAGTTTATCAACGACAATGA
pAS4_F	TCACATTTCGCGGACCTATTGCGCATATGTTTCGAAACAATACTCGACGAGGAAGAGCGCGGTCAAGTCCG
pAS5_F	TCACATTTCGCGGACCTATTGCGCATATGTTTCGAAACAATACTGAACGA
pAS5_R	CATGTGGTGGTGGTGGTGGTGCATATTAGAGCCGGACCGCTT
pAS6_R	CATGTGGTGGTGGTGGTGGTGCATATTACAGCCTCACCGAGG
pAS7_F	TCACATTTCGCGGACCTATTGCGCATATGTTTCGAGTTTATCAACAACGACA
pTA1228_seqF	GTCCTCGAAAGTGACATCGCTC
pTA1228_seqR	GGCCGCTCTAGAAGTAGTGGAT

containing appropriate NaCl concentrations (0%–20%) at a concentration of 0.5–1.0 mg/ml. SDS-PAGE was performed using 9%–12% acrylamide gels. The proteins were stained with Coomassie brilliant blue G 250. To prepare samples for microcalorimetry, the archaeal filaments were also dialyzed against the abovementioned buffer solutions. Protein concentrations were determined using the Coomassie Plus Protein Assay Reagent Kit (Pierce, IL) according to the manufacturer's protocol. ImageJ software (NIH) was used to scan stained acrylamide gels and estimate FlaB1/FlaB2 molar ratios based on the measured densities of the corresponding bands and FlaB1/FlaB2 molecular weights.

## 2.4 | Chromatography mass spectrometry analysis

Protein bands were excised and treated with proteinase K (Promega) and trypsin (Sigma) at 37°C in a Thermo Mixer thermo shaker (Eppendorf, Germany). To stabilize proteinase K, CaCl<sub>2</sub> was added to the solution to a final concentration of 5 mM. The molar ratio of enzyme-to-protein was 1/50. The reaction was stopped by adding trifluoroacetic acid to the solution. Before mass spectrometric analysis, the peptides were separated by reversed-phase high-performance liquid chromatography using an Easy-nLC 1,000 Nanoliquid chromatography (Thermo Fisher Scientific). The separation was carried out in a homemade column 25 cm in length and 100 µm in diameter packed with a C18 adsorbent, with an adsorbent particle size of 3 µm, and pore size of 300 Å. The column was packed under laboratory conditions at a pressure of 500 atm. The peptides were eluted in a gradient of acetonitrile from 3% to 40% for 180 min; the mobile phase flow rate was 0.3 µl/min. Mass spectra of the samples were obtained using an OrbiTrap Elite mass spectrometer (Thermo Scientific, Germany). The peptides were ionized by electrospray at nano-liter flow rates with 2 kV ion spray voltage; ion fragmentation was induced by collisions with an inert gas (collision-induced dissociation in a high-energy cell). The mass spectra were processed, and peptides were identified using Thermo Xcalibur Qual Browser and PEAKS Studio (ver. 7.5)

programs based on the sequences of UniRef-100. Parent Mass Error Tolerance was 2.0 ppm, and fragment Mass Error Tolerance was 0.1 Da. Only peptides were taken into account with a "10 L gP." threshold value higher than 15.

## 2.5 | Electron microscopy

The archaeal filament specimens were prepared by negative staining with 2% uranyl acetate on Formvar-coated copper grids. A grid was floated on a 20-µl drop of filament solution (about 0.01 mg/ml, in 20% NaCl, 10 mM Na-phosphate, pH 8.0) for 2 min, blotted with filter paper, placed on top of a drop of 2% uranyl acetate, and left for 1–1.5 min. Excess stain was removed by touching the grid to filter paper, and the grid was air-dried. Samples were examined on a Jeol JEM-1400 transmission electron microscope (JEOL, Japan) operated at 120 kV. Images were recorded digitally using a high-resolution water-cooled bottom-mounted CCD camera.

## 2.6 | Scanning microcalorimetry

Scanning microcalorimetry experiments were performed on a differential scanning microcalorimeter SCAL-1 (Scal Co., Pushchino, Russia) with a 0.33 glass cell at a heating rate of 1 K/min, under a pressure of 2.5 atm (Senin, Potekhin, Tiktopulo, & Filimonov, 2000). The measurements and necessary calculations were performed according to Privalov and Potekhin (1986) and described in detail in Tarasov, Kostyukova, Tiktopulo, Pyatibratov, and Fedorov (1995).

## 2.7 | Limited proteolysis

Limited proteolysis by trypsin (Sigma) was performed in 10 mM Na-phosphate buffer, pH 8.0 at 21°C. 20 µl aliquots for electrophoresis were taken at defined periods. The reaction was terminated by adding an equimolar amount of trypsin inhibitor from ovomucoid (Sigma).



## 2.8 | Phylogenetic reconstruction

Sequences were aligned with MAFFT (Katoh & Standley, 2013) or PRANK (Löytynoja & Goldman, 2008). For some analyses, unreliably aligned sites were removed using guidance (Sela, Ashkenazy, Katoh, & Pupko, 2015). Search for the best model to describe sequence evolution and search for the maximum likelihood tree were performed in IQ-TREE (Nguyen, Schmidt, von Haeseler, & Minh, 2015) using the Bayesian information criterion (BIC). The only difference between phylogenetic reconstruction from the different alignments is that in case of alignments filtered for conserved sites, the branches are shorter (due to the removal of variable sites), the bootstrap support values are lower, and the most appropriate models determined with IQ-TREE are simpler, because the removal of the variable more difficult to align sites also removes phylogenetic information.

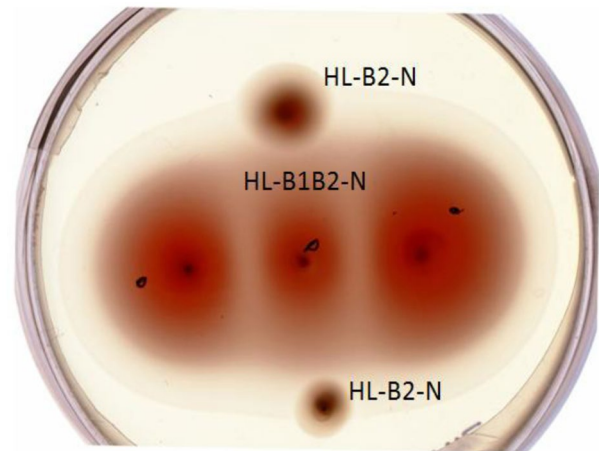
## 3 | RESULTS

### 3.1 | Comparison of natural *Hrr. lacusprofundi* strains DL18 and ACAM 34

Both *Hrr. lacusprofundi* strains were isolated from the relict hypersaline Deep Lake in Antarctica (Franzmann et al., 1988; Liao et al., 2016). Due to its high salinity, this lake never freezes and its surface temperature ranges from  $-20^{\circ}\text{C}$  to  $+10^{\circ}\text{C}$  depending on the season. In the laboratory, *Hrr. lacusprofundi* cells can grow at temperatures ranging from  $-1^{\circ}\text{C}$  to  $+44^{\circ}\text{C}$  (optimum temperature is  $33^{\circ}\text{C}$ ) (Franzmann et al., 1988).

In *Hrr. lacusprofundi* DL18, two strongly diverged archaellin genes *flaB1* and *flaB2* are arranged tandemly and shared the same promoter. The stop codon of gene *flaB1* and the start codon of *flaB2* are separated by a two nucleotide spacer CA. *Hrr. lacusprofundi* ACAM 34 genome contains only one archaellin gene *flaB2*. First, we analyzed the archaellin amino acid sequences of both *Hrr. lacusprofundi* strains. Remarkably, this analysis showed that the signal peptide MFETILNNEERG encoded by the ACAM 34 *flaB2* gene is different from the DL18 FlaB2 signal peptide MFEFINNDKDRG and identical to that of DL18 FlaB1 (Figure A1). At the same time, the nucleotide sequences of the *flaB2* genes of both strains, except the region encoding the signal peptide, are completely identical (Figure A2). Thus, the disappearance of the *flaB1* gene in the ACAM 34 strain can be explained by recombination and subsequent deletion event of a DNA segment in an ancestral two archaellin operon.

*Hrr. lacusprofundi* DL18 FlaB1 and FlaB2 archaellins are strongly diverged (identity 36%), as in other *Halorubrum* species (Figure A1). This strikingly distinguishes them from other haloarchaeal systems. For example, the identity between the two *Hfx. volcanii* archaellins (FlgA1 and FlgA2) is 60%; between those of *Har. hispanica* (FlgA1, FlgB, FlgA2), it is at least 55%; and in *Hbt. salinarum* (FlgA1, FlgA2, FlgB1-FlgB3), the identity is over 80%. *Hrr. lacusprofundi* DL18 FlaB1 and FlaB2 also differ significantly in molecular weight: 19,796.67 and 23,593.45, respectively.



**FIGURE 1** Comparison of cell motility of *Hrr. lacusprofundi* ACAM 34 (HL-B2-N) (in the top and bottom spots) and DL18 (HL-B1B2-N) strains (in three central spots), St-HL medium, 0.19% agar,  $25^{\circ}\text{C}$ , 28 days

Earlier, we showed that the cells of the *Hrr. lacusprofundi* ACAM 34 strain are motile on semisolid media (Syutkin et al., 2012). We compared the motility of both strains on semisolid 0.19% agar under the same conditions and found that the DL18 strain shows significantly higher motility (Figure 1). It should be noted that at the same time, the growth rate of the DL18 in liquid media is higher than that of the ACAM 34 (Figure A3). The maximum archaella yield in the late stationary phase was approximately 10 mg per 1-L culture for both strains. In contrast to the ACAM 34, the cells of the DL18 strain demonstrate a more stable motility and archaella production. To obtain the relatively motile ACAM 34 cells with high archaella yield that were used in the above experiment, it was necessary to pass cells through semisolid (0.19%) agar with 2–3 cycles of a selection of the most motile cells (Syutkin et al., 2012). When ACAM 34 cells that were kept for a long time (about 1 month) in a liquid medium were used to inoculate, the archaella yield decreases dramatically.

### 3.2 | Heterologous expression of *Hrr. lacusprofundi* archaellins in *Haloflex volcanii*

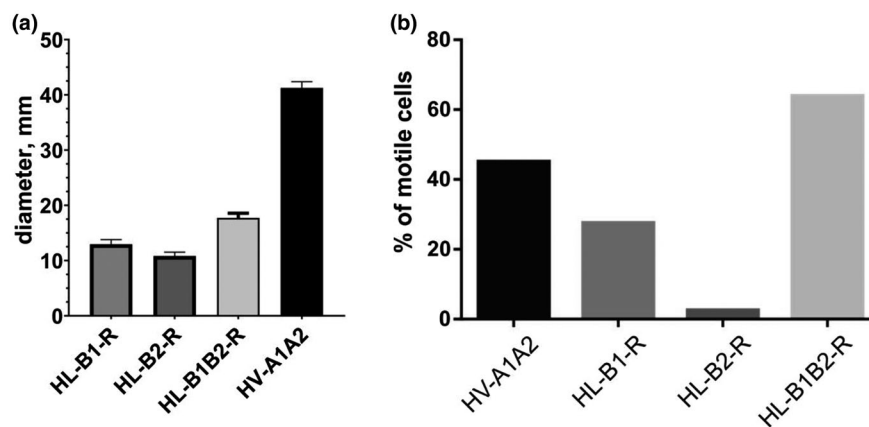
The analysis of natural *Halorubrum* strains allowed us to isolate archaellar filaments consisting of FlaB1/FlaB2 and FlaB2 archaellins. Next, we aimed to study whether the FlaB1 archaellin is capable of producing functional archaella. To this end, we expressed the different archaellin genes in the heterologous host *Hfx. volcanii* MT2 in which the *flgA1flgA2* archaellin operon was removed (Tripepi et al., 2010). The *flaB1* and *flaB2* genes (separately and together as an operon) were cloned into corresponding plasmids based on the pTA1228 vector (*Amp<sup>r</sup>*, *pyrE2* and *hdrB* markers, inducible *ptna* promoter) (Allers, Barak, Liddell, Wardell, & Mevarech, 2010). After transformation, *Hfx. volcanii* cells were grown in the presence of tryptophan as an inducer of the archaellin expression.

Expression of *Hrr. lacusprofundi* archaeellins in nonmotile *Hfx. volcanii*  $\Delta flgA1flgA2$  leads to the restoration of motility (Figure 2a and Figure A4). Motility was demonstrated by cells with each of the three archaeella types (HL-B1-R, HL-B2-R, HL-B1B2-R, R—recombinant). The HL-B1B2-R strain had the best motility on semisolid agar (measured by the diameter of the motility ring) compared with the HL-B1-R and HL-B2-R strains. At the same time, the motility on semisolid agar in all three strains was markedly less than in *Hfx. volcanii* strain expressing their natural archaeellins (FlgA1/FlgA2) (Figure 2a and Figure A8). Electron microscopy confirmed the presence of archaeella bundles in HL-B1-R, HL-B2-R, and HL-B1B2-R cells. HL-B1-R and HL-B1B2-R cells were noticeably more archaeellated than HL-B2-R cells, for which specimens without archaeella were often observed (Figure A5). Also, we analyzed the swimming behavior of the three strains in liquid medium with live-cell microscopy and compared this with a *Hfx. volcanii* strain with a pMT21 plasmid containing *flgA1flgA2His* (Tripepi et al., 2010) that synthesizes the functional archaeella when cells are grown in the presence of tryptophan. There were no significant differences in the velocity or the frequency of reversals between any of the strains (Figure A4). However, in the *flaB2* expressing strain, an extremely low percentage of cells was motile (<5%) (Figure 2b; Movie S1, <https://doi.org/10.5281/zenodo.3723268>). This is in contrast with the other strains, for which a large fraction of cells is motile in liquid medium in the early exponential phase (Movie S2–S4, <https://doi.org/10.5281/zenodo.3723268>). The expression of *flaB1* alone results in motile cells, corresponding with the analysis on the semisolid agar plate. However, in the liquid medium, the percentage of motile cells is significantly lower (~30%) than for the *flaB1flaB2* expression strain (~60%) (Figure 2b). Thus, the expression of FlaB1 only archaeellum filaments might lead to slightly less motile cells in comparison with the two-archaeellin filaments.

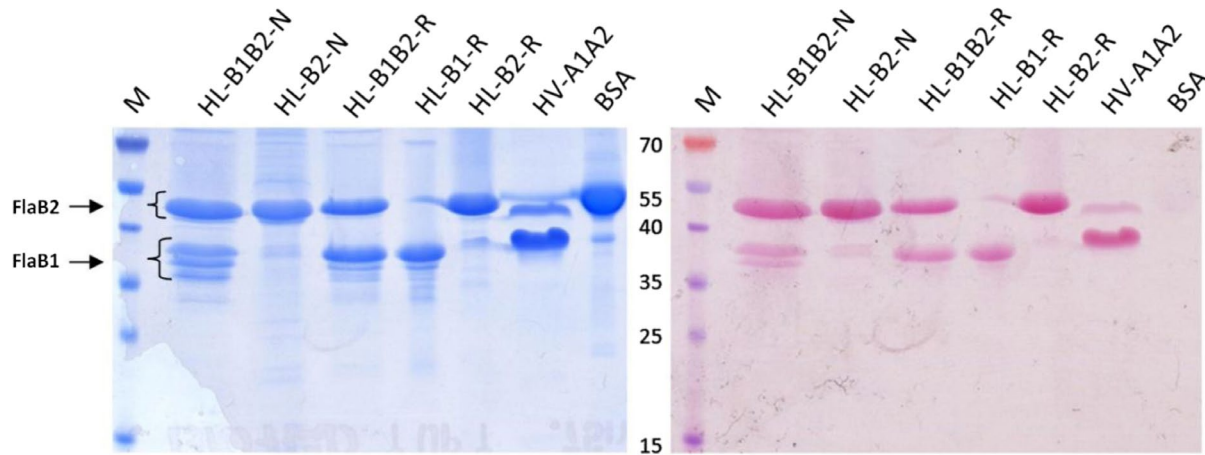
### 3.3 | Comparison of natural and recombinant *Hrr. lacusprofundi* archaeella

The archaeella, isolated from natural DL18 and ACAM 34 strains, were designated as HL-B1B2-N and HL-B2-N—respectively. SDS-PAGE of *Hrr. lacusprofundi* ACAM 34 archaeellum filaments (Figure 3) showed a single major band, corresponding to a molecular mass of ~50 kDa. For the DL18 strain, the same and additional major bands of ~37 kDa were observed (Figure 3). Mass spectrometry analysis (Figure A13) confirmed that the isolated proteins were *Hrr. lacusprofundi* archaeellins FlaB2 (~50 kDa) and FlaB1 (~37 kDa). The apparent molecular masses determined by SDS gel electrophoresis are higher than the true values (23.6 and 19.8 kDa for FlaB1 and FlaB2, respectively), which is typical for halophilic archaeellins due to high content of carbonic acids and posttranslational modifications (Fedorov, Pyatibratov, Kostyukova, Osina, & Tarasov, 1994; Gerl & Sumper, 1988; Pyatibratov et al., 2008). The *Hrr. lacusprofundi* DL18 archaeellum consists of both FlaB1 and FlaB2 archaeellins, present in comparable quantities. The molar ratio FlaB1/FlaB2, determined by measuring the density of the protein bands on SDS-PAGE gels is approximately 1:1 ( $0.90 \pm 0.04$ ; Figure 3).

For isolation of archaeella, we grew *Hfx. volcanii* cells in the modified medium used by Guan et al. as described in the Material and Methods (Guan, Naparstek, Calo, & Eichler, 2012). This modified medium in combination with moderately high growth temperature (37°C) allowed us to obtain a high yield of recombinant archaeella from *Hfx. volcanii*. The electrophoretic mobilities of recombinant archaeellins differed a little from those of natural archaeellins (Figure 3). The synthesis of recombinant archaeella was confirmed by mass spectrometry (Figures A14 and A15) and electron microscopy (Figure 4 and Figure A5). The molar ratio of FlaB1/FlaB2 determined by scanning of the SDS-PAGE gels is approximately  $1.20 \pm 0.04$ , which is



**FIGURE 2** Swimming behavior of *Hfx. volcanii* strains expressing different combinations of *Hrr. lacusprofundi* archaeellins. (a) Comparison of swarm sizes on the semisolid agar plates. The swarming diameters were measured 120 hr after inoculation: *Hfx. volcanii* MT2 transformed with pMT21 (HV-A1A2), pAS5 (HL-B1B2-R), pAS6 (HL-B1-R), and pAS7 (HL-B2-R), Mod-HV medium containing 0.5 mg/ml tryptophan, 0.24% agar. (b) Swimming behavior in liquid media analyzed using thermomicroscopy. Percentage of motile cells in liquid cultures in early exponential phase. Cells were analyzed by light microscopy, and time-lapse images were taken for 15 s. All strains were analyzed at the same time during at least three independent experiments, which included >100 cells. The number of cells that showed motility during a 15-s time-lapse movie, was calculated as the percentage of the total observed cells. HL, archaeellin from *Hrr. lacusprofundi*. HV, archaeellin from *Hfx. volcanii*



**FIGURE 3** Archaeellin staining with Schiff's reagent in 12.5% polyacrylamide gel (right); the same gel stained with Coomassie G250 (left). M—prestained protein standard, HL-B1B2-N, and HL-B2-N—natural archaeella of *Hrr. lacusprofundi* DL18 and ACAM 34, HL-B1B2-R, HL-B1-R, HL-B2-R—recombinant archaeella isolated from *Hfx. volcanii* MT2 transformed with appropriate plasmids: HV-A1A2—*Hfx. volcanii* MT45 archaeella (positive control), BSA—bovine serum albumin (negative control)

higher than for natural archaeella ( $0.90 \pm 0.04$ ) (Figure 3). Both FlaB1 and FlaB2 sequences contain putative N-glycosylation sites, N-X-S(T) (Figure A1). Glycosylation of natural and recombinant archaeellins was confirmed by staining with Schiff's reagent. Note that in both native and recombinant filaments, several bands appear below the main FlaB1 band. These could represent different glycoforms, as they are differently stained by Schiff's reagent. The lowest band may correspond to the nonglycosylated form (Figure 3). The structure of *Hfx. volcanii* archaeellin oligosaccharides is well known (Tripepi et al., 2010, 2012). This is in contrast to the glycans of *Hrr. lacusprofundi*. In other *Halorubrum* species, unique sialic acid-like saccharides not characteristic of the *Hfx. volcanii* were found (Zaretsky, Roine, & Eichler, 2018). It remains unclear whether the same Asn residues are modified in the native and recombinant archaeellin version. While N-glycosylation is necessary for halophilic archaeella assembly (Tripepi et al., 2012; Zaretsky, Darnell, Schmid, & Eichler, 2019), we demonstrate that the *Hrr. lacusprofundi* archaeellins can still assemble into functional archaeella while being likely being decorated by the non-native N-glycosylation from the *Hfx. volcanii*.

The archaeella isolated from the natural DL18 and ACAM 34 strains were observed by electron microscopy, which showed that the HL-B2-N archaeella are quite flexible and often twisted in loops and tangles (Figure 4). The HL-B1B2-N archaeella, on the other hand, were generally longer and less flexible (Figure 4). The thickness of the HL-B1B2-N and the HL-B2-N archaeella were both  $\sim 10$  nm. No structures resembling hooks or basal bodies were observed. The HL-B1B2-N filaments were homogeneous; no filaments twisted into loops characteristic of the HL-B2-N archaeella were observed. It may indirectly indicate the absence of FlaB2 homopolymers in the DL18 archaeella samples.

Comparison of electron microscopic images of native (HL-B2-N, HL-B1B2-N) and recombinant (HL-B2-R, HL-B1B2-R) archaeella shows that recombinant archaeella, in general, have the same features as natural ones (Figure 4). In electron micrographs, *Hrr. lacusprofundi*

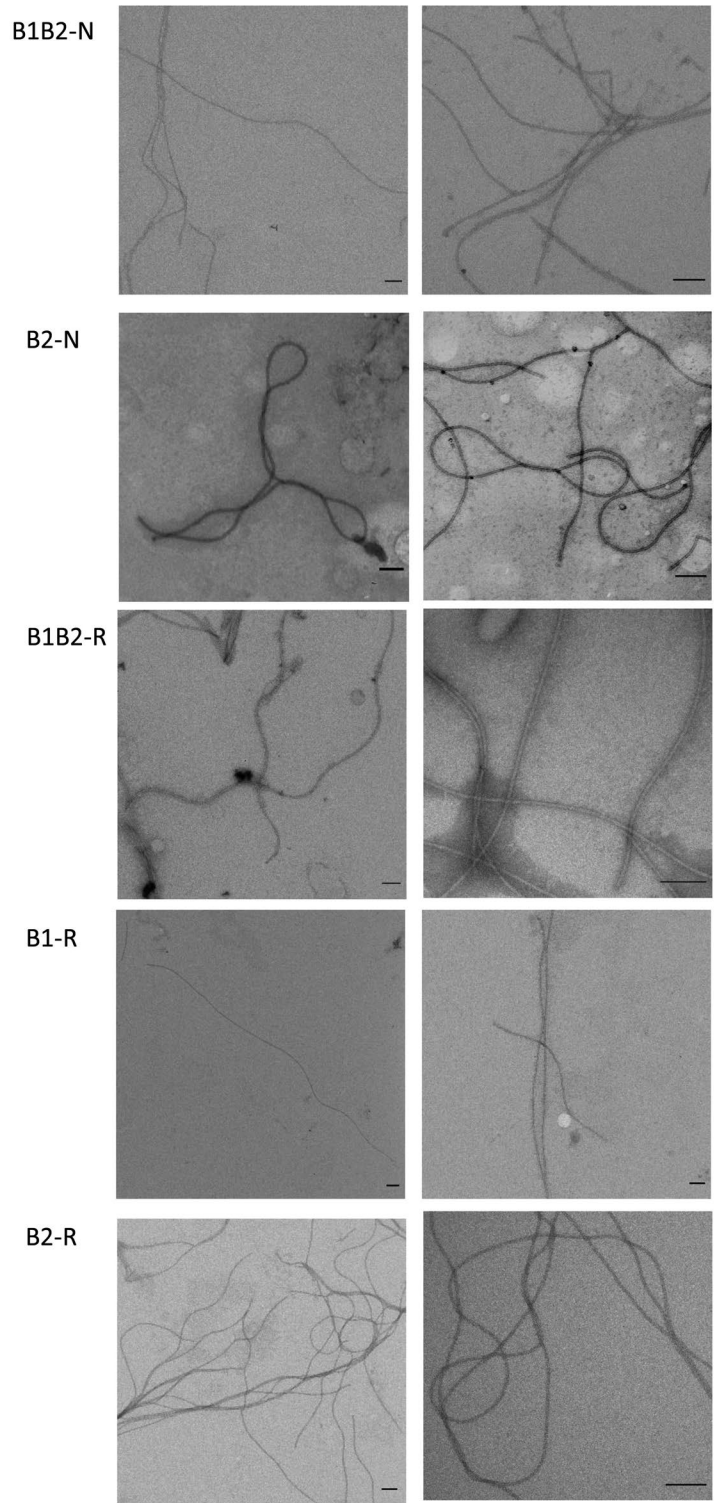
archaella appear as semiflexible filaments in contrast to halobacterial archaeella (Alam & Oesterhelt, 1984), which appear to have a more rigid supercoiled shape. The HL-B1B2-R filaments have a wavy shape, characteristic for "classical" archaeella (Figure 4). HL-B2-R archaeella are also very flexible and often fold into loops, as is the case for the natural B2-N archaeella. The HL-B1-R filaments appear straight and inflexible in comparison with the HL-B1B2-R archaeella (Figure 4). This altered morphology of the HL-B1-R filaments might lead to motility problems. Indeed, live-cell imaging showed that the percentage of motile cells expressing HL-B1-R is slightly reduced compared to HL-B1B2-R. The diameter of the recombinant and natural filaments is about the same ( $\sim 10$  nm).

Since the *Hfx. volcanii* can exist in a fairly wide range of salt concentration (1.5–4 M NaCl) (Mullakhanbhai & Larsen, 1975), we decided to compare the functionality of various types of recombinant archaeella depending on the content of sodium chloride in the growth medium. Motilities on semisolid agar were compared at 10, 15, 20, and 25% NaCl (1.71–4.28 M) for *Hfx. volcanii* strains synthesizing the three archaeella types (HL-B1-R, HL-B2-R, and HL-B1B2-R). Interestingly, the HL-B1B2-R strain has an advantage in motility at 15 and especially at 10% NaCl (Figure 5). Thus, the two-component archaeella appear to be better adapted to environmental salinity changes than the one-component archaeella, allowing the species to occupy a larger ecological niche.

*Halorubrum saccharovorum* ATCC 29,252 (Tomlinson & Hochstein, 1976) is closely related to *Hrr. lacusprofundi* DL18 and has the same organization of the archaeellin genes, encoding proteins with high similarity to those of DL18 (the identity of their FlaB1 archaeellins is 70%, and FlaB2—73%). We isolated the archaeella from *Hrr. saccharovorum* and established that, similar to *Hrr. lacusprofundi* DL18, in this organism both archaeellins are present in an approximate 1:1 ratio also (Figure A6). We were able to successfully express the *Hrr. saccharovorum* flaB1 gene and the flaB1/flaB2 archaeellin operon in *Hfx. volcanii* MT2.

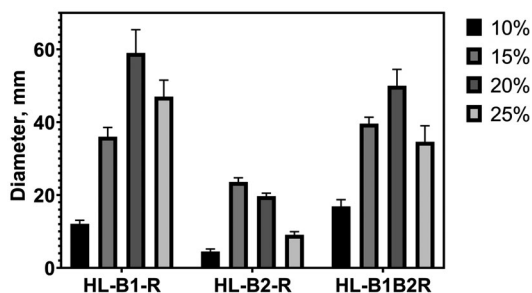


**FIGURE 4** Negatively stained (1% uranyl acetate) preparations of *Hrr. lacusprofundi* ACAM 34 and DL18 archaeellar filaments (HL-B2-N and HL-B1B2-N, respectively) and preparations of recombinant *Hrr. lacusprofundi* archaeellar filaments HL-B1B2-R, HL-B1-R, and HL-B2-R isolated from *Hfx. volcanii* MT2 in 20% NaCl, 10 mM Na-phosphate, pH 8.0. Scale bar—100 nm



Unfortunately, an attempt to heterologously express the *Hrr. saccharovororum* *flaB2* gene did not lead to the archaeella synthesis. An attempt was made to express the modified *Hrr. saccharovororum* *flaB2* gene in which the nucleotide sequence encoding the signal peptide was replaced with that for the FlaB1, but this also was unsuccessful. Possibly, FlaB2 of *Hrr. saccharovororum* can only form a stable filament in the presence of FlaB1. Probably, the *Hrr. saccharovororum* FlaB2 is less adaptable to heterogeneous assembly and

a divergent glycosylation system. Synthesis of recombinant HS-B1-R and HS-B1B2-R archaeella was confirmed by mass spectrometry (Figures A16 and A17). The results of heterologous archaeellin expression were similar to those found for the recombinant expression of *Hrr. lacusprofundi* archaeellins: Expression of *flaB1/flaB2* and *flaB1* restored motility (Figures A7 and A8), and the synthesized archaeella had a similar morphology to the native filaments (Figure A9). Again, the HS-B1-R filaments appeared straighter and



**FIGURE 5** Comparison of cell motility of *Hfx. volcanii* strains expressing *Hrr. lacusprofundi* archaeellins on semisolid media at different salinities. *Hfx. volcanii* MT2 strain was transformed with pAS5 (HL-B1B2-R), pAS6 (HL-B1-R), and pAS7 (HL-B2-R), Mod-HV medium containing 10, 15, 20, and 25% NaCl with 0.5 mg/ml tryptophan, 0.24% agar, 37°C, 7 days

less flexible as the HS-B1B2-R filaments (Figure A9). Glycosylation of natural and recombinant archaeellins of *Hrr. saccharovororum* was confirmed by specific staining with Schiff's reagent (Figure A6). Interestingly, the staining intensities of natural *Hrr. saccharovororum* archaeellins are noticeably less than that of recombinant archaeellins. At the same time, both are glycosylated less than natural *Hrr. lacusprofundi* archaeellins.

### 3.4 | Scanning microcalorimetry experiments

To obtain additional information regarding archaeella of different composition, we applied differential scanning microcalorimetry (DSC). Isolated *Hrr. lacusprofundi* archaeella were heated in near-natural (20% NaCl) and low (10% NaCl) salt conditions. We found that at 20% NaCl the two-component *Hrr. lacusprofundi* DL18 archaeella (HL-B1B2-N) are much more stable than the natural HL-B2-N archaeella of *Hrr. lacusprofundi* ACAM 34. The temperature of the heat absorption peak maximum ( $T_m$ ) of the archaeella consisting of two different archaeellins (97.5°C) was substantially higher than those of archaeella build of a single type of archaeellin (80.0°C) (Figure 6b; Table 3). For both types of archaeella, only a single heat absorption peak was observed in the 20% NaCl buffer. This indicates that the DL18 strain presents only heteropolymeric archaeella (consisting of both FlaB1 and FlaB2). If two different types of homopolymeric filaments would be present, we would expect two different melting peaks one of which corresponds to the melting curve of the HL-B2-N archaeella. A decrease in the NaCl concentration (10%) resulted in a decrease in the  $T_m$ . In 10% NaCl, we observed a melting curve with two peaks at physiological temperature (39 and 45°C) for the one-component archaeella HL-B2-N and the single peak corresponding to a significantly higher temperature (81.5°C) for the two-component archaeella HL-B1B2-N (Figure 6a; Table 3). Thus, two-component archaeella are much more resistant to lower salinity. These data suggest cooperative interactions and a close relationship between FlaB1 and FlaB2 subunits in the archaeellar structure.

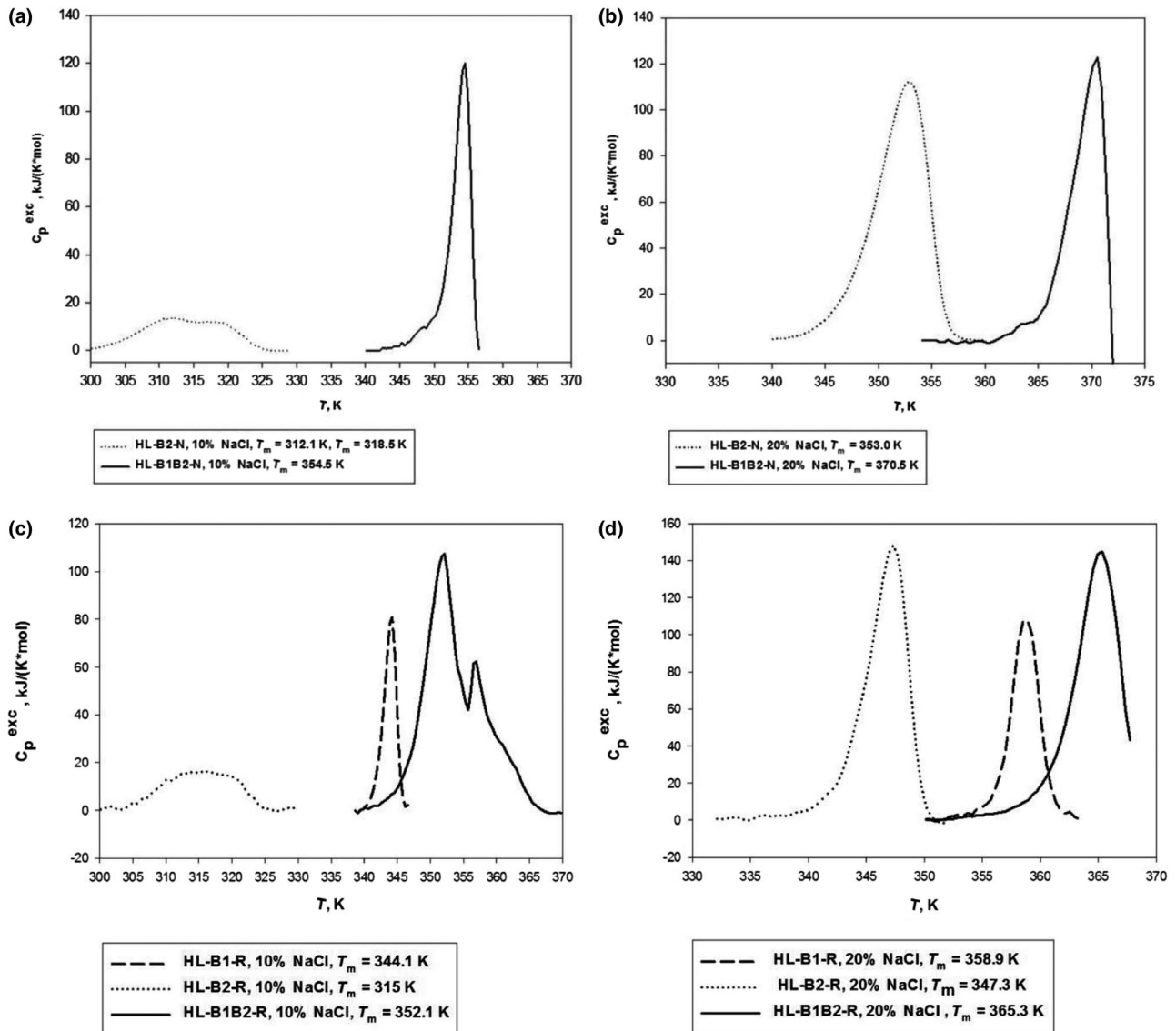
DSC data for recombinant filaments are similar to results obtained for the natural archaeella. The temperature of the heat absorption peak maximum of HL-B1B2-R archaeella (92°C) under 20% NaCl was noticeably higher than that of recombinant filaments consisting either only of FlaB1 (86°C) or FlaB2 (74°C) (Figure 6d; Table 3). These temperatures were all slightly lower than those observed for the natural filaments isolated from *Hrr. lacusprofundi* (HL-B1B2-N, 97.5°C and HL-B2-N, 80°C). On melting at 10% NaCl, extended heat absorption peak with a maximum of about 42°C was observed for the HL-B2-R archaeella (in comparison with two peaks at 39 and 45°C for HL-B2-N). The HL-B1-R and HL-B2-R melting curves are different (Figure 6c; Table 3). The combination of both subunits in one archaeellar filament (HL-B1B2) leads to structural changes that are reflected in a new melting curve, both natural and recombinant archaeella.

Similar experiments were carried out on *Hrr. saccharovororum* archaeella. Here, we have compared three archaeella types: HS-B1B2-N, HS-B1B2-R, and HS-B1-R. The experiments were carried out at 5 and 10% NaCl, since at 20% NaCl the natural *Hrr. saccharovororum* archaeella melted near the upper limit of the experimental temperature range. In this case, the  $T_m$  of all three types of filaments was very similar, which is in line with the findings for *Hrr. lacusprofundi* archaeella (Figure A10).

Thus, in general, the FlaB1FlaB2 filaments are slightly more stable than the FlaB1 and much more stable than the FlaB2 filaments.

### 3.5 | Limited proteolysis confirms the interaction of *Hrr. lacusprofundi* archaeellins in the filaments

To probe conformational features of two-component and one-component *Hrr. lacusprofundi* archaeellar filaments, we used limited trypsinolysis. When the NaCl concentration was <8%, FlaB2 archaeellin from HL-B2-N archaeella was digested with trypsin, while both FlaB1 and FlaB2 archaeellins from HL-B1B2-N filaments were protected from trypsin digestion under the same conditions (Figure 7). This effect was also observed for recombinant archaeella. The HL-B1B2-R and HL-B1-R filaments are more resistant to trypsin digestion than the HL-B2-R (Figure 7). When the HL-B1-R and HL-B2-R archaeellar filaments were mixed, the bands on the SDS gel indicated the digestion of FlaB2, suggesting that the stabilizing role of the intermolecular FlaB1/FlaB2 interactions only occurs when a mixed filament is built (Figure 7). Trypsin cleaves peptides on the C-terminal side of lysine and arginine residues. These residues are rare in *Hrr. lacusprofundi* archaeellins (5 and 3 arginines, no lysines, respectively, in processed FlaB1 and FlaB2). From the distribution of arginines in the FlaB2 protein (Figure A1) (R61, R83, R230), it can be concluded that in the presence of FlaB1, the R61 and R83 sites of FlaB2 are protected from trypsin attack. From a comparison with the known archaeella structures, it can be expected that R61 and R83 are localized between N-terminal -helix and -strand 1, and within in -strand 2, respectively (Poweleit et al., 2016). The presence of FlaB1 seems to protect these sites, possibly by shielding them for trypsin.



**FIGURE 6** Temperature dependence of excess heat capacity of *Hrr. lacusprofundi* archaellar filaments. *Hrr. lacusprofundi* ACAM34 and DL18 archaellar filaments at two salinities: (a)–10% (1.7 M) and (b)–20% (3.4 M) NaCl, 10 mM Na-phosphate, pH 8.0. Recombinant *Hrr. lacusprofundi* B1B2, B1, and B2 archaellar filaments at two salinities: (c)–10% (1.7 M) and (d)–20% (3.4 M) NaCl, 10 mM Na-phosphate, pH 8.0

### 3.6 | Bioinformatical analysis of *Halorubrum* archaellins

After analyzing the *Halorubrum*s genomes available on insert date (50 in total), we found that in most species the organization of archaellin genes is similar to that of *Hrr. lacusprofundi* DL18: In 47 species, there are two strongly diverged genes *flaB1* and *flaB2*, organized as a single operon, among them 9 species have an additional archaellin operon, which is likely to result from gene duplication events. Only *Hrr. lacusprofundi* ACAM 34 genome contains one archaellin gene. Each of the two *Halorubrum* archaellin types (FlaB1 and FlaB2) is characterized by a high degree of conservation. It is possible to identify genetic signatures unique to each of the two groups, for example, in internal (50–75 a. a.) and C-terminal partially conserved regions (Figure A11). FlaB1 sequences range

from 187 to 210 amino acid residues, which is significantly shorter than the FlaB2 sizes from 201 to 456 a.a. Approximately 2/3rds of the *Halorubrum* FlaB2 sequences have no more than 250 residues (Table A1). In most *Halorubrum* species, the stop codon of gene *flaB1* and the start codon of *flaB2* are separated by two nucleotide spacer CA (in 42 cases out of 53). The CC and CG spacers were found 3 times and AA, AC, AT, and TT once.

In two *Halorubrum* species (*Hrr. halodurans* and *Hrr. vacuolatum*), the organization of archaellin genes is fundamentally different: They contain several diverged archaellin genes that cannot be classified as *flaB1* or *flaB2*. The corresponding proteins do not have signatures typical for archaellins of other *Halorubrum* species. Based on phylogenetic analysis, these archaellins should be probably attributed to the FlaB1 branch (Figure 8). The archaellin paralogs of these two species are more similar than the FlaB1 and FlaB2 paralogs in other

*Halorubrum* species. Thus, identities between three *Hrr. halodurans* archaeellins are >55%, and >60% for four *Hrr. vacuolatum* archaeellins. *Hrr. halodurans* archaeellin genes constitute a single operon, the *flaB<sub>a</sub>* and *flaB<sub>b</sub>* genes are separated by CG spacer, and the start codon of *flaB<sub>c</sub>* gene immediately follows the *flaB<sub>b</sub>* stop codon. The *Hrr. vacuolatum* genome contains three archaeellin operons, one of them consists of two genes separated by a spacer of four nucleotides (GACC).

Interestingly, for the other haloarchaeal genera (*Halopiger*, *Natrialba*, *Halobiforma*, *Natronolimnobius*, and *Natrarchaeobius*) the situation with archaeellins is very similar to that of *Halorubrum*.

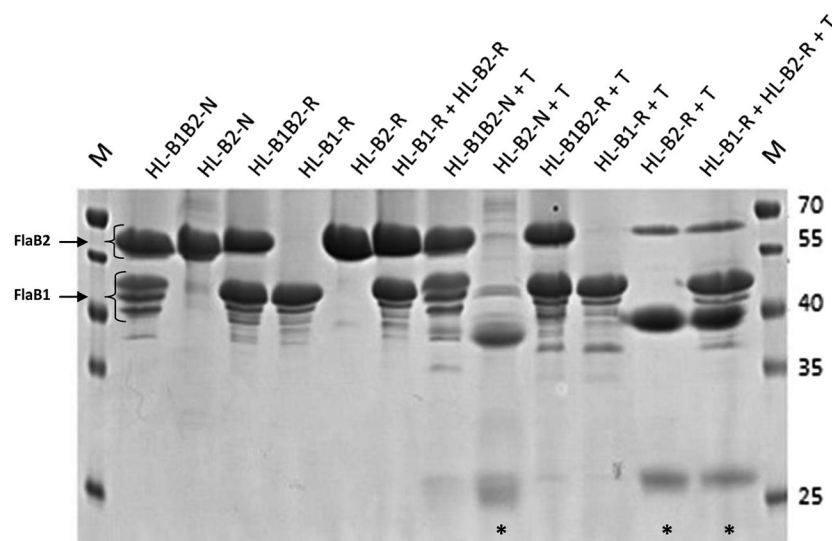
**TABLE 3** The melting points of different filament types, determined using scanning microcalorimetry

% NaCl	Archaeella	T <sub>m</sub> (°C)
<i>Hrr. lacusprofundi</i>		
20	B1B2-N	97.5
20	B2-N	80.0
20	B1B2-R	92.3
20	B1-R	85.9
20	B2-R	74.3
10	B1B2-N	81.5
10	B2-N	39.1/45.5
10	B1B2-R	79.1
10	B1-R	71.4
10	B2-R	42.0
<i>Hrr. saccharovororum</i>		
10	B1B2-N	85.2
10	B1B2-R	82.3
10	B1-R	81.9
5	B1B2-N	73.2
5	B1B2-R	71.9
5	B1-R	70.9

Despite the rather high similarity of their archaeellins with that of the *Halorubrum* species, all these haloarchaea belong to a clade (the order *Natrialbales*, the family *Natrialbaceae*) evolutionarily distinct from *Halorubrum* (order *Haloferocales*, family *Halorubraceae*) (Amoozegar, Siroosi, Atashgahi, Smidt, & Ventosa, 2017). It is likely that operons from two highly diverged paralogs, having a common origin with *Halorubrum flaB1* and *flaB2* genes, were exchanged between these taxa by horizontal gene transfer. The type of archaeellin gene organization, characteristic for *Halorubrum*, predominates in genomes of the most representatives of these groups. It is interesting that, unlike *Halorubrum*, these haloarchaeal groups are characterized by large FlaB1 archaeellins (from 193 to 481 a.a., for half of them the archaeellin size is >300 a.a.) and relatively small FlaB2 (from 208 to 261 a.a.). In genomes of some species (*Hpg. djelfmassiliensis* and *Hbf. nitratireducens*), archaeellin genes have a type of organization, similar to that in *Hrr. halodurans* and *Hrr. vacuolatum*. The corresponding protein products have their characteristics, which are not typical of most *Halorubrum* archaeellins.

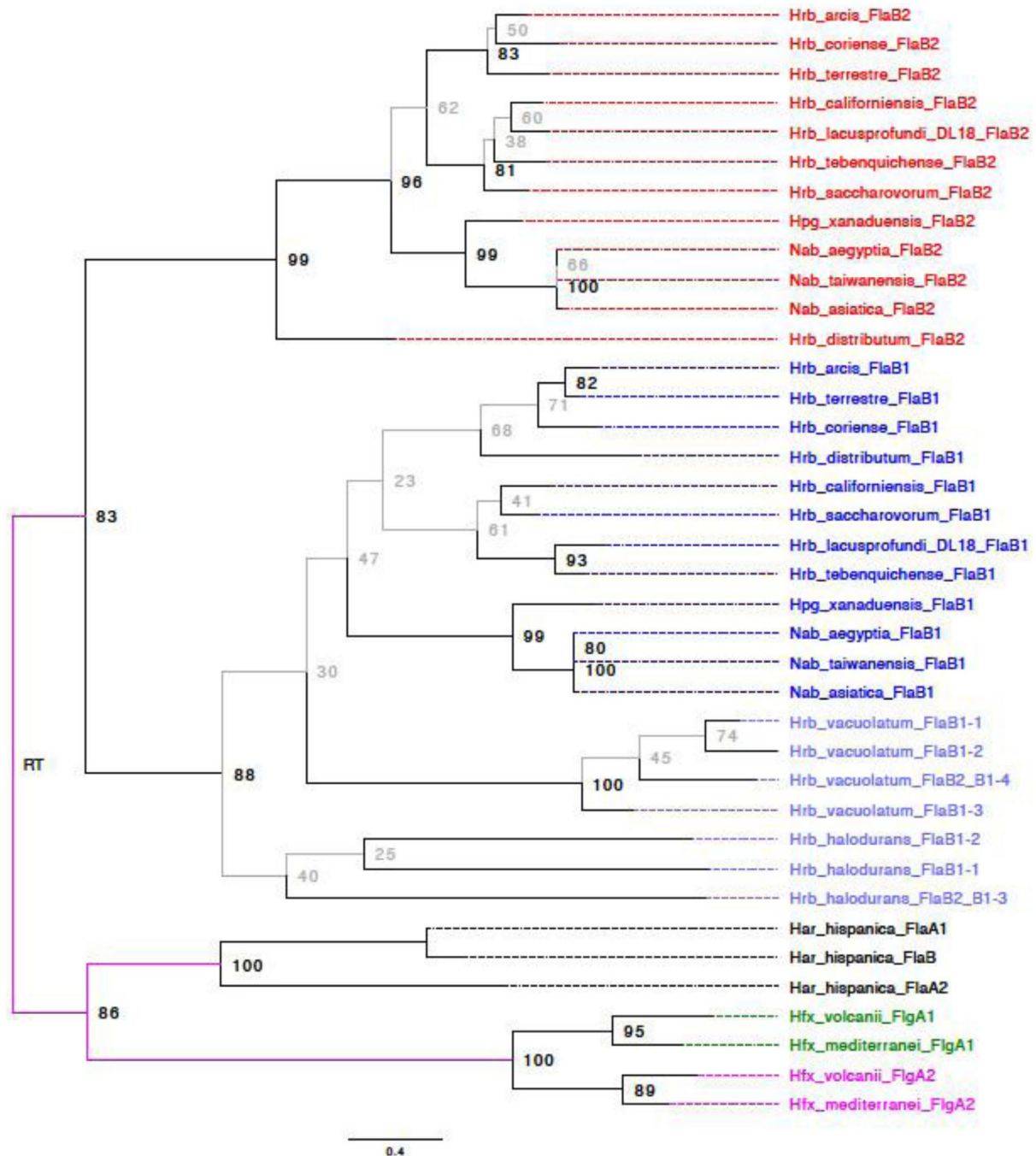
*Halorubrum* archaeellins reveal similarity with archaeellins of the evolutionarily distant (Becker et al., 2014) haloarchaeal groups *Halobiforma*, *Halopiger*, *Natrialba*, and *Natronolimnobius*. The evolutionary history of the archaeellins does not reflect organismal relationships inferred from genome comparisons (Gupta, Naushad, Fabros, & Adeolu, 2016). *Halorubrum* is considered a member of the *Haloferocales*, whereas *Halobiforma*, *Halopiger*, *Natrialba*, and *Natronolimnobius* are placed in the *Natrialbales*. Figure 8 and Figure A12 show that for both *flaB1* and *flaB2* these divergent microorganisms form well-supported clades in the archaeellin phylogeny. Of particular interest is the recently described haloarchaea *Natrarchaeobius chitinivorans* (Sorokin et al., 2019) having two archaeellin operons. One operon consists of two genes closely related to *Hrr. lacusprofundi flaB1* and *flaB2*, and the other distant operon consist of three genes having a different origin.

It should be emphasized that, as can be seen from the evolutionary tree (Figure 8), the divergence between *flaB1* and *flaB2*



**FIGURE 7** Results of limited trypsinolysis of isolated natural and recombinant *Hrr. lacusprofundi* archaeellar filaments. Intact natural (HL-B1B2-N, HL-B2-N) and recombinant (HL-B1B2-R, HL-B1-R, HL-B2-R) archaeella before and after trypsinolysis in 10 mM Na-phosphate, pH 8.0 containing 2% NaCl. Archaeella were treated with trypsin 60 min at room temperature; the protein/enzyme ratio was 100:1. M—MW standards. Asterisks indicate samples where the FlaB2 is sensitive to trypsinolysis





**FIGURE 8** Phylogenetic tree of archaeellins in the selected haloarchaea species (11 *Halorubrum* species, 2 species of *Halopiger* and *Natrialba*, *Hfx. Volcanii*, and *Har. hispanica*). Red—FlaB2 archaeellins, blue—FlaB1 archaeellins. The depicted maximum likelihood phylogeny and bootstrap support values (traditional nonparametric) were calculated from a MAFT alignment without filtering for conserved sites using a WAG + F + G4 substitution model. Branches with less than 80% bootstrap support are given in gray. Branches from which the root cannot be excluded assuming that gene duplications occurred distal from the root are given in fuchsia. The depicted phylogeny was rooted using generalized midpoint optimization (Maljkovic Berry et al., 2009)

genes is a more ancient compared to the divergence of the corresponding genes in such haloarchaea as *Hrr. halodurans*, *Hrr. vacuolatum*, *Hfx. volcanii*, *Har. hispanica* and *Hbt. salinarum*. *Hfx. volcanii* and *Hrr. lacusprofundi* DL18 represent two distinct types: The first of them (*Hfx. volcanii*) is characterized by archaeellins that have a close relationship, and the second group is characterized

by highly divergent archaeellins. The occurrence of the second type among divergent groups of *Halobacteria* is likely due to horizontal gene transfer between *Natrialbales* and *Haloferacales*. It can be assumed that the principles of the structural organization of the archaeella have significant differences between these two types.



## 4 | DISCUSSION

The archaeal motility structure, the archaeum, consists of thousands of copies of N-terminally cleaved archaeum subunits. While crenarchaea usually encode a single type of archaeum, the euryarchaea are characterized by the presence of multiple types of archaeum encoding genes. Recently, high-resolution structures of archaeum filaments of methanogens and hyperthermophilic euryarchaea became available (Daum et al., 2017; Meshcheryakov et al., 2019; Poweleit et al., 2016). In these structures, only a single type of archaeum is present in the filament, even though several archaeum genes are present in the genome. Daum et al. suggested that these other archaeums either (a) are minor, and form specific basal or terminal segments of the filament, or (b) that each of the different types of archaeums forms individual filaments (Daum et al., 2017).

We aimed to understand the biological relevance of archaeum multiplicity by using the halophilic euryarchaeon *Hrr. lacusprofundi*, which encodes two divergent archaeums, FlaB1 and FlaB2, that are easily distinguishable from each other in amino acid sequences and sizes. Both archaeums were present in the archaeum filaments in comparable amounts. We used natural *Hrr. lacusprofundi* strains encoding FlaB1 and FlaB2 or only the FlaB2 protein. Also, we expressed the FlaB1, FlaB2, and a combination of the two proteins in *Hfx. volcanii*. This allowed us to study the role of the individual archaeums, indicating that a combination of FlaB1 and FlaB2 is required to provide stability to the archaeum. In *flaB1*- and *flaB2*-containing strains, microcalorimetry and SDS-PAGE analysis confirmed that both archaeums are part of each filament. Differential melting curves in scanning microcalorimetry experiments and protection against trypsin digestion indicate that the FlaB1 and FlaB2 proteins tightly interact within the archaeum and as such provide stability to the filament. In the absence of one of the archaeums, the archaeum either become over flexible (FlaB2 filaments) or quite stiff (FlaB1 filaments). These archaeums consisting of single components are still functional and provide motility, although that of filaments consisting only of FlaB2 is strongly reduced. Comparison of the motility of *Hfx. volcanii* strains expressing different archaeum types indicate that two-component archaeums are better adapted to stress caused by both extra low and extra high salt concentrations.

We suppose that FlaB2 adopts a final more stable conformational state by interacting with FlaB1. It can also be assumed that FlaB1 and FlaB2 form a stable heterodimer that then assembles into the archaeum. The pairwise interaction between FlaB1 and FlaB2 is compatible with the single higher melting point. The microcalorimetry experiments show that the archaeum consisting of two types of archaeums are more stable under varying conditions, such as low salt stress. In a comparative study of one- and two-component filaments, we have found that the presence of FlaB1 substantially stabilizes the filament structure. Instead of a superposition of FlaB1 and FlaB2 peaks, we observe a new peak of heat absorption that melting point was slightly higher than for FlaB1 filaments and significantly higher than for FlaB2 filaments.

Thus, the two-component composition of *Hrr. lacusprofundi* archaeum filaments contributes to additional stabilization of the archaeum structure and adaptation to a wider range of external conditions and it is not required for archaeum supercoiling.

By applying the heterologous expression of the *Hrr. lacusprofundi* archaeums in *Hfx. volcanii*, we demonstrated that archaeums can assemble in functional archaeum, even in species that possess highly divergent archaeum genes. This suggests that foreign archaeum genes captured via horizontal transfer can quite easily adapt to the assembly and glycosylation system of the new host. Exchange of archaeums could provide an evolutionary advantage as it might allow adaptation to new environments or block the attachment of archaeum specific viruses (Pyatibratov et al., 2008; Tschitschko et al., 2018).

Euryarchaea are characterized by the genomic presence of multiple different archaeum genes. Most euryarchaea have two archaeum genes, but in some cases, the number of different archaeums is even higher (such as *Hht. litchfieldiae*, which has seven archaeum genes) (Tschitschko et al., 2015).

Several explanations for the existence of multiple archaeums have been proposed. Firstly, some archaeums might form minor components of the archaeum. This hypothesis is consistent with the results of the work (Chaban et al., 2007) when the archaeum hook segment is built of the minor FlaB3 archaeum in the methanogenic archaeum. Also in *Hfx. Volcanii*, functional archaeum can be formed only from the major archaeum FlgA1, and cells with such archaeum are hypermotile compared to cells of the natural strain, whose archaeum consist of two archaeums (Tripepi et al., 2013). Secondly, multiple archaeums were shown to act as ecoparalogs. *Har. marismortui* is capable of assembling functional (i.e., supercoiled) filaments from either one of the two encoded archaeums (Syutkin et al., 2014). The type of archaeums incorporated in the filament is dependent on environmental conditions (such as ionic strength) (Syutkin et al., 2019).

Archaeum of several euryarchaea consist of multiple archaeums in comparable amounts, which is not in correspondence with either of the two abovementioned explanations. For example, the archaeum filaments of *Hbt. salinarum* (formerly *Hbt. halobium*) contain the products of all five archaeum genes, and the proportion of archaeums FlgA1 and FlgB2 is comparable with the proportion of archaeums FlgA2, FlgB1, and FlgB3 (Gerl et al., 1989). Earlier, it had been suggested that archaeum multiplicity may cause the archaeum to become supercoiled (Tarasov et al., 2000). This hypothesis was based on an analogy with the bacterial flagellar filaments, where two conformational flagellin states provide filament supercoiling (Calladine, 1978). It was shown that inactivation of the *Hbt. salinarum* archaeum genes led to a disruption of the archaeum assembly, while only straight filaments could be formed from the product of a single archaeum gene (*flgA1* or *flgA2*) (Tarasov et al., 2000, 2004). In the case of *Hrr. lacusprofundi*, the presence of two archaeums is not required for supercoiling, as functional filaments can be formed from each of the single archaeum types. However, the presence of two archaeums provides extra stability to the filament, causing it to better withstand ionic stress conditions and to provide the highest level of motility.

Thus, with this work, we add another aspect of encoding multiple archaellins to the other previously discovered mechanisms. Besides forming specialized minor components of the filament, or acting as ecomparalogs, we now show that multiple archaellins can also be important to form the filament with optimal properties in terms of flexibility and stability. Also, we provide evidence that the exchange of archaellins between different species can result in functional archaellum filaments. Together, these findings sketch an evolutionary picture in which accidental duplication of archaellins was used to the advantage of several euryarchaea, specifically the haloarchaea. Frequent horizontal gene transfer of archaellins promotes the evolutionary adaptability of different species. In light of these evolutionary advantages, one might wonder how crenarchaea can be successful with only a single type of archaellin. Possibly, the variation of environmental conditions of haloarchaea (such as ionic strength) promotes the presence of multiple archaellins.

## ACKNOWLEDGMENTS

This work was supported by a RFBR grant No. 19-04-00613 A to M.P., an Emmy Noether grant (411069969) from the Deutsche Forschungsgemeinschaft to T.E.F.Q. and in part by a RFBR grant No. 19-04-01327 A to A.S. We thank Dr. T. Allers for pTA1228 plasmid, Dr. R. Cavicchioli and Dr. M. Pohlschröder for the kindly presented strains, and M. Suvorina for the help with MS analysis. Electron microscopy was performed with the support of the Moscow State University development program (PNR 5.13). The mass spectrometric analysis was performed in facilities of the United Pushchino Center "Structural and functional studies of proteins and RNA" (584307).

## CONFLICT OF INTEREST

None declared.

## AUTHORS' CONTRIBUTIONS

**Mikhail G. Pyatibratov:** Conceptualization (lead); Data curation (equal); Formal analysis (equal); Funding acquisition (lead); Investigation (lead); Methodology (equal); Project administration (lead); Resources (lead); Supervision (lead); Writing-original draft (lead); Writing-review & editing (lead). **Alexey S. Syutkin:** Conceptualization (lead); Data curation (lead); Formal analysis (equal); Funding acquisition (supporting); Investigation (lead); Methodology (lead); Project administration (equal); Resources (equal); Software (equal); Supervision (equal); Validation (equal); Visualization (equal); Writing-original draft (lead); Writing-review & editing (equal). **Tessa E. F. Quax:** Conceptualization (equal); Data curation (equal); Formal analysis (equal); Funding acquisition (equal); Investigation (lead); Methodology (equal); Resources (equal); Software (equal); Visualization (lead); Writing-original draft (lead); Writing-review & editing (lead). **Tatjana N. Melnik:** Data curation (equal); Formal analysis (equal); Investigation (equal); Methodology (equal); Resources (equal); Software (equal); Validation (equal); Visualization (equal); Writing-original draft (supporting); Writing-review & editing (supporting). **R. Thane Papke:** Formal analysis (equal); Investigation (equal); Methodology (equal); Software (equal); Validation (equal); Writing-original draft (supporting);

Writing-review & editing (supporting). **Johann Peter Gogarten:** Formal analysis (equal); Investigation (equal); Methodology (equal); Software (equal); Validation (equal); Writing-original draft (equal); Writing-review & editing (supporting). **Igor I. Kireev:** Data curation (equal); Formal analysis (equal); Methodology (equal); Resources (equal); Software (equal); Validation (equal); Visualization (equal). **Alexey K. Surin:** Data curation (equal); Formal analysis (lead); Funding acquisition (equal); Investigation (equal); Methodology (equal); Resources (equal); Software (equal); Validation (lead); Visualization (equal). **Sergei N. Beznosov:** Investigation (equal); Methodology (equal); Software (equal); Visualization (equal). **Anna V. Galeva:** Formal analysis (equal); Investigation (equal); Writing-original draft (supporting); Writing-review & editing (supporting). **Oleg V. Fedorov:** Conceptualization (equal); Funding acquisition (equal); Project administration (lead); Resources (lead); Supervision (lead); Writing-original draft (equal).









## ETHICS STATEMENT

None required.

## DATA AVAILABILITY STATEMENT

All data are provided in the results section and appendices (Table A1 and Figures A1–A17) of this paper. The movie files demonstrating swimming behavior of *Hfx. volcanii* strains (Movie S1–S4) are deposited in the Zenodo repository at <https://doi.org/10.5281/zenodo.3723268> [Movie S1. Swimming behavior of *Hfx. volcanii* MT2 transformed with pAS7 (HL-B2-R); Movie S2. Swimming behavior of *Hfx. volcanii* MT2 transformed with pAS6 (HL-B1-R); Movie S3. Swimming behavior of *Hfx. volcanii* MT2 transformed with pAS5 (HL-B1B2-R); Movie S4. Swimming behavior of *Hfx. volcanii* MT2 transformed with pMT21 (HV-A1A2)].

## ORCID

Mikhail G. Pyatibratov  <https://orcid.org/0000-0003-4341-3529>  
 Alexey S. Syutkin  <https://orcid.org/0000-0001-5151-4632>  
 Tessa E. F. Quax  <https://orcid.org/0000-0001-5516-5871>  
 Tatjana N. Melnik  <https://orcid.org/0000-0002-9081-4678>  
 Johann Peter Gogarten  <https://orcid.org/0000-0001-6459-6518>  
 Igor I. Kireev  <https://orcid.org/0000-0001-9252-6808>  
 Alexey K. Surin  <https://orcid.org/0000-0001-9246-7369>  
 Sergei N. Beznosov  <https://orcid.org/0000-0001-9423-1595>  
 Anna V. Galeva  <https://orcid.org/0000-0002-1893-6160>

## REFERENCES

- Alam, M., & Oesterhelt, D. (1984). Morphology, function and isolation of halobacterial flagella. *Journal of Molecular Biology*, 176, 459–475. [https://doi.org/10.1016/0022-2836\(84\)90172-4](https://doi.org/10.1016/0022-2836(84)90172-4)
- Allers, T., Barak, S., Liddell, S., Wardell, K., & Mevarech, M. (2010). Improved strains and plasmid vectors for conditional overexpression of his-tagged proteins in *Haloferax volcanii*. *Applied and Environmental Microbiology*, 76, 1759–1769. <https://doi.org/10.1128/AEM.02670-09>
- Amoozegar, M. A., Siroosi, M., Atashgahi, S., Smidt, H., & Ventosa, A. (2017). Systematics of haloarchaea and biotechnological potential of

- their hydrolytic enzymes. *Microbiology*, 163, 623–645. <https://doi.org/10.1099/mic.0.000463>
- Bardy, S. L., Mori, T., Komoriya, K., Aizawa, S., & Jarrell, K. F. (2002). Identification and localization of flagellins FlaA and FlaB3 within flagella of *Methanococcus voltae*. *Journal of Bacteriology*, 184, 5223–5233. <https://doi.org/10.1128/jb.184.19.5223-5233.2002>
- Becker, E. A., Seitzer, P. M., Tritt, A., Larsen, D., Krusor, M., Yao, A. I., ... Facciotti, M. T. (2014). Phylogenetically driven sequencing of extremely halophilic archaea reveals strategies for static and dynamic osmo-response. *PLOS Genetics*, 10, e1004784. <https://doi.org/10.1371/journal.pgen.1004784>
- Braun, T., Vos, M. R., Kalisman, N., Sherman, N. E., Rachel, R., Wirth, R., ... Egelman, E. H. (2016). Archaeal flagellin combines a bacterial type IV pilin domain with an Ig-like domain. *Proceedings of the National Academy of Sciences of the United States of America*, 113, 10352–10357. <https://doi.org/10.1073/pnas.1607756113>
- Brendel, J., Stoll, B., Lange, S. J., Sharma, K., Lenz, C., Stachler, A. E., ... Marchfelder, A. (2014). A complex of Cas proteins 5, 6, and 7 is required for the biogenesis and stability of clustered regularly interspaced short palindromic repeats (crispr)-derived rnas (crnas) in *Haloferax volcanii*. *Journal of Biological Chemistry*, 289, 7164–7177. <https://doi.org/10.1074/jbc.M113.508184>
- Calladine, C. R. (1978). Change of waveform in bacterial flagella: The role of mechanics at the molecular level. *Journal of Molecular Biology*, 118, 457–479. [https://doi.org/10.1016/0022-2836\(78\)90285-1](https://doi.org/10.1016/0022-2836(78)90285-1)
- Chaban, B., Ng, S. Y., Kanbe, M., Saltzman, I., Nimmo, G., Aizawa, S., & Jarrell, K. F. (2007). Systematic deletion analyses of the *fla* genes in the flagella operon identify several genes essential for proper assembly and function of flagella in the archaeon, *Methanococcus maripaludis*. *Molecular Microbiology*, 66, 596–609. <https://doi.org/10.1111/j.1365-2958.2007.05913.x>
- Coq, N., Du Roure, O., Marthelot, J., Bartolo, D., & Fermigier, M. (2008). Rotational dynamics of a soft filament: Wrapping transition and propulsive forces. *Physics of Fluids*, 20, 051703. <https://doi.org/10.1063/1.2909603>
- Crooks, G. E., Hon, G., Chandonia, J. M., & Brenner, S. E. (2004). WebLogo: A sequence logo generator. *Genome Research*, 14, 1188–1190. <https://doi.org/10.1101/gr.849004>
- Daum, B., Vonck, J., Bellack, A., Chaudhury, P., Reichelt, R., Albers, S.-V., ... Kühlbrandt, W. (2017). Structure and *in situ* organisation of the *Pyrococcus furiosus* archaeellum machinery. *Elife*, 6, e27470. <https://doi.org/10.7554/eLife.27470>
- Desmond, E., Brochier-Armanet, C., & Gribaldo, S. (2007). Phylogenomics of the archaeal flagellum: Rare horizontal gene transfer in a unique motility structure. *BMC Evolutionary Biology*, 7, 106. <https://doi.org/10.1186/1471-2148-7-106>
- Erdmann, S., Tschitschko, B., Zhong, L., Raftery, M. J., & Cavicchioli, R. (2017). A plasmid from an Antarctic haloarchaeon uses specialized membrane vesicles to disseminate and infect plasmid-free cells. *Nature Microbiology*, 2, 1446–1455. <https://doi.org/10.1038/s41564-017-0009-2>
- Fedorov, O. V., Pyatibratov, M. G., Kostyukova, A. S., Osina, N. K., & Tarasov, V. Y. (1994). Protofilament as a structural element of flagella of haloalkalophilic archaeobacteria. *Canadian Journal of Microbiology*, 40, 45–53. <https://doi.org/10.1139/m94-007>
- Franzmann, P. D., Stackebrandt, E., Sanderson, K., Volkman, J. K., Cameron, D. E., Stevenson, P. L., ... Burton, H. R. (1988). *Halobacterium lacusprofundi* sp. nov., a halophilic bacterium isolated from deep lake, Antarctica. *Systematic and Applied Microbiology*, 11, 20–27. [https://doi.org/10.1016/S0723-2020\(88\)80044-4](https://doi.org/10.1016/S0723-2020(88)80044-4)
- Gerl, L., Deutzmann, R., & Sumper, M. (1989). Halobacterial flagellins are encoded by a multigene family. Identification of all five gene products. *FEBS Letters*, 244, 137–140. [https://doi.org/10.1016/0014-5793\(89\)81179-2](https://doi.org/10.1016/0014-5793(89)81179-2)
- Gerl, L., & Sumper, M. (1988). Halobacterial flagellins are encoded by a multigene family. Characterization of five flagellin genes. *Journal of Biological Chemistry*, 263, 13246–13251.
- Guan, Z., Naparstek, S., Calo, D., & Eichler, J. (2012). Protein glycosylation as an adaptive response in Archaea: Growth at different salt concentrations leads to alterations in *Haloferax volcanii* S-layer glycoprotein N-glycosylation. *Environmental Microbiology*, 14, 743–753. <https://doi.org/10.1111/j.1462-2920.2011.02625.x>
- Gupta, R. S., Naushad, S., Fabros, R., & Adeolu, M. (2016). A phylogenomic reappraisal of family-level divisions within the class *Halobacteriia*: 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. <https://doi.org/10.1007/s10482-016-0660-2>
- Jarrell, K. F., & Albers, S. V. (2012). The archaeellum: An old motility structure with a new name. *Trends in Microbiology*, 20, 307–312. <https://doi.org/10.1016/j.tim.2012.04.007>
- Katoh, K., & Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Molecular Biology and Evolution*, 30, 772–780. <https://doi.org/10.1093/molbev/mst010>
- Li, M. Z., & Elledge, S. J. (2012). SLIC: A method for sequence- and ligation-independent cloning. *Methods in Molecular Biology*, 852, 51–59. [https://doi.org/10.1007/978-1-61779-564-0\\_5](https://doi.org/10.1007/978-1-61779-564-0_5)
- Liao, Y., Williams, T. J., Walsh, J. C., Ji, M., Poljak, A., Curmi, P. M. G., ... Cavicchioli, R. (2016). Developing a genetic manipulation system for the Antarctic archaeon, *Halorubrum lacusprofundi*: Investigating acetamidase gene function. *Scientific Reports*, 6, 34639. <https://doi.org/10.1038/srep34639>
- Löytynoja, A., & Goldman, N. (2008). Phylogeny-aware gap placement prevents errors in sequence alignment and evolutionary analysis. *Science*, 320, 1632–1635. <https://doi.org/10.1126/science.1158395>
- Lukeš, J., Archibald, J. M., Keeling, P. J., Doolittle, W. F., & Gray, M. W. (2011). How a neutral evolutionary ratchet can build cellular complexity. *IUBMB Life*, 63, 528–537. <https://doi.org/10.1002/iub.489>
- Maljkovic Berry, I., Athreya, G., Kothari, M., Daniels, M., Bruno, W., Korber, B., ... Leitner, T. (2009). A simple method for optimizing the root in phylogenetic trees with longitudinal data reveals an evolutionary rate that tracks the epidemiological dynamics of HIV-1. *Epidemics*, 1, 230–239. <https://doi.org/10.1016/j.epidem.2009.10.003>
- Meshcheryakov, V. A., Shibata, S., Schreiber, M. T., Villar-Briones, A., Jarrell, K. F., Aizawa, S. I., & Wolf, M. (2019). High-resolution archaeellum structure reveals a conserved metal-binding site. *EMBO Reports*, 20, e46340. <https://doi.org/10.15252/embr.201846340>
- Mullakhanbhai, M. F., & Larsen, H. (1975). *Halobacterium volcanii* spec. nov., a dead sea halobacterium with a moderate salt requirement. *Archives of Microbiology*, 104, 207–214. <https://doi.org/10.1007/bf00447326>
- Nguyen, L.-T., Schmidt, H. A., von Haeseler, A., & Minh, B. Q. (2015). IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Molecular Biology and Evolution*, 32, 268–274. <https://doi.org/10.1093/molbev/msu300>
- Poweleit, N., Ge, P., Nguyen, H. H., Loo, R. R., Gunsalus, R. P., & Zhou, Z. H. (2016). CryoEM structure of the *Methanospirillum hungatei* archaeellum reveals structural features distinct from the bacterial flagellum and type IV pilus. *Nature Microbiology*, 2, 16222. <https://doi.org/10.1038/nmicrobiol.2016.222>
- Privalov, P. L., & Potekhin, S. A. (1986). Scanning microcalorimetry in studying temperature-induced changes in proteins. *Methods in Enzymology*, 131, 4–51. [https://doi.org/10.1016/0076-6879\(86\)31033-4](https://doi.org/10.1016/0076-6879(86)31033-4)
- Pyatibratov, M. G., Beznosov, S. N., Rachel, R., Tiktopulo, E. I., Surin, A. K., Syutkin, A. S., & Fedorov, O. V. (2008). Alternative flagellar filament types in the haloarchaeon *Haloarcula marismortui*. *Canadian Journal of Microbiology*, 54, 835–844. <https://doi.org/10.1139/w08-076>

- Quax, T. E. F., Altegoer, F., Rossi, F., Li, Z., Rodriguez-Franco, M., Kraus, F., ... Albers, S.-V. (2018). Structure and function of the archaeal response regulator CheY. *Proceedings of the National Academy of Sciences of the United States of America*, 115, E1259–E1268. <https://doi.org/10.1073/pnas.1716661115>
- Sela, I., Ashkenazy, H., Katoh, K., & Pupko, T. (2015). GUIDANCE2: Accurate detection of unreliable alignment regions accounting for the uncertainty of multiple parameters. *Nucleic Acids Research*, 43, W7–W14. <https://doi.org/10.1093/nar/gkv318>
- Senin, A. A., Potekhin, S. A., Tiktupulo, E. I., & Filimonov, V. V. (2000). Differential scanning microcalorimeter SCAL-1. *Journal of Thermal Analysis and Calorimetry*, 62, 153–160. <https://doi.org/10.1023/A:1010171013669>
- Sievers, F., & Higgins, D. G. (2018). Clustal Omega for making accurate alignments of many protein sciences. *Protein Science*, 27, 135–145. <https://doi.org/10.1038/msb.2011.75>
- Sorokin, D. Y., Elcheninov, A. G., Toshchakov, S. V., Bale, N. J., Sinninghe Damsté, J. S., Khijniak, T. V., & Kublanov, I. V. (2019). *Natrarchaeobius chitinivorans* gen. nov., sp. nov., and *Natrarchaeobius halalkaliphilus* sp. nov., alkaliphilic, chitin-utilizing haloarchaea from hypersaline alkaline lakes. *Systematic and Applied Microbiology*, 42, 309–318. <https://doi.org/10.1016/j.syapm.2019.01.001>
- Syutkin, A. S., Pyatibratov, M. G., Beznosov, S. N., & Fedorov, O. V. (2012). Various mechanisms of flagella helicity formation in haloarchaea. *Microbiology*, 81, 573–581. <https://doi.org/10.1134/S0026261712050153>
- Syutkin, A. S., Pyatibratov, M. G., Galzitskaya, O. V., Rodríguez-Valera, F., & Fedorov, O. V. (2014). *Haloarcula marismortui* archaeellin genes as ecoparalogs. *Extremophiles*, 18, 341–349. <https://doi.org/10.1007/s00792-013-0619-4>
- Syutkin, A. S., van Wolferen, M., Surin, A. K., Albers, S. V., Pyatibratov, M. G., Fedorov, O. V., & Quax, T. E. F. (2019). Salt-dependent regulation of archaeellins in *Haloarcula marismortui*. *Microbiologyopen*, 8, e00718. <https://doi.org/10.1002/mbo3.718>
- Tarasov, V. Y., Kostyukova, A. S., Tiktupulo, E. I., Pyatibratov, M. G., & Fedorov, O. V. (1995). Unfolding of tertiary structure of *Halobacterium halobium* flagellins does not result in flagella destruction. *Journal of Protein Chemistry*, 14, 27–31. <https://doi.org/10.1007/bf01902841>
- Tarasov, V. Y., Pyatibratov, M. G., Beznosov, S. N., & Fedorov, O. V. (2004). On the supramolecular organization of the flagellar filament in archaea. *Doklady Biochemistry and Biophysics*, 396, 203–205. <https://doi.org/10.1023/B:DOBI.0000033530.66078.a6>
- Tarasov, V. Y., Pyatibratov, M. G., Tang, S. L., Dyll-Smith, M., & Fedorov, O. V. (2000). Role of flagellins from A and B loci in flagella formation of *Halobacterium salinarum*. *Molecular Microbiology*, 35, 69–78. <https://doi.org/10.1046/j.1365-2958.2000.01677.x>
- Tomlinson, G. A., & Hochstein, L. I. (1976). *Halobacterium saccharovorum* sp. nov., a carbohydrate-metabolizing, extremely halophilic bacterium. *Canadian Journal of Microbiology*, 22, 587–591. <https://doi.org/10.1139/m76-087>
- Tony, S. Y., Lauga, E., & Hosoi, A. (2006). Experimental investigations of elastic tail propulsion at low Reynolds number. *Physics of Fluids*, 18, 091701. <https://doi.org/10.1063/1.2349585>
- Tripepi, M., Esquivel, R. N., Wirth, R., & Pohlschröder, M. (2013). *Haloferax volcanii* cells lacking the flagellin FlgA2 are hypermotile. *Microbiology*, 159, 2249–2258. <https://doi.org/10.1099/mic.0.069617-0>
- Tripepi, M., Imam, S., & Pohlschröder, M. (2010). *Haloferax volcanii* flagella are required for motility but are not involved in PibD-dependent surface adhesion. *Journal of Bacteriology*, 192, 3093–3102. <https://doi.org/10.1128/JB.00133-10>
- Tripepi, M., You, J., Temel, S., Önder, Ö., Brisson, D., & Pohlschröder, M. (2012). N-glycosylation of *Haloferax volcanii* flagellins requires known Agl proteins and is essential for biosynthesis of stable flagella. *Journal of Bacteriology*, 194, 4876–4887. <https://doi.org/10.1128/JB.00731-12>
- Tschitschko, B., Erdmann, S., DeMaere, M. Z., Roux, S., Panwar, P., Allen, M. A., ... Cavicchioli, R. (2018). Genomic variation and biogeography of Antarctic haloarchaea. *Microbiome*, 6, 113. <https://doi.org/10.1186/s40168-018-0495-3>
- Tschitschko, B., Williams, T. J., Allen, M. A., Páez-Espino, D., Kyrpidis, N., Zhong, L., ... Cavicchioli, R. (2015). Antarctic archaea-virus interactions: Metaproteome-led analysis of invasion, evasion and adaptation. *ISME Journal*, 9, 2094–2107. <https://doi.org/10.1038/ismej.2015.110>
- Wolgemuth, C. W., Powers, T. R., Goldstein, R., & E., (2000). Twirling and whirling: Viscous dynamics of rotating elastic filaments. *Physical Review Letters*, 84, 1623–1626. <https://doi.org/10.1103/PhysRevLett.84.1623>
- Zaretsky, M., Darnell, C. L., Schmid, A. K., & Eichler, J. (2019). N-Glycosylation is important for *Halobacterium salinarum* archaeellin expression, archaeellum assembly and cell motility. *Frontiers in Microbiology*, 10, 1367. <https://doi.org/10.3389/fmicb.2019.01367>
- Zaretsky, M., Roine, E., & Eichler, J. (2018). Sialic acid-like sugars in Archaea: Legionaminic acid biosynthesis in the halophile *Halorubrum* sp. PV6. *Frontiers in Microbiology*, 9, 2133. <https://doi.org/10.3389/fmicb.2018.02133>

**How to cite this article:** Pyatibratov MG, Syutkin AS, Quax TEF, et al. Interaction of two strongly divergent archaeellins stabilizes the structure of the *Halorubrum* archaeellum. *MicrobiologyOpen*. 2020;9:e1047. <https://doi.org/10.1002/mbo3.1047>

## APPENDIX A

**TABLE A1** A list of selected haloarchaea with known genomic sequences: 50 species of the *Halorubrum* genus and species of other genera having archaeal genes, whose origin and genome organization are close to the *Halorubrum* archaealins. Proteins whose amino acid sequences are significantly different from the FlaB1, typical of *Halorubrum* species, are marked in red

#	Archaea	FlaB1		FlaB2		Identity		Nucleotide spacer between archaeal genes	Number of archaeal genes	notes
		Accession No	Size (aa)	Accession No	Size (aa)	FlaB1/FlaB2 (%)	FlaB1/FlaB2 (%)			
<i>Halorubrum</i> (taxid:56,688)										
1	<i>Hrr. aethiopicum</i>	WP_066413746.1	199	WP_066413747.1	241	40	4	CA		*—very divergent FlaB2 with an insert in the central part, contain AidA superfamily domain with probable adhesive properties
		WP_066413749.1	202	Unnamed*	335			CA		
2	<i>Hrr. aidingense</i>	WP_008001643.1	196	WP_008001645.1	225	42	4	CA		Archaeal operon duplicated
		WP_008001647.1	190	WP_008001649.1	217	45		CG		
3	<i>Hrr. arcis</i>	WP_007996033.1	199	WP_007996034.1	215	42	2	CA		
4	<i>Hrr. californiensis</i> DSM 19,288	WP_008440986.1	196	WP_008440988.1	239	32	2	CA		
5	<i>Hrr. chaoviator</i>	SNR51654.1	202	SNR51643.1	456	24	2	CA		
6	<i>Hrr. coriense</i>	WP_006114527.1	206	WP_006114526.1	217	36	2	AA		
7	<i>Hrr. distributum</i> E8	OYR82934.1	199	OYR82935.1	215	43	2	CA		
8	<i>Hrr. distributum</i> JCM 9,100	WP_004597585.1	193	WP_004597583.1	207	44	2	CA		
9	<i>Hrr. ezzemoulense</i> DSM 17,463	WP_049933480.1	198	Unnamed	412	No data	2	AC		
10	<i>Hrr. ezzemoulense</i> Ec15	WP_094494443.1	204	Unnamed	418	No data	2	CA		
11	<i>Hrr. ezzemoulense</i> Fb21	WP_100050495.1	198	Unnamed	>283*	No data	2	CA		*—incomplete data
12	<i>Hrr. ezzemoulense</i> G37	WP_094582235.1	205	WP_094582237.1	229	42	2	CC		
13	<i>Hrr. ezzemoulense</i> Ga2p	OYR66521.1	207*	OYR66519.1	361	No data	2	CG		*—corrected data
14	<i>Hrr. ezzemoulense</i> Ga36	WP_094553369.1	206	WP_094553368.1	225	45	2	CA		

(Continues)



TABLE 1 (Continued)

#	Archaea	FlaB1		FlaB2		Identity		Nucleotide spacer between archaeal genes	Number of archaeal genes	notes
		Accession No	Size (aa)	Accession No	Size (aa)	FlaB1/FlaB2 (%)	FlaB1/FlaB2 (%)			
15	<i>Hrr. ezzemoulense</i> LD3	WP_094579523.1	197	Unnamed	423	No data		CA	2	
16	<i>Hrr. ezzemoulense</i> LG1	WP_094521168.1	193	WP_094521167.1	203	43		CA	4	
17	<i>Hrr. halodurans</i>	WP_094520318.1	197	Unnamed	418	No data		CA	3	All 3 genes are transcribed as one operon.
		WP_094529811.1	190							
18	<i>Hrr. halophilum</i>	WP_094529813.1	187							All 3 archaealins belong to the FlaB1 branch
		WP_094529815.1*	224							
19	<i>Hrr. hochstenium</i>	WP_050033245.1	196	WP_050033244.1	217	43		CC	2	*—combines the properties both FlaB1 and FlaB2
20	<i>Hrr. kocurii</i>	WP_008585051.1	199	WP_008585048.1	227	43		CA	2	
21	<i>Hrr. lacusprofundi</i> ATCC 49,239	WP_008847356.1	192	WP_008847355.1	233	41		CA	2	
22	<i>Hrr. lacusprofundi</i> DL18	none	—	WP_015911241.1	243	—		—	1	
23	<i>Hrr. lipolyticum</i>	WP_088901573.1	206	WP_088901574.1	243	36		CA	2	
24	<i>Hrr. litoreum</i>	WP_008008131.1	191	WP_008008132.1	235	41		CA	2	
25	<i>Hrr. persicum</i>	WP_008367009.1	210	WP_008367012.1	239	40		CA	2	
26	<i>Hrr. saccharovororum</i> DSM1137	WP_099253835.1	204	Unnamed	407	No data		CA	2	
27	<i>Hrr. saccharovororum</i> H3	WP_004047439.1	196	WP_004047440.1	236	42		CA	2	
28	<i>Hrr. sodomense</i>	WP_050026362.1	196	WP_050026360.1	217	45		CC	2	
29	<i>Hrr. sp. 48_1_W</i>	WP_092919746.1	198	WP_092919748.1	211	43		CA	2	
30	<i>Hrr. sp. AJ67</i>	WP_112080473.1	197	Unnamed	327	No data		CG	2	*—related to SNR51643.1 of <i>Hrr. chaoviator</i>
31	<i>Hrr. sp. Atlit-8R</i> (Atlit-9R)	WP_048077909.1	197	WP_048077910.1	237	41		CA	3	*—3rd distant archaeal gene related to SNR51643.1 of <i>Hrr. chaoviator</i>
		none	—	WP_048075921.1*	331	—		—		
32	<i>Hrr. sp. Atlit-26R</i>	WP_121563110.1	196	WP_121563109.1	214	43		CA	2	
32	<i>Hrr. sp. Atlit-26R</i>	WP_121598119.1	198	WP_121598118.1	227	43		CA	3	*—partial, diverged significantly from 2nd paralog. Incomplete data.
		?	—	WP_121598601.1* (206?)	189	—		—	—	

(Continues)

TABLE 1 (Continued)

#	Archaea	FlaB1		FlaB2		Size (aa)	Identity	Number of archaeallin genes	Nucleotide spacer between archaeallin genes	notes
		Accession No	Size (aa)	Accession No	Size (aa)					
33	<i>Hrr. sp. Atlit-28R</i>	WP_121599988.1	204	Unnamed <sup>*</sup>	421	No data	2	CA		
34	<i>Hrr. sp. BY1</i>	WP_049983855.1	189	WP_049983854.1	201	46	2	TT		
35	<i>Hrr. sp. C191</i>	WP_099288816.1	203	Unnamed <sup>*</sup>	455	No data	2	CA	*—closely related to SNR51643.1 of <i>Hrr. chaoviator</i>	
36	<i>Hrr. sp. CBA1229</i>	WP_123113424.1	209	WP_123113425.1	239	38	2	CA		
37	<i>Hrr. sp. CSM-61</i>	WP_123623808.1	194	WP_123623809.1	235	43	2	CA		
38	<i>Hrr. sp. Ea1</i>	WP_094557020.1	205	WP_094557021.1	310	25	4	CA		
		WP_094557025.1	204	WP_094557019.1	243	31		CA		
39	<i>Hrr. sp. Ea8</i>	WP_094527332.1	205	Unnamed <sup>*</sup>	451	No data	2	CA	*—closely related to SNR51643.1 of <i>Hrr. chaoviator</i>	
40	<i>Hrr. sp. Eb13</i>	WP_094524745.1	205	Unnamed <sup>*</sup>	454	No data	2	CA	*—closely related to SNR51643.1 of <i>Hrr. chaoviator</i>	
41	<i>Hrr. sp. Hd13</i>	WP_094590524.1	202	Unnamed <sup>*</sup>	332	No data	2	AT	*—related to SNR51643.1 of <i>Hrr. chaoviator</i>	
42	<i>Hrr. sp. lb24</i>	WP_094580884.1	204	Unnamed <sup>*</sup>	>336 <sup>*</sup>	No data	2	CA	*—incomplete data	
43	<i>Hrr. sp. SD612</i>	WP_086220603.1	199	Unnamed <sup>*</sup>	418	No data	2	CA	*—Contains a DUF4397 domain of unknown function (it is presented in algininate O-acetyl transferase AlgF).	
44	<i>Hrr. sp. T3</i>	WP_017343754.1	204	WP_017343755.1	414	No data	2	CA		
45	<i>Hrr. sp. WN019</i>	WP_095636868.1	195	WP_095636867.1	234	36	4	CA	Archaeallin operon duplicated	
46	<i>Hrr. tebenquichense</i>	WP_095636866.1	190	WP_095636865.1	223	43	2	CA		
47	<i>Hrr. terrestris</i>	WP_006628949.1	208	WP_006628948.1	239	40	2	CA		
48	<i>Hrr. trapanicum</i>	WP_007345485.1	198	WP_007345484.1	215	40	2	CA		
49	<i>Hrr. tropicale</i>	WP_096394437.1	197	WP_096394438.1	220	42	2	CA		
		WP_053770455.1	196	WP_053770456.1	226	42	4	CA	Archaeallin operon duplicated	
		WP_053770457.1	191	WP_053770458.1	223	41		CG		

(Continues)

TABLE 1 (Continued)

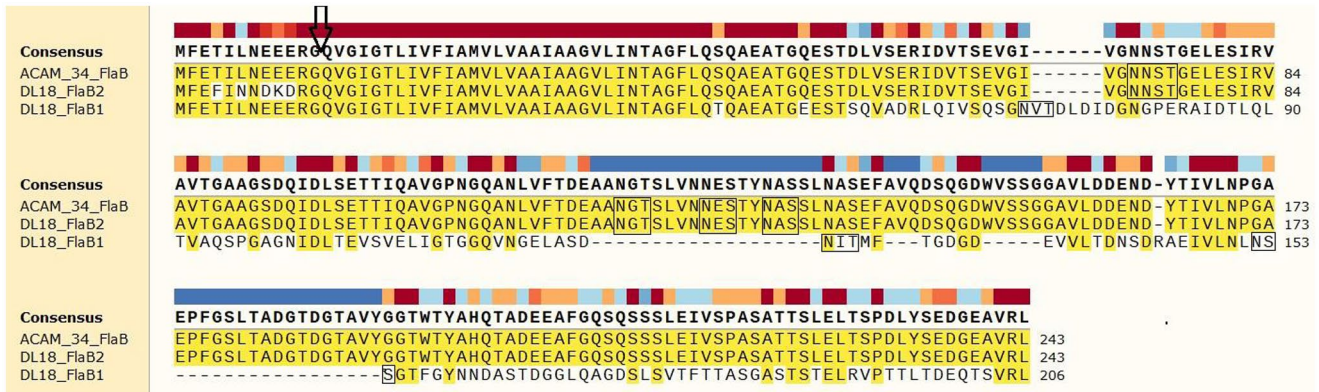
#	Archaea	FlaB1		FlaB2		Identity		Nucleotide spacer between archaeal genes	notes
		Accession No	Size (aa)	Accession No	Size (aa)	FlaB1/FlaB2 (%)	Number of archaeal genes		
50	<i>Hrr. vacuolatum</i>	WP_089384684.1	239				4		All 4 archaealins belong to the FlaB1 branch. *—are transcribed as one operon.
		WP_089384625.1	260						
		WP_089384626.1*	266						
		WP_089384627.1*	235						
<i>Halopiger</i> (taxid:387342)									
1	<i>Hpg. aswanensis</i>	WP_120245279.1*	199	WP_120245070.1	209	43	3	CA	*—belong to the FlaB1 branch.
		WP_120243198.1*	236						
2	<i>Hpg. djelfmassiliensis</i>	WP_081661498.1*	213				4		*—are transcribed as one operon. All 4 archaealins belong to the FlaB1 branch.
		WP_049922547.1*	243						
		WP_081661562.1*	213						
		WP_049923757.1	238						
3	<i>Hpg. goleiensis</i>	WP_081655481.1*	402	WP_049925873.1*	208	No data	3	CA	*—are transcribed as one operon.
				WP_049925872.1	208				
4	<i>Hpg. xanaduensis</i>	WP_013881028.1	193	WP_013881027.1	208	44	2	TC	
<i>Natrialba</i> (taxid:1.644,060)									
1	<i>Nab. aegyptia</i>	WP_006664776.1	195	WP_006664775.1	222	46	2	CC	
2	<i>Nab. asiatica</i>	WP_006110336.1	195	WP_006110334.1	222	46	2	CC	
3	<i>Nab. chahannaensis</i>	WP_006169645.1	431	WP_006169644.1	212	No data	2	CA	
4	<i>Nab. hulunbeirensis</i>	WP_006653761.1	481	WP_006653762.1	212	No data	2	CA	
5	<i>Nab. magadii</i>	WP_004267190.1	201	WP_004267189.1	259	33	4	CC	
		WP_012996744.1	395	WP_004267187.1	261	No data		CC	
6	<i>Nab. sp. SSL1</i>	WP_071401430.1	445	WP_071401429.1	219	No data	4	CA	*—incomplete data
		WP_071400338.1*	>147	WP_084777668.1	219	No data		CC	
7	<i>Nab. taiwanensis</i>	WP_006664776.1	195	WP_006826473.1	222	47	2	CC	
<i>Natronolimnobius</i> (taxid:253,106)									
1	<i>Nln. aegyptiacus</i>	WP_086887356.1*	199	WP_086887355.1*	205	43	3	TC	*—are transcribed as one operon.

(Continues)

TABLE 1 (Continued)

#	Archaea	FlaB1		FlaB2		Identity		Nucleotide spacer between archaeellin genes	notes
		Accession No	Size (aa)	Accession No	Size (aa)	FlaB1/FlaB2 (%)	Number of archaeellin genes		
2	<i>Nln. sulfireducens</i> AArc1	WP_086887354.1	208	WP_086887354.1	208	42			
		AXR78766.1	475 (487)	AXR78764.1	220	No data	5	?	
		AXR78768.1	410	AXR78765.1	216	No data		CC	
3	<i>Nln. sulfireducens</i> AArc-Mg	WP_117370007.1	460	WP_117367910.1	217	No data	5	CA	
		Unnamed*	466	WP_117367912.1	218			?	
		WP_117368415.1	233						*~99% identity with WP_117370007.1
<i>Halobiforma</i> (taxid:203,193)									
1	<i>Hbf. haloterrestis</i>	WP_089789912.1	416	WP_089789910.1	219	No data	2	CA	
2	<i>Hbf. laciisalsi</i>	WP_007139842.1	462	WP_007139843.1	214	No data	2	CA	
3	<i>Hbf. nitratireducens</i>	WP_006673863.1	312						
		WP_006673864.1	334						
		WP_006673865.1	332						
		WP_006673866.1	310						
		WP_006673867.1	240						
<i>Natrarchaeobius</i> (taxid:2,501,796)									
1	<i>N. chitinivorans</i>	WP_124196377.1	201	WP_124196378.1	213	47		CA	
		WP_124192235.1*	208						*—the archaeellin operon is adjacent to an additional fla-gene cluster
2	<i>N. halalkaliphilus</i>	WP_124192236.1*	232						
		WP_124192237.1*	248						
		WP_124179648.1	188	WP_124179647.1	215	42		CA	

APPENDIX B



**FIGURE A1** Alignment of the sequence of the ACAM 34 and DL18 *Hrr. lacusprofundi* archaellins. Sequence alignment was performed using Clustal Omega (Sievers & Higgins, 2018; <http://www.clustal.org/omega/>). In each sequence, potentially N-glycosylation sites, N-X-T(S), are boxed. The signal peptide cleavage site is indicated by an arrow



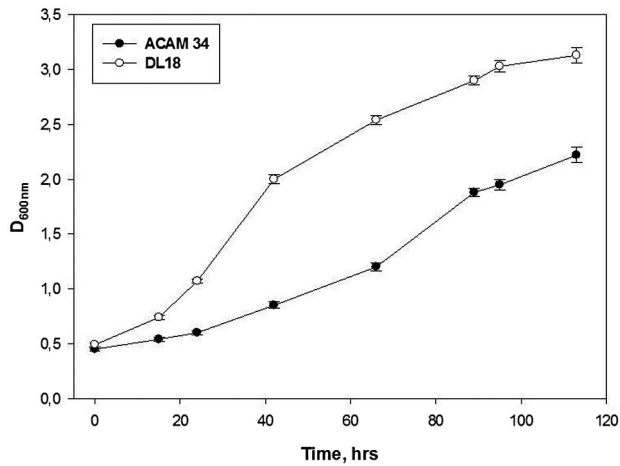
>NZ\_MKFG01000003.1:83451-85227 Halorubrum lacusprofundi strain DL18 contig003, whole genome shotgun sequence

GGCGGGTCGGAACCTCCGCGGAGCGGCCGAAGCGATCGGCGGCGGACCAGCGGATTCGGGAAGCGACTGTCCGACGATCGCACGCGAGAGCGTCCGGG  
 ATCCGCGTGCCGACCGCGCTCGATCCGACCGTCGATACGGAACCTCTACAGATCCAACGCCGATAATCCGATCAGCACTTTGTTGATATATTCAAAA  
 ATCATTGCTAGAAAATAATCCGATAATCGCAACTGCAGCGTTTTATCAATTTGGATAATCTGCACGGAGTCTTTAATAGGGAAATTACACTTCG  
 TAGCGATAAACCCCGCGGGCAGGCCGGGGCCAACGCAACACACAATGTTTCGAACAATACTGAACGAGGAAGAGCGCGGTCAAGTGGGGATCGGG  
 ACCCTCATCGTGTTCATCGCGATGGTGTGGTGGCGCGATCGCCGCCGGCTCTGATCAACACGGCCGGCTTCCTTCAGACGCAGGCGGAAGCAA  
 CGGGCGAAGAGAGTACGAGTACGGTGGCCGACCGGCTGCAGATCGTCAGCCAGTCGGGGAACGTGACTGATCTTGATATTGACGGAAACGGTCCCGA  
 ACGCGCGATCGATACCCTTCAGCTCACGGTTCGCGCAGTCAACCGGTTGCCGAAACATCGATCTCACGGAAGTGAAGCGTCAACTGATCGGCACCGGT  
 GGGCAGGTCAACGGAGAAGTGGCCAGCGATAACATCACGATGTTACCGGAGACGGTGACGAAGTCTGTTCTCACCGACAACAGCGACCGCGCTGAGA  
 TTGTGCTCAACCTCAATAGTAGCGGAACGTTCCGGTACAACAACGATGCAAGCACTGACGGCGGTCTACAGGCTGGCGACAGCCTCTCGGTGACGTT  
 CACGACCGCGTCCGGTGCATCGACGAGTACCGAAGTGCAGCTCCCGACGACGCTGACTGACGAGCAAACCTCGGTGAGGCTGTAAACAATGTTTCGAAGT  
 TTATCAACAAACGACAAGACCGCGGTACGGTTGGCATCGGTACGCTCATCGTGTTCATCGCGATGGTACTGGTGGCGCGATCGCCGCCGGTGTCT  
 GATCAACACGGCCGGCTTCCTCCAGTTCGACGGCGGAAGCAACCGGACAGGAGAGTACGGATCTCGTCTCCGAACGGATCGACGTGACGAGCGAGGTT  
 GGTATCGTCCGGAAACAGCACCGGCGAAGTTGAGTTCGATCCGCGTCCCGGTTACCGGTGCGCGCGGTCCGATCAGATCGACTTATCAGAGACGA  
 CGATCCAAGCGGTCCGTCCGAACGGACAGGCGAAGCTCGTGTTCACCGACGAGGCTGCAAAATGGTACATCTCTAGTCAATAACGAGAGCACATACAA  
 TGGCAGCAGCCTCAATGCGAGTGAAGTTCGCTGTGCAAGATTCCCAAGGCGATTGGGTGACGAGTGGCGGTGCAAGTGTGGACGACGAGAACGATTAC  
 ACCATCGTCTCAACCTGGCGCAGAACCGTTCCGGAGCCTCACTGCGGACGGTACCGATGGCACAGCAGTCTACGGTGGAACTGGACCTACGCTC  
 ACCAAACTGCAGACGAAGAGGCTTCGGACAGAGCCAATCTCGTTCGATCGAGATCGTCTCGCCCGCTCGGGCAGCACCTCACTCGAATCACTTC  
 GCCCGACTCTACAGCGAAGACGGCGAAGCGGTCCGGCTCTAACGACCTCGAATAGCCCCCTAACTACCTTCCCGCGGATCCCTCCGCCCGAAC  
 ACCTATTTCTCGCACCAACGGCTACGCC

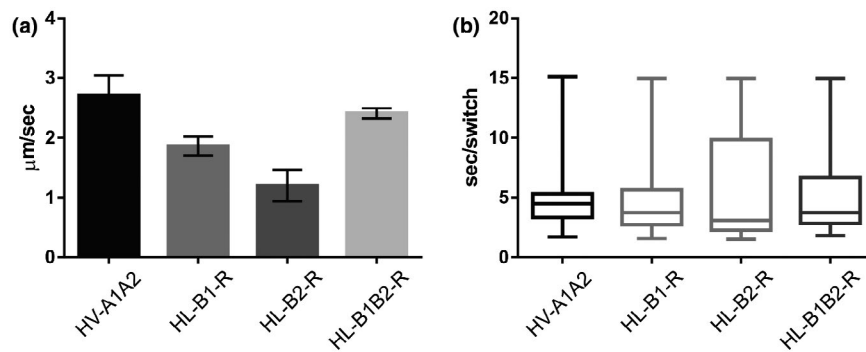
>CP001365.1:2532506-2533659 Halorubrum lacusprofundi ATCC 49239 chromosome 1, complete sequence

GGCGGGTCGGAACCTCCGCGGAGCGGCCGAAGCGATCGGCGGCGGACCAGCGGATTCGGGAAGCGACTGTCCGACGATCGCACGCGAGAGCGTCCGGG  
 ATCCGCGTGCCGACCGCGCTCGATCCGACCGTCGATACGGAACCTCTACAGATCCAACGCCGATAATCCGATCAGCACTTTGTTGATATATTCAAAA  
 ATCATTGCTAGAAAATAATCCGATAATCGCAACTGCAGCGTTTTATCAATTTGGATAATCTGCACGGAGTCTTTAATAGGGAAATTACACTTCG  
 TAGCGATAAACCCCGCGGGCAGGCCGGGGCCAACGCAACACACAATGTTTCGAACAATACTGAACGAGGAAGAGCGCGGTCAAGTGGGCATCGGT  
 ACGCTCATCGTGTTCATCGCGATGGTACTGGTGGCGCGATCGCCGCCGGTGTCTGATCAACACGGCCGGCTTCCTCCAGTTCGACGGCGGAAGCAA  
 CCGGACAGGAGAGTACGGATCTCGTCTCCGAACGGATCGACGTGACGAGCGAGGTCGGTATCGTCCGGAAACAACAGCACCGGCGAAGTTGAGTTCGAT  
 CCGCGTCCCGGTTACCGGTGCGCGCGGTCCGATCAGATCGACTTATCAGAGACGACGATCCAAGCGGTCCGTCCGAACGGACAGGCGAAGCTCGTG  
 TTCACCGACGAGGCTGCAAAATGGTACATCTCTAGTCAATAACGAGAGCACATACAAATGCGAGCAGCCTCAATGCGAGTGAAGTTCGCTGTGCAAGATT  
 CCCAAGGCGATTGGGTGACGAGTGGCGGTGCAAGTGTGGACGACGAGAACGATTACACCATCGTCTCAACCTGGCGCAGAACCGTTCCGGAGCCT  
 CACTGCGGACGGTACCGATGGCACAGCAGTCTACGGTGGAACTGACCTACGCTCACCAAACTGCAGACGAAGAGGCTTCGGACAGAGCCAATCC  
 TCGTTCGCTCGAGATCGTCTCGCCCGCTCGGCGACGACCTCACTCGAATCACTTCGCCCGACTCTACAGCGAAGACGGCGAAGCGGTCCGGCTCT  
 AACGACCTCGAATAGCCCCCTAACTACCTTCCCGCGGATCCCTCCGCCCGAACACCTATTTCTCGCACCAACGGCTACGCC

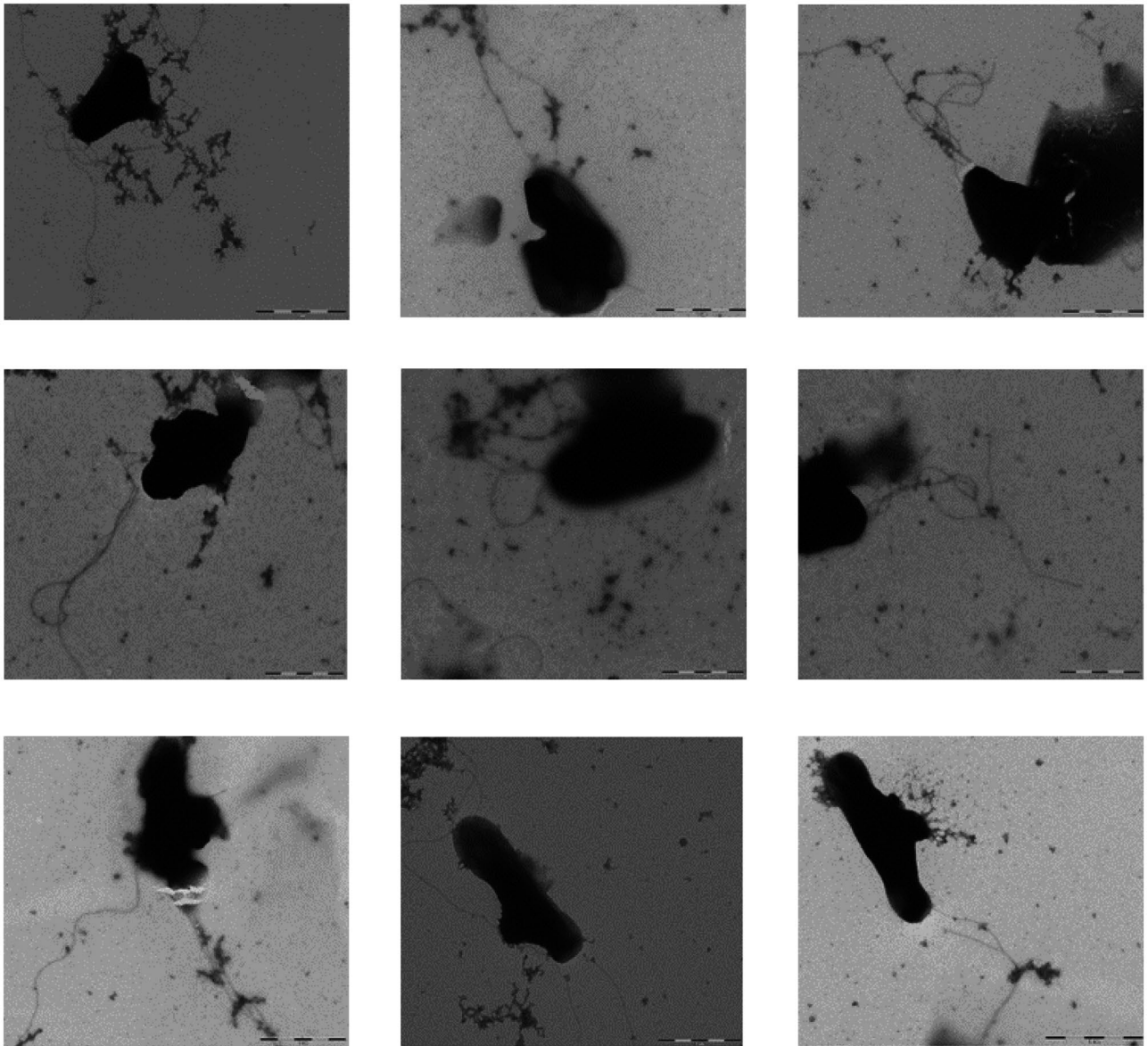
**FIGURE A2** Nucleotide sequences of archaellin genes of *Hrr. lacusprofundi* DL18 and ACAM 34 strains. The sequence of the *flaB1* gene is highlighted in yellow and *flaB2* gene 2—in green. The sequences that are only in *flaB2* of the DL18 and are absent in that of ACAM 34 are highlighted in red. Sequences common to both *flaB1* and *flaB2* of the ACAM 34 are highlighted in blue



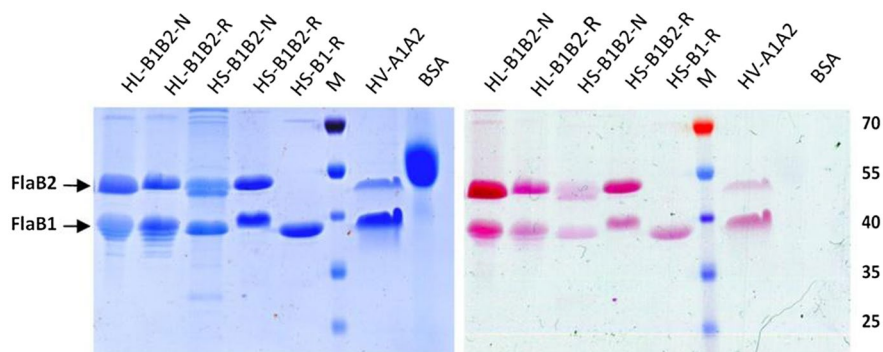
**FIGURE A3** Growth curves of *Hrr. lacusprofundi* ACAM 34 and DL18 strains (in three central spots), St-HL liquid medium, 30°C



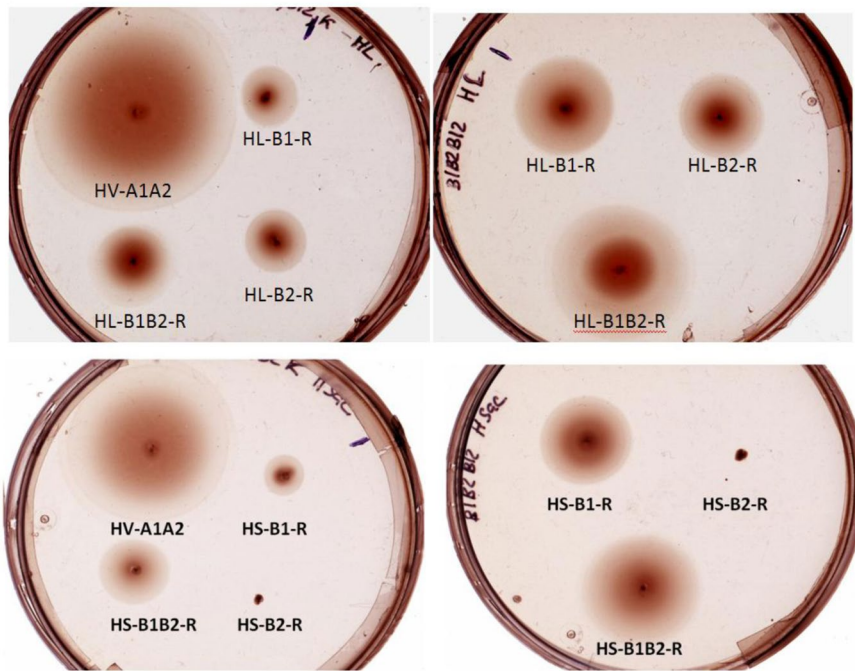
**FIGURE A4** Swimming behavior of *Hfx. volcanii* strains expressing different combinations of archaellins. Liquid cultures were analyzed by light microscopy at 45°C. (a) The average velocity of different strains. Bars indicate *SD*. (b) Tukey box plot of the number of seconds between two subsequent turns of  $> 90^\circ$ . For each strain, the middle line in the box displays the median. Boxes display the 25–75th percentile, and lower and upper bars represent the minimum and maximum time. HL, archaellin from *Hrr. lacusprofundi*. HV, archaellin from *Hfx. volcanii*



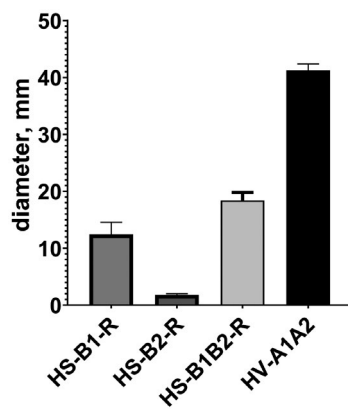
**FIGURE A5** Electron micrographs of *Hfx. volcanii* MT2 transformed with pAS5 (HL-B1B2-R) (top), pAS6 (HL-B1-R) (middle), and pAS7 (HL-B2-R) (bottom) negatively stained with 2% uranyl acetate. Scale bars = 1  $\mu$ m



**FIGURE A6** Archaellin staining with Schiff's reagent in 12.5% polyacrylamide gel (right); the same gel stained with Coomassie G250 (left). M—prestained protein standard, HL-B1B2-N, and HL-B1B2-R—natural and recombinant *Hrr. lacusprofundi* DL18 archaella, HS-B1B2-N—natural archaella of *Hrr. saccharovororum*, HS-B1B2-R, HS-B1-R—recombinant *Hrr. saccharovororum* archaella isolated from *Hfx. volcanii* MT2; A1A2—*Hfx. volcanii* MT45 archaella (positive control), BSA—bovine serum albumin (negative control)



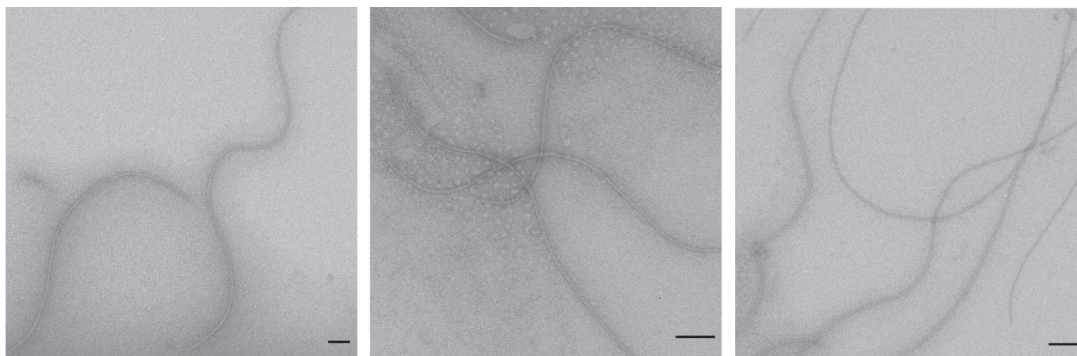
**FIGURE A7** Comparison of cell motility of *Hfx. volcanii* strains expressing *Hrr. lacusprofundi* and *Hrr. saccharovororum* archaeellin genes. Photographs of swarming plates were taken at different times after inoculation. Top: left—*Hfx. volcanii* MT2 transformed with pMT21 (HV-A1A2), pAS6 (HL-B1-R), pAS7 (HL-B2-R), and pAS5 (HL-B1B2-R), Mod-HV medium containing 0.5 mg/ml tryptophan, 0.24% agar, 37°C, 5 days; right—the same, without pMT21, 8 days. Bottom: left—*Hfx. volcanii* MT2 transformed with pMT21 (FlgA1FlgA2), pAS2 (HS-B1-R), pAS3 (HS-B2-R), and pAS1 (HS-B1B2-R), the abovementioned medium, 5 days; right—the same, without pMT21, 8 days



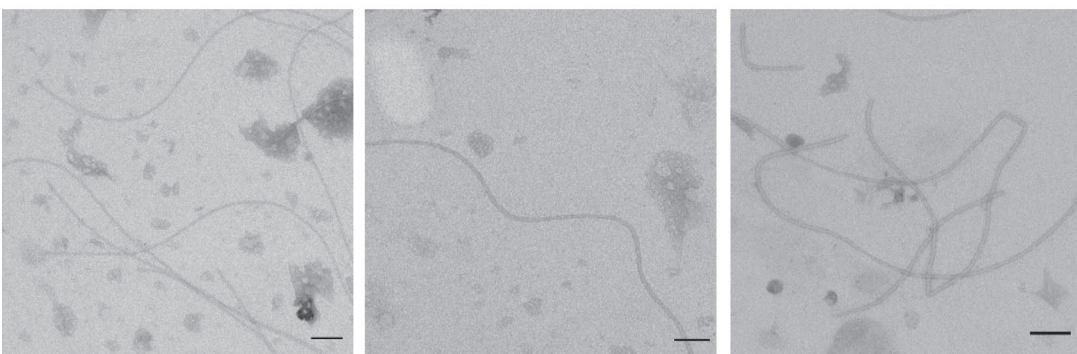
**FIGURE A8** Comparison of cell motility of haloarchaeal strains expressing *Hrr. saccharovororum* archaeellins. The swarming diameters were measured 120 hr after inoculation: *Hfx. volcanii* MT2 transformed with pMT21 (HV-A1A2), pAS1 (HS-B1B2-R), pAS2 (HS-B1-R), and pAS3 (HS-B2-R), Mod-HV medium containing 0.5 mg/ml tryptophan, 0.24% agar



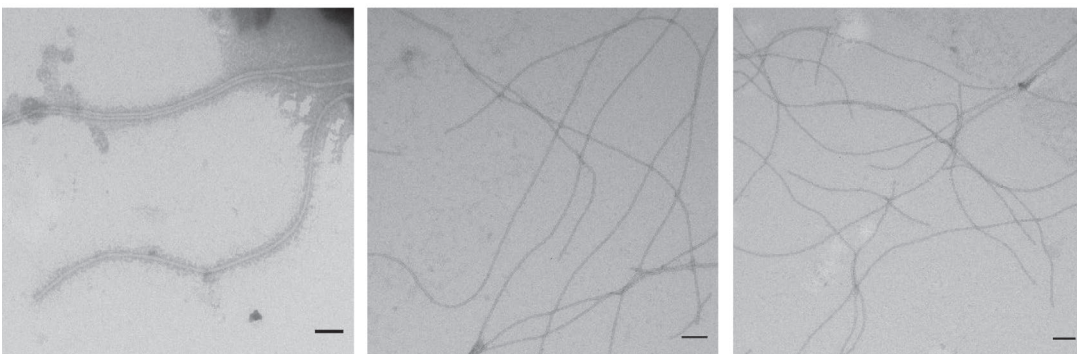
HS-B1B2-N



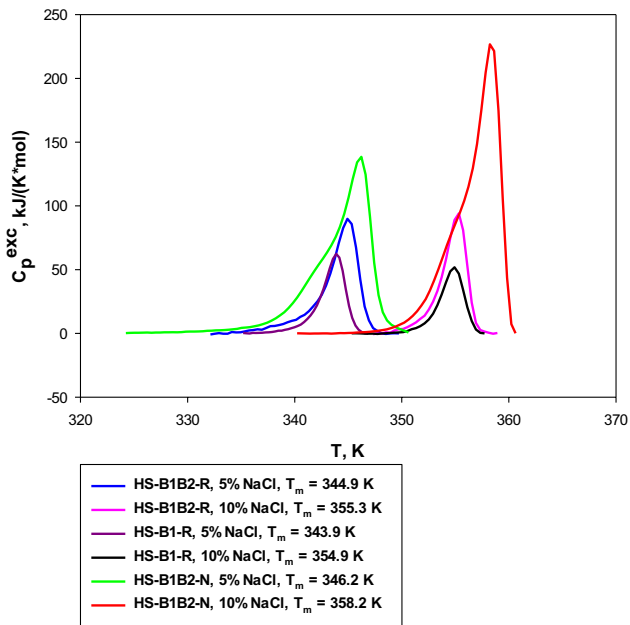
HS-B1B2-R



HS-B1-R

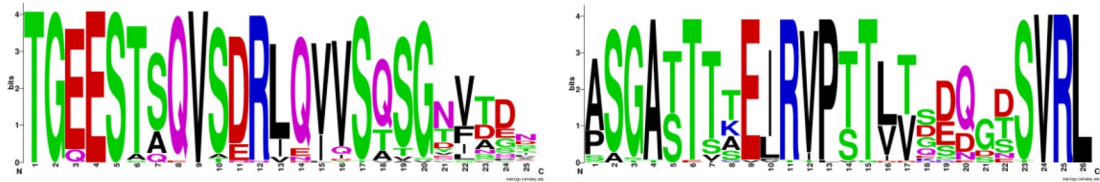


**FIGURE A9** Negatively stained (1% uranyl acetate) preparations of natural (HS-B1B2-N) and recombinant (HS-B1B2-R and HS-B1-R) archaellar filaments of *Hrr. saccharovororum* in 20% NaCl, 10 mM Na-phosphate, pH 8.0. Scale bar—100 nm

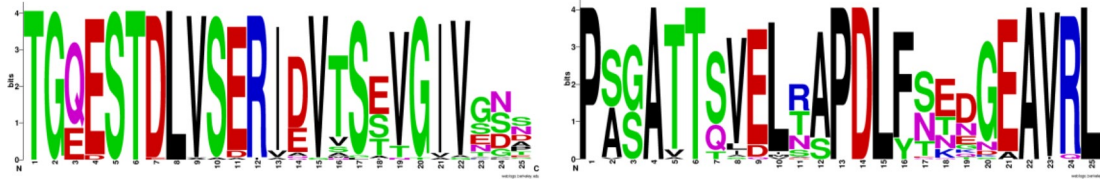


**FIGURE A10** Temperature dependence of excess heat capacity of *Hrr. saccharovorum* recombinant HS-B1B2-R and HS-B1-R archaeellar filaments in comparison with native (HS-B1B2-N) filaments at two salinities: 10% (1.7 M) and 5% (0.85 M) NaCl, 10 mM Na-phosphate, pH 8.0





FlaB1



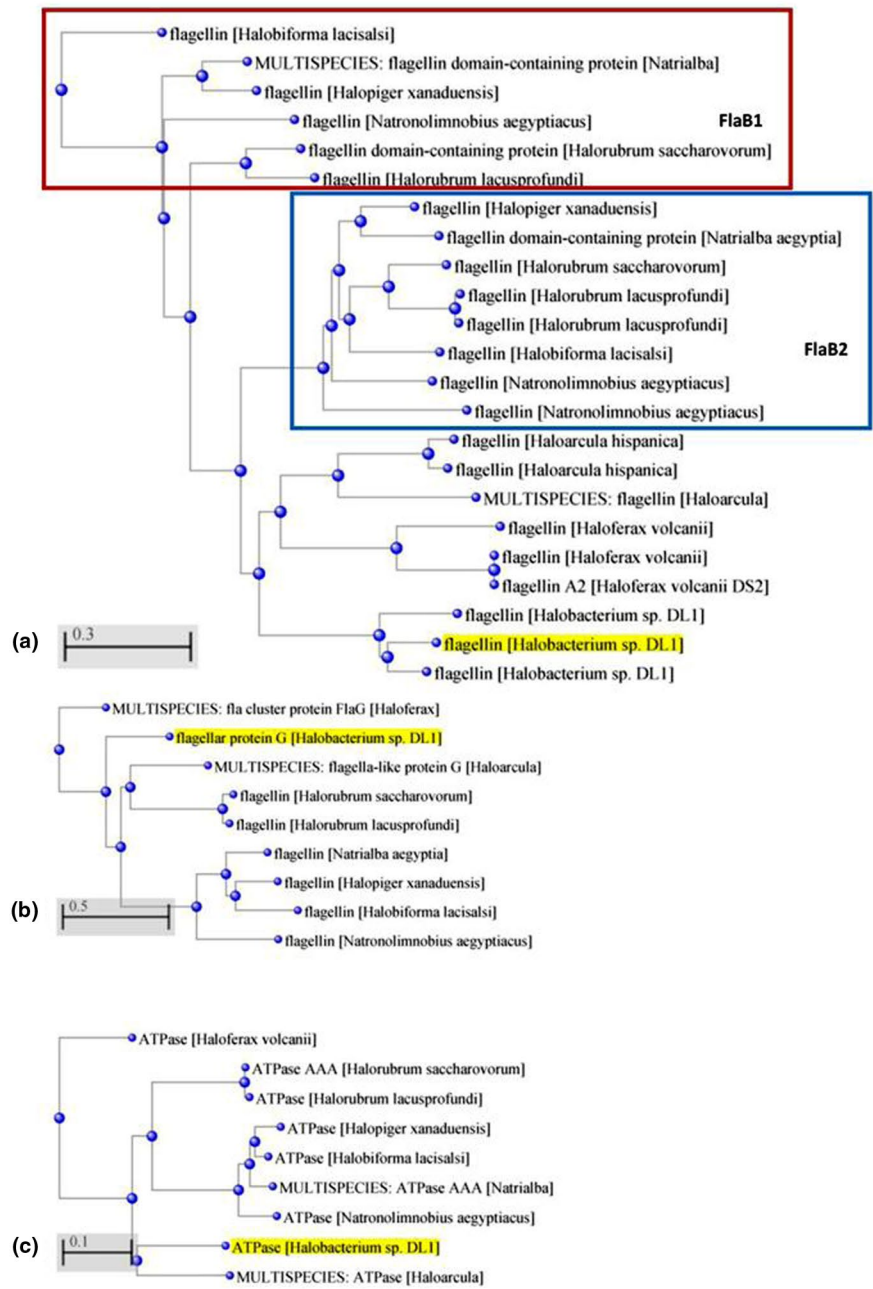
FlaB2

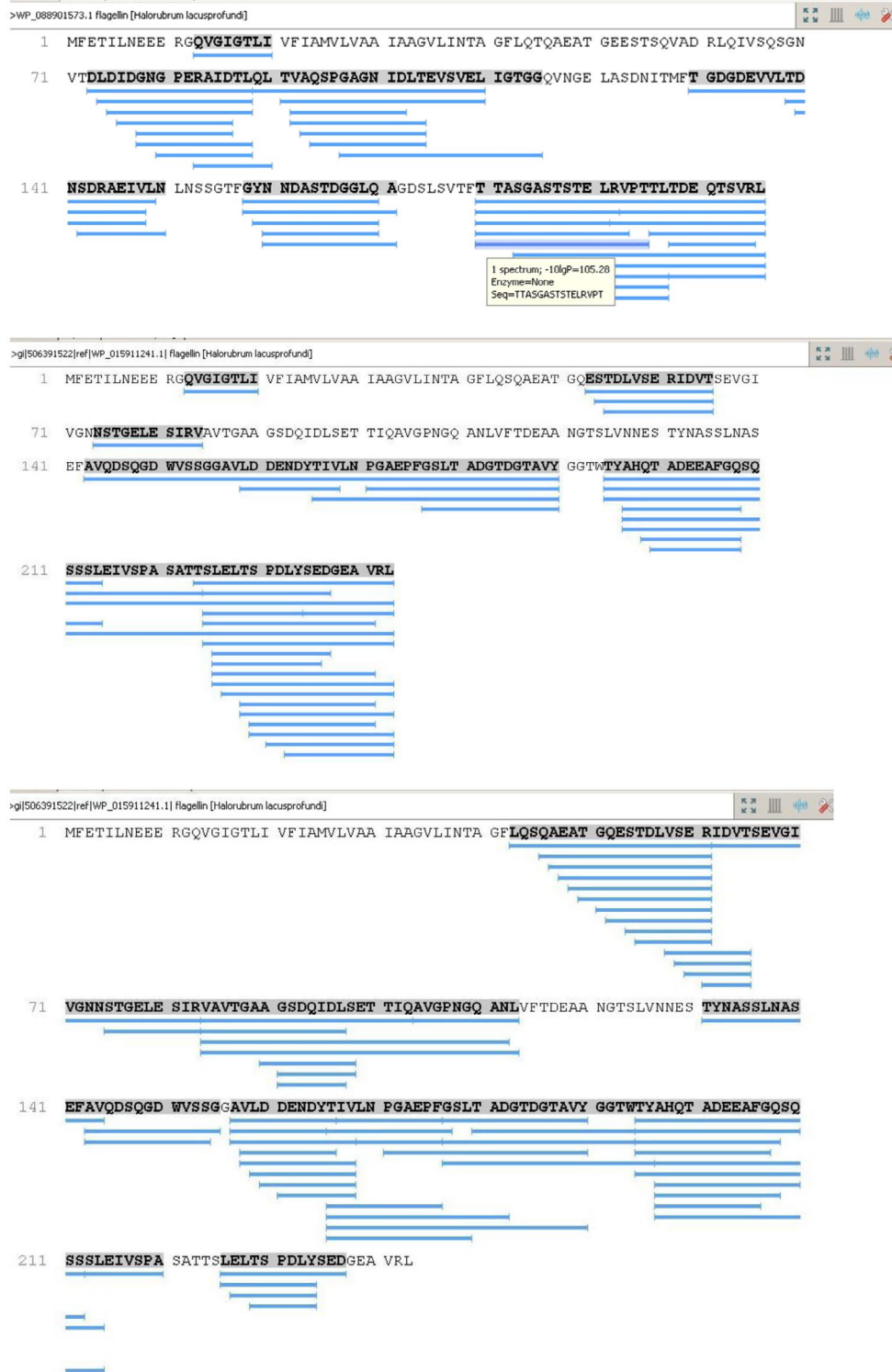
	54	55	56	57	58	59	60	61	62	63	64	65	66	66	67	68
FlaB1	S	T	S[A]	Q	V	S	D[E]	R	L[I]	Q	V[I]	V	S	Q[TA]	S	G
FlaB2	S	T	D	L	V	S	E[D]	R	I[V]	D[E]	V	T	S	E[ST]	V	G

	17	16	15	14	13	12	11	10	9	8	7	6	5	4	3	2	1
FlaB1	E	I[L]	R	V	P	T[S]	T	L[V]	T[V]	D[E]	Q[D]	G[D]	D[TS]	S	V	R	L
FlaB2	E	L	R[TN]	A[S]	P	D	L	F[Y]	S[N]	E[TN]	D[NE]	G	E	A	V	R	L

**FIGURE A11** Weblogo representation of the alignment of internal sequences of *Hrr. lacusprofundi* FlaB1 and FlaB2 archaeellins. In this representation, the overall height of a stack indicates the sequence conservation at that position, while the height of symbols within the stack indicates the relative frequency of each amino acid at that position (Crooks, Hon, Chandonia, & Brenner, 2004). Left column: Amino acid residues 50–75; 53 *Halorubrum* FlaB1 sequences and 55 *Halorubrum* FlaB2 sequences were used. Right column: C-terminal sequences (~25 amino acid residues from C-termini); 50 FlaB1 sequences and 53 FlaB2 sequences were used. The tables below show the corresponding characteristic signatures for FlaB1 and FlaB2. Conservative residues for both archaeellins are colored red, conservative residues different for FlaB1 and FlaB2 are blue, and slightly conservative residues are black

**FIGURE A12** Schematic phylogenetic trees obtained using the Standard Protein BLAST <https://blast.ncbi.nlm.nih.gov/Blast.cgi> for selected archaeellins (a), FlaG (b), and FlaH (c) of *Har. hispanica*, *Hbt. DL1*, *Hbf. lacisalsi*, *Hfx. volcanii*, *Hpg. xanaduensis*, *Hrr. lacusprofundi*, *Hrr. saccharovororum*, *Nab. Aegyptia*, and *Nln. aegyptiacus*





**FIGURE A13** Mass spectrometry analysis of most prominent proteins in archaella isolated from wild-type *Hrr. lacusprofundi* DL18 (top and middle) and ACAM 34 (bottom) strains. Protein coverage of archaellins B1 (WP\_088901573.1) and B2 (WP\_088901574.1/WP\_015911241.1) is shown. The full protein sequence is shown in gray. The unique peptides are depicted in blue



**FIGURE A14** Mass spectrometry analysis of most prominent proteins in archaella isolated from *Hfx. volcanii* MT2 transformed with pAS5 (HL-B1B2-R). Protein coverage of archaellins B1 (WP\_088901573.1) (top) and B2 (WP\_088901574.1/WP\_015911241.1) (bottom) is shown. The full protein sequence is shown in gray. The unique peptides are depicted in blue

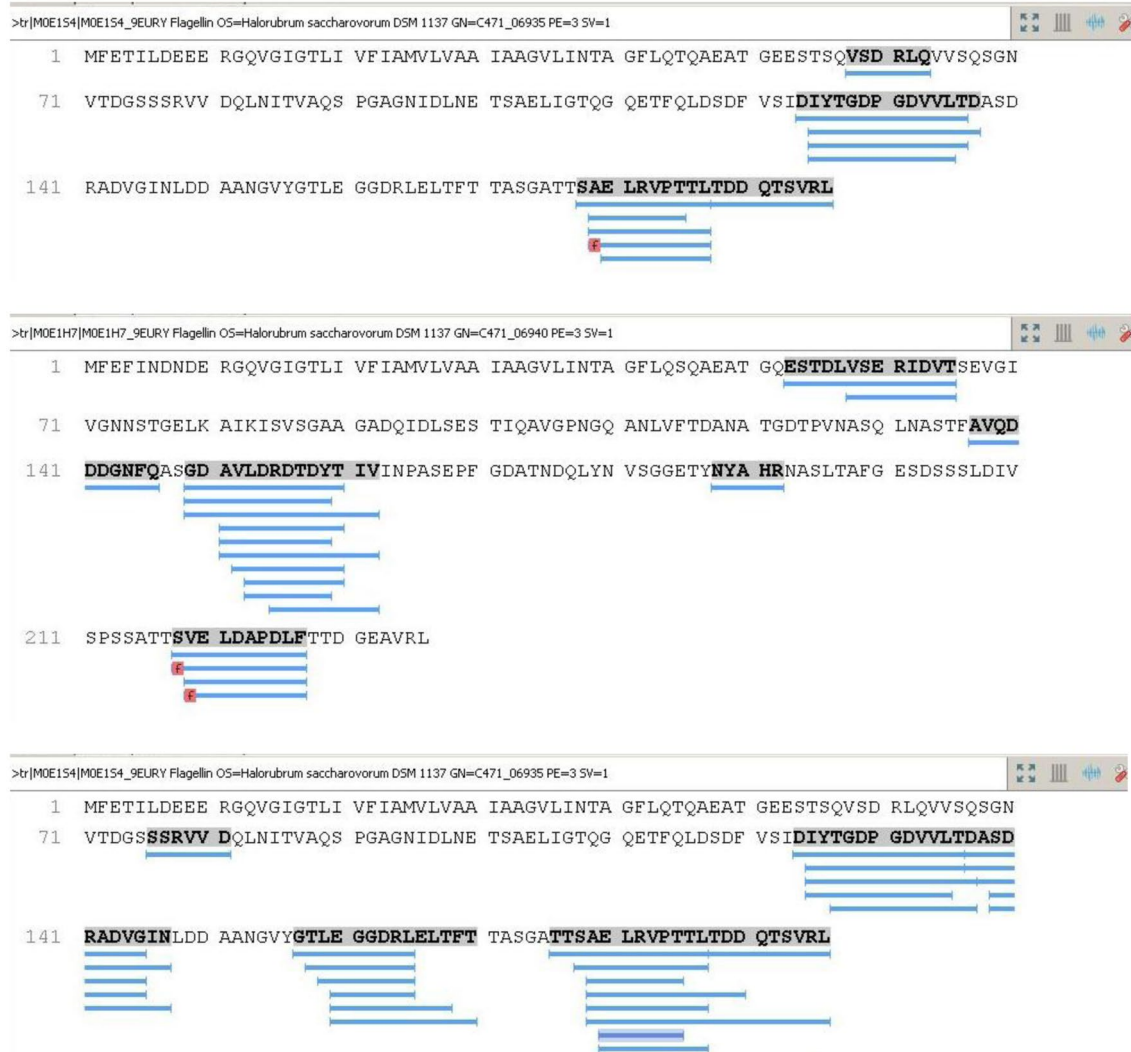


**FIGURE A15** Mass spectrometry analysis of most prominent proteins in archaella isolated from *Hfx. volcanii* MT2 transformed with pAS6 (HL-B1-R) (top) and pAS7 (HL-B2-R) (bottom). Protein coverage of archaellins B1 (WP\_088901573.1) and B2 (WP\_088901574.1/ WP\_015911241.1) is shown. The full protein sequence is shown in gray. The unique peptides are depicted in blue. The red square corresponds to the possible formylation





**FIGURE A16** Mass spectrometry analysis of most prominent proteins in archaella isolated from wild-type *Halorubrum saccharovororum* ATCC 29,252. Protein coverage of archaellins B1 (WP\_004047439.1) (top) and B2 (WP\_004047440.1) (bottom) is shown. The full protein sequence is shown in gray. The unique peptides are depicted in blue



**FIGURE A17** Mass spectrometry analysis of most prominent proteins in archaella isolated from *Hfx. volcanii* MT2 transformed with pAS1 (HS-B1B2-R) (top and middle) and pAS2 (HS-B1-R) (bottom). Protein coverage of archaellins B1 (WP\_004047439.1) and B2 (WP\_004047440.1) is shown. The full protein sequence is shown in gray. The unique peptides are depicted in blue